Data Provided: None

DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING

Spring Semester 2009-2010 (2 hours)

Multimedia Systems 1

Solutions

1. An audio signal with 20 kHz bandwidth is sampled at the Nyquist rate into 5 distinct voltage levels as follows: {-2, -1, 0, 1, 2}. The corresponding probabilities of occurrence for these symbols are {0.08, 0.1, 0.55, 0.12, 0.15}, respectively.

Answer the following questions based on the above scenario showing all steps involved in your computations.

a.

p	0.0800	0.1000	0.5500	0.1200	0.1500	
$-\log_2(p)$	3.6439	3.3219	0.8625	3.0589	2.7370	
$-p.\log_2(p)$	0.2915	0.3322	0.4744	0.3671	0.4105	=1.88 bits/sample

(3)

b.

	p											Code	
0	0.55							1				1	
2	0.15	1	0.27			1	0.45	0				011	
1	0.12	0										010	
-1	0.1	1	0.18			0						001	
-2	0.08	0										000	**(6)**

c.

p	0.08	0.1	0.55	0.12	0.15	
n	3	3	1	3	3	
p.n	0.24	0.3	0.55	0.36	0.45	=1.9 bits/sample

efficiency = 1.88/1.9 = 99%

(4)

d. The smaller code (1) does not appear as the prefix of any of the longer codes (000, 001, 010 and 011) **(2)**

e. Fixed length code length is 3 bits top represent 5 symbols

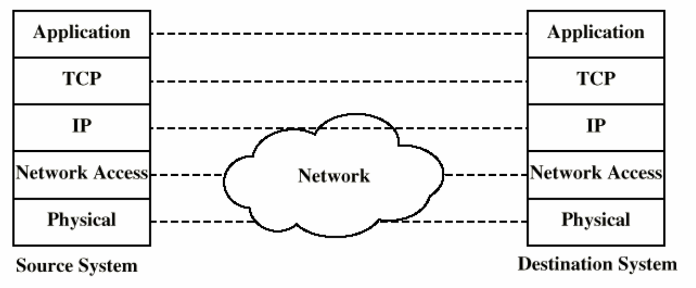
Compression ratio= 3 : 1.9 = 1.58: 1 **(2)**

f. Sampling rate = 40 kHz

Data rate = 40k x 1.9 bits/s

Audio length = (128 M x 8) / (40k x 1.9 x 60) = 235.5 min **(3)**

2. **a.**



(i) Network layer

Concerned with handling packets along individual links, access to local network, the physical aspects of transmitting and receiving bit streams.

(ii) Internet Protocol (IP) Layer

Concerned with locating destination node and overall routing of packets across networks

(iii) Transport Control Protocol (TCP) Layer

Concerned with the reliable transmission of complete message (collection of packets) across networks from specified source and destination

(iv) Application Layer

Concerned with receiving and delivering messages between remote applications **(6)**

b. Propagation time p= 210k / (3×10^8) = 0.7 mili sec

Transmit time = (20 M x 8) / (120 M) = 1.33 sec

Total time = 1.334 sec = 1.33 sec

(3)

c. Circuit switching: carry data on dedicated channel (circuit)

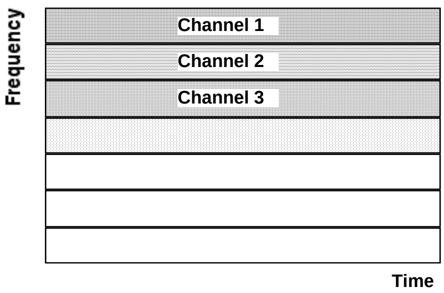
e.g., original telephone network

Packet switching: take one hop at a time through switches (routers). Can use any available link. Needs to know the destination address

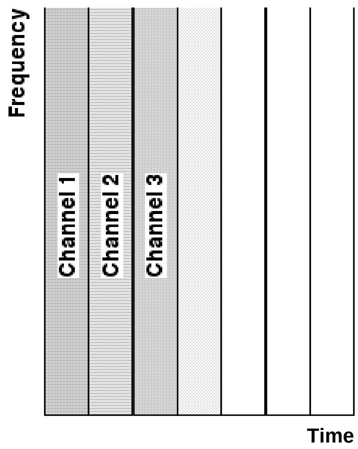
e.g., Internet **(4)**

d. Frequency Division Multiplexing -- Dedicated band of frequencies for each channel - available for all time (e.g., traditional analogue radio transmissions)

Time Division Multiplexing -- Dedicated period of time for each channel - all share same band of frequencies (e.g., digital phone network, terrestrial digital tv) **(5)**

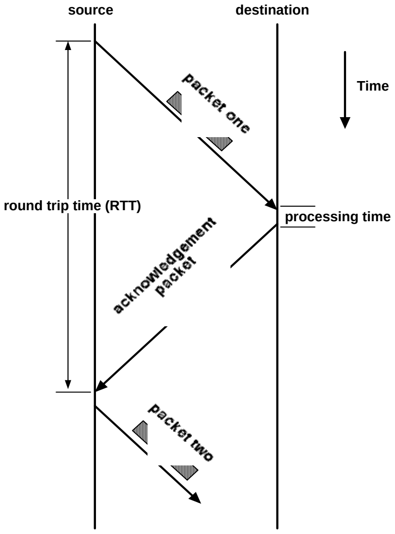

Frequency
Channel 1
Channel 2
Channel 3
Time

FDM


Frequency
Channel 1
Channel 2
Channel 3
Time

TDM

e.


source
destination
Time
packet one
round trip time (RTT)
processing time
acknowledgement packet
packet two

(2)

3. **a.** State why data (such as music or video) compression is possible and why is it often necessary?

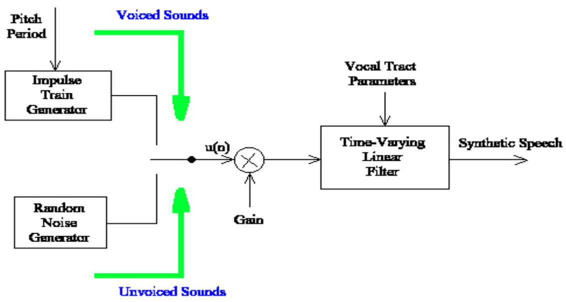
-because data representation is redundant in terms of correlations in the neighbouring samples and due to fixed length codes

- also due to irrelevant data with respect to human perception

- therefore it is possible to compress by removing redundant and irrelevant data

-It is necessary due to limitations in storage and transmission capacity **(4)**

b.



Speech Synthesis model based on LPC model

Voiced sounds: sounds are produced by forcing the air stream through vocal cords while vocal cords are forced to open and close rapidly to produce a series of periodic puffs which has a fundamental frequency - same as the vocal cord vibration frequency. The Formant frequencies of a vowel are determined by the parameters of the vocal tract configuration such as the length of vocal tract, the position of tongue, and the shape of lips. The Formant frequencies for vowels and the vocal tract parameters as sown in the figure are used.

Unvoiced sounds: produced without vibration of the vocal cord. Can be fricatives or plosives. Can be characterized by noise. Random noise generator and noise variance are used as in the figure to synthesize these sounds. **(5)**

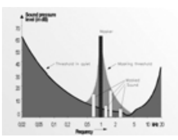
c.

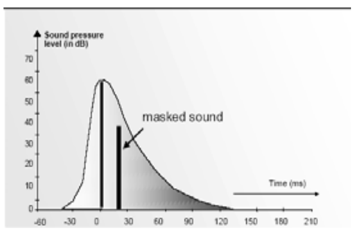
- sampling rate reduction causes the sounds very telephonic as only frequencies up to 5.5 kHz is represented.

- quantisation bits reduction causes distortion. Thus, signal to noise ratio is lower. Noise can be heard in sounds rather than in quiet passages. Less loud sounds are more distorted. **(3)**

d. Temporal masking – For a strong signal, there is a short time afterwards while we cannot hear a quieter sound

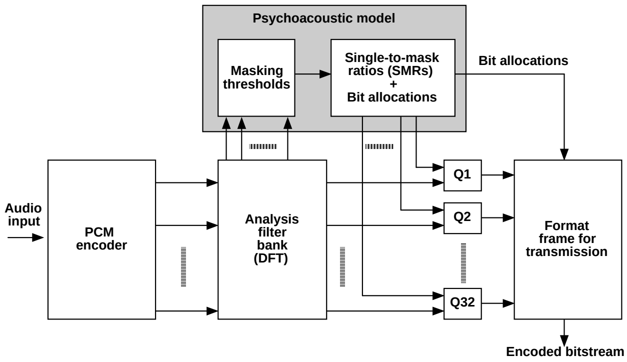
Frequency masking - When ear hears a signal at one frequency it may reduce the level of sensitivity of another signal at a similar frequency





(4)

e.



DFT – Decompose the data into frequency sub bands

Psychoacoustic model – Analyse the frequency sub-bands and frequency masking for bit allocation.

Q1- Q32 represent quantisation and Huffman coding for each of the frequency subbands

Finally format frame module represents creating a frame (inserting the frame header) to output the encoded bitstream. **(4)**

4. **a.**

Lossless image compression:

- Exact recovery of pixel values.
- Mostly DPCM and Entropy coding based compression. No quantisation.
- Only small compression ratios. 1.2:1 to 2:1.
- Example:- JPEG-LS, TIFF (LZW)
- Application:- Medical imaging, remote sensing imaging

Lossy image compression:

- Not exact recovery of pixel values.
- Quantisation (mostly based on HVS models) is used
- Large compression ratios. Up to 20:1(for acceptable visual quality)
- Example:- JPEG, JPEG 2000

Application:- web imaging, digital cameras **(4)**

b. In R-G-B format, each colour plane contains the same amount of data and would have the same data rate and bits per sample. But in Y Cb Cr, we can exploit the human eye's tolerance to reduced resolution in chrominance layers. Therefore they can de down sampled and the data rates can be reduced.

Y Cb Cr format also useful in transition from back and white tv transmissions to colour tv transmissions. **(3)**

c. Artificial contours are visible in homogeneous regions when fewer bits are used. But not in high detailed regions, which are less affected. **(3)**

d. W:H= 16:9 H=810; Therefore, W= 810x16/9 = 1440

Y resolution: 1440 x 810

Cb and Cr resolution = 720 x 405

Frame rate = 25

Bit rate= (Y resolution + 2xCb resolution) x colour depth x frame rate

= (1440 x 810 + 2 x 720 x 405)x8 x 25

= 349920000 bits/second

to store a 1-hour HDTV programme:

File size = Bit rate x 60 x60 x 1

= 1.26 x 10^{12} bits

= 146.7 G Bytes. **(5)**

e. 1 x I frames

1 x P frames

4 x B frames

6 frames in total in a GOP

This is compressed to : 1/10 + 1/40 + 4/80 frames = 0.175 frames

Compression ratio is 6 : 0.175 = 34.3 : 1

Uncompressed data rate = 349920000 bits/second

Compressed data rate = 349920000/34.3 = 10.2 Mbits/sec **(5)**

GCKA