Data Provided: None

DEPARTMENT OF ELECTRONIC AND ELECTRICAL ENGINEERING

Spring Semester 2008-2009 (2 hours)

Multimedia Systems 1 SOLUTIONS

1.

a.

Gray Levels	0	1	2	3	4	5		
Frequencies of occurrence	600	1000	1200	4000	2000	1200		
P	0.06	0.1	0.12	0.4	0.2	0.12		
$-\log_2(P)*P$	0.24	0.33	0.37	0.53	0.46	0.37		

Shannon's Entropy = 2.3 bits per pixel **(3)**

b.

3	0.4									1		1	
4	0.2					1	0.36	1	0.6	0		011	
2	0.12			1	0.24			0				001	
5	0.12			0								000	
1	0.1	1	0.16			0						0101	
0	0.06	0										0100	

(6)

c. Average code length

P	0.06	0.1	0.12	0.4	0.2	0.12
n	4	4	3	1	3	3

$$\sum_{i=0}^{5} n_i P_i = (4\text{x}0.16)+3\text{x}0.44+1\text{x}0.4 = 2.36 \text{ bits per pixel}$$

Efficiency = 2.3 /2.36 = 97.5% **(3)**

d. Compression ratio = 3/2.33 = 1.27:1 **(2)**

e. Data rate = 2.36 bits per pixel

Total data = 2.36 x 100 x 100 =23600

Total time = propagation time + transmit time

= distance /velocity + datasize/bandwidth

$$= \frac{150 \times 10^3}{3 \times 10^8} + \frac{23600}{1024 \times 1024} = 23 \text{ milliseconds.}$$

(4)

f. Advantage – low data rate, close to the theoretical minimum, high efficiency

Drawbacks – errors can propagate, need to know the probability of occurrence **(2)**

2. **a.** (Any 3 of the following)

- Common format for all signal/data types
- Can use full capabilities of digital computers
- Common storage format (memory chips, hard discs, etc.)
- Possible to reconstruct perfectly if effected by noise
- Can include additional codes to detect and correct errors
- Can modify the time to transmit/playback signals

Generally cheaper to transmit - can combine (**multiplex**) different channels **(3)**

b. (i) 10 kHz

(ii) Sampling freq = 20 kHz

Bit rate = 240 kbps

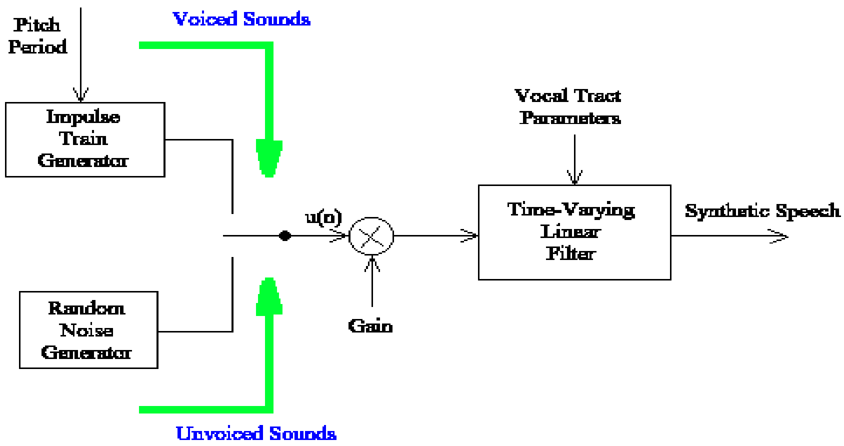
Therefore, Bits per sample = 240/20 = 12 bits

(i) Quantisation interval: $Q = \dfrac{V_{\max} - V_{\min} + 1}{2^N} = \dfrac{(2500 + 1500) + 1}{2^{12}}$

= 0.98 mV **(5)**

c. Signal to noise ratio is lower. Noise can be heard in sounds rather than in quiet passages. Less loud sounds are more distorted. **(2)**

d.



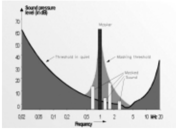
Speech Synthesis model based on LPC model

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)

e. (i) Frequency masking - When the human ear hears a signal at one frequency it may reduce the level of sensitivity of another signal at a similar frequency.



(ii) Decompose the audio signal into frequency sub bands using the MDCT. Then frequency masking is exploited in the psychoacoustic modelling block to discard sub bands with inaudible audio components due to masking effect of the neighbouring sub bands. Finally the remaining components are quantized and entropy encoded. **(5)**

3. **a.** Unicast – source message to single destination node (e.g., private phone call, personal letter, email)

Broadcast – source message to all nodes on a network (e.g., television channel)

Multicast – source message to subset of nodes on a network (e.g., junk mail, spam email) **(3)**

b. 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) **(3)**

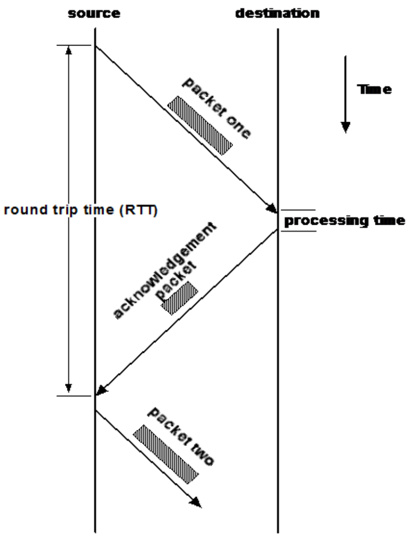
c. Circuit switching: carry data on dedicated channel (circuit)

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.



Round trip time is the total time from sending a packet till receiving the acknowledgement packet. This is usually twice the propagation time.

(3)

e. (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

(i) Concerned with receiving and delivering messages between remote applications **(4)**

f. UDP-IP focuses only on best-effort delivery. UDP provides information of source and destination ports to determine the application programmes. On the other hand TCP-IP deals with reliable delivery.

UDP-IP application ---- live streaming of video and audio.. **(3)**

4. a. Data (such as music or video) compression is possible due to:

- Redundancy in source data (inter-pixel redundancy due to correlation and coding redundancy due to methods used in representation)
- Irrelevant data due to limitations of the human visual and hearing system

It is often necessary due to

- Uncompressed data consists of a lot of bits. But limited bandwidth in networks and limited space in storage devices. Therefore, compression is necessary **(4)**

b.

I	B	B	P	B	B	I	B	B	P
1	2	3	4	5	6	7	8	9	10

Figure 1: Frame ordering in video coding.

Figure 1 shows the frame number and the coding type of frames, arranged in display order, in a video sequence.

(i) 6

(ii) 1 4 2 3 7 5 6 10 8 9

(iii) 2 **(3)**

c. Frame 7 is an I frame. That means an Intra-only frame. Such frames are coded as a still image. No reference (or prediction) with respect to other codes is made.

The frame is transformed, quantised and entropy encoded.

They are quantised only moderately as they are used as reference frames.

The decoded I frame is kept in the buffer **(3)**

d. - The number of horizontal TV lines: 2160

- Aspect ratio (width : height): 16:9

- Colour format: YCbCr 4:2:0

- Colour depth: 8 bits per each colour component sample

- Frame rate: 50 frames per second (Non-interlaced)

(i) Compute the bit rate required to transmit an uncompressed Ultra-HDTV video sequence with above specifications.

W:H= 16:9 H=2160; Therefore, W= 2160x16/9 = 3840

Y resolution: 2160 x 3840

Cb and Cr resolution = 1080 x 1920 **(9)**

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

= 2160 x 3840+ 2x1080 x 1920) x 8 x 50

= 4,976,640,000 bits/second

= 4.635 G bits/second

(ii) The overall compression ratio:

In the GOP there are 1-I, 1-P and 4-B frames.

Considering the compression ratios,

Average frame number after compression = 1/30 + 1/90 + 4/180

= 12/180 frames

compression ratio 6/(12/180) = 90:1

(iii) What would be the bit rate of the compressed video?

(4.635 / 90) G bits per second

= 52.73 M bits per second

e. H.264 or MPEG-4 **(1)**

GCKA