Multimedia Exam 2015

1a.

Classes of modulation:

Amplitude Modulation – wah wah

Frequency Modulation – audio synthesis

Phase Modulation – vibrato

b.

Classes of audio effects that use fourier based signal processing methods

Equalisation/Tone Filtering

Reverb

Spectral Enhancers

ci.

Remove a constant mains hum from an audio signal

use a bandpass filter with limits 50-60hz

compute fourier transform

low pass filter at 60hz

high pass filter at 50hz

inverse fourier transform

ii.

cross synthesis

short time fourier transform

assume two signals: modulator and carrier

compute spectral envelope of each time frame on each signal

divide spectrum of each carrier frame by its own spectral envelope (flattens it)

multiply flattened spectral frame by envelope of corresponding modulator frame, replacing carriers envelope with modulators envelope

iii.

pitch synchronous and asynchronous granular synthesis

application of short time fourier transform similar to phase vocoder

pitch synchronous granular synthesis assumes overlapping grains of sounds

asynchronous granular synthesis allows for grains to be scattered in a statistically controlled manner

essentially implement stft with very small windows, allow for overlapping of windows

manioulate grains in a variety of frequency processes ways e.g. psgc, detect pitch of grain, spectral analysis

resynthesize via inverse stft

2a.

Eye is sensitive to colour and intensity

Retina has neurons on which light is focused, each neuron is either a rod or cone

Rods are not sensitive to colour, they sense intensity in monochrome

3 types of cones, long wavelengths – red/yellow colours, medium wavelength peak at green, short wavelength blue

Each respond differently, non linear and not equal for rgb differently to various frequency of light

Compression uses the fact that intensity can be modelled in high resolution and colour in lower resolution and non-linearly wrt colour sensitivity

b.

cmyk colour model uses cyan magenta yellow and black as primary componenets, commonly used in printing because colour pigments on the paper absorb certain colours so a subtractive model is suitable, black is used to produce a darker black than just mixing cmy

convert rgb to cmy 111 – rgb = 111 – 0.4 0.5 .03 = 0.6 0.5 0.7 k = min(cmy) = 0.5 c = 0.1 m =0 y =0.2

c.

range of colours produces half tone printing, standard cmyk is limited, printing a dark colour the printer will saturate an area with more coloured ink dots, fewer ink dots create a light colour, this can produce harsh effects especially with the wider range of colours in photographs

including lighter cm inks in addition to cmyk gives a n better distribution of light colour saturation for lighter colours with less half toning, given a more even saturation overall

yellow is already perceived as a light colour by humans and the difference in a light yellow would be negligible

d.

chroma sub sampling is a method that stores colour information at a lower resolution than intensity information

reduces badnwith for colour detail with almost no perceivable visual difference

appropriate because human visual system is less sensitive to colour variations than brightness/intensity

e.

sub sampling 4:2:2 first column and 3rd column, 4:1:1 1st column 4:2:0 every other value 1,1 1,3 3,1 3,3

3a

Lossy methods can produce much smaller files than any lossless method but the loss may be imperceptible to a human

b.

huffman coding assumes an integer (k) of bits for each symbol so k is never less than 1, arithmetic coding can represent fractional number of bits and can achieve better compression ratios

ci.

shannon fano codes, order symbols by size c - 00, a-01, d-10, b-110, e-111

entropy sum of Pi log 1/P

d.

w = nil

while (read a character k) {

if wk exists in dictionary

w = wk

else{

add wk to the dictionary

output code for w

w=k

}

}

Starts at 256 0-255 ascii codes

4a.

Jpeg quantisation

Eyes is most sensitive to low frequencies (upper left corner), less sensitive to high frequencies (bottom right corner) hence we quantise more aggressively in the high frequency range

This is a result of perceptual experiments, the experiments determined perception thresholds in these frequency bands

The quality would be reduced by doubling values as this makes quantisation more aggressive

b.

major steps in jpeg coding

colour space transform and sub sampling

dct (discrete cosing transform)

quantisation

zig zag scan

discrete pulse code modulation (dpcm) on dc component

run length encoding on ac components

entropy coding – Huffman or arithmetic

c.

macro block of NxN pixels, algorithm search range +- R pixels in both directions, estimate how many operations to perform motion compensation for a frame M1xM2 pixels

for each block (2R+1)^2 candidate matching blocks

computing MAD for one candidate bloxk requires O(N^2) operations

to estimate motion vector for one block requires O((2R+1)^2 x N^2) operations, there are M1M2/N^2 blocks in the frame, total operations for entire frame O((2R+1)^2xM1xM2)

d.

Key difference between I P and B frames in mpeg 2, advantages and disadvantages of b frames, most and least bits to encode, why insert I frames

I frame – basic reference frame for each group of pictures, essentially jpeg compressed image

P frame – coded forward difference from wrt last I or p frame

B frame coded backward difference frame wrt last I or p frame

B frame advantages – improve code efficiency/quality, improvement In the case of moving objects that reveal hidden areas, better error propogation as b frames are not used to predict future frames

Disadvantages – frame reconstruction requires memory buffers within encoder and decoder to accommodate anchor frames, potentially more delays for online applications

B frames are the most compact and I frames are typically the largest

Inserting I frames frequently, differences between frames get too large, hard to track fast blocks so need to restart card with a new I frame.