UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2016

IS52020B
Perception & Multimedia Computing

Duration: 3 hours

Date and time:

This paper is in two parts: part A and part B. You should answer ALL questions from part A and THREE questions from part B. Part A carries 40 marks, and each question from part B carries 20 marks. The marks for each part of a question are indicated at the end of the part in [.] *brackets.*

There are 100 marks available on this paper.

This is a practical examination; each answer requiring code or other computational material should be named according to question number, part and sub-part: for example, `Q5_b_2.pde` *for a Processing sketch in answer to part (b) sub-part (ii) of question 5. Save your answer to the exam submission folder. You are responsible for ensuring that your answers have been saved in the correct location.*

THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM

Part A

Multiple Choice

Question 1

Answer all questions; **more than one answer to each question might be correct, and all must be given for full credit**. [10]

(a) In a normal eye, cone cells have how many pigments?

i. two
ii. three
iii. four
iv. five
v. something else

(b) How many primary colours does the opponent model of colour perception suggest that there are?

i. two
ii. three
iii. four
iv. five
v. something else

(c) On a typical consumer-level digital screen, the image updates at a rate of approximately what?

i. 3Hz
ii. 12Hz
iii. 60Hz
iv. 140Hz
v. something else

(d) Which of the following is *not* a device-dependent colour space?

i. RGB
ii. HSB
iii. Pantone
iv. CIE xyY
v. none of the above

(e) Which of the following is a good choice for a colour-picking widget?

i. RGB
ii. HSB
iii. CIE xyY
iv. CMYK
v. none of the above

(f) In descending order of brightness, the sRGB primaries are

i. green, red, blue
ii. red, green, blue
iii. green, blue, red
iv. they are all the same brightness
v. none of the above

(g) Physical modelling is a suitable technique for

i. animating lifelike characters
ii. simulating simple physical systems
iii. choreographing character interactions
iv. visualising particle simulations
v. none of the above

(h) The ear and brain locate sources for *low*-frequency sound waves through

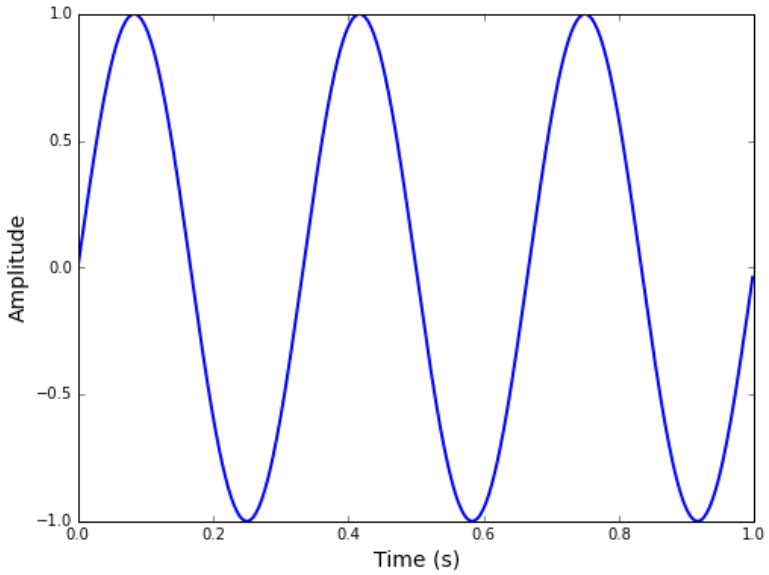
i. amplitude difference
ii. wavelength difference
iii. timbre difference
iv. phase difference
v. none of the above

Question 2

Answer all questions; **more than one answer to each question might be correct, and all must be given for full credit**. [10]

(a) What is the frequency of the waveform below?

i. 2Hz
ii. 3Hz
iii. 5Hz
iv. 6Hz
v. none of the above



(b) What is the amplitude of the sine wave with the equation below?

$y = 0.3\sin(2\pi 100t + pi/2)$

i. 0.3
ii. 2
iii. 100
iv. $\pi/2$
v. none of the above

(c) A note is played on a pitched instrument, and its fundamental frequency is 200Hz. Which of the following frequencies are also likely to be present in the sound? Select all that apply.

i. 400 Hz
ii. 600 Hz
iii. 800 Hz
iv. 1000 Hz
v. 1200 Hz

(d) Which of the following statements are true about audio perception? Select all that apply.

i. Increasing the frequency of a sine wave will increase its pitch. (Assuming the sound is audible.)
ii. Increasing the amplitude of a sine wave will increase its volume. (Assuming the sound is audible.)
iii. Changing the phase of a sine wave will change its pitch. (Assuming the sound is audible.)
iv. Changing the shape of a waveform will change its timbre/tone colour. (Assuming the sound is audible.)
v. None of the above

(e) The process of measuring an analog signal at discrete, evenly-spaced points in time is called _____.

i. The Fast Fourier Transform (FFT)
ii. The Inverse Fast Fourier Transform (IFFT)
iii. Quantizing
iv. Sampling
v. None of the above

Question 3

Answer all questions; **more than one answer to each question might be correct, and all must be given for full credit**. [10]

(a) Which of the following statements are true about the Fourier Transform? Select all that apply.

i. Taking the Fourier transform of 1024 audio samples will result in a spectrum with 1024 frequency bins, only the first 512 of which are relevant

ii. The Fast Fourier Transform is fastest when applied to a number of samples which is a power of 3.

iii. The Fourier Transform is only relevant to audio signals, not to images or video.

iv. When plotting magnitude spectrum output by applying the FFT to an audio signal, frequency will appear on the x-axis of the plot.

v. When plotting magnitude spectrum output by applying the FFT to an audio signal, time will appear on the x-axis of the plot.

(b) What will be the effect of the image kernel below?

i. Gaussian blur

ii. Edge detector

iii. Motion blur

iv. Emboss

v. None of the above

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(c) Which of the following would probably be a good distance metric to use to compare two sounds in a sound effect database? Select all that apply.

i. Levenshtein

ii. Euclidean

iii. Hamming

iv. Kilometres

v. None of the above

(d) Which of the following statements is true? Select all that apply.

i. A multimedia fingerprint is a number
ii. A good multimedia fingerprint system will produce the same fingerprint for the same piece of media (e.g., the same photo, even if it is compressed or saved in a different file format).
iii. A good multimedia fingerprint system will produce a similar fingerprint for similar pieces of media (e.g., two songs by the same band)
iv. None of the above

(e) Which of the following audio features would be most useful if you wanted to build a search engine to find sounds with a similar timbre/tone colour? Select the **one** best answer.

i. RMS
ii. Spectral centroid
iii. Zero-crossing rate
iv. Fundamental frequency
v. Pitch histogram

Question 4

Answer all questions; one and ONLY one answer to each question is correct. [10]

(a) When considering reflection from 3D objects, which statement is true about diffuse and specular reflections?

i. Diffuse reflection is viewing angle dependent, specular is not.
ii. Specular reflection is viewing angle dependent, diffuse is not.
iii. Both diffuse and specular reflection are viewing angle dependent.
iv. None of diffuse and specular reflection is viewing angle dependent.

(b) Vector maths: for unit vector x (1,0) and y (0,1), we have:
a = x.dot(x)
b = x.dot(y)
Which one of the following statements returns true?

i. a == 1 && b == 1
ii. a == 0 && b == 0
iii. a == 1 && b == 0
iv. a == 0 && b == 1

(c) Vector maths: given two vectors, A(-10,10), B(1,0), what is the projection vector of A onto B?

i. 10
ii. -10
iii. (10,0)
iv. (-10,0)

(d) Which one of the following statements is true?

i. Parallel projection is used by architects and engineers to create drawings that preserve scale and shape. Perspective projection is similar to human visual system.
ii. Perspective projection is used by architects and engineers to create drawings that preserve scale and shape. Parallel projection is similar to human visual system
iii. Both parallel and perspective projections are used by architects and engineers to create drawings that preserve scale and shape
iv. None of the above.

(e) Which one of the following statements is true?

i. Ray casting uses object-order rendering, rasterisation uses image-order rendering.
ii. Ray casting uses image-order rendering, rasterisation uses object-order rendering
iii. Both ray casting and rasterisation use object-order rendering
iv. None of the above

Part B

Answer three questions

Question 5 Animation

(a) Describe the flipbook animation technique, and name the phenomenon which leads us to perceive moving image in a flipbook animation. [4]

(b) Describe how keyframing is used in computer-based animation, giving an example of a scenario where keyframing would be an appropriate technique for the animation. [5]

(c) A lead animator specifies key frames for an object as follows:

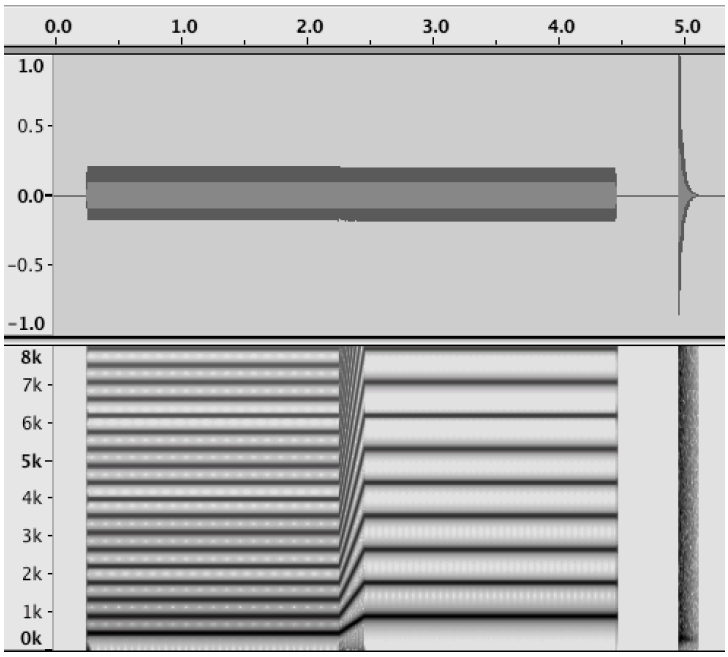
key frame time	x	y
0s	100	100
1s	60	120
3s	20	100

Using linear interpolation, calculate the positions of the object in the in-between frames, given that the frame rate is 2Hz. [4]

(d) Write a *Processing* sketch to visualise the above animation. You may if you choose use the 2Hz frame rate, or an appropriate frame rate of your choice. [7]

Question 6 Sound and the FFT

(a) The diagram below shows the waveform (top) and spectrogram (bottom) for the same audio signal. Time is on the x-axis, labeled in seconds on the top. Describe as precisely as possible what you will hear if you listen to this sound. For instance, you might comment on pitch, volume, tone quality/timbre, possible sound source(s), and how these all change over time. [6]



(b) Describe how humans are able to identify the location of a sound we hear. [3]

(c) A waveform has the following equation:

$y = 0.1\sin(2\pi 100t) + \sin(2\pi 300t) + 0.5\sin(2\pi 400t)$

It is sampled at 1000Hz. Sketch its spectrum, with frequency on the x-axis and magnitude on the y-axis. (Do label the x-axis with specific frequencies, but don't worry about the exact values on the y-axes; just try to get the relative magnitudes approximately correct.) [5]

Specifically, it should have peaks at 100/300/400Hz [3], with heights of small/high/medium [1]. X-axis should go from 0 to 500[1] (or it should go from 0 to 1000, but spectrum should be mirrored in that case; -[0.5] if it's not).

(d) Aliasing

i. What is aliasing, and why does it occur? [3]

ii. Describe one of the possible perceptual consequences of aliasing for audio, image, or video (pick one). [3]

Question 7 Multimedia signals, analysis, and effects

(a) Describe one possible use of a low-pass filter for audio, image, or sensor data. [4]

(b) Impulse Response

i. What is an impulse response? [3]

ii. Why is it useful to know the impulse response of a system? [3]

(c) Rank the following from smallest file size to largest file size. Defend your answer. [4]
A: A WAV file of a song
B: The same WAV file, converted to MP3
C: A .zip file of the same WAV file
D: The same WAV file, converted to FLAC

(d) You are building a search engine for paint colours that allows people to search for images with similar colours to an example colour (the query).

i. What features would you choose to use to represent the query and colours in the database, and why? [3]

ii. What distance measure would you use, and why? [3]

Question 8 Graphics

(a) Why do we often use triangle meshes? [2]

(b) What is surface normal and how is it normally calculated? [6]

(c) How is surface normal used in lighting calculation? [6]

(d) How is vertex normal used and how is it normally calculated? [6]