Examiners' commentary 2015–16 – Zone B

CO2227 Creative computing II: interactive multimedia

General remarks

Overall performance on this paper was good, with a few candidates showing a excellent understanding of the subject, and a few candidates obtaining a mark above 90%. The pass rate was good too, with around 80 per cent of candidates achieving a pass. However of the remaining 20 per cent, there were some very weak answer papers, showing little more than an ability to remember some facts.

In any examination, it is very important to read and address the question that is asked. Little or no credit is given to correct but irrelevant material in an answer, because that irrelevant material does not demonstrate understanding of the material, and actually gives an indication that the candidate may not fully understand some of the concepts.

It is also very important to answer all parts of a question, as there are marks assigned to each part. For example, a question that asks you to describe some concept and give examples requires that you do both: describe, and give examples. An answer that omits the examples will not score as highlyl.

Candidates were asked to choose four out of six questions to answer; all candidates answered four questions, which the examiners were pleased to see.

What follows is a brief discussion of the individual questions on this paper, with some explanations of the answers expected by the examiners.

Comments on specific questions

Question 1: Colour and light

This was a popular question, and many of those who chose this in general gave good responses. There were also a few excellent answers given.

Part (a) required an understanding that the CMY are subtractive primaries, and that these are printed on top of each other. This makes them act as filters, thereby removing certain colours.

Printing using magenta and yellow will subtract green and blue, leaving red. Finally, many candidates knew that the K refers to the Key, which is used to print black. Some explained that using all three of CMY results in a rather muddy brownish black, which is why K is preferred; however, partial credit was given if candidates did say that CMY is a way to print black.

Part (b) was less well answered. Candidates were asked for the **purpose** of the CIE LAB colour space, and some just gave a description of how it is used or implemented.

The main point is that it places colours into a coordinate system or space where Euclidian distance is related to perceptual similarity; candidates who described aspects of how it works without including this were given partial credit.

Reasonable answers were received for part (c). There is a location at the back of the eye which contains no photoreceptive cells, due to the nerve cells connected to the optic nerve. This is the blind spot. Some candidates mentioned that it is possible to perceive the blind spot by moving a small object towards one's face until the light from the object is in that part of the eye where the optic nerve is.

For part (d), a good explanation would say that when the disk is spun, people perceive colours. In addition, the colours change if disk direction changes, and colours may vary among people (though these are usually reds and blues). Other relevant points were also given credit, though a few candidates did not mention the connection with visual perception at all.

Finally, for part (e), candidates were expected to know most of the following: different cones are sensitive to different wavelengths (in particular 570, 540, 430nm); these sensitivities are due to different pigments; and they fire in different proportions according to frequency of light. Some candidates described the opponent theory or mentioned visual anomalies that relate to the cone cells, which were credited with marks.

Question 2: Animation

This was another popular question, answered by most candidates. While there were some weak responses, there were also excellent answers from some candidates.

Part (a) is a straightforward question requiring a description of the process of interpolation, and an explanation of how it is used in animation. A figure was required, which not all candidates provided. Answers often focused excessively on tweening, which is relevant but not an explanation of what interpolation itself.

Part (b) required calculation of the position of a circle at specific frames in an animation. Though there was not an explicit requirement to show working, those candidates who did were awarded partial marks even if their calculation was not completely correct. The calculation is a straightforward one, requiring knowledge of linear interpolation, and obtaining the answers $(220, 20)$ at frame 5, and $(230, 40)$ at frame 12.

For part (c), some candidates were able to describe what stop-motion animation is; this was straightforward knowledge, though not all candidates understood that it requires a sequence of still images of the positions of objects, with small changes between each still. Many candidates were able to say that we would have to view the stills at a rate of at least 16 per second; some said that 24 was a better rate (though some also said that it would have to be at least 24, which is not strictly true).

Part (d) was the least well answered part of this question, though there were also some excellent answers. Weaker answers did not mention persistence of vision, which is central to the perception of motion from stills. Credit was given for discussing beta motion, the phi phenomenon, gestalt principles, and the need for relatively small changes between frames, as well as a high enough rate of change.

Question 3: Audio and music perception

Very few candidates chose this question, though reasonable answers were given by those who did.

Part (a) required an understanding that rhythm is a complex perceptual phenomenon involving the perceptual grouping of sequences of beats or of stronger and weaker pulses.

Most candidates could correctly identify the auditory canal, the eardrum (or tympanic membrane) and the cochlea in response to part (b). A couple said that the cochlea was the basilar membrane, and only got half marks for this.

Part (c) responses should mention that the tympanic membrane converts air pressure level oscillations into mechanical vibrations, and thereby transmits vibrations into the inner ear (and the rest of the hearing system). For part (d), slightly more was expected, and a good response would include that the basilar membrane is responsible for perception of different frequencies of sounds, because different frequencies excite different locations along the membrane. Some candidates also mentioned the role of the hair cells, and discussed dissonance or masking.

Part (e) was generally answered well, and most candidates were able to identify that pitch will change from low to high, and that sound will not be heard below 20Hz or above approximately 20kHz. Most candidates were able to identify that volume will change with frequency, though not all noted that it would become louder around 3–4 kHz and then softer again. It was essential to address timbre specifically, as this was asked for: even though timbre should not change appreciably, it was important to mention this.

The final two parts were more difficult for some candidates, though others gave very good answers. For part (f), the string is vibrating at 300Hz; some candidates also mentioned harmonics, or listed correct examples of harmonics. Full marks could only be obtained if it was clear that the candidate understood that it is not possible to know the exact number of harmonics.

Not many sophisticated responses were received for part (g), where examiners were looking for some understanding that both beats and dissonance are the relevant concepts.

Question 4: Digital media signals and their representations

This was a popular question, answered by most candidates.

Part (a) required simply reading off the frequency (3Hz), amplitude (0.5) and phase ($\pi/2$). These values should then have been used in forming the equation for $y(t)$, with the result $y(t) = 0.5sin(2\pi 3t + \pi/2)$.

Some candidates confused sampling with quantisation for part (b). While there are similarities, quantisation is the process of representing an audio sample (which would have been obtained via sampling) in a specific number of bits. More bits will give higher quality, while fewer bits take less storage; explicit examples were expected which not all candidates gave. Often, answers given were really more aligned to the issues of compression, rather than sampling.

Part (c) was actually quite an easy question, though some candidates answered it weakly. $n-2, n-1$ and n are points in time, rather than actual values; most candidates who answered this question did not understand this, even though it is very obvious from the diagram. $e_2(n)$ is the residual, or the error. It is the difference between the predicted value and the actual value, at time n.

The process is linear prediction, and it is useful for compression because the residuals are typically smaller in magnitude than the original signal, so fewer bits are required to represent the values with the same accuracy.

Part (d)(i) required a calculation to determine that the file size would be 1722.7 KB; most candidates could do this. The second sub-part would arrange the size as MP3, FLAC, ZIP, WAV. Many candidates were able to do this too.

Part (e) was answered reasonably well; correct examples of lossless formats include FLAC and ZIP, and lossless format is preferred in circumstances where further editing of an image might be required, or if the purpose of compression is for archive. Candidates lost marks if they gave only a very general answer, such as needing high quality; a specific situation was asked for.

Question 5: Signals and systems

About half of the candidates answered this question, and generally gave good answers.

Part (a) required the drawing of a signal that has a value of 1 at time 0, and a value of 0 everywhere else on the x-axis. Many candidates managed this well.

Part (b) required straightforward knowledge of the material in the subject guide, in order to complete the blanks. Most candidates knew that convolving is equivalent to multiplying in the context given, and that to obtain the LTI, an input signal must be convolved with the system's impulse response. For part (iii), not all candidates understood that convolving with the unit impulse produces the same signal. Linearity was expected for (iv), and time-invariance is the property that is described by the equation in (v). A system that exhibits both linearity and time invariance is called an LTI (or linear, time-invariant) system.

Part (c) was a reasonably straightforward question, though some candidates seemed to find it confusing. Examiners were looking for a spectrum (rather than a signal), with peaks at 1000, 3000 and 4000. The relative heights of the peaks (examiners were not looking for exact heights) should be longer, shorter and then medium, in that order.

For part (d), the kernel produces a vertical echo or blur, with the total image brightness remaining unchanged. Answers that said that the image will be shifted to the right and left and then superimposed on each other were also acceptable.

For part (e), most candidates were able to name effects such as echo, edge detection, motion blur, low-pass filter, high-pass filter, etc.

Question 6: Information retrieval

This was a reasonably popular question, answered by about three-quarters of candidates.

Part (a) required a simple calculation to establish that there are 35 true positives and 5 false negatives for the query.

For part (b), though there were some good answers, a significant number of candidates did not relate their answers to the particular context given, and provided very general comments about what high precision and high recall mean. The question specifically asked what these meant for this particular system. A good example answer for part (iii) might be that in this system, both precision and recall are important. Low precision will be frustrating, because a user may get excited to see a film that is not in fact being shown. But low recall will also be a problem, because it may lead a user to think that a film he or she is interested in is not showing when, in fact, it is.

Most candidates correctly explained for part (c) that the Levenshtein distance is a measure of the similarity of two text strings; and some mentioned that there are particular ways to assign cost to insertion, deletion or editing. IR applications that are about text-based search usually find Levenshtein an appropriate measure, and most candidates gave good examples of these.

Part (d) was generally answered weakly. Candidates had to calculate the Euclidian distance for the co-ordinates (respectively 122, 58, 76.6 and 47.0) and then identify that the file `fourth.png` has the smallest distance, and should therefore be the one chosen. Weak answers generally did not identify that Euclidian distance was an appropriate metric.

The final part was a more open question, and candidates who demonstrated thought and understanding, even if they did not choose standard features, obtained full marks, if they included an explanation of why the feature they chose is relevant to musical similarity.