Examiners' commentaries 2016–17

CO2227 Creative computing II: interactive multimedia – Zone B

General remarks

Overall performance on this paper was good, with a couple of candidates showing a strong understanding of the subject, with marks well into the first class range. Although there were a few failures, the overall pass rate was higher than the previous year.

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 well.

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 gave good responses in general. There were also a few excellent answers given.

Most candidates were able to correctly identify, for part (a) that the labels indicated the vitreous humour, the lens and the pupil. Similarly for part (b), candidates understood that the cornea is where light enters the eye; it refracts light and helps to focus it.

Part (c) was also answered well. The retina is tissue lining the inner surface of the eye; it contains photoreceptive cells, which are called rods and cones.

The questions about Grassman's axioms of colour perception were not always well answered. Examiners were happy to accept, for part (d) either correct algebraic responses, or clear descriptive answers. Part (e) required demonstration of an understanding that the laws hold under strong illumination, or when the cones dominate.

Part (f) was reasonably answered. The laws matter because there are many methods for digital colour production that assume that we can represent any desirable colour as the mixture of primary colours.

Part (g) showed some weakness in understanding. Not all candidates knew that the

dashed line represents a mixture line (or the line of colours that can be made by mixing red and cyan light). Point G represents green, while the name for the set of points circumscribed by the triangle is the gamut. It is the set of colours that are expressible by all possible mixtures of these primaries.

Finally, part (h) was reasonably answered, and only a few candidates made the correct choices for each option. The first, and the last three options are true of rods, while the second and third are true of cones.

Question 2: Animation

This was also a very popular question, answered by all candidates. There were some excellent answers from some candidates, though other candidates showed lack of basic knowledge and misunderstandings.

Examiners were concerned to note that a number of candidates could not explain what key-framing is, as required for part (a). This is very basic knowledge, and it should have been a straightforward question.

For part (b) examiners accepted all well-reasoned answers. Aspects mentioned included aesthetics, labour and the suitability of digital tools.

The responses for part (c) were weak in the sketching of key-frames, and many candidates demonstrated inappropriate use of λ notation. Examiners expected two sketches, with the car in the correct location in each. Not all candidates included the location information. The next sub-part is a simple calculation. Candidates were asked to show their work, and in this and the next, many did not. Many, but not all, obtained the correct answers of 70 and 25 for the co-ordinates. For the final sub-part, the optimal equation is $50 + (f - 20)/80 * 80$; partial correctness was given partial credit.

Part (d) was a question about splines. Many candidates failed to say what cubic Hermite splines are, even though they described their use in animation, which is usually to provide a smoother animation than linear interpolation. A variety of responses, not always correct, were given for the second sub-part. The effect is that the car will start and stop smoothly, rather than abruptly. Some candidates focused on overall speed which is not correct; others showed a curved movement of the car itself. Finally, most candidates knew that $P(t_{k+1})$ is the position of the car at the next key-frame.

For the final part, part (e), candidates demonstrated an overall understanding, though sometimes detail was confused or lacking. The work is from the subject guide. Persistence of vision relates to the fact that the response to a visual stimulus can persist after that stimulus has disappeared; another take is that it is the overall processing by the visual system which allows the perception of motion from a series of still images shown in rapid succession. Examples could include perception of motion from a sequence of stills/projector, or the 'trail' of a torch or sparkler seen in the dark, or anything else reasonable.

Question 3: Audio and Music Perception

A less popular question, answered by a few candidates.

A basic misunderstanding that some candidates made for part (a) was to assume that the two waveforms being compared in each sub-part were ones that were to be played simultaneously. This created confusion, resulting in incorrect answers. Most candidates knew that A is louder than B, and that F is higher in pitch than E; however not all were able to say that there is no difference in how the sounds of C and D would be perceived. It is important to understand the impact of a phase shift on the resulting auditory signal.

Part (b) was similar in requirement to part (a), but required comparison of FFT signals. Answers to this part were weaker; H is one octave higher than G; J is quieter than I, and for K and L, the timbre will be different.

For part (c), examiners expected candidates to understand that the outer ear plays a role in this, and that time difference in soundwaves reaching each of the ears is a factor. A relevant diagram was required, which was not always given.

For part (d), candidates were expected to either agree with, or disagree with, the statement, and to justify and discuss their response, which was not always done. In general, higher amplitudes do increase perceived volume. An explanation of how loudness depends on frequency was also relevant. Other issues, such as that we don't hear anything under 20Hz or above 20k Hz; how the amplitude of a sound at our ears is different to amplitude at source and is impacted by our head position; the distance to source, and so on, could also be included.

In the main, part (e) was answered correctly, though not all candidates included the explanation required. Frequencies that are at integer multiples of 600 Hz would be present, because the string vibrates with harmonically related frequencies of a fundamental of 600Hz.

Finally, for part (f), all reasonable responses were accepted. One answer might be that melody is a (usually continuous) time sequence of tones, which is perceived as a single musical entity.

Question 4: Digital Media Signals and their Representations

A moderately popular question with those who chose to answer it giving reasonable or good answers. The question is one with many sub-parts, each worth a few marks, so it is important to collect a reasonable score by making sure to respond to each sub-part.

Part (a) was about audio representations. The first two sub-parts required calculations, and with any calculation it is important that you show your working. Not all candidates did this; it was particularly important in the first sub-part, which has a detailed set of calculations to obtain the correct answer of around 42.1 MB. The compression ratio is a simple division, resulting in the answer of about 20%. For the latter sub-parts,

mentioning the fact that MP3 is a lossy format which removes sounds that are masked according to a psychoacoustic model was required. The correct order of size for the given formats was MP3, FLAC, ZIP, WAV, which most candidates were able to explain.

The questions about compression in part (b) were generally answered well. A lossless compression algorithm allows the original digital signal to be perfectly reconstructed; in contrast, a lossy compression algorithm removes information from the signal and does not allow perfect reconstruction. Examples of when a lossless representation is preferred over a lossy one might be if the compression is for archival purposes, or when the media in the file will undergo further processing or manipulation.

Part (c) was a very straightforward calculation; doubling the frequency being digitised gave an answer of 140 kHz.

For part (d), the responses required were directly related to coverage in the subject guide. For any frequency above the Nyquist frequency, which is half the sampling rate, there exists a unique frequency between 0Hz and the Nyquist rate that is indistinguishable from it (its alias). When the sample rate is not high enough — at least twice the rate of the highest frequency — aliasing will occur, resulting in artefacts. Not all candidates included the required diagram. For the second sub-part, candidates need to understand that aliasing can be avoided by applying a low-pass filter with a cutoff frequency at or below the Nyquist frequency, before sampling.

Again for part (e), a diagram was required but not always provided. Quantisation refers to the process of representing each audio sample with a finite number of bits.

Question 5: Signals and systems

An extremely unpopular question, answered by only a handful of candidates; however those that chose it gave reasonable to good answers.

Part (a) required the drawing of a unit impulse, showing labels. This is straightforward book-work and very basic knowledge. The impulse happens at time 0, and has a value of 1, and it was essential to show this. The impulse response is simply the output of a system when the unit impulse is input. It allows us to compute the output of the system for any new input, by convolving the impulse response with the new input. Some candidates confused the unit impulse with the impulse response.

Part (b) required a calculation of the output of a particular signal, showing the calculation for four different times, and also showing the working. There was unfortunately an error in the question paper, in the specification of the impulse response. Most candidates worked out that the second line should have been for time index 1 (rather than 0), but all candidates were given full marks if they either noted the error, made the correct assumption, or tried to work with the values as given. For the correct specification, the different outputs should be $[0, 1, 2, 3]$, and the system is one that implements a unit delay. Justification of the response was also important, whatever was chosen.

For part (c), examiners were expecting candidates to realise that a high-pass filter or a band-stop/band-reject filter should be used. Following this, it would be necessary to figure out what frequency the buzz is, for example using an FFT or STFT/spectrogram, and listening to the result to check that the buzz has been removed and undesired artefacts have not been introduced. The filter drawn needed to look like a frequency response filter, with a reasonable shape (such as high-pass or bandstop) and be realistic; for example a brick-wall filter would not be a good example.

Most candidates understood that the kernel given in part (d) will essentially overlay two copies of the image: one at 3 pixels above the original position, and one 3 pixels below. It creates a sort of small vertical echo, which might look something like camera shake. Additional credit was given for noting that its effect will be relatively small or localised, and that it will have square edges or be non-Gaussian.

All candidates answered part (e) correctly; acceptable answers included echo, edge detection, motion blur, low-pass filter, high-pass filter and Gaussian blur.

Question 6: Information retrieval

This was a moderately popular question, answered by more than two thirds of the candidates.

A good answer for part (a) might be that a feature is a value computed from a piece of media. This value provides relevant information about that piece of media (for example information about how it might be perceived by a person), such that items that are similar will have similar feature values.

Part (b) was about specific features. CIE L*a*b* would be used for image or animation, and two objects with similar values would appear similar in colour. Cepstrum would most usually be used for music or melody (sound, audio, speech would also be acceptable); two pieces of media having similar cepstrum values would have similar timbre, or similar spectral shape. Term-frequency would most usually be used for text, or web pages, and similar term-frequencies would indicate that the documents have similar rates of occurrence of the same term. It is important to note the distinction between rates of occurrence — which is essentially proportion — and actual number of occurrences, which not all candidates did.

Part (c) was usually answered well. The Euclidean distance is 50, and B is more similar to A because it has a lower distance. Not all candidates explained that the lower distance indicates similarity. Finally, other metrics could include Manhattan distance and p-norm; metrics that require binary features would not be appropriate.

For the final part, (d), most candidates were able to calculate the precision (66.66%) and recall (50%) correctly, though some confused them with each other. Responses for the third sub-part varied, but all reasonable and justified answers were acceptable. One approach could take the view that it is more important to have high precision, since the journalist will not want to waste time looking at things that are not cars. It is not so

important that they find the single best photo of a car. However, if recall is very poor this could also be bad, since none of the cars that are returned might be suitable.