Examiners' commentaries 2016–17

CO2227 Creative computing II: interactive multimedia – Zone A

General remarks

Overall performance on this paper was reasonable, with a couple of candidates showing a very good understanding of the subject, with marks well into the first class range. In contrast there were also some weak papers, including a few failures.

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 cornea and the optic nerve. Similarly for part (b), candidates understood that the retina is tissue lining the inner surface of the eye, and that it contains photoreceptive cells (or rods and cones).

A variety of answers were given for part (c); some described the pupil without explaining its function, which is to allow light to pass into the eye.

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 B represents blue, 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 three options are true of cones, while the last three are true of rods.

Question 2: Animation

This was also a popular question, and there were some excellent answers.

Part (a) required a straightforward response that relates to the relevant aspects of animation. Layering is a time-saving technique. A background may remain static, and only figures, or certain figures, may move. In this case, the same background layer can be used for multiple frames.

For part (b) examiners accepted all well-reasoned answers. For example, some candidates mentioned the choice of whether clay or hand drawing is the appropriate aesthetic for the animation concerned; whether flat backgrounds are appropriate; ways in which workload could be distributed. Another aspect is the reality of physics, which stop-motion animation must adhere to more than flat animation. Some candidates felt that the choice was linked to effort; however both are manually intensive processes, so that isn't relevant to the choice of one or the other.

The responses for part (c) were weak in the sketching of keyframes. 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 50 and 110 for the co-ordinates. For the final sub-part, the optimal equation is $30 + (f - 30)/50 * 100$; 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. Finally, for this part, $m(t_k)$ is the velocity at time k. Some candidates mentioned that this is the tangent, which was given credit by the examiners.

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 very popular question, answered by many candidates.

For part (a), most candidates knew that D is louder than C; however not all were able to say that A is higher in pitch than B, and that there is no difference in how the sounds of E and F 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 quieter than G; J is one octave lower than I, and for K and L, the timbre will be different.

Part (c) was very straightforward, and could be answered well simply using knowledge from the subject guide. It was important though, that the parts chosen were actually related to the inner ear; a couple of candidates chose parts that weren't.

For part (d), various kinds of responses were given. Credit was given for mentioning that in general, closer frequencies result in more dissonance. Further discussion was expected on topics such as critical bands, masking, and other complexities (for example, the fact that most sounds have multiple frequency components, or the fact that critical bandwidth is dependent on frequency). Full credit was given for an appropriate subset of the above.

In the main, part (e) was answered correctly, though not all candidates included the explanation requested. Frequencies that are at integer multiples of 400 Hz would be present, because the string vibrates with harmonically related frequencies of a fundamental of 400Hz.

Finally, for part (f), all reasonable responses were accepted. One example might be that rhythm is a complex perceptual phenomenon involving the perceptual grouping of sequences of beats or of stronger and weaker pulses.

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 34.3 MB. Working out how many seconds in length the song is, and how many bits are required for each second's representation are the initial steps. For the latter sub-parts, mentioning the fact that MP3 is a lossy format was essential, and describing how this impacts on the compression rate was also 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 lossy algorithms may be preferable to lossless ones included contexts such as streaming over the internet or listening on personal music players, where having a smaller file size is more important than being able to perfectly reconstruct the original signal.

Part (c) required a very straightforward calculation; doubling the frequency being digitised gave an answer of 320 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. The different outputs should be $[0.5, 1, 1.5, 0]$. The system is one that performs scaling. Other acceptable answers might

be that it changes volume, or changes gain; however it was important to also justify the answer given.

For part (c), examiners were expecting candidates to realise that a low-pass filter or a band-stop/band-reject filter should be used. Following this, it would be necessary to figure out what frequency the alarm is, for example using an FFT or STFT/spectrogram, and listening to the result to check that the alarm 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 and be realistic; for example a brick-wall filter would not be a good example.

Most candidates understood that part (d) showed a blur kernel. 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 half 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. Parson's code would most usually be used for music or melody; two pieces of media having similar Parson's code values would have similar melodic shape or melodic contour. 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. Vague responses, such as *'similar in content'* received fewer marks. Finally, for this part, CIE L*a*b* would be used for image or animation, and two objects with similar values would appear similar in colour.

Part (c) was usually answered well. The Euclidean distance is 60.53, and C 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 and recall 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.