# Examiners' commentary 2017–2018

## 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. There were also a few failures, but fewer than last 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.

In quite a few instances, it seemed as though candidates were giving learned answers, without deeper understanding of what they were writing down. If the learning approach is to learn model answers to anticipated questions, then candidates are likely to do less well than if they learn in order to understand the material. This is especially true if the question asked is similar to but not exactly the same as the anticipated one; if candidates give answers that are those for a slightly different question, it is very easy for examiners to identify that this is the approach being taken, and candidates will earn fewer marks for such responses.

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 in general gave good responses. 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 iris and the retina.

For part (b), candidates did not always know that the ciliary muscles can deform the lens, which changes its focal characteristics.

Part (c), however, was usually answered well, with candidates understanding that the pupil is the opening that allows light into the eye.

The descriptions of hue, saturation and brightness given for part (d) were usually good. One mark was given for each of these.

Part (e) was a straightforward calculation, though a number of candidates calculated h incorrectly. The correct equation to choose was the one for *max* = *r* — which most candidates managed — and it is simply a matter of filling in the values of *r*; *g*; *max* and *min*, and then working out what this comes to — which many candidates calculated incorrectly. Most candidates were able to correctly calculate *s* and *b* as around 0.9 and 1 respectively.

Most candidates were able answer part (f) correctly. The first, third and fifth options are true of cones, while the others are true of rods.

Part (g) was also generally well answered. Most candidates knew that point R represents red. However not all candidates knew that E represents white, and is the point of equal energy. Almost all candidates could explain that the name for the set of points circumscribed by the triangle is the gamut, and that this is the set of colours that are expressible by all possible mixtures of these primaries.

## Question 2

### Animation

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

Examiners were concerned, for part (a) to note that a number of candidates produced sketches that were not actually keyframes, showing that they did not understand what these are. Those that did, did not always indicate the correct locations of the car in these. Most candidates could, however, calculate the *x* and *y* co-ordinates of the car in frame 66. For part (a)(iv), the equation given for calculating the *x* co-ordinate at any time sometimes was one that made use of a previous position (from a previous frame), which does not appropriately allow calculation of position for any frame.

Part (b) was a question about splines. Most candidates were able to say what cubic Hermite splines are, and in general to describe 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, for part (b)(iii) many candidates did not know what *m(tk)* is, and sometimes said – incorrectly - that it was the position of the car at the current key-frame.

Most candidates gave reasonable descriptions of stop-motion animation for part (c), though sometimes not enough detail was included. There are many reasonable answers for part (c)(ii) which include aesthetic considerations, and ways in which workload is to be distributed. Some candidates claimed that stop motion is slower than flat animation, but did not give evidence or reasons for this claim.

For the final part, part (d), candidates demonstrated an overall understanding, though sometimes detail was confused or lacking. The work is taken directly from the subject guide, and explanations should be clear and straightforward. 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

This was a less popular question, answered by a few candidates, but usually with good answers.

Part (a) was answered well. Most candidates know that A is higher in pitch than B, and that D is louder than C. Many were also able to say that there is no difference in how E and F will sound.

Part (b) was similar in requirement to part (a), but required comparison of FFT signals.

This was also answered well; H is louder than G while J is one octave lower than I.

However, most were unable to say that for K and L, the timbre will be different.

Part (c) was a very straightforward question, and could be answered well simply by using knowledge obtained 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 were not.

In the main, part (d) was answered correctly, though not all candidates included the explanation required. Frequencies that are integer multiples of 160 Hz would be present, because the air column vibrates with these harmonically related frequencies.

For part (e), all reasonable responses were accepted. Examiners expected mention of perceptual grouping of sequences of beats, or of stronger or weaker pulses. One candidates claimed that rhythm is a sequence of highly memorable stimuli, giving the impression they had learned a pre-prepared answer and not recalled it correctly.

Finally, for part (f), examiners expected candidates to know that the harmonic series is a set of harmonically related frequencies, which are multiples of a single fundamental frequency; and that the fundamental frequency typically corresponds to pitch while the timbre is influenced by the harmonic components. Many candidates did not include a discussion of how this relates to consonance and dissonance, nor was there any mention of the partials, giving the impression that learned answers to a similar question were being used.

## Question 4

### Digital Media Signals and their representations

This was a 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 possible to collect a reasonable score just 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; this was particularly important in the first sub-part, which has a detailed set of calculations to obtain the correct answer of around 45.9 MB. The compression ratio is a simple division, resulting in the answer of about 17.6%. 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. Some candidates were only able to say that "information is removed", but could not explain on what basis this is done. The correct order of size for the given formats was MP3, FLAC, ZIP, WAV, which most candidates were able to give.

The questions about compression in part (b) were generally answered well. Only one example of either audio or image lossy compression was asked for, making this part an easy question. 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. The descriptions of how FLAC uses LPC for part (b)(iii) often did not note that LPC is only useful when the signal is nearly linear. Then, instead of storing the value at time t, we can store just the error between the true value at time t and the value predicted by times t-1 and t-2. It was essential to include a diagram. Some candidates noted that for FLAC, residuals are typically smaller in magnitude than the original signal, so fewer bits are required to represent the values with the same degree of precision.

Part (c)(i) is a very straightforward calculation, where doubling the frequency being digitised gives an answer of 24 kHz. For part (c)(ii), the responses required were directly related to 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.

Again for part (d), a diagram was required but not always provided. Quantisation refers to the process of representing each audio sample with a finite number of bits. Part (d)(iii) highlighted a number of misunderstandings, again with candidates giving answers that were correct but not always related to the question. Some candidates confused quantisation with aliasing; others thought that a higher sampling rate would always result in a larger bit depth, although these aspects are not directly related to each other. Finally, candidates did not always understand why more bits give higher fidelity.

## Question 5

### Signals and Systems

This was an unpopular question, answered by only a small handful of candidates, with varying quality.

Part (a) required the drawing of a unit impulse, showing labels. This is straightforward bookwork 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, while others were unable to say how the output of the system is calculated.

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 system is a low-pass filter, and marks were awarded for answers that stated this; however it was important to also justify the answer given in order to obtain full marks.

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 (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) is basically a sharpen kernel which emphasises differences in adjacent pixel values, making the image look more vivid.

All candidates answered part (e) correctly; acceptable answers included echo, blur, 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 who entered the examination.

A straightforward calculation was required for part (a)(i) as well as for part (a)(ii). For the final sub-part, some candidates presented an argument along the lines of it being more important to have high recall, since the user really wants to find the right song.

High precision is less important, because the user can ignore bad results. However, if precision is very low this could also be bad. One candidate did point out, in this case, that it is harder to ignore bad results because you have to listen to the results in order to decide, and this takes time with audio, more than with images. Unfortunately many of the answers were generic, not related to the particular context at all, and some candidates said that the size of the database is an important factor.

Part (b) was about specific features. Most candidates correctly identified that CIE L*a*b is for image, and edit distance is for text, but some thought that a chromagram relates to image and in particular colour.

Part (c) was usually answered well. The Euclidean distance is 131, and document 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, part (d), candidates were awarded marks for discussing any appropriate feature, providing the explanation of how it is relevant to musical similarity was clear.