

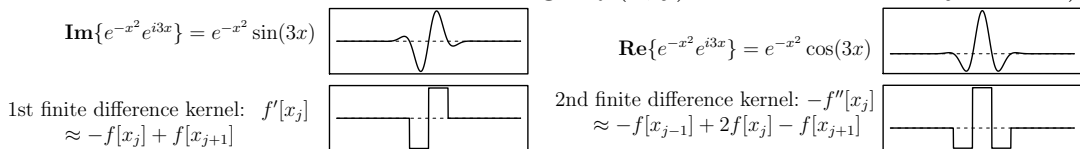
4 Computer Vision (jgd1000)

(a) In the context of automated detection and interpretation of affective expressions using FACS (Facial Action Coding System), define the following concepts:

- (i) facial muscle action unit (AU)
- (ii) action descriptor (AD)
- (iii) valence
- (iv) arousal
- (v) “Pan-Am smile”
- (vi) cross-cultural universal

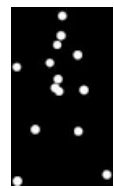
[6 marks]

(b) Gabor wavelets offer a unifying framework for many tasks in computer vision, including edge detection, finding facial features, and pattern matching. The complex wavelet components (upper row) are parameterised to approximate the convolution kernels for computing 1st and 2nd derivatives (lower row, given as functions of x but rotatable in images $f(x, y)$ into functions of y as well):



Explain how Gabor wavelets can estimate the gradient vector field $\vec{\nabla}f(x, y)$ in edge detection, extracting both edge strength and edge direction. Also describe how they can be used in a demodulation network to localise facial features. Identify one application of Gabor wavelets in pattern matching. [5 marks]

(c) What can we learn from the perceptual experiments of the Swedish psychologist Johansson, involving sparse dot patterns such as shown on the right? How might his findings be useful in computer vision for data fusion, integration of motion cues in object recognition, and general aspects of scene understanding?



[4 marks]

(d) Biological neurones are notoriously noisy, are apparently random in their connections and their firing patterns, and sluggish, with maximum firing rates around 100 Hz. Yet biological vision systems are wonderfully capable. Is there really any need for computer vision systems to use double precision arithmetic and GHz clock speeds? Give three examples of tasks in machine vision whose execution appears to require double precision arithmetic and high FLOPS, and for each example, explain this contrast with biological solutions. [5 marks]