Computer Vision

(a) Name four Gestalt laws of perceptual organisation. What is the main theme of Gestalt Psychology, and in what way is it relevant to automated object recognition in images and video sequences? [4 marks]

(b) Many computer vision algorithms such as SIFT (scale-invariant feature transform) seek to detect and analyse features at multiple scales of analysis.

(i) What is “scale-space” and what is a “scale-space fingerprint”? [2 marks]

(ii) Briefly describe how SIFT achieves scale and rotation invariance. [3 marks]

(iii) Show that the difference-of-Gaussian operation $\text{DoG}$ given by

$$\text{DoG}(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

on image data $I$ as used by SIFT is an approximation of the scale-normalised Laplacian of Gaussian given by $\sigma^2 \nabla^2 G$. You may assume that

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G$$

Comment on a good choice for the value of $k$ for practical applications of SIFT. [6 marks]

(c) The curvature map of a closed image contour is a bandlimited signal and can be described by its zero-crossings. Explain how this property can be exploited for 2D shape description. What properties make such a representation suitable for classification and recognition of shapes? [5 marks]