2009 Paper 8 Question 4

Computer Vision

(*a*) Discuss the significance of the fact that mammalian brains send almost ten times as many neural fibres back down the visual pathway (from cortex to thalamus), as there are ascending neural fibres bringing visual data from the retina up to the thalamus and cortex. Does this massive neural feedback projection support the thesis of "vision as hypothesis testing," and if so, how? Try to marshall other evidence supporting the view that in human vision "what you see is your own graphics" rather than the retinal image as faithfully recorded by photoreceptors in the eye. [5 marks]

(*b*) Write a block of pseudo-code for convolving an image with a feature-detecting kernel. (You may ignore out-of-bounds issues at the image array boundaries.) [3 marks]

(*c*) The Differentiation Theorem of Fourier analysis asserts that, for an image $f(x, y)$ whose 2D Fourier Transform (*2DFT*) is $F(\mu, \nu)$, taking derivatives of various orders m, n has the following spectral consequence:

$$\left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial y}\right)^n f(x,y) \stackrel{2DFT}{\Longrightarrow} (i\mu)^m (i\nu)^n F(\mu,\nu)$$

In particular when the 2nd-derivative Laplacian operator ∇^2 is applied to an image, its action is to multiply the Fourier transform of the image by a paraboloid:

$$\nabla^2 f(x,y) \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(x,y) \stackrel{2DFT}{\Longrightarrow} -(\mu^2 + \nu^2) F(\mu,\nu)$$

What therefore best describes the filtering operation that corresponds to taking derivatives of an image: lowpass, bandpass, or highpass? If noise in an image resides mainly in the higher spatial frequencies, should you use higher or lower order derivatives for edge detection to be less sensitive to the noise? What else could you do with differentiating filters to reduce their noise sensitivity? [3 marks]

(*d*) In pattern classification with two classes, explain how an ROC curve is derived from the underlying distributions. Define a threshold-independent performance metric based on the distributions' moments. [4 marks]

(*e*) When visually inferring a 3D representation of a face, it is useful to extract separately both a shape model, and a texture model. Explain the purposes of these steps, their use in morphable models for pose-invariant face recognition, and how the shape and texture models are extracted and later re-combined. [5 marks]