Model Answer – Computer Vision 2004, Paper 8 Question 11 (JGD)

(a) For solving classification problems, visual features should be chosen which minimise the within-class variability and maximise the between-class variability. This allows the diameters of clusters in feature space to be small compared with the spacings amongst the clusters, thereby minimising overlap and thus classification errors. [2 marks]

(b) (i) The two kernels form a quadrature filter pair because they have a 90 degree phase offset. The first is even-symmetric (in fact a cosine Gabor), and the second is odd-symmetric (in fact a sine Gabor). The two kernels are orthogonal to each other (their inner product = 0). [2 marks]

(ii) The DC response of each kernel is 0. This means they give no response to uniform areas of an image (where brightness is constant). [1 mark]

(iii) These filters are most response to horizontal structures such as edges, or other modulations (such as fingers) that are horizontal. [1 mark]

(iv) The kernels would be used by convolving them with an image. Positioned over each pixel in the image, the sum of the products of each tap in the filter with each corresponding pixel in the image would become the new pixel at that point in a new image: the filtered image. [1 mark]

(v) Alternatively, the same result could be obtained just by multiplying the Fourier Transform of each kernel with the Fourier Transform of the image, and then taking the inverse Fourier Transform of the product. [1 mark]

(vi) Taking the modulus (the sum of the squares) of the results from convolving a facial image with the two kernels yields peaks of energy at locations corresponding to the eyes and the mouth, when the scale is appropriate, since such facial features are local wavelet-like undulations. [2 marks]

(c) A problem is defined (by Hadamard) as "well-posed" if all of these three conditions apply: (1) a solution exists; (2) the solution is unique; and (3) the solution depends continuously on the data. In general, the problem of infering object properties from image properties violates one or more of these three conditions. The problems only become soluble by adding ancillary assumptions, or other data (such as past knowledge learned from experience or from other modalities). The task of infering the spectral reflectances of object surfaces from image colours (the wavelengths of light received at the sensor) is only soluble if one knows the wavelength composition of the illuminant, since that distribution is multiplied by the spectral reflectance function of the object surface. The problem of colour inference becomes well-posed and almost instantly soluble if the illuminant is known; but in the absence of that information or of strong assumptions, it is ill-posed. [5 marks]

(*d*) The "Paradox of Cognitive Penetrance" refers to the fact that the visual tasks that humans are particularly skilled at, such as face recognition, visual learning, navigation, and solving Correspondence Problems, are performed without our having an understanding of how we do them. In contrast, the tasks for which we have an in-depth theoretical understanding and which we know how to write algorithms for, are often tasks that we humans are rather poor at performing, such as numerical operations and mathematical transformations.

The systematic geometrical illusions which occur in the human visual system suggest that fidelity to image properties is not always a goal of biological visual algorithms. We are aware of the illusions but we don't know why they occur. In machine vision today, it is difficult to imagine trying to design algorithms which would intentionally make systematic errors; and yet arguably the human visual illusions are consequences of valuable adaptive strategies.

The significance of the Paradox of Cognitive Penetrance is that the prospects for "reverse engineering" human visual faculties may be reduced by the difficulty of gaining insight into how we actually do what we do. A further implication is that machine vision algorithms, even if successful, are likely to adopt quite different strategies than the biological ones. [5 marks]