1999

P8q 12
JGD

Model Answer, Computer Vision, Question 1.

1. One example of the use of complex variables in computer vision is the Fourier transform, which is the basis for efficient implementations of filtering as well as for certain pattern classification tasks. Here the image $I(x, y)$ is represented as a linear combination of complex exponentials:

$$I(x, y) = \sum_k a_k \exp(i(\mu_k x + \nu_k y))$$

The purpose of the transform is to obtain the set of coefficients a_k for every spatial frequency and orientation in the 2D Fourier domain spanned by the 2D frequency variables (μ_k, ν_k). These coefficients may be computed by the following integral:

$$a_k = (2\pi)^{-2} \int_X \int_Y \exp(-i(\mu_k x + \nu_k y)) I(x, y) dx dy$$

Being complex-valued, they are usefully resolved into polar complex form as amplitude and phase. The modulus of such coefficients (the sum of the squares of their real and imaginary parts) gives the power spectrum, which is shift-invariant.

A second example is the wavelet transform. This differs from the Fourier transform mainly in that the domain of description remains the image domain, with frequency being merely a parameter, and the spectral analysis is local rather than global. Thus one extracts the structure in the signal (i.e. image) at particular "scales of analysis" in the (x, y) domain. Wavelet expansions can play a role in many computer vision tasks, including edge extraction and pattern classification. Indeed, they include the Fourier transform as just a special case! Wavelet representations allow a continuous deformation between space-domain and frequency-domain descriptions of a pattern, allowing one to enjoy the benefits of both worlds simultaneously.

[10 marks]

2. When the real and the imaginary parts of wavelet representations are resolved into their complex polar form as modulus and phase, generally the modulus is useful for pattern classification whilst the phase is useful for pattern identification. Thus, for example, in the wavelet representation for a human face, the modulus would help to detect that this pattern is a face and not some other category of object. However, the modulus would not be very useful for identifying *whose* face it was. The phase structure, in contrast, would be useful for identification of the face as one particular face. Moreover, specific facial expressions can be described as *phase modulations* of an underlying canonical face. These efficient ways of encoding the details of how the face in question is "configured" illustrate how the two parts of complex variables when resolved into polar form as modulus and phase can play very different roles in image understanding.

[10 marks]