P 9 q 12
JC-D

Model Answer, Computer Vision, Question 2. 2000

1. Stereo vision requires that corresponding object points, in two images acquired from slightly different vantage points, be associated to each other in order to make possible the measurement of their relative disparity in the image plane and thereby a computation of their depth in space relative to the focal plane.

   Motion vision requires that corresponding object points, in two images acquired from the same vantage point but at slighly different moments in time, be associated to each other in order to make possible a measurement of their relative displacement and thereby a calculation of the motion vector over this interval of time.

2. These two cases of the Correspondence Problem are symmetrical, with the roles of space and time simply interchanged. In stereo vision, the two image frames are simultaneous in time but displaced in space. In motion vision, the two image frames are from the same vantage point (coincident in space) but displaced in time.

3. The complexity of the computation depends, in the worst case, quadratically on the number of features that constitute the data. This is because in principle, every feature in one image frame could be associated with every possible feature in the other image frame. Hence N features may need up to NxN candidate hypotheses about what goes with what. N could even be as large as the number of pixels in the image.

4. One way to make this computation more efficient is by stochastic relaxation, in which large-deviation (large displacement) correspondence hypotheses no longer need to be considered once enough evidence has accumulated for a more conservative solution. The amplitude of the deviations may be slowly diminished in a way that corresponds to decline of temperature in annealing algorithms. Another approach, more radical (but less reliable statistically), would be a competitive winner-take-all neural net.