Homework #4 solutions

Q8.2: The instantaneous epipole is the focus of expansion (or contraction). The terminology is inherited from stereo. Here's the basic idea: Transport the translation vector from the moving point to the cener of projection. The perspective projection of the translation vector is the instantaneous epipole.
Q8.3: The motion field and the optical flow are conceptually different and yet obviously related. Simply put, the motion field is the apparent motion in the image plane of actual 3D motion. The optical flow is the displacement field at each pixel which brings two consecutive images into register. The two are one and the same for instance if intensity-based segmentation is capable of segmenting rigid objects. Then the displacement field required to register two images is the same as the motion field.
Q8.4: The image brightness is usually a continuous function except at object or feature boundaries. Consequently, assuming the intensity function to be differentiable is not egregious. However, at object boundaries or at boundaries where texture differences occur, the image intensity may not be continuous or differentiable.
E8.2: If a visible corner can be detected, the aperture problem is solvable. This is because a unique optical flow vector can be obtained by tracking the corner. Even if the corner is attached to two separate line segments (which is typically the case), the corner's optical flow is unique even if the two line segments' optical flow is not. After obtaining the optical flow from the corner and assuming that line segments immediately connected and adjacent to the corner are from the same rigid object, we can trivially propagate the missing tangential component of the optical flow to the two line segments. Note that the tangential component propagated will be different for the two line segments.
E8.3: Let's assume that the optical flow is not a constant in the patch but can be approximated by the linear equation $o_x(x,y) = ax + by + c$ and $o_y(x,y) = dx+ey+f$. This gives us six unknowns per patch as opposed to the two unknowns per patch in the CONSTANT FLOW case. Having parameterized the optical flow in this manner, the constraint equation can be turned into a least-squares equation as follows:

$$E_{least-squares}(a,b,c,d,e,f) = \sum_{i=1}^{N\times N} |\nabla E_i^T o_i + Et_i|^2$$

where $o_i = [ax_i + by_i + c, dx_i + ey_i + f]^T$ and the point (x_i, y_i) is set based on the index i in the $N \times N$ block. The cost function $E_{least-squares}$ is minimized

w.r.t. the 6 unknown parameters (a, b, c, d, e, f). You need at least 6 pixels in the $N \times N$ block and usually quite a few more in order to get reasonable estimates of the linearly varying optical flow.