3DCV HW2

r12521601

Q1:

The original intrinsic matrix K is:

$$K = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

When cropping the image to J, the focal lengths f_u and f_v, and the skew s, remain the same. The principal point (u_0, v_0) will shift to a new position due to the cropping. This shift in u_0 and v_0 accounts for the new location of the principal point in the cropped image.

The new intrinsic matrix for J, denoted K_J, is:

$$K_J = \begin{bmatrix} f_u & s & u_0' \\ 0 & f_v & v_0' \\ 0 & 0 & 1 \end{bmatrix}$$

Where:

$$u_0' = u_0 - u_1, \quad v_0' = v_0 - v_1$$

Q2:

SuperPoint uses a tensor of downsized maps ($W/8$ and $H/8$) with 65 channels instead of deconvolution layers and a binary mask of the original image size to encode the keypoints. The advantage of this design is Computation Efficiency. Since deconvolution layers are computationally expensive, SuperPoint avoids deconvolution, making the network simpler and more efficient. In addition, with 65 channels, it can represent 64 possible locations and 1 dustbin (no key point), which can also perform a good detection of keypoints.