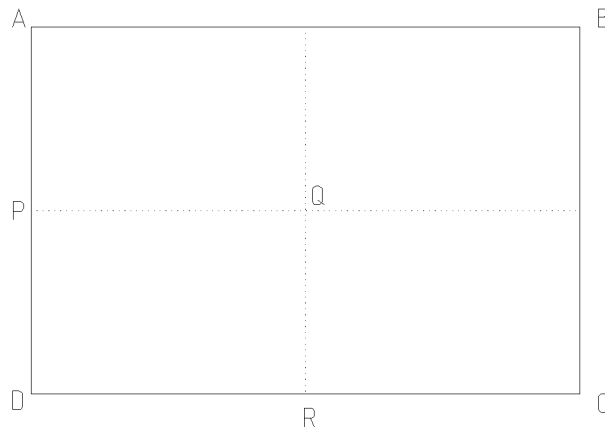


Exercises in 3D Computer Vision

Exercise 14 Homographies

A surveillance camera views a rectangular parking lot to ensure that only authorized vehicles occupy spaces in the lower-left quadrant, defined by points P , Q , R , and D in the figure below. The four corners of the parking lot, A , B , C , and D , perspective project into the image as corners of a quadrilateral with image coordinates $a = (1, 1)$, $b = (2, 1)$, $c = (3, 0)$, and $d = (0, 0)$, respectively. Assume the parking lot is flat, and the ratio of its “length” (i.e., distance from A to B) to its “width” (i.e., distance from B to C) is $3/2$. Using a plane projective camera model, compute the real-valued image coordinates of the scene points P , Q , and R .



Exercise 15 Singular Value Decomposition (SVD)

Let \mathbf{A} be a $(m \times n)$ real matrix with $m \geq n$. Then, for any such matrix \mathbf{A} , there exists the so-called *SVD decomposition* of \mathbf{A} which can be written as

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$$

The matrix \mathbf{U} is defined as an $(m \times m)$ matrix, \mathbf{S} as an $(m \times n)$ matrix, and \mathbf{V} as an $(n \times n)$ matrix.

The matrices \mathbf{U} and \mathbf{V} are orthogonal matrices and the entries of the matrix \mathbf{S} are null except on the diagonal (on the first super-diagonal). The entries σ_i of the diagonal of the matrix \mathbf{S} are called the singular values of the matrix \mathbf{A} . These entries are positive and always arranged in descending order: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

We admit the following proposition (based on Gram-Schmidt orthogonalization theorem):

If a $(m \times n)$ matrix \mathbf{V}_1 has orthonormal columns ($m > n$), then there exists a $(m \times (m - n))$ matrix \mathbf{V}_2 such that $\mathbf{V} = [\mathbf{V}_1 \ \mathbf{V}_2]$ is orthogonal, where the orthogonal complement of the span of column vectors of the matrix \mathbf{V}_1 is equal to the span of the column vectors of the matrix \mathbf{V}_2 .

Prove the existence of the SVD decomposition for any $(m \times n)$ real matrix \mathbf{A} .

Exercise 16 Direct Linear Transformation (DLT)

The Direct Linear Transformation Algorithm computes a homography for a projective transformation by using at least 4 point correspondences $\{\mathbf{x}'_i \leftrightarrow \mathbf{x}_i\}$ and minimizing the norm $\|\mathbf{A}\mathbf{h}\|$, where \mathbf{A} contains the stacked-up constraints given by the homography equation $\mathbf{x}'_i \propto \mathbf{H}\mathbf{x}_i$ for all point correspondences, and \mathbf{h} is a (9×1) vector consisting of the entries of \mathbf{H} . The DLT algorithm is given by:

- (i) For each point correspondence $\mathbf{x}'_i \leftrightarrow \mathbf{x}_i$ compute the (2×9) matrix \mathbf{A}_i given by the homography equation $\mathbf{x}'_i \propto \mathbf{H}\mathbf{x}_i$.
 - (ii) Assemble the matrices \mathbf{A}_i into a single $(2n \times 9)$ matrix \mathbf{A} .
 - (iii) Obtain the Singular Value Decomposition (SVD) of \mathbf{A} . The unit singular vector corresponding to the smallest singular value is the solution \mathbf{h} .
 - (iv) Reshaping the (9×1) vector \mathbf{h} gives the homography \mathbf{H} .
- a) Explain the last two steps. Develop the DLT algorithm from the point correspondences $\{\mathbf{x}'_i \leftrightarrow \mathbf{x}_i\}$ and the homography equation $\mathbf{x}'_i \propto \mathbf{H}\mathbf{x}_i$. What is the homography normalization in this case?
 - b) (**H**) In the last exercise sheet you were asked to determine the 2D homography for a surveillance camera. Implement a program in MatLab that computes the homography out of the 4 point correspondences using the DLT.

Exercise 17 **Algebraic vs. geometric error and Maximum Likelihood Estimation**

The DLT algorithm minimizes the algebraic error, i.e. the norm $\|\mathbf{A}\mathbf{h}\|$. However, this is not always suitable since it sometimes has no geometric meaning and thus makes it hard to be interpreted and error-prone. Moreover, the algebraic cost function disregards measurement errors (noise) in the point correspondences. Hence, the **geometric error** based on the Euclidean distance $d(\mathbf{x}_i, \mathbf{x}'_i)$ between two points \mathbf{x} and \mathbf{x}'_i is introduced and used as a cost function to determine a homography. The cost function can now be altered such that it also considers noise, i.e. measurement errors in the point correspondences in both images. The cost function most frequently used is called the **reprojection error**,

$$\sum_i (d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2),$$

where $\mathbf{x}_i, \mathbf{x}'_i$ are measured points and $\hat{\mathbf{x}}'_i = \hat{\mathbf{H}}\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_i = \hat{\mathbf{H}}^{-1}\hat{\mathbf{x}}'_i$ are the estimated points.

- a) What is the relationship between algebraic distance ($\|\mathbf{A}_i\mathbf{h}\|$) and Euclidean distance?
- b) What do the estimated point correspondences $\{\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i\}$ represent?
- c) (**H**) The Maximum likelihood estimation (MLE) imposes a model on the measurement error in order to obtain an optimal estimate. The assumption made is that the noise obeys a Gaussian distribution with zero mean and variance σ^2 . The MLE then maximizes the log-likelihood of a random variable consisting of all point correspondences. Explain what error function the ML estimate minimizes.