Exercises

Processamento de Imagem e Visão (PIV)

draft version, Sept. 4, 2014

This list of exercises aims to help students learn the main concepts discussed in PIV course and to be a useful tool to prepare the final exam. Please do not deliver the problem solutions to other students, since you would be preventing them from the most important benefit: solving the problems by their own and improving their problem solving capabilities.
Try to be as precise as possible in your answers and include all mathematical details when possible.

1 Least Squares Basics

These techniques are used in later sections (4, 7, 8, and 10).

1. Consider a linear model
$$\hat{y} = f(x;\theta) = \theta_1\phi_1(x) + \cdots + \theta_m\phi_m(x)$$
where $\phi_i(x), i = 1, \ldots, m$ are known basis functions and $\theta = (\theta_1, \ldots, \theta_m)$ is a vector of unknown coefficients to be estimated. Given a set of input-output pairs $D = \{(x_1, y_1), \ldots, (x_N, y_N)\}, N > m$, determine the vector of coefficients θ^*, that minimizes the quadratic cost functional
$$E(\theta) = \sum_{i=1}^{N}(y_i - f(x_i;\theta))^2$$

2. Show that the previous problem can be formulated using matrix notation, by defining $\hat{y} = M\theta$, with $\hat{y} = [\hat{y}_1, \ldots, \hat{y}_N]^T$, M being an appropriate $N \times m$ matrix. In this case, the cost function becomes $E(\theta) = \|y - M\theta\|^2$, with $y = [y_1, \ldots, y_N]^T$.

3. Given a cost function $E(\theta) = \|y - M\theta\|^2$ where y, M, θ are defined as before, determine the vector θ^* which minimizes $E(\theta)$.

4. Suppose $y = 0$ and $E(\theta) = \|M\theta\|^2$. However, we are not interested in the trivial solution $\theta = 0$ and impose an additional constraint on the solution $\|\theta\|^2 = 1$. Minimize $E(\theta)$ under the constraint $\|\theta\|^2 = 1$. (Hint: use Lagrangean multipliers).

5. Consider the previous minimization problem assuming two unknown variables (m=2). Sketck a possible set of level curves (ellipses) and the constraint $\|\theta\|^2 = 1$. Try to sketch the optimal solution that minimizes $E(\theta)$.

2 Homogeneous Coordinates

1. Prove that a line in a 2D plane can be writen as $\tilde{l}^T\tilde{x} = 0$, in homogeneous coordinates, where $\tilde{x} = \lambda(x, y, 1)$ and $\tilde{l} = (a, b, c)$.

2. Prove that the intersection of two 2D lines, $\tilde{l}_1, \tilde{l}_2$, in homogeneous coordinates, can be obtained by $\tilde{x} = \tilde{l}_1 \times \tilde{l}_2$.

3. Prove that a 2D line that contains two points $\tilde{x}_1, \tilde{x}_2$, is given by $\tilde{l} = \tilde{x}_1 \times \tilde{x}_2$.

4. Consider a 3D line that contains the points $\tilde{x}_1, \tilde{x}_2$. Prove that this line can be described by $\tilde{x} = \alpha\tilde{x}_1 + \beta\tilde{x}_2$.

5. Consider a rotation in the plane, characterized by a matrix

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Show that R is an orthogonal matrix *i.e.*, $R^T R = RR^T = I$ (and therefore $R^{-1} = R^T$). Show also that that $\det R = 1$.

6. Consider an homography in the plane $\tilde{x}' \sim H\tilde{x}$, where H is a non-singular 3×3 matrix ($\tilde{x} \sim \tilde{y}$ means that $\exists \lambda \neq 0 : \tilde{x} = \lambda \tilde{y}$). Show that

   (a) $\tilde{x}' \sim M\tilde{x}$, with $M = \alpha H$, and $\alpha \neq 0$, represents the same homography.

   (b) the homography can be expressed in Cartesian coordinates (x, y), (x', y') by

$$\begin{cases} \tilde{x}' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}} \\ \tilde{y}' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}} \end{cases}$$

7. An homography in the plane transforms points according to $\tilde{x}' \sim H\tilde{x}$ where H is a 3×3 matrix. Does the homography tansforms lines parameters l in the same way as points *i.e.*, $\tilde{l}' \sim H\tilde{l}$?

8. Consider a 3D rotation defined by $x' = y, y' = z, z' = x$. Determine the rotation matriz R and prove that R it is an orthogonal matrix ($R^T R = RR^T = I$) and furthermore $\det R = 1$.

9. Consider a 3D rotation matrix R and let r_i^T denote the i-th row of R. Show that r_i are orthonormal vectors *i.e.*, $r_i^T r_j = \delta_{ij}$ ($\delta_{ij} = 1$, if $i = j$ and $\delta_{ij} = 0$, otherwise).

10. Write the following 2D geometric transformations in homogeneous coordinates: rigid body transformation, afinne transformation and homography.

11. Repeat the previous exercise for 3D transformations.

3 Camera Model

1. Consider a rectangular room of size $2 \times 4 \times 3$ m and a camera located at the center of the ceiling, pointing towards the floor. Define the world coordinate system located at a corner of the floor. Determine the rigid body transformation that converts the world coordinates into the camera coordinates.

2. Suppose you wish to monitor a highway using a camera on the top of a pole. Assume that the camera points towards the highway and the angle between the optical axis of the camera and the pole is 45°. Deine the world coordinate frame and the camera coordinate frame. Determine the rigid body transformation that converts the world coordinates into the camera coordinates.

3. Given a camera with projective model $P = K[R|t]$, (with known K, R, t), determine the coordinates of the camera optical center.

4. A camera has a focal length of 50mm, pixel size of $4 \times 4 \mu m^2$ and principal point $(200, 250) pixel$. Determine the matrix K of intrinsic parameters. Suggestion: convert metric units to m.

5. Consider a camera model in homogeneous coordinates $\tilde{x} \sim P\tilde{p}$ where $\tilde{x} = (x, y, 1)$ and $\tilde{p} = (p_x, p_y, p_z, 1)$. Prove that the model can be written in Cartesian coordinates as follows

$$\begin{cases} x = \frac{\pi_0^T \tilde{p}}{\pi_3^T \tilde{p}} = \frac{P_{00}p_x + P_{01}p_y + P_{02}p_z + P_{03}}{P_{20}p_x + P_{21}p_y + P_{22}p_z + P_{23}} \\ y = \frac{\pi_1^T \tilde{p}}{\pi_3^T \tilde{p}} = \frac{P_{10}p_x + P_{11}p_y + P_{12}p_z + P_{13}}{P_{20}p_x + P_{32}p_y + P_{22}p_z + P_{23}} \end{cases}$$

6. Consider a point p moving on a straight line with constant velocity; assume that p is projected on the image plane by a projective camera. Determine the coordinates of the vanishing point using the projective model of the camera in Cartesian coordinates.

7. Consider a point p moving on a 3D plane, defined by $p_z = 0$ and its projection performed by a projective camera $\tilde{x} \sim P\tilde{p}$. Show that the coordinates of the point in the plane $p = (p_x, p_y)$ are related to the coordinates of the projection (x, y) by a homography $\tilde{x} \sim H\tilde{p}$.

4 Camera Calibration

1. Given a set of 3D points and their projections on the image plane, $(p_i, x_i), i = 1, \ldots, N$, derive an algorithm for the estimation of the camera matrix P, by solving a quadratic least squares problem (DLT calibration method).
2. Explain why the error defined in the previous exercise is not a gometric error and therefore does not have a geometric meaning. Discuss what happens if you wish to minimize the geometric error.
3. Consider a camera matrix P, decomposed into the product of an upper triangular matriz of intrinsic parameters by a matrix of extrinsic parameters $P = K[R|t]$. Show that we can estimate K, R, t from P, using the QR decomposition of matrix calculus.

5 Color

1. Explain trichromatic theory of color.
2. Explain why two light sources with different spectra $S(\lambda), S'(\lambda)$ may produce the same color.
3. Consider two electromagnetic spectra produced by the superposition of two sets of primary sources $S(\lambda) = c_1P_1(\lambda)+c_2P_2(\lambda)+c_3P_3(\lambda)$ and $S'(\lambda) = c'_1P'_1(\lambda)+c'_2P'_2(\lambda)+c'_3P'_3(\lambda)$. Determine the relationship between both sets of coefficients so that both spectra represent the same color.
4. Explain why some colors may not be synthesized by the superposition of three primary colors.

6 Image Processing

7 Feature Detection and Matching

1. Describe Canny edge detection algorithm.
2. Describe Harris corner detector.
3. Describe SIFT keypoints detection.
4. Describe the compuatation of SIFT descritors and explain why are they approximately invariant with respect translation, rotation and scaling.
5. Describe an algorithm for matching SIFT feaatures detected in two different images.
6. Describe the detection of a line in an edge images using the RANSAC algorithm. Assume that you know a set of edge points $x_1, \ldots, x_N$, some of them associated to the line you wish to detect. Discuss the detection of multiple lines.
7. Repeat the previous exercise for the detection of circles.
8. Suppose you wish to detect line segments in an image and specifically you wish to detect the beginning and end of each line. How would you proceed?

8 Image Alignment

1. We wish to align two images based on two sets of matched keypoints, $(x_i, x'_i), i = 1, \ldots, N$. Define a least squares criterion to measure the alignment error. Derive an optimization algorithm for the following transformations:
   (a) translation;
   (b) rigid body transformation;

(c) affine transformation;
(d) homography.

Please discuss which transformations are computed in a single iteration of the algorithm and which transformations require an iterative algorithm.

2. Given a pair of images $I_0(x), I_1(x)$ we wish to transform the second image in order to align it with the first. The corresponding points should have similar intensity. Assuming the geometric transform is a translation $f(x;t) = x + t$, we wish to estimate the parameter t in such a way that $I_0(x) \approx I_1(f(x;t))$.
   (a) define a quadratic cost functional $E(t)$ for this problem;
   (b) discuss if the minimization of E is a linear or a nonlinear optimization problem;
   (c) define an optimization strategy to obtain the global optimum (with a tolerance Δ)
   (d) derive a recursive algorithm to minimize $E(t)$.
3. Compare the algorithms obtained in the previous problem with the Lucas-Kanade algorithm. Discuss the differences.
4. Assume that you have multiple images and you wish to align them all using appropriate geometric transformations $f(x;\theta)$. How would you formulate the problem?
5. Define the optical flow problem. Compare the alignment of a pair of images with the optical flow. What are the similarities and the differences?
6. Derive the optical flow equation $I_x u + I_y v + I_t = 0$.

9 Image Segmentation

1. Describe the region growing algorithm.
2. Describe pixel-based segmentation. Discuss its advantages and disadvantages.
3. Describe the Snake algorithm, assuming a discrete shape model.

10 3D Structure from Motion

1. Consider two calibrated cameras with projective matrices

$$P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad P_1 = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & \sqrt{2} & 0 & 0 \\ -1 & 0 & 1 & 1 \end{bmatrix}$$

   and the projections of a point p on the two image planes $x_0 = [0\ 0]^T, x_1 = [0\ 0]^T$.
   (a) Determine the optical centers of both cameras in world coordinates;
   (b) Determine the coordinates of the 3D point p from the projections x_0, x_1 ;
2. Solve the same problem geometrically
   (a) Draw the camera local axis, p_{cx}, p_{cz}, in the p_x, p_z plane.
   (b) Draw the optical axis associated to x_0, x_1 and intersect them in the p_x, p_z plane. Check if you obtained the same result.
3. Repeat problem 1, assuming that the 3D point p is projected by the two cameras into the points $x_0 = [0.5\ 0]^T, x_1 = [1\ 0]^T$.
4. Determine the essential matrix, E, for the camera matrices P_0, P_1 defined in problem 1. Check if the two pairs of projections $\tilde{x}_0, \tilde{x}_1$ defined above (converted into homogeneous coordinates) verify the fundamental property $\tilde{x}_1^T E \tilde{x}_0^T = 0$.
5. Assuming P_0, P_1 defined before, determine the epipolar lines associated to $\tilde{x}_0 = [0\ 1\ 1]^T$.

11 Object Recognition

1. Suppose you know the sillouettes of objects in images. Explain how can you characterize the silloutte and how can you use this information to separate different types of objects.

2. Describe a face recognition algorithm. Define the features extracted from the image. Explain how to learn the decision mechanism.

3. Explain the bag-of-words method for the analysis of text.

4. Explain the bag-of-features method for the classification of images. Explain what is the difference with respect to the bag-of-words method.