# Homework II

Instructor : Gabriel Taubin*
TA : Silvio Savarese†

## 1 Transformations

**Exercise 1.** An *Euclidean transformation* is a transformation of the plane that preserves the distance. A *Similar transformation* is an Euclidean transformation followed by a scaling (same scaling for each component). Such a transformation can be represented as $p' = Sp$, where $S$ is a 3x3 matrix, $p = [x,\, y,\, 1]^T$ is a point expressed in homogeneous coordinated and $p' = [x',\, y',\, 1]^T$ is its transformation. Consider $S$ as a block matrix and gives an expression and a geometric interpretation of each block. (Hint: start from a vector $e_1 = [1,\, 0]^T$ and see how such vector is transformed according to the similar transformation).

**Exercise 2.** An *Affine transformation* is a non singular linear transformation followed by a translation . Such transformation can be represented as $p' = Ap$, where $A$ is a 3x3 matrix, $p$ and $p'$ are defined as above. Consider $A$ as a block matrix and gives an expression and a geometric interpretation of each block. (Hint: start from 2 vectors $e_1 = [1,\, 0]^T$ and $e_2 = [0,\, 1]^T$ and see how such vectors are transformed according to the affine transformation. If $p = e_1 + e_2$ then the transformation of $p$ is the sum of the transformation of $e_1$ plus the transformation of $e_2$.)

**Exercise 3.** i) Prove that, given 3 non-collinear points $p_a$, $p_b$ and $p_c$ belonging to a plane $\pi$, any point $p \in \pi$ can be written in a unique way as an affine combination of $p_a$, $p_b$ and $p_c$:

$$p = p_a\lambda_a + p_b\lambda_b + p_c\lambda_c, \forall p \qquad (1)$$

where $\lambda_a$, $\lambda_b$, $\lambda_c$ are called the affine coordinates of $p$ with respect to $p_a$, $p_b$ and $p_c$ and where the lambda's satisfy the constraint $\lambda_a + \lambda_b + \lambda_c = 1$.

ii) Consider the set of points defined by positive affine coordinates ($\lambda_a, \lambda_b, \lambda_c > 0$). Locate those points in the plane $\pi$.

iii) Compute the affine coordinates attached to the points $p_a$, $p_b$ and $p_c$.

iv) Give an expression of $\lambda_a$, $\lambda_b$, $\lambda_c$ as function of $p$, $p_a$, $p_b$ and $p_c$.

v) Prove that affine coordinates are invariant with respect to an affine transformation.That is, show that if $p' = Ap$ and $p = p_a\lambda_a + p_b\lambda_b + p_c\lambda_c$, then $p'$ is an affine combination of $p'_a$, $p'_b$ and $p'_c$ with the same $\lambda$'s

**Exercise 4.** A *Perspective transformation* is a non singular linear transformation of homogeneous coordinates. Such transformation can be represented as $p' = Hp$, where $H$ is a 3x3 matrix, $p = [x,\, y,\, 1]^T$ is a point expressed in homogeneous coordinated and $p' = [x',\, y',\, z']^T$ is its transformation. Consider $H$ as a block matrix and gives an expression and a geometric interpretation of each block. (Hint: start from an unitary square (defined by the points $[0,\, 0]^T$, $[1,\, 0]^T$, $[0,\, 1]^T$ and $[1,\, 1]^T$ and see how the unitary square is transformed according to the perspective transformation.)

**Exercise 5.** A 3D scene is distorted under the perspective projection into the image plane. The image in *whiteboard.jpg* does not appear as a rectangular board even though the original is so. The goal of this exercise is to undo the perspective transformation $H$ by computing an inverse transformation $H^{-1}$. If we apply $H^{-1}$ to *whiteboard.jpg* we get a synthesized image which represents the whiteboard in its correct geometric shape. This operation is called *rectification*. Note that we are not interested in estimating a scale factor but only the shape geometry according to its real proportions. The matrix $H$ can be estimated as follows. Let us call $a'$, $b'$, $c'$ and $d'$ the 4 corners of the whiteboard in *whiteboard.jpg* and $a$, $b$, $c$ and $d$ the corresponding corners in the corrected image. $a$, $b$, $c$ and $d$ must be the corners of a rectangle whose height/width ratio is $1.35$. Remember that we do no need the estimate the scale factor. For each pair of corners between the 2 images, we can write the following relationship (in particular for the $1^{st}$ corner):

$$\lambda a' = Ha \qquad (2)$$

where $\lambda$ is multiplicative factor, $a' = [x',\, y',\, 1]^T$, $H$ is the 3x3 perspective transformation matrix, $a = [x,\, y,\, 1]^T$. Furthermore assume that the entry $h_{3,3}$ of $H$ is 1 (what is the geometrical meaning of this assumption?). By writing the above relationship for the $4$ corners and by eliminating $\lambda$, we can get a system of $8$ equations in $8$ unknowns. By solving the system we can estimate $H$.

The matlab code rectificator.m allows you to load an image (*whiteboard.jpg*) and to select the 4 corners $a'$, $b'$, $c'$ and $d'$. Write the routine to compute $H$. Once $H$ is known, synthesize the correct image from *whiteboard.jpg*. That is, write a routine that for each pixel $p$ of the correct image computes the coordinates of the corresponding pixel $p'$ in *whiteboard.jpg* and assigns to $p$ the color information attached to $p'$. This is basically a color mapping transformation. Turn in the code, the estimated $H$ and a print out of the rectified image.

## 2 Calibration

As discussed in class, the calibration procedure allows to estimate intrinsic and extrinsic parameters of the camera in the least square sense. It is important to have a good starting point for the least square minimization process. Exercise 6 and 7 provide a technique to compute a good approximation for both intrinsic and extrinsic parameters which can be used as a starting point for the full calibration procedure.

**Exercise 6.** In this exercise we want to estimate the distortion factor present in the image *chessboard-dist.jpg*. We assume to deal with a radial distortion centered in the middle of the image. Write a dist-estimator.m program in order to: i) load *chessboard-dist.jpg* and select 4 corners from the image (in order to define a rectangle in the real chessboard); ii) use the 4 corners to estimate $H$ (each square of the chessboard has a height/width ratio = 1). iii) by using $H$, map points of the rectified grid (defined by the 4 corners) into points $p_m$ in the distorted image; iv) by using $p_m$, the corresponding grid points of the distorted image and the radial distortion

*taubin@caltech.edu
†savarese@vision.caltech.edu

correction equation used in last lecture, try to estimate $k$ in the least square sense.

**Exercise 7.** In this exercise we perform full camera calibration using the linear estimation method. The result of this estimation can be used as initial estimates for the non-linear least squares minimization process. We denote the projection matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} = \begin{pmatrix} p_1^t & | & p_{14} \\ p_2^t & | & p_{24} \\ p_3 & | & p_{34} \end{pmatrix}$$

where $p_1$, $p_2$, and $p_3$ are 3 dimensional vectors. Since the matrix $P$ is determined modulo a multiplicative factor, it has 11 degrees of freedom. Each calibration point $(x_i, y_i, z_i)^t$ projects onto an image plane point with coordinates $(u_i, v_i)^t$ determined by the equation

$$\lambda_i \begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix} = P \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix}$$

where $\lambda_i$ is a non-zero multiplicative factor. The following equation produces two linear equations with the elements of $P$ as unknowns

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -u_i & -v_i \end{pmatrix} P \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

How many calibration points are needed to estimate P? Estimate P in the least squares sense by imposing the additional linear constraint $p_{34} = 1$. Is it acceptable to impose this additional constraint? If not, what other constraint can be imposed?

Ignoring lens distortion, the projection matrix $P$ can be factorized as follows

$$\begin{aligned} P &= \begin{pmatrix} f_u & 0 & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1^t & | & t_1 \\ r_2^t & | & t_2 \\ r_3^t & | & t_3 \end{pmatrix} \\ &= \begin{pmatrix} (f_u r_1 + c_u r_3)^t & | & f_u t_1 + c_u t_3 \\ (f_v r_2 + c_v r_3)^t & | & f_v t_2 + c_v t_3 \\ r_3^t & | & t_3 \end{pmatrix} \end{aligned}$$

where $r_1$, $r_2$, $r_3$ are 3 dimensional vectors so that

$$R = (\, r_1 \quad r_2 \quad r_3 \,)^t \ ,$$

$t = -RO$, $R$ is the camera orientation matrix, and $O$ is the center of projection of the camera in object coordinates. Prove that

$$P \begin{pmatrix} O \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and use the equation to estimate $O$ from $P$. Prove that

$$\begin{cases} p_3^t(p_1 - c_u p_3) & = & 0 \\ p_3^t(p_2 - c_v p_3) & = & 0 \end{cases}$$

and use the equation to estimate $c_u$ and $c_v$. Prove that if $P$ is normalized so that $|p_3| = 1$, then

$$\begin{cases} |f_u| & = & |p_1 - c_u p_3| \\ |f_v| & = & |p_2 - c_v p_3| \end{cases}$$

and use the equations to estimate $f_u$ and $f_v$. What can you say about the sign of $f_u$ and $f_v$ ?

By using the calibration pattern in *calib.jpg* and the above equations, write a routine in order to estimate the extrinsic ($R$ and $O$) and intrinsic ($f_u$, $f_v$, $c_u$ and $c_v$) parameters. The size of the square in *calib.jpg* is 2x2cm.