Topics in Algebraic Geometry (Mutiple View Geometry, 2020-1) Homework 2

Jemin You

April 30, 2020

Problem 1. Let P be a finite camera and let $\mathsf{x} = \mathsf{P}\mathsf{X}$ be the projection of the point $\mathsf{X} \in \mathbb{R}^3$. Express the back projection of the point x via P in terms of x and the camera center C.

Solution.
Recall that a *general camera* is a projection map $\mathbb{P}^3 \dashrightarrow \pi$ to an embedded plane from a point, called the *center* of the camera. When $\pi \neq \pi_\infty$ (the plane at infinity) and the center $\mathsf{C} \notin \pi_\infty$, the camera is called a *finite camera*.

In this situation, P is represented by a 4×3 matrix up to scaling (of which we denote by the same letter). The first 3 columns of P is linearly independent due to the finiteness conditions. The null space of the linear map L_P (the left multiplication by P) is the space generated by C.

Therefore, $\mathsf{Y} \in \mathbb{R}^3$ projects to x iff $\mathsf{P}(\mathsf{Y} - \mathsf{X}) = 0$ iff

$$\mathsf{Y} = \mu\mathsf{X} + \lambda\mathsf{C}$$

for some $\lambda, \mu \in \mathbb{R}$, $\mu \neq 0$. If we have a X projecting to x, this is the right conclusion. When we do not *a priori* have a X, then one finds such X by using the fact that P has rank 3:

$$\mathsf{X}_0 = \mathsf{P}^t(\mathsf{P}\mathsf{P}^t)^{-1}\mathsf{x}$$

maps to x via P. In this case, we have the final answer in the form

$$\mathsf{Y} = \mu\mathsf{P}^t(\mathsf{P}\mathsf{P}^t)^{-1}\mathsf{x} + \lambda\mathsf{C}$$

where $\lambda, \mu \in \mathbb{R}$, $\mu \neq 0$. Note that since we are using homogeneous coordinates, the above is actually a 1-dimensional family.

□

Problem 2. Let P, P' be finite cameras with image planes π, π' respectively.

(a) Express a homography H satisfying $\mathsf{H}(\mathsf{PX}) = \mathsf{P}'\mathsf{X}$ for all $\mathsf{X} \in \{Z = 0\}$ in terms of internal/external parameters of P, P'.

(b) Explain how $\mathsf{H}(\mathsf{PX})$ is obtained when $Z \neq 0$.

Here, by *internal parameter* we mean the camera calibration matrix K and by *external parameters* the rotation matrix R and the camera center $\widetilde{\mathsf{C}} \in \mathbb{R}^3$.

Solution.
(a) One expresses $\mathsf{X} = \begin{pmatrix} X & Y & Z & W \end{pmatrix}^t$. Then unless one of the C of C' lie on the plane $\{Z = 0\}$, the cameras P, P' restrict to bijective homography between the image planes and the plane $\{Z = 0\}$. As the cases where one of the centers lie in the plane $\{Z = 0\}$ do not have unique solutions, we will exclude those cases and assume that $\mathsf{C}, \mathsf{C}' \notin \{Z = 0\}$. Therefore, there exists a unique homography H satisfying $\mathsf{H}(\mathsf{PX}) = \mathsf{P}'\mathsf{X}$ for all $\mathsf{X} \in \{Z = 0\}$.

As the uniqueness is guaranteed, we just need to find H from some constraints. Letting $\mathsf{X} = 0 \in \mathbb{R}^3$, the origin, we have

$$\mathsf{HP} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \mathsf{P}' \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Using the parameters, we obtain

$$\mathsf{HKR}(-\widetilde{\mathsf{C}}) = \mathsf{K}'\mathsf{R}'(-\widetilde{\mathsf{C}}').$$

On the other hand, the equation for the points at infinity of $\{Z = 0\}$ is

$$\mathsf{HP} \begin{pmatrix} X \\ Y \\ 0 \\ 0 \end{pmatrix} = \mathsf{P}' \begin{pmatrix} X \\ Y \\ 0 \\ 0 \end{pmatrix}.$$

This is equivalent to

$$\mathsf{HKR} \begin{pmatrix} X \\ Y \\ 0 \end{pmatrix} = \mathsf{K}'\mathsf{R}' \begin{pmatrix} X \\ Y \\ 0 \end{pmatrix}.$$

We conclude that

$$\mathsf{HKR} \begin{pmatrix} \mathsf{e}_1 & \mathsf{e}_2 & \widetilde{\mathsf{C}} \end{pmatrix} = \mathsf{K}'\mathsf{R}' \begin{pmatrix} \mathsf{e}_1 & \mathsf{e}_2 & \widetilde{\mathsf{C}}' \end{pmatrix}$$

implying

$$\mathsf{H} = \mathsf{K}'\mathsf{R}' \begin{pmatrix} \mathsf{e}_1 & \mathsf{e}_2 & \widetilde{\mathsf{C}}' \end{pmatrix} \begin{pmatrix} \mathsf{e}_1 & \mathsf{e}_2 & \widetilde{\mathsf{C}} \end{pmatrix}^{-1} \mathsf{R}^{-1}\mathsf{K}^{-1}.$$

Note that we have used the equations of homogeneous coordinates as the equations of genuine numbers in all the equations. However, the uniqueness of the solutions show that this possible solution is the unique solution in homogeneous coordinates. Another way of seeing this is recognizing that the same equations hold for linear combinations of two equations that appeared above: the homogeneous proportionality constant must be the same for all equations for linear relations to be transferred via H.

(b) Let $\mathsf{X} \notin \{Z = 0\}$. We have assumed that $\mathsf{C} \notin \{Z = 0\}$. Then there is a unique intersection

Y of the line $\overleftrightarrow{\mathsf{CX}}$ and the plane $\{Z = 0\}$ in the projective 3-space $\mathbb{P}^3$. As $\mathsf{C}, \mathsf{X}, \mathsf{Y}$ are colinear, $\mathsf{P}(\mathsf{X}) = \mathsf{P}(\mathsf{Y})$. Therefore, one just applies H with X replaced by Y:

$$\mathsf{H}(\mathsf{PX}) = \mathsf{H}(\mathsf{PY}) = \mathsf{P}'\mathsf{Y}.$$

Computing Y is easy: if the Z-coordinates of C, X are the same, we just take the direction vector $\overrightarrow{CX}$ to obtain a point at the line at infinity of $\{Z = 0\}$. Otherwise, one uses the linear parametrization $\lambda\mathsf{C} + \mu\mathsf{X}$ and calculate the ratio $\lambda : \mu$ where $\lambda\mathsf{C} + \mu\mathsf{X} \in \{Z = 0\}$. $\square$

Problem 3 (Exercise 6.5.2 (i))**.** Let I_0 be a projective image and let another camera take image of the image I_0. Denote the composite image by I'. Show that the apparent camera center of I' is the same with that of I_0's, while the other parameters may be different. Speculate on how this explains why a portrait's eyes 'follow you around the room'.

Solution.
Let I_0 be the projective image of $I \subseteq \mathbb{P}^3$ with the center C and the image plane π via camera P, and I' the image of $I_0 \subseteq \pi$ by the projective camera of center C' and the image plane π' via camera P'. We will assume that P' maps π bijectively to π' for our purposes. Then, such restriction is a homography, of which we will denote by H.

Then, the image I' is the result of applying a general projection

$$\mathsf{HP}.$$

Hence, I' is mapped to π' as if it were directly observed by the camera with matrix HP. As

$$\mathsf{HP} = \mathsf{HKR}[I| - \widetilde{\mathsf{C}}],$$

the projective camera HP is a *finite camera* with center C. However, as the 'KR'-decomposition of HKR may be arbitrary, we can only ensure that the camera center is C.

For the case of the portraits, if the drawn individual looked the painter while being painted, the drawn image appears to gaze the painter. This is because the painter drew what 'appeared in his views'. Then, when a different person looks into the painting, the mechanism explained in this problem happens in his eyes. Therefore, the image that the viewer sees appears is the picture that apparently has the center C, which is the painter's position when the painting was drawn. Hence, the viewer sees the image that would have appeared when the painting object would have gazed him, with distortion (variance of other parameters).

$\square$

Problem 4 (Exercise 6.5.2 (ii))**.** Show that the dual Plücker coordinates of the ray back-projected from an image point x by a projective camera P is

$$\mathsf{L}^* = \mathsf{P}^t[\mathsf{x}]_\times \mathsf{P}$$

where

$$[\begin{pmatrix} a \\ b \\ c \end{pmatrix}]_\times = \begin{pmatrix} & c & -b \\ -c & & a \\ b & -a & \end{pmatrix}.$$

Solution.
Let $\mathsf{P} : \mathbb{P}^3 = \mathbb{P}(\mathsf{V}) \to \mathbb{P}(\mathsf{W}) = \pi$ be a general camera projection where the bases of the vector spaces V, W are already chosen. We will denote the pre-projectivized linear map by the same letter P. Then a point $\mathsf{x} \in \pi$ corresponds to a line $l = \mathbb{R}\mathsf{w} \subseteq \mathsf{W}$. As W is 3-dimensional, if we fix a volume element ω, the perfect 'wedge' pairing $\mathsf{W} \otimes \bigwedge^2 \mathsf{W} \to \bigwedge^3 \mathsf{W} = \mathbb{R}\omega$ gives us an identification $\mathsf{W} \simeq (\bigwedge^2 \mathsf{W})^\vee$.

On the other hand, the functoriality of the wedge constructions induce a linear map

$$\mathsf{P}_* : \bigwedge^2 \mathsf{V} \to \bigwedge^2 \mathsf{W}$$

and its dual

$$\mathsf{P}^* : (\bigwedge^2 \mathsf{W})^\vee \to (\bigwedge^2 \mathsf{V})^\vee.$$

We compose this map with the upper one so we have a linear map

$$\mathsf{F} : \mathsf{W} \to (\bigwedge^2 \mathsf{V})^\vee.$$

This projectivizes to

$$\mathbb{P}(\mathsf{F}) : \mathbb{P}(\mathsf{W}) \to \mathbb{P}((\bigwedge^2 \mathsf{V})^\vee).$$

I claim the map $\mathbb{P}(\mathsf{F})$ is the back projection where the image lies on the Grassmannian $\mathrm{Gr}(2, \mathsf{V}$ parametrizing codimension 2 subspaces of V via dual Plücker coordinates. Indeed, if $\mathsf{h} \in \bigwedge^2 V$, we have

$$\mathsf{F}(\mathsf{w})(\mathsf{h})\omega = \mathsf{w} \wedge \mathsf{P}_*(\mathsf{h})$$

from the definitions. As P is surjective, so is P_*, which means that the right-hand side is 0 for h's in a 5-dimensional subspace of $\bigwedge^2 \mathsf{V}$. Such only happens for h's representing subplanes of V (well-known in wedge linear algebra). Therefore, $\mathsf{F}(\mathsf{w})$ represents a plane. If h represents a plane in V, then $\mathsf{F}(\mathsf{w})(\mathsf{h}) = \mathsf{w} \wedge \mathsf{P}_*(\mathsf{h}) \neq 0$ iff $\mathsf{h} + \mathsf{P}^{-1}(\mathbb{R}\mathsf{w}) = \mathsf{V}$, as we need $\mathsf{w} + \mathsf{P}_*(\mathsf{h}) = \mathsf{W}$. This exactly means that $\mathsf{F}(\mathsf{w})$ represents the plane $\mathsf{P}^{-1}(\mathbb{R}\mathsf{w})$.

Finally, let bases of V, W be chosen. Then if $\mathsf{w} = a\mathsf{w}_1 + b\mathsf{w}_2 + c\mathsf{w}_3$, we have

$$\mathsf{w} \wedge \mathsf{w}_1 \wedge \mathsf{w}_2 = c\omega, \mathsf{w} \wedge \mathsf{w}_1 \wedge \mathsf{w}_3 = -b\omega, \mathsf{w} \wedge \mathsf{w}_2 \wedge \mathsf{w}_3 = a\omega$$

where $\omega = \mathsf{w}_1 \wedge \mathsf{w}_2 \wedge \mathsf{w}_3$. The $[\mathsf{x}]_\times$ notation is a matrix representation of this fact: $\mathsf{w} \wedge \mathsf{w}_i \wedge \mathsf{w}_j = ([\mathsf{x}]_\times)_{ij}\omega$. Moreover, if $\mathsf{P}(\mathsf{v}_j) = \sum_i p_{ij}\mathsf{w}_i$, then

$$\begin{aligned} \mathsf{F}(\mathsf{w})(\sum_{k,l} h_{kl}\mathsf{v}_k \wedge \mathsf{v}_l)\omega &= \mathsf{w} \wedge \big(\sum_{k,l} h_{kl}(\sum_i p_{ik}\mathsf{w}_i) \wedge (\sum_j p_{jl}\mathsf{w}_j)\big) \\ &= \sum_{i,j,k,l} p_{ik}h_{kl}p_{jl}\mathsf{w} \wedge \mathsf{w}_i \wedge \mathsf{w}_j \\ &= \sum_{i,j,k,l} p_{ik}h_{kl}p_{jl}([\mathsf{x}]_\times)_{ij}\omega. \end{aligned}$$

Therefore, the transformation behaviour shows that the matrix representation required holds:

$$(\mathsf{L}^*)_{kl} = \sum_{i,j} p_{ik}([\mathsf{x}]_\times)_{ij} p_{jl} = \sum_{i,j} (\mathsf{P}^t)_{ki}([\mathsf{x}]_\times)_{ij}(\mathsf{P})_{jl} = (\mathsf{P}^t[\mathsf{x}]_\times \mathsf{P})_{kl}.$$

This completes the proof.

□

* The following ‘calculus’ proof is more shorter and intuitive, instead of the above linear algebra toolkits involving wedges: since

$$[\mathsf{x}]_\times \mathsf{y} = 0 \Leftrightarrow \mathsf{y} = \mathsf{x},$$

we see that

$$\mathsf{P}^t[\mathsf{x}]_\times \mathsf{P}\mathsf{v} = 0 \Leftrightarrow \mathsf{P}\mathsf{v} = \mathsf{x}$$

(be careful of the homogeneous coordintates!) because P^t has trivial null space.

Therefore, L^* and $\mathsf{P}^t[\mathsf{x}]_\times \mathsf{P}$ have the same null space, and both are dual Plücker coordinates as both of the ranks are 2. Hence, the two must be the same, as the null space represents the points in the plane for such matrices.

Problem 5 (Exercise 7.5.2). (i) Given 5 world-to-image point correspondences $\mathsf{X}_i \leftrightarrow \mathsf{x}_i$, show that there are in general four solutions for a camera matrix P *with zero skew* where $\mathsf{P}(\mathsf{X}_i) = \mathsf{x}_i$.

(ii) Given 3 world-to-image point correspondences $\mathsf{X}_i \leftrightarrow \mathsf{x}_i$, show that there are in general four solutions for a camera matrix P *with known calibration* K where $\mathsf{P}(\mathsf{X}_i) = \mathsf{x}_i$.

Solution.
The situations are identical in both cases: let

$$\mathsf{P} = \mathsf{K}\mathsf{R}[I| - \widetilde{\mathsf{C}}]$$

be the camera matrix, where the calibration matrix has the form

$$\mathsf{K} = \begin{pmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{pmatrix}.$$

The world-to-image correspondence gives the constraints

$$\mathsf{x} \times \mathsf{P}\mathsf{X} = 0$$

which give *two* linear relations for elements of P. As $\mathsf{K}, \mathsf{R}, \widetilde{\mathsf{C}}$ have the degrees of freedoms $5, 3, 3$ respectively, P have 11. Therefore, $11/2$ pairs of world-to-image points in general positions are required to compute P.

In (i), we have the additional constraint $s = 0$ which decreases the degree of freedom to 10. We will then need 5 pairs in general. In (ii), we already know K so we have the degree of freedom 6. We will then need 3 pairs.

If the constraints given for P was *linear* in P's entries, we would have had 1 exact solution. The issue is that *the given relations are non-linear.* We will look into this separately.

(ii) We have

$$[\mathsf{R}| - \mathsf{R}\widetilde{\mathsf{C}}] = \mathsf{K}^{-1}\mathsf{P}$$

so fixing the calibration amounts to giving the rotational relations for the first three columns of $\mathsf{K}^{-1}\mathsf{P}$. If the first three columns are $\mathsf{u}, \mathsf{v}, \mathsf{w}$, the rotation relation is

$$\mathsf{u}^t\mathsf{v} = \mathsf{u}^t\mathsf{w} = \mathsf{v}^t\mathsf{w} = \mathsf{u}^t\mathsf{u} - \mathsf{v}^t\mathsf{v} = \mathsf{u}^t\mathsf{u} - \mathsf{w}^t\mathsf{w} = 0.$$

When three pairs of world-to-image correspondences are given, we are then left with six linear relations on entries of $[\mathsf{u}\ \ \mathsf{v}\ \ \mathsf{w}|\mathsf{t} = -\mathsf{R}\widetilde{\mathsf{C}}]$. As t have the degree of freedom 3, what remains is 3 linear relations on the R part.

Therefore, the problem boils to finding rotations with given linear relations up to scale. As the orthogonality was defined by 5 equations $\mathsf{u}^t\mathsf{v} = \mathsf{u}^t\mathsf{w} = \mathsf{v}^t\mathsf{w} = \mathsf{u}^t\mathsf{u} - \mathsf{v}^t\mathsf{v} = \mathsf{u}^t\mathsf{u} - \mathsf{w}^t\mathsf{w} = 0$, it suffices to show that there are examples of three linear relations that give 4 solutions (This is well-known in intersection theory: when such thing happens, intersection product of the two cycle classes corresponding to the codimension 3 linear subspace and the orthogonal subspace in the Chow ring of $\mathbb{P}^8$ is $4 \in CH_0(\mathbb{P}^8)$. If the number is actually realized by an example, it implies that the number is realized for a general choice of codimension 3 linear subspace. The theorem works for).

Here is an example: we use three linear relations to fix

$$\mathsf{R} = \begin{pmatrix} 3 & 5 & * \\ 4 & a & * \\ 5 & b & * \end{pmatrix}.$$

The last column is the exterior product of the first to columns up to fixed scaling, so there are two solutions for each solution of the second column. The second column has two solutions since it has to satisfy

$$a^2 + b^2 = 25, 15 + 4a + 5b = 0.$$

(i) We have

$$\mathsf{K} = \begin{pmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{pmatrix}$$

in the skew zero case. Then,

$$\mathsf{P} = \mathsf{K}[\mathsf{R}|\mathsf{t}] = \left[\begin{pmatrix} \alpha_x \mathsf{R}^1 + x_0 \mathsf{R}^3 \\ \alpha_y \mathsf{R}^2 + y_0 \mathsf{R}^3 \\ \mathsf{R}^3 \end{pmatrix} \middle| \begin{pmatrix} \alpha_x \mathsf{t}^1 + x_0 \mathsf{t}^3 \\ \alpha_y \mathsf{t}^2 + y_0 \mathsf{t}^3 \\ \mathsf{t}^3 \end{pmatrix} \right].$$

Now, we can recover x_0, y_0 from the inner product of the rows of KR. Hence, giving a matrix of the right-hand side form lets us recover

$$\left[\begin{pmatrix} \alpha_x \mathsf{R}^1 \\ \alpha_y \mathsf{R}^2 \\ \mathsf{R}^3 \end{pmatrix} \middle| \begin{pmatrix} \alpha_x \mathsf{t}^1 \\ \alpha_y \mathsf{t}^2 \\ \mathsf{t}^3 \end{pmatrix} \right].$$

Then, by comparing the size of the rows of the left part, we recover the positive numbers α_x, α_y. From those, other elements R, t are immediately recovered.

When 5 pairs of world-to-image correspondences are given, P is brought down to a 1-dimensional family $p_{ij} = p_{ij}(r)$. The numbers x_0, y_0 are then given by $x_0 = (\mathsf{KR}^1)^t \mathsf{R}^3 ((\mathsf{R}^3)^t \mathsf{R}^3)^{-1}, y_0 = (\mathsf{KR}^2)^t \mathsf{R}^3 ((\mathsf{R}^3)^t \mathsf{R}^3)^{-1}$. As these are fractions with quadratic terms in both the denominator and the numerator, we involve solving a quartic equation of r to calculate the necessary parameters: the rows $\alpha_x \mathsf{R}^1, \alpha_y \mathsf{R}^3$ should be orthogonal, so we have a sum of product of such fractions $= 0$. Therefore, we have 4 solutions in general.

$\square$

Problem 6 (Exercise 7.5.2 (iii))**.** Find a linear algorithm for computing the camera matrix P under each of the following conditions:

(a) The camera location, but not the orientation, is known.

(b) The direction of the principal ray of the camera is known.

(c) The camera location and the principal ray of the camera are knwon.

(d) The camera location and the complete orientation of the camera are known.

(e) The camera location and the orientation are known, as well as some subsets of the internal camera parameters are known.

Solution.
Recall that

$$\mathsf{P} = \mathsf{KR}[I| - \widetilde{\mathsf{C}}]$$

where

$$\mathsf{K} = \begin{pmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{pmatrix}.$$

A single pair of world-to-image correspondence gave the relation

$$\mathsf{PX} \times \mathsf{x} = 0$$

which reduces the degree of freedom of P by 2. When there are no a priori conditions, we need $11/2$ general pairs. Using a half pair means that we are using one of the two linear relations arising from the above exterior product relation.

(a) Since we know the camera location $\widetilde{\mathsf{C}}$, the world-to-image correspondence reduces to homography KR. Since the degree of freedom is 8 for this matrix, one finds 4 points where the directions from $\widetilde{\mathsf{C}}$ are general, then find KR from the constraints

$$\mathsf{KR}(\widetilde{\mathsf{X}} - \widetilde{\mathsf{C}}) \times \mathsf{x} = 0.$$

Then since the 'KR'-decomposition is unique, one can find the matrices K, R from KR.

(b) The direction of the principal ray is the first three entries of the third row of P. After fixing an exact tuple numbers amongst the possible homogeneous coordiantes, 9 degrees of freedom are left. Therefore, we need $9/2$ general pairs of world-to-image correspondences.

This process can be made more explicit in this case: one first finds $\mathsf{X} \in \mathbb{R}^3$ so from the third component of the equation

$$\mathsf{x} = \mathsf{PX}$$

we can fix the third row completely (de-homogenized, after fixing first three numbers exactly). Such amounts to the half pair part.

Afterwards, we find 4 general world-to-image correspondences where X does not lie in the null space of the third row of P. Then the relations we obtain are

$$\mathsf{x}^i = \frac{\mathsf{x}^3}{\mathsf{P}^3\mathsf{X}} \mathsf{P}^i X$$

in *exact numbers.*

(c) If we know $\widetilde{\mathsf{C}}$ and the first three numbers of the third row of P (up to scale), then since

$$\mathsf{P} = [\mathsf{KR} | - \mathsf{KR}\widetilde{\mathsf{C}}],$$

we know the third row of KR and the vector $\widetilde{\mathsf{C}}$. Fix (de-homogenize) the third row of KR. Since we know also the center $\widetilde{\mathsf{C}}$, one applies the dehomogenized world-to-image correspondence relations

$$\mathsf{x}^i = \frac{\mathsf{x}^3}{\mathsf{KR}^3(\widetilde{\mathsf{X}} - \widetilde{\mathsf{C}})} \mathsf{KR}^i(\widetilde{\mathsf{X}} - \widetilde{\mathsf{C}})$$

to obtain 2 relations for each pairs. Since there are 6 unknown numbers, 3 pairs suffice to find the remaining numbers.

(d)(e) In these cases, we know R, C so we need $5/2$ pairs of world-to-image correspondences, as we apply the homography correspondences

$$\mathsf{K}(\mathsf{R}(\widetilde{\mathsf{X}} - \widetilde{\mathsf{C}})) \times \mathsf{x} = 0$$

to find K. When n linear conditions between the 5 parameters of K are known, we reduce the required pairs to $(5 - n)/2$ pairs. $\square$