POLITECNICO
MILANO 1863

Camera Calibration

Lecture 11

Re-projection to 3D space from sensor(s) undistorted data and sensor center corrected: collinearity equations (7 parameters)

$$\tilde{\mathbf{p}} = \mathbf{K}\mathbf{R}^T[\mathbf{I}|-\mathbf{C}]\,{}^g\tilde{\mathbf{P}}$$

Note: $C=X_0$
${}^c x = x$ on the camera

$$\mathbf{K}\mathbf{R}^T[\mathbf{I}|-\mathbf{X}_0] = \lambda \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & -R\mathbf{X}_O \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix}$$

first element of the rotation matrix

translated X

same on Y

*original point is traslated, rotated, and finally scaled

$${}^c x = c\frac{r_{11}(X - X_O) + r_{12}(Y - Y_O) + r_{13}(Z - Z_O)}{r_{31}(X - X_O) + r_{32}(Y - Y_O) + r_{33}(Z - Z_O)}$$

2nd eq

$${}^c y = c\frac{r_{21}(X - X_O) + r_{22}(Y - Y_O) + r_{23}(Z - Z_O)}{r_{31}(X - X_O) + r_{32}(Y - Y_O) + r_{33}(Z - Z_O)}$$

third equation at the denominator

*Z rappresentata come coordinata dal punto di vista della CAMERA. Infatti poi viene yc = Ycam * (c/Z)

Parameter estimation: DLT – classical 9 parameters coll. eqn.

Marzan, G.T. & Karara, H.M. (1975). A computer program for direct linear transformation solution of the collinearity condition, and some applications of it. *Proceedings of the Symposium on Close-Range Photogrammetric Systems (pp. 420-476)*. Falls Church, VA: American Society of Photogrammetry.

Considering only **c** and $\mathbf{x_0}$, $\mathbf{y_0}$ as parameters, the collinearity equations can be written as:

*principal points?

$$x = x_0 - c\frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

*proiettati

$$y = y_0 - c\frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

anche "+" però diobello

Rewrite collinearity equations re-grouping known parameters

same equation of slide before

$$x = \boxed{x_0} - \boxed{c}\,\frac{\boxed{r_{11}}(X - \boxed{X_0}) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

$$y = y_0 - c\,\frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

DLT: Direct Linear Transform

I call it l1, it's just one parameter

$$x = \frac{l_1 X + l_2 Y + l_3 Z + l_4}{l_5 X + l_6 Y + l_7 Z + \boxed{1}}$$

$$y = \frac{l_8 X + l_9 Y + l_{10} Z + l_{11}}{l_5 X + l_6 Y + l_7 Z + 1}$$

*proiettati

This is a NL problem to be solved, I'd ike to obtain 3 angles (omega phi key) of the rotation matrix, and the translation, and focal length, and position of camera center (principal point) I'd like to compute them, but it's a big NL system!
One way is to use this mathematical reelaboration.
I take everything is multiplying X (r11 by c) and I call it l1, it's just one parameter. Then also for all the others, Y and Z. And for all the equations.

I obtain this eq. linear into the parameters "l", not linear in x0, r, etc. but linear in these new pararmeters. This way I have no longer any physical meaning. Before I had elements of translation (y0) rotation (r) focal point (..). Now no meaning at all (11 parameters in l), but the eqs are linear, and if I compute these parameters then I can use them to invert the projection problem and try to estimate the 3D coordinates in space.

Easy way to compute the eleven unknown parameters

I can rearrange these equations in order to put in evidence the known points (x, y, z) (and l1?). For each points I write two equations like this.11 unknowns. In the end if I "survey" 5 points I'll have 10 equations. If I use 9 it's a linear system of eqs that can be solved easily. But ofc you'll have noises.. if you increase the number of points you'll have better results (better accuracy). I mean, more than 5 points. Increasing the number of eqs your linear system won't have a square coefficient matrix but a rectangular so probably no solution to that problem, but you can impose a constrain: the sum of the overall square error is minimized. Indeed if you have more than 9 eqs the others will be conflicting, if you have noise. So you say: its' not equal to zero the set of eqs, it's equal to an error. Error related to that eqs that it's not perfect, because the system is not perfect. I'll put the residual error there, and I'll try to minimize the invert product of that error vector. wt*w, scalar value, sum of all the squared errors

$$(l_1 - l_5 x)X + (l_2 - l_6 x)Y + (l_3 - l_7 x)Z + l_4 - x = 0$$

[=w1] error, in case of >9 eqs

$$(l_8 - l_5 y)X + (\boxed{l_9} - l_6 y)Y + (l_{10} - l_7 y)\boxed{Z} + l_{11} - \boxed{y} = 0$$

to be estimated

known during calibration

known point on camera plane (*sensor?)

For each point surveyed, 2 equations are given in the 11 unknowns l_i; 6 points are enough to determine the parameters. A least squares approach can be used:

coefficienti di L, per esempio per L1 abbiamo X

$$A[L]=B$$

either x, or x+w1

it gives you the unknown terms: the parameters. You can estimate this way the unknown parameters of the collinearity eqs: with some points of known coordinates

$$[L]=(A^TA)^{-1}A^TB$$

solution: instead of using the inverse matrix we use the pseudo inverts

Inverting DLT for space intersection (X, Y, Z from x,y,L)

*camera a

$$(l_{1a} - l_{5a} x_a)\boxed{X} + (l_{2a} - l_{6a} x_a)\boxed{Y} + (l_{3a} - l_{7a} x_a)\boxed{Z} + l_{4a} - x_a = 0$$

$$(l_{8a} - l_{5a} y_a)X + (l_{9a} - l_{6a} y_a)Y + (l_{10a} - l_{7a} y_a)Z + l_{11a} - y_a = 0$$

these now are the known parts!

$$(l_{1b} - l_{5b} x_b)X + (l_{2b} - l_{6b} x_b)Y + (l_{3b} - l_{7b} x_b)Z + l_{4b} - x_b = 0$$

$$(l_{8b} - l_{5b} y_b)X + (l_{9b} - l_{6b} y_b)Y + (l_{10b} - l_{7b} y_b)Z + l_{11b} - y_b = 0$$

*camera b

With at least two cameras, 4 equations in 3 unknowns can be solved by least squares technique

$$A\,[X,Y,Z]=B$$

*known l coefficients and projections

$$[X,Y,Z]=(A^TA)^{-1}A^TB$$

DLT problems

This equation was called originally (60s) DLT, direct linear transform. Because we have converted the NL problem into L problem.

drawbacks: parameters l have no phyisical meaning :(they are translation/rotation/focal length. Is some knd of abstract parameter, if you do the calibration of a camera and you look at the parameter (after), you can see if you have done errors. You cannot make obvious tests with these parameters.

When you have more parameters then d.of. the system will work well close the points of where you have estimated parameterse, if you go away from these anything can happen. Usually these overdetermined system tries to use their dof (11 par) in order to optimize very well in the points in which they can actually compute the error, but this increases the error outside.

L parameters have not a physical meaning. Quite hard to identify calibration failure from those parameters (easyer if parameters have a physical meaning as focal length, position in space, orientation)

11 parameters are estimated from 9 real degrees of freedom. Parameters can correlate well on observed points, but go ‘weard’ outside.

Back to our days. DLT mapping (11 parameters).

The homogeneous projection matrix

qst non è DLT(2) ma è DLT(1) considerando tutti gli 11DoF, che semplicemente risulta matematicamente equivalente

$$\mathsf{P} = \mathsf{K}R[I_3 | -\boldsymbol{X}_O]$$ Here $X_0 = C$

contains **11 parameters**

- 6 extrinsic parameters: $R, \boldsymbol{X}_O$
- 5 intrinsic parameters: c, x_H, y_H, m, s

Euclidean world:

$$^s x = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + \boxed{p_{34}}} \qquad x = \frac{l_1 X + l_2 Y + l_3 Z + l_4}{l_5 X + l_6 Y + l_7 Z + 1}$$

$$^s y = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}} \qquad y = \frac{l_8 X + l_9 Y + l_{10} Z + l_{11}}{l_5 X + l_6 Y + l_7 Z + 1}$$

these are no longer the elements of the rotation matrix, but just parameters. The same as those we computed for DLT, but here we have p34.

Divide by p_{34}

indeeed we are working in a projection world. We obtain the same formulation as before

Inversion - Goal: map from ${}^{a}\mathbf{x}$ back to $\mathbf{X}$

1st step: ${}^{a}\mathbf{x} \rightarrow {}^{s}\mathbf{x}$ from sensor to ideal proj.

2nd step: ${}^{s}\mathbf{x} \rightarrow \mathbf{X}$ from ideal proj. to space

How do we invert the eqs to compute starting from these coordinates the 3D coordinates?

2 step:
1st: solving the problem of distortion correction. Once the projection is linear, you can project the points into the space.

If we use only one camera and try to make these projections we are not addressing exactly the measurements in one point in space but we ar emeasuring something different

UN altro motivo sul perché con una camera non siamo in grado di ricostruire. Io penso che si potesse fare LSqmethod come con DLT da 9 dof semplicemente.

From ideal projection to space (Inversion of the projective mapping)

[23.00] -> what?

We cannot reconstruct the 3D point but the ray though the 3D point
With the known matrix **P**, we can write

*[xs ys 1], IDEAL projections

$$
\begin{aligned}
\lambda\tilde{\mathbf{x}} &= \mathsf{P}\mathbf{X} = \mathsf{K}R[I_3 | -\boldsymbol{X}_O]\tilde{\mathbf{X}} \\
&= [\mathsf{K}R | -\mathsf{K}R\boldsymbol{X}_O] \begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix} \\
&= \mathsf{K}R\boldsymbol{X} - \mathsf{K}R\boldsymbol{X}_O
\end{aligned}
$$

translation (was c)

calibration matrix K

rotation matrix

Finding the ray through X

$$\begin{aligned}\lambda\tilde{\mathbf{x}} &= \mathsf{P}\mathbf{X} = \mathsf{K}R[I_3| - \boldsymbol{X}_O]\tilde{\mathbf{X}} \\ &= [\mathsf{K}R| - \mathsf{K}R\boldsymbol{X}_O]\begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix} \\ &= \mathsf{K}R\boldsymbol{X} - \mathsf{K}R\boldsymbol{X}_O\end{aligned}$$

knowing x lets you compute only the line that passes through X0

*in the global coordinates

General X determined by a projection.

$$\begin{aligned}\boldsymbol{X} &= (\mathsf{K}R)^{-1}\mathsf{K}R\boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x} \\ &= \boldsymbol{X}_O + \lambda(\mathsf{K}R)^{-1}\mathbf{x}\end{aligned}$$

it starts from X0 and goes in a given direction, given by projection and rotation matrix

varying lambda you go all over the line

gives direction

λ Defines a ray passing through $\mathbf{X_0}$ and $\mathbf{x}$

"definitions you can read them"

DEFINITIONS:

- If the intrinsics are **unknown**, we call the camera **uncalibrated**
- If the intrinsics are **known**, we call the camera **calibrated**
- If the intrinsics are known and do not change, it is a **metric camera**
- The process of obtaining the intrinsics is called **camera calibration**

Recap calibration matrices

all possible approaches in which you increase the number of parameters.

camerea coincides with the ref. frame, and unit focal lenght, no skew, right position of center

D.O.F.

unit $\quad {}^{0}\mathsf{K} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ 6 (6+0)

Ideal but with focal length

ideal $\quad {}^{k}\mathsf{K} = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ 7 (6+1)

you add the sensor position

Euclidean $\quad {}^{p}\mathsf{K} = \begin{bmatrix} c & 0 & x_H \\ 0 & c & y_H \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ 9 (6+3)

inlcuding scaling

affine $\quad \mathsf{K} = \begin{bmatrix} c & cs & x_H \\ 0 & c(1+m) & y_H \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ 11 (6+5)

general $\quad {}^{a}\mathsf{K} = \begin{bmatrix} c & cs & x_H + \Delta x \\ 0 & c(1+m) & y_H + \Delta y \\ 0 & 0 & 1 \end{bmatrix}$ $\quad$ 11+N

also distortion correction

modern approach
that's similar to DLN

Camera Calibration:

Direct Linear Transform (2) and Zhang's Method

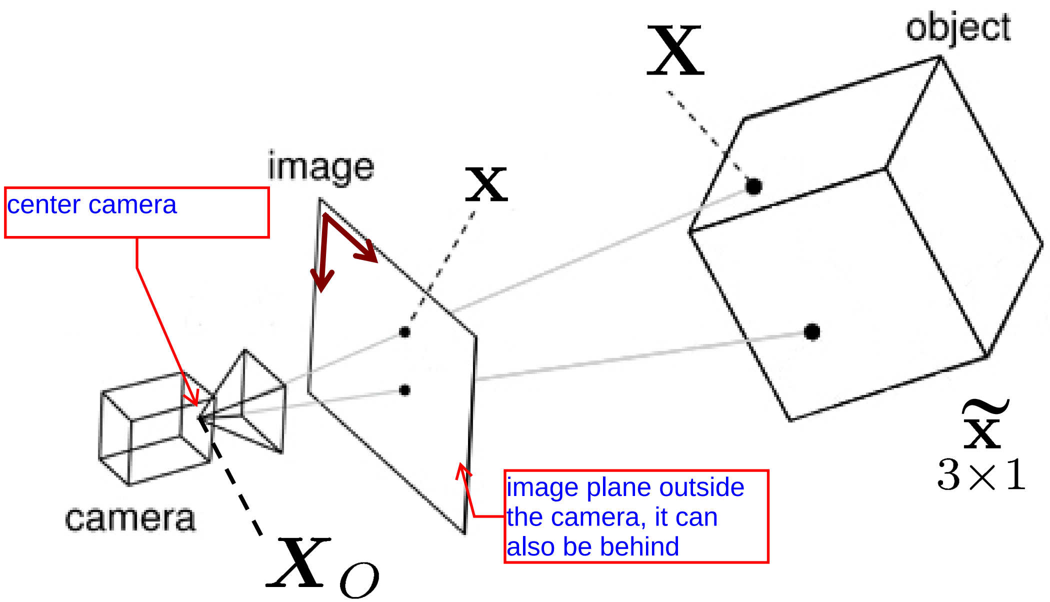
Images of checker board, we can observe the xy coordinate in the image plane.

Wanted: Extrinsic and intrinsic parameters of a camera
Given: Coordinates of object points (control points) checker
Observed: Coordinates x, y of those object points in an image

Mapping

Direct linear transform (DLT) maps any object point **X** to the image point x

Projection on a plane of the point over an object, project lowercase x toward the center. Already discussed.



$$\underset{3\times1}{\tilde{\mathbf{x}}} = \underset{3\times3}{\mathsf{K}}\ \underset{3\times3}{R}\ \underbrace{\left[\underset{3\times3}{I_3} \,\middle|\, -\underset{3\times1}{\boldsymbol{X}_O}\right]}_{3\times4}\ \underset{4\times1}{\tilde{\mathbf{X}}}$$

$$= \underset{3\times4}{\mathsf{P}}\ \underset{4\times1}{\tilde{\mathbf{X}}}$$

Spatial resection

Summary of what's needed and what's given.

Spatial resection: find the position of the camera in space.

Then space intersectioin: computing coordinates of point in 3D starting from stereometric system.

$$\tilde{\mathbf{x}} = KR[I_3| - \boldsymbol{X}_O]\tilde{\mathbf{X}}$$

observed image point

c, s, m, x_h, y_h

3 rotations

3 translations

control point coordinates (given)

Spatial Resection: given the intrinsic parameters, compute the **6 extrinsic parameters**

c, s, m, x_h, y_h (given)
X0, Y0, Z0, ω, ϕ, κ (needed)

How Many Points Are Needed?

already made this considerations

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathsf{P} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Each point gives **two** observation equations, one for each image coordinate

$$x = \frac{p_{11}X + p_{12}Y + p_{13}Z + p_{14}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

$$y = \frac{p_{21}X + p_{22}Y + p_{23}Z + p_{24}}{p_{31}X + p_{32}Y + p_{33}Z + p_{34}}$$

We are trying to estimate these.

Rearringing of collinearity eqs of the extended camera calibration matrix.

Space Resection vs. DLT

We are trying to estimate these [omography] coefficient of transformation between 3D space to 2D space on the image plane. If we use a calibrated camera we have 6 uknowns (just a position in space), meaning we need at least 3 points. 11 unknown includes position of center of the sensor, the skew, etc.

it means we know the 5 intrinsics.

Calibrated camera

- 6 unknowns
- We need at least **3 points**
- Problem solved by **spatial resection**

just position in space of the ref frame of the camera

to do spatial resection, knowing the position of the camera to the global ref frame, we only need at least 3 point in space

this includes parameters like the position of the center of the sensor, the skew etc.

Uncalibrated camera

- 11 unknowns (distortions supposed to be corrected – N parameters)
- We need at least **6 points** each point gives us couple of equations
- Assuming the model of an **affine camera** rederistibuted parameters (pij elements)
- Problem solved by **DLT** slightly different method for computing the parameters to respect the simple LSQ system. In that case you are able to compute the parameters with no meaning.. wIth this dlt you compute parameters with a physical meaning.

DLT (2) solution

Find 11 elements of P by surveying at least N>=6 known points P in space and measuring x, y coordinates of projections on the sensor.

$$\mathbf{x}_i = \mathsf{P}\,\mathbf{X}_i \qquad i = 1, \ldots, I$$

global point, homogeneous coordinates

affine transformation in matrix form

$$\mathbf{x}_i = \underset{3\times 4}{\mathsf{P}}\,\mathbf{X}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \mathbf{X}_i$$

it means to estimate these. Not really 12 but 11 since we are in the homogeneous space. We can normalize dividing everything by p34

We can solve the problem of estimating parameters using DLN, meaning eestimate them. Not 12, but only 11 since we are in the homogeneous space, we can normalize the 4x3 matrix

DLT (2) solution

This is where we start from

The big picture

this part here is related to the position of camera in space, P0 is the translation of the camera

$$\tilde{\mathbf{p}} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x(p,q) \\ 0 & 1 & \Delta y(p,q) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & s & x_c \\ 0 & 1+m & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{R}\left[\mathbf{I}_3 \mid -\mathbf{P}_0\right] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

rotation matrix

Known

chained multiplications gives rise to our "K", calibration matrix

Rearrange DLT equations

[X Y Z 1]

projection

At

$$\mathbf{x}_i = \underset{3\times 4}{\mathsf{P}}\,\mathbf{X}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \mathbf{X}_i$$

first line

$$= \begin{bmatrix} \mathbf{A}^\top \\ \mathbf{B}^\top \\ \mathbf{C}^\top \end{bmatrix} \mathbf{X}_i$$

Solve the problem: we rearrange the equations. I'll call At the first line, Bt the second etc.

still a 3x4, but I grouped each vectors in the variables

in homogeneous world

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = \begin{bmatrix} \mathbf{A}^\top \mathbf{X}_i \\ \mathbf{B}^\top \mathbf{X}_i \\ \mathbf{C}^\top \mathbf{X}_i \end{bmatrix}$$

NOT the X coordinate, it is a vector XYZ1, point in the homogeneous space

btw this is the focal length, if we want to know the physical relationship. The position of the image plane in to the ref frame of the camera

Rearrange the DLT equations

$$\mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = \text{lambda*} \begin{bmatrix} \mathbf{A}^\top \mathbf{X}_i \\ \mathbf{B}^\top \mathbf{X}_i \\ \mathbf{C}^\top \mathbf{X}_i \end{bmatrix}$$

[new slide]

[audio]

To eliminate this arbitrary coefficient I divide.. (I obtain 3 equations ui = AtXi and so on. wi is the position of the image plane into the ref. frame of the camera = focal length).

I'm solving for x, I get rid of the lambda factor dividing for wi

in the homogeneous space

I can rearrange the equation:

$$x_i = \frac{u_i}{w_i} = \boxed{\frac{\mathbf{A}^\top \mathbf{X}_i}{\mathbf{C}^\top \mathbf{X}_i}} \Rightarrow x_i \mathbf{C}^\top \mathbf{X}_i - \mathbf{A}^\top \mathbf{X}_i = 0$$

[5.00] Just like slides before, to get rid of lambda I divide.

$$y_i = \frac{v_i}{w_i} = \boxed{\frac{\mathbf{B}^\top \mathbf{X}_i}{\mathbf{C}^\top \mathbf{X}_i}} \Rightarrow y_i \mathbf{C}^\top \mathbf{X}_i - \mathbf{B}^\top \mathbf{X}_i = 0$$

Linear system of equations (remind DLT(1) and collinearity eqns.)

Linear equations system

we can build a linear system of equations

$$x_i \mathbf{C}^\top \mathbf{X}_i - \mathbf{A}^\top \mathbf{X}_i = 0$$
$$y_i \mathbf{C}^\top \mathbf{X}_i - \mathbf{B}^\top \mathbf{X}_i = 0$$

fin ora direi sia letteralmente uguale a DLT(1), se non che abbiamo p11 -> p34; e non abbiamo diviso per p34

with the 2 eq. I can build the linear system of equations

0: coeff of B for the first eq.

$$\begin{array}{cccl} -\mathbf{X}_i^\top \boxed{\mathbf{A}} & \boxed{0} & & +x_i \mathbf{X}_i^\top \boxed{\mathbf{C}} = 0 \\ \boxed{0} & & -\mathbf{X}_i^\top \boxed{\mathbf{B}} & +y_i \mathbf{X}_i^\top \boxed{\mathbf{C}} = 0 \end{array}$$

coefficient of A

No "A"

vectors of 4 elements each one

linear in the unknown parameters A, B and C

vector of 4 elements each one

Estimate the elements of P

Create a full unkown vector: I put A B and C in column, I obtain a 12 element vector, called p. It represents the unknowns of the problem.

Collect the elements of P within a vector p

$$\boldsymbol{p} = (p_k) = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \end{bmatrix} = \mathrm{vec}(\mathsf{P}^\mathsf{T})$$

rows of P as column-vectors, one below the other (12x1)

Rewrite

$$-\mathbf{X}_i^\mathsf{T}\mathbf{A} \qquad\qquad +x_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0$$
$$-\mathbf{X}_i^\mathsf{T}\mathbf{B} \quad +y_i\,\mathbf{X}_i^\mathsf{T}\mathbf{C} = 0$$

as:

$$\boldsymbol{a}_{x_i}^\mathsf{T}\boldsymbol{p} = 0$$
$$\boldsymbol{a}_{y_i}^\mathsf{T}\boldsymbol{p} = 0$$

$$\begin{aligned}
\boldsymbol{p} &= (p_k) = \mathrm{vec}(\mathsf{P}^\mathsf{T}) \\
\boldsymbol{a}_{x_i}^\mathsf{T} &= (-\mathbf{X}_i^\mathsf{T}, \boxed{\mathbf{0}^\mathsf{T}}, x_i\mathbf{X}_i^\mathsf{T}) \\
&= (-X_i, -Y_i, -Z_i, -1, 0, 0, 0, 0, x_iX_i, x_iY_i, x_iZ_i, x_i) \\
\boldsymbol{a}_{y_i}^\mathsf{T} &= (\boxed{\mathbf{0}^\mathsf{T}}, -\mathbf{X}_i^\mathsf{T}, y_i\mathbf{X}_i^\mathsf{T}) \\
&= (0, 0, 0, 0, -X_i, -Y_i, -Z_i, -1, y_iX_i, y_iY_i, y_iZ_i, y_i)
\end{aligned}$$

expanded

we can rewrite the equations in function of the vector p (*unknowns)

alfa x and alfa y are coefficient of the linear system of equation (Xit, 0, xiXit) *known values

Estimate the elements of P

So we have a new system, with a single vector of unkonwns.
For each equation(in this case x coordinate on image plane):
12 coefficients which multiply a vector of 12 unknowns.
You need to have at least 12 equations to solve:
you must image at least 6 points, each gives you 2 eq.

$$
\begin{array}{c}
\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}
\end{array}
\begin{array}{c} \\ \boldsymbol{A} \\ \\ \\ \\ \boldsymbol{B} \\ \\ \\ \\ \boldsymbol{C} \\ \\ \end{array}
$$

$$
\boldsymbol{a}_{x_i}^{\top}\boldsymbol{p} = 0 = (\boxed{-X_i,\ -Y_i,\ -Z_i,\ -1},\ \boxed{0,\ 0,\ 0,\ 0},\ \boxed{x_iX_i,\ x_iY_i,\ x_iZ_i,\ x_i})
$$

$$
= \left(\quad -\mathbf{X}_i^{\top}, \quad \mathbf{0}, \quad x_i\mathbf{X}_i^{\top} \quad \right) \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{bmatrix}
$$

Estimate the elements of P by I observations

$$\boldsymbol{a}_{x_i}^{\mathsf{T}}\boldsymbol{p} = 0$$
$$\boldsymbol{a}_{y_i}^{\mathsf{T}}\boldsymbol{p} = 0$$

I can write this for each i point

$$\begin{bmatrix} \boldsymbol{a}_{x_1}^{\mathsf{T}} \\ \boldsymbol{a}_{y_1}^{\mathsf{T}} \\ \cdots \\ \boldsymbol{a}_{x_i}^{\mathsf{T}} \\ \boldsymbol{a}_{y_i}^{\mathsf{T}} \\ \cdots \\ \boldsymbol{a}_{x_I}^{\mathsf{T}} \\ \boldsymbol{a}_{y_I}^{\mathsf{T}} \end{bmatrix} \boldsymbol{p} = \underset{2I\times 12}{\mathsf{M}} \; \underset{12\times 1}{\boldsymbol{p}} = 0$$

If the number of points is greater then 6 I have an overdetermined system. If there was no noise it would be ok, instead with noise on each line of the equation, you can, if you have more eqs, minimize the sum of the square error that you have. Meanin each eqs will turn to M, p =>
Mp = w, then you minimize the sum of square of w.

when 2*i = at least to 12 I can resolve it

or equal.

In case of redundant observations (I>6), we get contradictions

$$\mathsf{M}\ \boldsymbol{p} = \boldsymbol{w}$$

With ***w*** = vector of residual errors

Redundancy management: minimize error

Find p such that it minimizes (akin to pseudoinverse)

merit figure: sum of the square of w, vector of errors

$$\Omega = \boldsymbol{w}^{\mathsf{T}}\boldsymbol{w}$$

$$\Rightarrow \widehat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \boldsymbol{w}^{\mathsf{T}}\boldsymbol{w}$$

I want estimate p with the minimization of omega

$$= \arg\min_{\boldsymbol{p}} \boxed{\boldsymbol{p}^{\mathsf{T}}\mathrm{M}^{\mathsf{T}}}\mathrm{M}\boldsymbol{p}$$

*prima cosa effettivamente diversa

I substitute w with its definition. (transposed: invert the sequence). This in the end will be a scalar product, which I need to minimize.

*invece di dividere per p34

with $||P||_2 = \sum_{ij} p_{ij}^2 = \|\boldsymbol{p}\| = 1$

(homogeneity leaves one degree of freedom)

Solution of the least square problem

The matrix M is 2Ix12. Now we have I bigger than 6, so a rectangular matrix. It can be decomposed into the product of 3 matrix. S is a diagonal matrix, and contains all the eigenvalues of M (MtM).

The other two fullfill the properties.

U and V are orthonormal matrices that contains ortonormal vectors, when you multiply one to the other you obtain an Identity matrix. [18.00]

- Singular value decomposition (SVD)

M can be decomposed in the product of 3 matrixes

$$\underset{2I\times 12}{\mathsf{M}} = \underset{2I\times 12}{U} \; \underset{12\times 12}{\mathsf{S}} \; \underset{12\times 12}{V^{\mathsf{T}}}$$

diagonal matrix: it contains all the eigen values of M [audio]

the other matrixes fulfill these conditions

with properties $U^{\mathsf{T}}U = I_{12}\,,\; V^{\mathsf{T}}V = I_{12}$

S diagonal and $s_1 \geq s_2 \geq \ldots \geq s_{12}$

identity matrix of dimension 12.

When we do SVD, usually these eigenvalues are ordered by size, s1 (s11) will be bigger than s22 etc.

Solution via SVD

Starting from that we have some computation: we start from omega, we substitute to M the SVD, and by the properties seen (identity matrix equivalences) some disappear, we multiply S by S, then since we have lowercase v, the components of matrix v, we know it's orthonormal, vivit is = 0 if i!=j, they are equal to 1 if it's the same vector i=j. So we arrive to the summation of vectors

Minimize:
$$\begin{aligned}\Omega &= \boldsymbol{p}^\top \mathsf{M}^\top \mathsf{M}\boldsymbol{p} \\ &= \boldsymbol{p}^\top V S U^\top U S V^\top \boldsymbol{p} \\ &= \boldsymbol{p}^\top V S^2 V^\top \boldsymbol{p} \\ &= \boldsymbol{p}^\top \left(\sum_{i=1}^{12} s_i^2 \boldsymbol{v}_i \boldsymbol{v}_i^\top \right) \boldsymbol{p}\end{aligned}$$

idnetity matrix, U disappear

you arrive to this summation

$$\boldsymbol{v}_i \boldsymbol{v}_j^\top = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

components of matrix V

Solution via SVD

Choosing $\boldsymbol{p} = \boldsymbol{v}_i$ as solution (anyone of the eigenvector, $\|\boldsymbol{p}\| = 1$)

$$\Omega = \boldsymbol{v}_i^{\mathsf{T}} (s_i^2 \boldsymbol{v}_i \boldsymbol{v}_i^{\mathsf{T}}) \boldsymbol{v}_i = s_i^2 \, \boldsymbol{v}_i^{\mathsf{T}} \boldsymbol{v}_i \boldsymbol{v}_i^{\mathsf{T}} \boldsymbol{v}_i = s_i^2$$

final result, scalar value.

this is equal to identity matrix

*SSE= minimum eigenvalue
wi = elemento del eigenvector corrispondente

Choosing the smallest eigenvalue s_{12} $\boldsymbol{p} = \boldsymbol{v}_{12}$

cascade of all these ^p1,...,^p12

Ω is minimized (M is rank 11)

$$\widehat{\boldsymbol{p}} = \begin{bmatrix} \widehat{\mathbf{A}} \\ \widehat{\mathbf{B}} \\ \widehat{\mathbf{C}} \end{bmatrix} = \boldsymbol{v}_{12} \qquad \widehat{\mathrm{P}} = \begin{bmatrix} \widehat{\mathbf{A}}^{\mathsf{T}} \\ \widehat{\mathbf{B}}^{\mathsf{T}} \\ \widehat{\mathbf{C}}^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \widehat{p}_1 & \widehat{p}_2 & \widehat{p}_3 & \widehat{p}_4 \\ \widehat{p}_5 & \widehat{p}_6 & \widehat{p}_7 & \widehat{p}_8 \\ \widehat{p}_9 & \widehat{p}_{10} & \widehat{p}_{11} & \widehat{p}_{12} \end{bmatrix}$$

Vector

Matrix

Critical known point distribution

consideration on positioning of the points. These were the coefficients of the linear eq system we started, before the minimization of the error.

No solution if all the control points are restricted on a plane.
Example Z=0 below: matrix rank is reduced

coefficients we started from

$$
\mathrm{M} = \begin{bmatrix} \cdots \\ \boldsymbol{a}_{x_i}^{\mathsf{T}} \\ \boldsymbol{a}_{y_i}^{\mathsf{T}} \\ \cdots \end{bmatrix}
= \begin{bmatrix} & & & & & \cdots & & & & & & \\ -X_i & -Y_i & \boxed{0} & -1 & 0 & 0 & \boxed{0} & 0 & x_iX_i & x_iY_i & \boxed{0} & x_i \\ 0 & 0 & \boxed{0} & 0 & -X_i & -Y_i & \boxed{0} & -1 & y_iX_i & y_iY_i & \boxed{0} & y_i \\ & & & & & \cdots & & & & & & \end{bmatrix}
$$

If I chose as Z coordinate Z=0, for all the possible equations I can write by imaging different points, the rank of the matrix of the coefficient is reduced.

Meaning there will be infinite solutions, we won't know which one to chose. We could do that to any plane tho', not just Z. If you have all the points in a single plane you won't find the solution to the problem. At least one has to be away. Obv the more the point spread the better.

Decomposition of P: X_0

We obtained this result by using the DLT(1?), the result is the same.
But now we are able to make a further step! NOt only to estimate these abstract value with no actual meaning, but to decompose the omography into the meaning of the phenomena: translation and rotation and calibration matrices from these we obtained.

We have $\widehat{\mathrm{P}}$, how to obtain $\widehat{\mathrm{K}}, \widehat{R}, \widehat{\boldsymbol{X}}_O$?

$$\widehat{\mathrm{P}} = \widehat{\mathrm{K}}\widehat{R}\,[I_3| - \widehat{\boldsymbol{X}}_O] = [\underset{3\times 3}{\widehat{\mathrm{H}}_\infty} | \underset{3\times 1}{\widehat{\mathbf{h}}}]$$

not just estimate these abstract with no meaning values

first part 3x3 we call Hinf

this vector will be ^h (NOT X0)

$$\begin{bmatrix} \widehat{p}_1 & \widehat{p}_2 & \widehat{p}_3 & \widehat{p}_4 \\ \widehat{p}_5 & \widehat{p}_6 & \widehat{p}_7 & \widehat{p}_8 \\ \widehat{p}_9 & \widehat{p}_{10} & \widehat{p}_{11} & \widehat{p}_{12} \end{bmatrix}$$

it's clear that:

With $\widehat{\mathrm{H}}_\infty = \widehat{\mathrm{K}}\widehat{R}$ and $\widehat{\mathbf{h}} = -\widehat{\mathrm{K}}\widehat{R}\mathbf{X}_O$

Projection centre is $\widehat{\boldsymbol{X}}_O = -\widehat{\mathrm{H}}_\infty^{-1}\,\widehat{\mathbf{h}}$

we can find the translation already (to the camera ref. to respect the global ref)

h^= $[p_4, p_8, p_{12}]^T$

H^=$[p_1, p_2, p_3$
p_5, p_6, p_7
$p_9, p_{10}, p_{11}]$

Decomposition of P: R and K

This allows to decompoose a generic matrix A into the product of a couple of matrices, which are called Q and R. Which includes an orthonormal matrix Q, in our case R (rotation matrix) and an upper triangular matrix R (the k matrix). We can use this by keeping into account that the we need the KR and not what the alg gives use. So instead of facturizing the matrix Hinf, we factorize the inverse of Hinf, so that is factorized in KR inverse, and we obtain the values of Rt and then we can compute K.

In the end we can normalize K matrix by normalizin for the third element

We want to obtain K and R now, decomposing, there's this algoritm

"the product of"

our k matrix?

QR factorization yields the decomposition of a matrix A (H infinite) into a product A = QR of an orthonormal matrix Q (R for us) and an upper triangular matrix R (K for us): A=RK, but we need KR, so we factorize (QR) the inverse of H infinite.

called like this because the alg. is QR

$$\widehat{H}_\infty^{-1} = (\underset{R}{\widehat{K}}\ \underset{Q}{\widehat{R}})^{-1} = \widehat{R}^{-1}\ \widehat{K}^{-1} = \widehat{R}^{T}\ \widehat{K}^{-1}$$

We obtain: R and K

Homogeneity allows to normalize K: $\widehat{\widehat{K}} = \dfrac{1}{\widehat{K}_{33}}\widehat{K}$

In the end: we calibrate a camera and.. unlike the previous slide, where we compute the abstract l values with no physical meaning, we geet all the physical components of the problem: rotation matrix, calibration matrix, translation.

in the end you normalize the k, by normalizing for the last element. Our upper diagonal matrix has a 1 in position 33, and we obtain that one by normalizing

Summary of DLT (2)

We can estimate the camera parameters given control points

$$\mathsf{P} \leftrightarrow (\mathsf{K}, R, \boldsymbol{X}_O)$$

Uncalibrated camera

- **DLT**
- Using **≥6 points**
- Direct solution
- Solution is unstable if the control points lie approximately on a plane

you start here

better more

or if the control points do not feel your control volume: your space volume where you intend to carry out the measurements. If I want to have reliable measurements all over the volumeIt is better to feel the volume with control points of known coordinates

If you feel only a cube of 10 cm x 10 cm x 10cm, you'll be very precise only on that cube. You'll have a lot of errors and noise if you are away from this. If you are working on a very small volume => very low errors there.[praticamente il sistema fa overfitting per ridurre l'errore IN QUEL volume, ergo fuori si potrebber oavere enormi errori]