Today

Perspective projection

Camera calibration

Recall: Full Perspective Camera Model

Projection matrix P

$$\lambda \begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \mathbf{r}_3^t & f_u t_X + u_0 t_z \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_z \\ \mathbf{r}_3^t & t_z \end{bmatrix}}_{\text{Projection matrix } \mathbf{P}} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

$$\lambda = Z^{(C)} = \mathbf{r}_3{}^{T} \mathbf{M}^{(W)} + t_Z$$

Recall Full Perspective Camera Model

$$\lambda \mathbf{m}^{(I)} = \mathbf{P}\mathbf{M}^{(W)}$$

Collinearity equation

$$u^{(I)} - u_0 = f_u \frac{r_{11}X^{(W)} + r_{12}Y^{(W)} + r_{13}Z^{(W)} + t_X}{r_{31}X^{(W)} + r_{32}Y^{(W)} + r_{33}Z^{(W)} + t_Z}$$

$$v^{(I)} - v_0 = f_v \frac{r_{21}X^{(W)} + r_{22}Y^{(W)} + r_{23}Z^{(W)} + t_Y}{r_{31}X^{(W)} + r_{32}Y^{(W)} + r_{33}Z^{(W)} + t_Z}$$

Weak Perspective Camera Model

Perspective projection:

$$u^{(C)} = f\frac{X^{(C)}}{Z^{(C)}} \approx f\frac{X^{(C)}}{\bar{Z}^{(C)}}$$

$$v^{(C)} = f\frac{Y^{(C)}}{Z^{(C)}} \approx f\frac{Y^{(C)}}{\bar{Z}^{(C)}}$$

$$\begin{bmatrix} u^{(C)} \\ v^{(C)} \end{bmatrix} = \frac{f}{\bar{Z}^{(C)}} \begin{bmatrix} X^{(C)} \\ Y^{(C)} \end{bmatrix}$$

$$\lambda \begin{bmatrix} u^{(C)} \\ v^{(C)} \\ 1 \end{bmatrix} = f \begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ \frac{\bar{Z}^{(C)}}{f} \end{bmatrix} \qquad \lambda = \bar{Z}^{(C)}$$

Weak Perspective Camera Model

Spatial sampling:

Since $$\begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u^{(c)} \\ v^{(c)} \\ 1 \end{bmatrix}$$

$$\lambda \begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = f \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ \dfrac{\bar{Z}^{(C)}}{f} \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 f \\ 0 & f_v & v_0 f \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ \dfrac{\bar{Z}^{(C)}}{f} \end{bmatrix}$$

Weak Perspective Camera Model

Object frame to camera frame:

$$\begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ Z^{(C)} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{3\times3} & \mathbf{T}_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^t & t_x \\ \mathbf{r}_2^t & t_y \\ \mathbf{r}_3^t & t_z \\ 0_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} X^{(C)} \\ Y^{(C)} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^t & t_X \\ \mathbf{r}_2^t & t_Y \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ \dfrac{\bar{Z}^{(C)}}{f} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^t & t_X \\ \mathbf{r}_2^t & t_Y \\ 0_{1\times3} & \dfrac{\bar{Z}^{(C)}}{f} \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

2x4 matrix

3x4 matrix

Weak Perspective Camera Model

Overall Weak Perspective Camera Model:

$$\lambda \begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 f \\ 0 & f_v & v_0 f \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ \frac{\bar{Z}^{(C)}}{f} \end{bmatrix} = \underbrace{\underbrace{\begin{bmatrix} f_u & 0 & u_0 f \\ 0 & f_v & v_0 f \\ 0 & 0 & f \end{bmatrix}}_{3\text{x}3} \underbrace{\begin{bmatrix} \mathbf{r}_1^t & t_X \\ \mathbf{r}_2^t & t_Y \\ 0_{1\times 3} & \frac{\bar{Z}^{(C)}}{f} \end{bmatrix}}_{3\text{x}4}}_{\text{Projection matrix } \mathbf{P}_{weak}} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

3x4

$$\mathbf{P}_{weak} = \begin{bmatrix} f_u \mathbf{r}_1^t & f_u t_X + u_0 \bar{Z}^{(C)} \\ f_v \mathbf{r}_2^t & f_v t_Y + v_0 \bar{Z}^{(C)} \\ 0_{1\times 3} & \boxed{\bar{Z}^{(C)}} \end{bmatrix}$$

What is the issue?

$$\lambda = \bar{Z}^{(C)} \approx \mathbf{r}_3{}^t \overline{\mathbf{M}}^{(W)} + t_Z$$

Perspective V.S. Weak Perspective

$$\mathbf{P} = \begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \boxed{\mathbf{r}_3^t} & f_u t_X + u_0 \boxed{t_Z} \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_Z \\ \mathbf{r}_3^t & t_Z \end{bmatrix}$$

$$\mathbf{P}_{weak} = \begin{bmatrix} f_u \mathbf{r}_1^t & f_u t_X + u_0 \bar{Z}^{(C)} \\ f_v \mathbf{r}_2^t & f_v t_Y + v_0 \bar{Z}^{(C)} \\ 0_{1\times 3} & \boxed{\bar{Z}^{(C)}} \end{bmatrix}$$

What's the major difference? $\bar{Z}^{(C)} \approx \mathbf{r}_3{}^t \bar{\mathbf{M}}^{(W)} + t_Z$

What is Camera Calibration?

Estimate the intrinsic parameters, which do not depend on camera position

1. $f_u = f k_u$
2. $f_v = f k_v$
3. u_0
4. v_0
5. $\gamma = f_u \cot\theta$ ← Skew of the two axes on the image plane

$$\mathbf{W}' = \begin{bmatrix} f_u & \gamma & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

In practice, we assume $\theta = 90^\circ \Rightarrow \cot\theta = 0$

Camera Calibration

Reliable and accurate camera calibration is a basic problem in computer vision

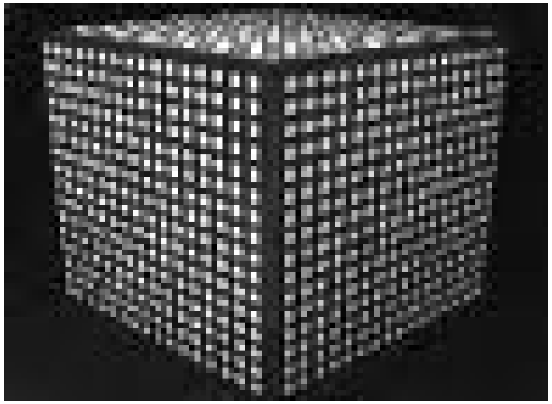
It has been investigated for many years

There are still many researchers working on it

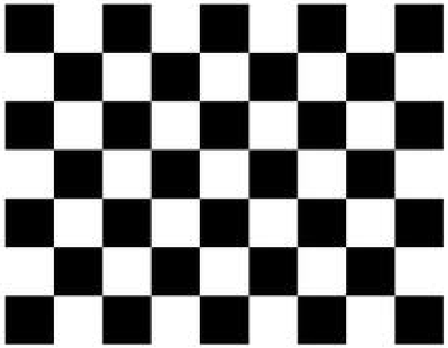
Calibration using features that are easier to achieve

Camera Calibration Methods

- **Conventional camera calibration method**
  - Input: pairs of 3D object points and 2D image points
  - May work for a single view of the object
- **Camera self-calibration**
  - Input: only 2D image points
  - Require multiple images taken from different views
  - Convenient but less constrained
- **Other methods using circles, balls, lines, vanishing points, etc.**



Pattern printed on a cubic



Planar pattern

Conventional Camera Calibration

General strategy:

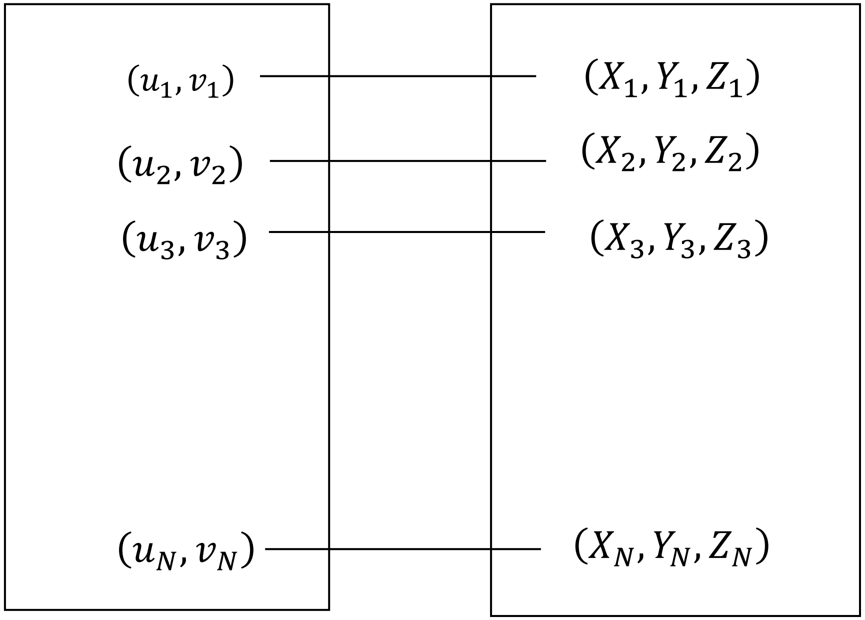
- Take pictures from calibration patterns with known coordinates of 3D features (e.g., points, lines, or curves)
- Identify the corresponding image features
- obtain the projection matrix by minimizing projection error
- obtain intrinsic parameters from the projection matrix

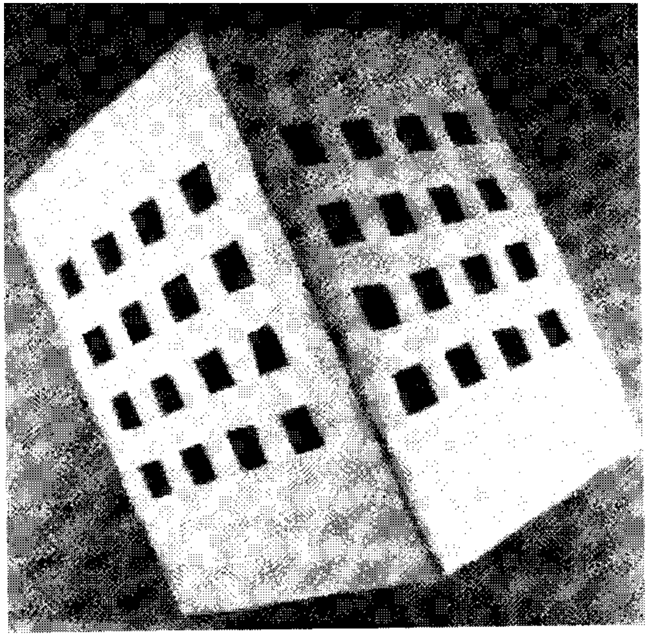
Projection matrix:

$$\mathbf{P} = \begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \mathbf{r}_3^t & f_u t_X + u_0 t_z \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_z \\ \mathbf{r}_3^t & t_z \end{bmatrix}$$

Estimating Camera Parameters

For N pairs of 3D point-- 2D image point





Recall: Full Perspective Camera Model

$$\lambda \mathbf{m}^{(I)} = \mathbf{W}\mathbf{M}^{(C)}, \text{and } \mathbf{M}^{(C)} = \mathbf{D}\mathbf{M}^{(W)} \Rightarrow \lambda \mathbf{m}^{(I)} = \mathbf{W}\mathbf{D}\mathbf{M}^{(W)}$$

$$\lambda \begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{3\times3} & \mathbf{T}_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

Projection matrix:

3x3 3x1

$$\mathbf{P}_{3\times4} = [\mathbf{W}' \quad 0_{3\times1}] \begin{bmatrix} \mathbf{R}_{3\times3} & \mathbf{T}_{3\times1} \\ 0_{1\times3} & 1 \end{bmatrix} = [\mathbf{W}'\mathbf{R} \quad \mathbf{W}'\mathbf{T}] = \begin{bmatrix} \mathbf{p}_1^t & p_{14} \\ \mathbf{p}_2^t & p_{24} \\ \mathbf{p}_3^t & p_{34} \end{bmatrix}$$

where $\mathbf{W}' = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ $\qquad \mathbf{p}_i = \begin{bmatrix} p_{i1} \\ p_{i2} \\ p_{i3} \end{bmatrix}$

Recall: Full Perspective Camera Model

Black box

$$\mathbf{P} = \begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \mathbf{r}_3^t & f_u t_X + u_0 t_z \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_z \\ \mathbf{r}_3^t & t_z \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^t & p_{14} \\ \mathbf{p}_2^t & p_{24} \\ \mathbf{p}_3^t & p_{34} \end{bmatrix}$$

$\Rightarrow$ $\mathbf{p}_3^t = \mathbf{r}_3^t$ and $p_{34} = t_z$

Camera Calibration from Projection Matrix P: A Linear Method

$$\lambda \begin{bmatrix} u^{(I)} \\ v^{(I)} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^t & p_{14} \\ \mathbf{p}_2^t & p_{24} \\ \mathbf{p}_3^t & p_{34} \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} \lambda u^{(I)} \\ \lambda v^{(I)} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^t \mathbf{M}'^{(W)} + p_{14} \\ \mathbf{p}_2^t \mathbf{M}'^{(W)} + p_{24} \\ \mathbf{p}_3^t \mathbf{M}'^{(W)} + p_{34} \end{bmatrix}$$

$$\lambda = \mathbf{p}_3^t \mathbf{M}'^{(W)} + p_{34} \qquad \mathbf{M}'^{(W)} = [X^{(W)} \quad Y^{(W)} \quad Z^{(W)}]^t$$

$$\Rightarrow \begin{aligned} \left(\mathbf{p}_3^t \mathbf{M}'^{(W)} + p_{34}\right) u^{(I)} &= \mathbf{p}_1^t \mathbf{M}'^{(W)} + p_{14} \\ \left(\mathbf{p}_3^t \mathbf{M}'^{(W)} + p_{34}\right) v^{(I)} &= \mathbf{p}_2^t \mathbf{M}'^{(W)} + p_{24} \end{aligned}$$

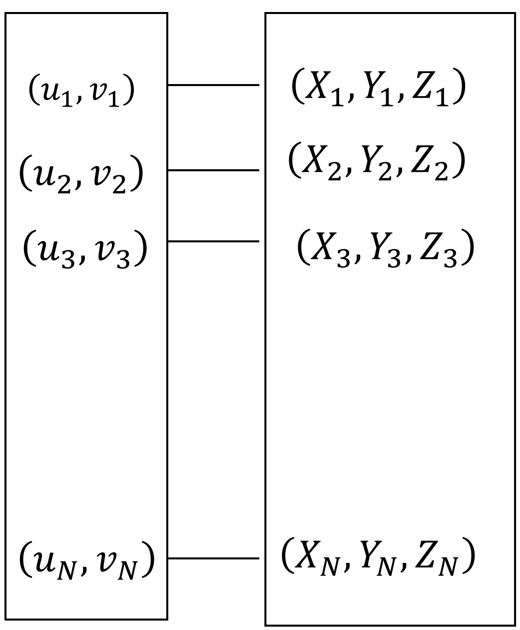
$$\Rightarrow \begin{aligned} \mathbf{p}_1^t \mathbf{M}'^{(W)} + p_{14} - \mathbf{p}_3^t \mathbf{M}'^{(W)} u^{(I)} - p_{34} u^{(I)} &= 0 \\ \mathbf{p}_2^t \mathbf{M}'^{(W)} + p_{24} - \mathbf{p}_3^t \mathbf{M}'^{(W)} v^{(I)} - p_{34} v^{(I)} &= 0 \end{aligned}$$

Camera Calibration from Projection Matrix P: A Linear Method

For every pair of 3D-2D points, we have a pair of linear equations:

$$\mathbf{p}_1^t \mathbf{M}'^{(W)} + p_{14} - \mathbf{p}_3^t \mathbf{M}'^{(W)} u^{(I)} - p_{34} u^{(I)} = 0$$
$$\mathbf{p}_2^t \mathbf{M}'^{(W)} + p_{24} - \mathbf{p}_3^t \mathbf{M}'^{(W)} v^{(I)} - p_{34} v^{(I)} = 0$$

N pairs



We can setup a linear equation system for N corresponding pairs with 2N equations:

- Known: 2D and 3D coordinates
- Unknown: 4x3 parameters of **P**

$$\mathbf{A}\mathbf{v} = 0$$

A is a $2N \times 12$ matrix containing the known
v is a 12×1 vector containing the unknown

$$\mathbf{v} = (\mathbf{p}_1^t \quad p_{14} \quad \mathbf{p}_2^t \quad p_{24} \quad \mathbf{p}_3^t \quad p_{34})^t$$

Camera Calibration from Projection Matrix P: A Linear Method

$$\mathbf{A} = \begin{bmatrix} \left(\mathbf{M'}_1{}^{(W)}\right)^t & 1 & 0_{1\times 3} & 0 & -\left(\mathbf{M'}_1{}^{(W)}\right)^t u_1{}^{(I)} & -u_1{}^{(I)} \\ 0_{1\times 3} & 0 & \left(\mathbf{M'}_1{}^{(W)}\right)^t & 1 & -\left(\mathbf{M'}_1{}^{(W)}\right)^t v_1{}^{(I)} & -v_1{}^{(I)} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ \left(\mathbf{M'}_N{}^{(W)}\right)^t & 1 & 0_{1\times 3} & 0 & -\left(\mathbf{M'}_N{}^{(W)}\right)^t u_N{}^{(I)} & -u_N{}^{(I)} \\ 0_{1\times 3} & 0 & \left(\mathbf{M'}_N{}^{(W)}\right)^t & 1 & -\left(\mathbf{M'}_N{}^{(W)}\right)^t v_N{}^{(I)} & -v_N{}^{(I)} \end{bmatrix}$$

How many pairs of points we need at least? 6

What is the requirement of the 3D points?

Cannot be coplanar

Estimate Projection Matrix by SVD

The projection matrix can be estimated as a solution to

$$\mathbf{Av} = 0$$

The solution lies in the null space of **A – the set of all vectors $\mathbf{x}$ that satisfies $\mathbf{Ax} = \mathbf{0}$**.

$$rank(A) + nullity(A) = \min(row(A), col(A))$$

- **Case 1**: If rank(**A**) = 11, the solution to **v** is unique up to a scaling factor → **Estimate v by performing SVD on A**
- **Case 2**: if rank(**A**) < 11, the solution to **v** is a linear combination of all null vectors of **A**

What about rank(A)=12?

Brief Review on SVD

Singular Value Decomposition:

- Any *mxn* matrix can be written as the product of three matrices

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\mathbf{D}_{m\times n}\mathbf{V}_{n\times n}{}^{T}$$

$\mathbf{u}_1$ $\qquad$ $\mathbf{v}_1{}^{T}$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ \vdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ u_{m1} & u_{m2} & \cdots & u_{mm} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & \sigma_n \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{n1} \\ v_{12} & v_{22} & \cdots & v_{n2} \\ \vdots & \cdots & \cdots & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{nn} \end{bmatrix}$$

- Singular values σ_i are fully determined by **A**
  - **D** is diagonal: $d_{ij} = 0 \; if \; i{\neq}j; \; d_{ii} = \sigma_i \; (i = 1,2,\dots,n)$
  - $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N \geq 0$; $N = rank(A)$
- Both **U** and **V** are not unique
  - Columns of each are mutual orthogonal vectors

Brief Review on SVD

- **U** contains the eigenvector of $\mathbf{A}\mathbf{A}^{\mathrm{T}}$
  - Columns of **U** called the left-singular vectors
- **V** contains the eigenvector of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$
  - Columns of **V** called the right-singular vectors
- Both **U** and **V** are unitary matrices:

$$\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{I}_{m\times m} \text{ and } \mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{I}_{n\times n}$$

Estimate Projection Matrix by SVD

When rank(**A**) = 11, **v** can be estimated by performing SVD on **A**

➡ **v** is the last column of the **V** matrix with an unknown scalar

$$\mathbf{v} = \alpha \mathbf{v}'$$

How to solve the scalar α?

$$\mathbf{p}_3^t = \mathbf{r}_3^t \quad \text{and} \quad \|\mathbf{r}_3\| = 1 \quad \Rightarrow \quad \|\mathbf{p}_3\| = 1$$

➡ $$|\alpha| = \frac{1}{\sqrt{v'^2_9 + v'^2_{10} + v'^2_{11}}}$$

The sign of α -- sign(α) is unknown!

➡ $$\mathbf{v} = |\alpha|\mathbf{v}'$$

The null vector of **A**

➡ Projection matrix **P** is solved!

Discussion on the Rank of A

$$\mathbf{A} = \begin{bmatrix} \left(\mathbf{M'}_1{}^{(W)}\right)^t & 1 & 0_{1\times 3} & 0 & -\left(\mathbf{M'}_1{}^{(W)}\right)^t u_1{}^{(I)} & -u_1{}^{(I)} \\ 0_{1\times 3} & 0 & \left(\mathbf{M'}_1{}^{(W)}\right)^t & 1 & -\left(\mathbf{M'}_1{}^{(W)}\right)^t v_1{}^{(I)} & -v_1{}^{(I)} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ \left(\mathbf{M'}_N{}^{(W)}\right)^t & 1 & 0_{1\times 3} & 0 & -\left(\mathbf{M'}_N{}^{(W)}\right)^t u_N{}^{(I)} & -u_N{}^{(I)} \\ 0_{1\times 3} & 0 & \left(\mathbf{M'}_N{}^{(W)}\right)^t & 1 & -\left(\mathbf{M'}_N{}^{(W)}\right)^t v_N{}^{(I)} & -v_N{}^{(I)} \end{bmatrix}$$

- **Rank(A)=11 with a proper configuration of the 3D points**
  - In practice, rank(A)=12 (full rank) may happen due to the image noise and error in location estimation
- **Rank(A)<11**
  - Rank(A)=8 if the 3D points are coplanar
  - Results in infinite solutions