CS-294.6

Today.

1. Bundle Adjustment

Given the noise model:
$$\begin{cases} \tilde{x}_1^j = x_1^j + w_1^j \\ \tilde{x}_2^j = x_2^j + w_2^j \end{cases}, \quad j = 1, \ldots, N.$$

The objective function is
$$\Phi(x_1, R, T, \lambda_1) = \sum_{j=1}^{n} \|\tilde{x}_1^j - x_1^j\|_2^2 + \|\tilde{x}_2^j - \pi(R\lambda_1^j x_1^j + T)\|_2^2$$

Conditioning that $x_1^{jT} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 1$, $R \in SO(3)$, $\|T\| = 1$

|| Recall: what would happen when $T \to 0$.

This optimization is highly nonlinear and expensive.

Solution:

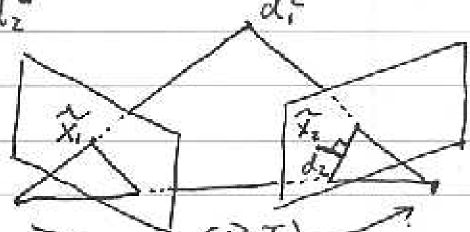
Iteration between optimizing (R, T) and (x_1, x_2)

Step 1: given $(\tilde{x}_1^j, \tilde{x}_2^j)_{j=1}^{N}$, solve for R, T using 8-point

Step 2: ~~Given~~ Optimize (R, T) by
$$\Phi_1(R, T) = \sum_{j=1}^{n} \underbrace{\frac{(\tilde{x}_2^{jT} \hat{T} R \tilde{x}_1^j)^2}{\|\hat{e}_3 \hat{T} R \tilde{x}_1^j\|^2}}_{d_2^2} + \underbrace{\frac{(\tilde{x}_2^{jT} \hat{T} R \tilde{x}_1^j)^2}{\|\tilde{x}_2^{jT} \hat{T} R \hat{e}_3^T\|^2}}_{d_1^2}$$

Geometric interpretation?

minimize the projected distances.



Step 3: Using (R, T) updated, optimize
$$\Phi_2(x) = \sum \|x_1^j - \tilde{x}_1^j\|^2 + \|x_2^j - \tilde{x}_2^j\|^2.$$

Step 4: if updates for R, T, x_1 are small, stop.
otherwise, go to step 2.

Further Reading: Section 5.A.

Implementations for the optimization:

Choice 1: Gauss-Newton Method

Choice 2: Levenberg-Marquardt Method

Gauss - Newton

Brief overview: let the cost function

$$\Phi(\vec{x}) = \frac{1}{2}\vec{r}(\vec{x})^T\vec{r}(\vec{x}). \quad \text{(sum of squares)}$$

$$\nabla\Phi(\vec{x}) = J^T(\vec{x})\vec{r}(\vec{x})$$

$$H_\phi(\vec{x}) = J^T(\vec{x})\vec{J}(\vec{x}) + \sum_{i=1}^{m} r_i(\vec{x}) H_{r_i}(x)$$

$$\simeq J^T(\vec{x})\vec{J}(\vec{x}) \quad \text{(~~linear~~ 1st-order approximation)}$$

Then apply Newton's method, starting with an initial $\vec{x}_0$,

$$\vec{x}_{k+1} = \vec{x}_k + \vec{S}_k, \text{ where}$$

$\vec{S}_k$ is computed by

$$(\vec{J}^T(\vec{x}_k) J(\vec{x}_k))\vec{S}_k = -J^T(\vec{x})\vec{r}(\vec{x})$$

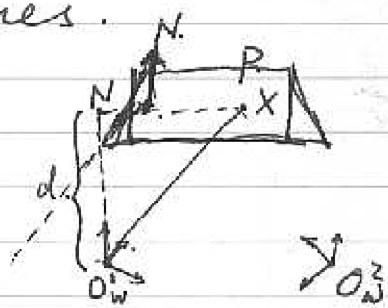
Recall: $f(\vec{x}+\vec{s}) \approx f(\vec{x}) + \nabla f(\vec{x})^T\vec{s} + \frac{1}{2}\vec{s}^T H_f(\vec{x})\vec{s}$

set $\frac{\partial f}{\partial \vec{s}} = 0 \Rightarrow H_f(\vec{x})\vec{s} = -\nabla f(\vec{x})$

Homography for planar scenes.

① for all points $X \in P$,

$$N^T X = d \Leftrightarrow \frac{1}{d}N^T X = 1$$



where N is the unit normal vector of P with respect to O_w^1.

$\therefore$ Given another vantage point O_w^2,

$$X_2 = RX_1 + T = RX_1 + T\frac{1}{d}N^T X_1 = (R + \frac{1}{d}TN^T)X_1$$

Call $H \doteq R + \frac{1}{d}TN^T \in \mathbb{R}^{3\times3}$ the homography matrix

⇒ The linear transformation between O_w^1 & O_w^2 is

$$X_2 = HX_1 \Leftrightarrow \lambda_2 x_2 = H\lambda_1 x_1$$
$$\Leftrightarrow x_2 \sim Hx_1 \quad \text{(equality up to a scalar)}$$
$$\Leftrightarrow \hat{x}_2 H x_1 = 0 \in \mathbb{R}^3$$

② Relations between homography and essential matrix.

Recall: To recover a correct E matrix, 3-D points have to be in general position

When all points lie on a plane. we consider the degenerate situation:

☆ Theorem 5.21. For $E = \hat{T}R$ and $H = R + Tu^T$ for $T, u \in \mathbb{R}^3$ with $\|T\| = 1$,

1. $E = \hat{T}H$
2. $H^TE + E^TH = 0$
3. $H = \hat{T}^TE + Tv^T$ for some $v \in \mathbb{R}^3$.

Proof: 1. $\hat{T}H = \hat{T}R + \hat{T}Tu^T = E$.

2. $H^TE = (R + Tu^T)^T\hat{T}R = R^T\hat{T}R$
$$= -(R^T\hat{T}R)^T$$
$$= -E^TH.$$

3. $H = \hat{T}^TE + Tv^T$ for some v.
$$\Leftrightarrow (H - \hat{T}^TE) = Tv^T \text{ for some } v.$$
$$\Leftrightarrow H - \hat{T}^TE = [v_1T \quad v_2T \quad v_3T]$$
$$\Leftrightarrow \hat{T}(H - \hat{T}^TE) = 0$$
$$\Leftrightarrow \hat{T}H = \hat{T}\hat{T}^TE \quad \Leftarrow$$

But $\hat{T}H = \underbrace{\hat{T}R \stackrel{?}{=} \hat{T}\hat{T}^T\hat{T}R}_{} = \hat{T}\hat{T}^TE$

|| HW 5.3

Note: $\hat{T}\hat{T}^T$ is called an orthogonal projection.

$$\hat{T}\hat{T}^Tv$$
$$= \hat{T}(-\hat{T}v)$$
$$= \hat{T}w = v_\perp, \quad \|v_\perp\| = \|v\|\sin\theta \text{ if } \|T\| = 1$$

⇒ $\boxed{\hat{T}\hat{T}^T = I - TT^T \quad \text{(Exer 5.3.1)}}$

☆ Corallary: Given a pair of points (x_1, x_2) and homography $x_2 \sim H x_1$, for $\forall u \in \mathbb{R}^3$,

$$x_2^T \hat{u} H x_1 = 0,$$

Therefore, a homography H corresponds to a 3-parameter family of essential matrix $E = \hat{u} H \in \mathbb{R}^{3\times3}$

proof: let $w = \hat{u} x_2$, $\therefore w \perp x_2 \sim H x_1$

$\therefore w^T H x_1 = 0 \Rightarrow x_2^T \hat{u} H x_1 = 0$.

Hence: eight-point algorithm does not apply to points from a planar scene.

③ Estimation of H.

$$\lambda_2 x_2 = H \lambda_1 x_1$$

$$\Leftrightarrow \hat{x}_2 H x_1 = 0$$

$$\Leftrightarrow (x_1 \otimes \hat{x}_2)^T H^s = 0$$

where $x_1 \otimes \hat{x}_2 \in \mathbb{R}^{9\times3}$, and $H^s \in \mathbb{R}^9$

To solve H up to a scale, we also need 8 linear constraints.

$\because$ rank $(\hat{x}_2) = 2$

$\therefore$ rank$(x_1 \otimes \hat{x}_2) = 2$

$\Rightarrow$ one pair of correspondence provides 2 constraints

Therefore, we need at least FOUR points in space to solve for H.

☆ four-point algorithm.

$$D = \begin{bmatrix} (x_1^1 \otimes \hat{x}_2^1)^T \\ \vdots \\ (x_1^4 \otimes \hat{x}_2^4)^T \end{bmatrix} \in \mathbb{R}^{12\times9}$$

rank$(D) = 8$

$[U, S, V] = \text{svd}(D)$, and $H^s = V(\text{end})$.

④ Normalization of H matrix

denote the "unstacked" version of H^s as H_L.

$$H_L \doteq \lambda H = \lambda\left(R + \frac{1}{d}TN^T\right).$$

Theorem 5.18: for $H_L = \lambda(R + \frac{1}{d}TN^T)$, $|\lambda| = \sigma_2(H_L)$.

Proof: Let $u = \frac{1}{d}R^T T \in \mathbb{R}^3$.

$$H_L^T H_L = \lambda^2 (I + uN^T + Nu^T + \|u\|^2 NN^T).$$

let $v = u \times N = \hat{u} N$, then $v \perp u$, $v \perp N$

$\Rightarrow H_L^T H_L v = \lambda^2 v$

$\therefore \lambda^2$ is an eigenvalue of $H_L^T H_L$.

$\Rightarrow |\lambda|$ is a singular value of H_L.

Next, we need to prove $|\lambda|$ is the second s.v.

Consider $Q = uN^T + Nu^T + \|u\|^2 NN^T$

$$= (u/\|u\| + \|u\|N)(u/\|u\| + \|u\|N)^T - (u/\|u\|)(u/\|u\|)^T.$$

(denoted) $= (w+v)(w+v)^T - ww^T$ || want to show $\sigma_2(Q) = 0$

(i) if $v \sim w$, let $w = av$ for $a \in \mathbb{R}$

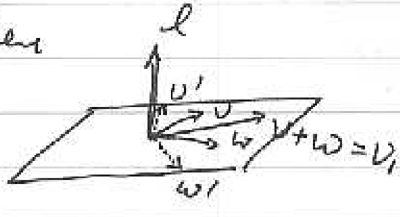
$$Q = (a+1)^2 vv^T - a^2 vv^T = (2a+1)vv^T$$

let a basis in $\mathbb{R}^3$ be $\{v, v_2, v_3\}$,

$$Qv_2 = 0, \quad Qv_3 = 0.$$

(ii) if $v \nsim w$, $\exists\, l$, $l \sim v \times w$. , then

$$Ql = 0 \in \mathbb{R}^3$$



denote $v_1 = w + v$, then

$$Q = v_1 v_1^T - ww^T$$

is a symmetric matrix.

it is easy to see that if $v_1 \nsim w$, then $\exists\, u \in \mathbb{R}^3$

$$u^T Q u > 0$$

and $\exists\, u' \in \mathbb{R}^3$ $u'^T Q u' < 0$.

$\therefore$ Q has one positive eigenvalue and one negative eigenvalue.

Summary. in both cases, 0 is $\sigma_2(Q)$.

$\therefore |\lambda| = \sigma_2(H_L)$

Hence, we set $H = H_L / \sigma_2(H_L) \Rightarrow \sigma_2(H) = 1$

This recovers H up to a sign difference
$$H = \pm(R + \tfrac{1}{d} T N^T).$$
To recover the sign, use the positive depths constraint
$$\lambda_2 x_2 = H_0 \lambda_1 x_1$$
$$\Rightarrow x_2^T H x_1 > 0 \quad \text{for all pairs } (x_1^j, x_2^j)$$

⑤ Decomposing H

Theorem 5.19. Given $H = R + \frac{1}{d} T N^T$, there are at most two possible solutions for a decomposition from 4 solutions of the following form: (Sketch)

(i) significance of $\sigma_2(H) = 1$

$\therefore \exists$ a vector $a \in \mathbb{R}^3$
$$\|Ha\|^2 = \|a\|^2$$
Preserve length

It is not surprising that for $\forall a \perp N$, $Ha = Ra$

Hence,
$$\operatorname{svd}(H^T H) = V \Sigma V^T, \quad \Sigma = \operatorname{diag}\{\sigma_1^2, \sigma_2^2, \sigma_3^2\}. \quad (\sigma_2^2 = 1)$$
$$H^T H v_1 = \sigma_1^2 v_1, \quad H^T H v_2 = v_2, \quad H^T H v_3 = \sigma_3^2 v_3.$$
$\therefore v_2 = V(:, 2)$ is $\perp N$

Construct $u_1 \doteq \dfrac{\sqrt{1-\sigma_3^2}\, v_1 + \sqrt{\sigma_1^2 - 1}\, v_3}{\sqrt{\sigma_1^2 - \sigma_3^2}}$

$$u_2 \doteq \frac{\sqrt{1-\sigma_3^2}\, v_1 - \sqrt{\sigma_1^2-1}\, v_3}{\sqrt{\sigma_1^2 - \sigma_3^2}}$$

we can check $\|H u_i\| = \|u_i\|$

further, H preserves any vector inside
$$S_1 = \operatorname{span}\{v_2, u_1\} \text{ and } S_2 = \operatorname{span}\{v_2, u_2\}$$
$\therefore N \perp S_1, \quad N \perp S_2$

$\therefore N = \widehat{v_2} u_1$ or $\widehat{v_2} u_2$ ✓

(ii) Define $U_1 = [v_2, u_1, \widehat{v_2} u_1]$; $W_1 = [H v_2, H u_1, \widehat{H v_2} H u_1]$

$U_2 = [v_2, u_2, \widehat{v_2} u_2]$; $W_2 = [H v_2, H u_2, \widehat{H v_2} H u_2]$

Then $RU_1 = W_1$, $RU_2 = W_2$
$\Rightarrow R = W_1 U_1^T$ or $W_2 U_2^T$

Hence: 4 Solutions:

$R_1 = W_1 U_1^T = R_2$
$N_1 = \hat{v}_2 u_1$, $\frac{1}{d} T_1 = (H - R_1) N_1$
$N_2 = -N_1$ $\frac{1}{d} T_2 = -\frac{1}{d} T_2$.

$R_3 = W_2 U_2^T = R_4$
$N_3 = \hat{v}_2 u_2$, $\frac{1}{d} T_3 = (H - R_3) N_3$
$N_4 = -N_3$, $\frac{1}{d} T_4 = -\frac{1}{d} T_3$.

Claim: Only two normals N_i have positive depth.

end of the four-point algorithm.

Homework #3: 5.3, 5.6, 5.11, 5.13, 5.19

Special Motions

① Pure rotation, $T=0$, $E = \hat{T}R = 0$, but $H = R$.
$\Rightarrow$ $X_2 = R X_1 \Leftrightarrow \hat{x}_2 R x_1 = 0$

meaning without translation, depth information is completely lost, the 3D scene can be interpreted to be planar.

② planar motion: the camera always moves on a plane: (Exercise 5.6)