1.

(a) advantage:

1. more robust to lighting, wether, dynamic objects, blur, occlusion and texture-less scene
2. Map scale not dependent on number of training images
3. Much faster (GPU) and requires smaller images than traditonal method

disadvantage:

1. coarse accuracy
2. difficult to learn both position and orientation

(b) β 不好調，若 position 準，orientation 就不準，反之亦然，很難使兩者皆很準

將 loss 改成

$$L = e^{-\hat{s}_x}\|x-\hat{x}\|_2 + e^{-\hat{s}_q}\|q-\hat{q}\|_2 + \hat{s}_x + \hat{s}_q$$

$$\hat{s}\downarrow \Rightarrow e^{-\hat{s}}\uparrow$$

$\hat{s}_x$，$\hat{s}_q$ 為可 train 的 parameter，將 loss 改成上方形式並學習 $\hat{s}_x$，$\hat{s}_q$ 來 minimize loss

2.

1. Midel represcntatiin

已給定3D point cloud，compute each points 的 descriptors，但一個 point cloud 可能在多張 view 看到，收集這些代表3D點的 descriptors 並做 mean-shift clustering 來 compress

2. Synthetic view

建立 virtual view，建立 descriptors 非影像的 view，希望將 descriptors 投射在 virtual view 上

建立 "synthetic" view 在額外位置，每個額外位置使用 12個 synthetic view，correspond to 30° rotation between camera

但這些 syntetic view 需滿足以下 criteria

- the projected feature must be in front of camera, and lie within the field of view
- the 3D scales ≥ 1 pixel，in terms of the respective DoG scale space extrema to ensure detectability

- the original image view of the associated descriptors is similar to the viewing direction

3. Vocabulary tree

因 scene 的 descriptors 數量多，用 tree 來增加效率，tree 的 leaves 為 hierarchical K-means clustering 得到的 quantized feature descriptors (visual words)

形成的 hierarchical K-means tree 即 vocabulary tree

3.

令 C_0 至 C_1 的 pose 為 R, t

$$t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

已知 $E = [t]_\times R$

$$[t]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$$

$$\det([t]_\times) = 0 + t_1t_2t_3 - t_1t_2t_3 - 0 - 0 - 0 = 0$$

$$\det(E) = \det([t]_\times)\det(R) = 0 \quad \#1$$

令 K 為 C_0 intrinsic, K' 為 C_1 intrinsic

$$F = K'^{-T} E K^{-1}$$

$$\det(F) = \det(K'^{-T}) \cdot \det(E) \cdot \det(K^{-1})$$

$\because \det(E) = 0$

$\therefore \det(F) = 0 \quad \#$

4.

(a) intrinsic : $\begin{bmatrix} f_x & s & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} fS_x & fS_\theta & O_x \\ 0 & fS_y & O_y \\ 0 & 0 & 1 \end{bmatrix}$

intrinsic matrix 由 f, S_x, S_y, S_θ, O_x, O_y 這些參數組合得到

f : CCS中原點到 image plane 距離,又叫 effective focal length

S_x, S_y : CCS中一單位和 pixel 一單位實際並不相等,兩者間的 scale,用 S_x, S_y 轉換, S_x, S_y 分別為 x, y 軸的 scale

S_θ : skew factor,若 pixel 非長方形,而是平行四邊形,用 S_θ 做修正

O_x, O_y : image plane 上 $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 的點轉至 Pixel cooridinate $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ 之間的 offset 關係 O_x, O_y 分別為 X, Y 至 u, v 的 offset

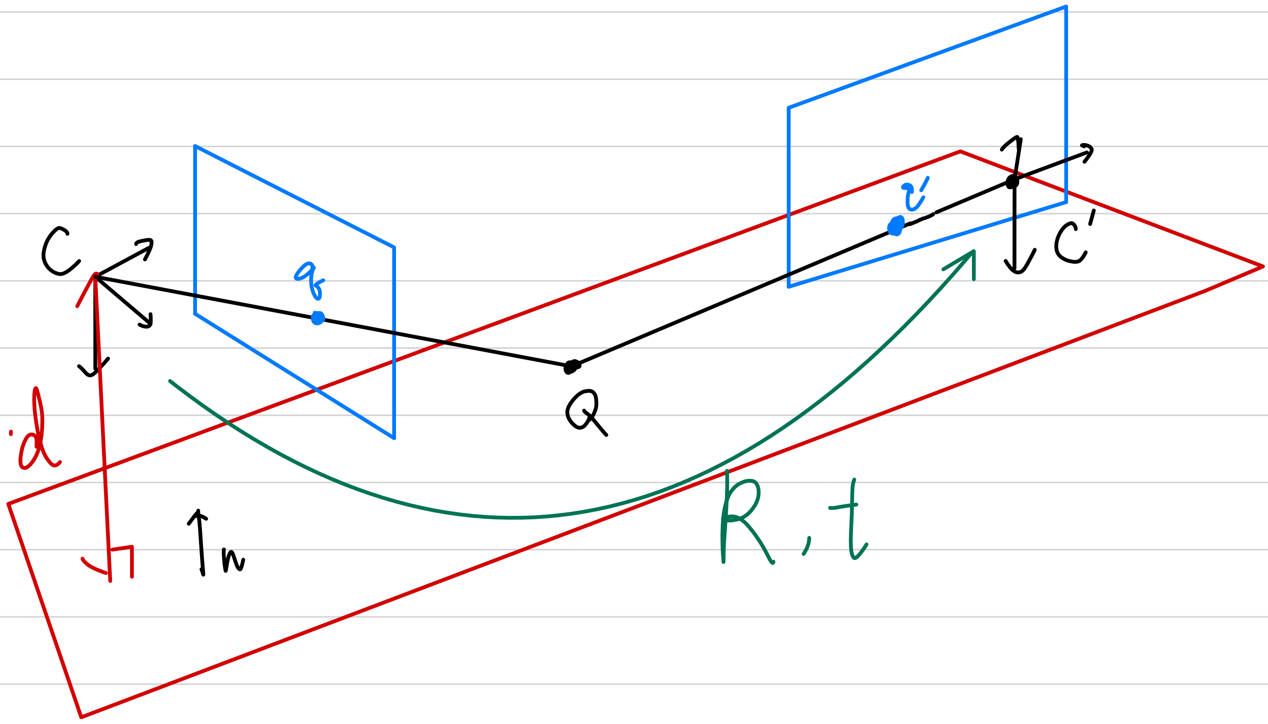
(b) 選 (c) Exposure,曝光改變和內參影響不大

(A) 直接影響 f

(B) 改變 O_x, O_y

(D) zoom in 改變 effective focal length, 同 (A)

5.



q 為 Q 在 C 的 pixel coordinate 下的投影點

q' 為 Q 在 C' 的 pixel coordinate 下的投影點

Q 為平面下的點，其座標為在 C 得之 3D 點

R, t 為 C 至 C' 的 pose

令 Q' 為 C' 觀測之 3D 座標，n 為平面對 C 的法向量

From geometry

$$d = n^T Q \Rightarrow \frac{1}{d} n^T Q = 1 \quad \text{—①}$$

由 C 轉至 C'

$$Q' = RQ + t \quad \text{—②}$$

①②得:

$$Q' = RQ + t\frac{1}{d}n^T Q = (R + \frac{1}{d}tn^T)Q \quad —③$$

3D to pixel:

$$\lambda_1 q = KQ \quad , \quad \lambda_2 q' = KQ'$$

$$\Rightarrow \lambda_1 K^{-1} q = Q \quad , \quad \lambda_2 K^{-1} q' = Q' \quad —④$$

③④得:

$$\lambda_2 K^{-1} q' = \lambda_1 (R + \frac{1}{d}tn^T) K^{-1} q$$

$$\Rightarrow \lambda q' = K(R + \frac{1}{d}tn^T)K^{-1} q$$

$$\lambda q' = Hq$$

$$H = K(R + \frac{1}{d}tn^T)K^{-1}$$

現知 d 很大

$$H_R = \lim_{\frac{t}{d} \to 0} H = \lim_{\frac{t}{d} \to 0} K(R + \frac{1}{d}tn^T)K^{-1}$$

$$= KRK^{-1} \#$$

6.

(a) projection matrix P 為 3×4 matrix，12個 entries，但有一個 scale factor，故 DoF 為 11，每對 correspondence 可產生 2 條 equation，故至少要 6 對 correspondence 才可滿足 11 個 DoF，故 $\min n = 6$

(b) $P = [P_1 \ p_2]$，$P_1 = \begin{bmatrix} 5 & -3 & 3 \\ 4 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

$$P_1^{-1} = \begin{bmatrix} 1/9 & 1/9 & -5/9 \\ -4/27 & 5/27 & 2/27 \\ 0 & 0 & 1 \end{bmatrix}$$

令 $u_1 = \begin{bmatrix} 1/9 \\ -4/27 \\ 0 \end{bmatrix}$，$u_2 = \begin{bmatrix} 1/9 \\ 5/27 \\ 0 \end{bmatrix}$，$u_3 = \begin{bmatrix} -5/9 \\ 2/27 \\ 1 \end{bmatrix}$

因 u_1, u_2 最後 entry 為 0，$\{u_1, u_3\}$ 和 $\{u_2, u_3\}$ 這兩組為 LI

又因 u_1, u_2 首尾 entries 相同，中間不同，$\{u_1, u_2\}$ 為 LI

$\{u_1, u_2, u_3\}$ 為 LI

由 Gram-Schmidt process 可得 orthogonal basis $\{v_1, v_2, v_3\}$

$$v_1 = \begin{bmatrix} 1/9 \\ -4/27 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 4/25 \\ 3/25 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

將$\{v_1, v_2, v_3\}$轉成orthonormal basis $\{q_1, q_2, q_3\}$

$$q_1 = \begin{bmatrix} 0.6 \\ -0.8 \\ 0 \end{bmatrix}, \quad q_2 = \begin{bmatrix} 0.8 \\ 0.6 \\ 0 \end{bmatrix}, \quad q_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$P_1^{-1} = Q\bar{R}, \quad Q = [\, q_1 \mid q_2 \mid q_3 \,]$$

$$\bar{R} = \begin{bmatrix} \langle u_1, q_1 \rangle & \langle u_2, q_1 \rangle & \langle u_3, q_1 \rangle \\ 0 & \langle u_2, q_2 \rangle & \langle u_3, q_2 \rangle \\ 0 & 0 & \langle u_3, q_3 \rangle \end{bmatrix} = \begin{bmatrix} 5/27 & -11/135 & -53/135 \\ 0 & 0.2 & -0.4 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(P_1^{-1})^{-1} = (Q\bar{R})^{-1} \Rightarrow P_1 = \bar{R}^{-1} Q^{-1}$$

$$\bar{R}^{-1} = \begin{bmatrix} 27/5 & 11/5 & 3 \\ 0 & 5 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q^{-1} = \begin{bmatrix} 3/5 & -4/5 & 0 \\ 4/5 & 3/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

camera intrinsic matrix $K = \bar{R}^{-1} = \begin{bmatrix} 27/5 & 11/5 & 3 \\ 0 & 5 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

extrinsic matrix $E = K^{-1} P = \bar{R} P = [R \mid t]$

$$= \begin{bmatrix} 5/27 & -11/135 & -53/135 \\ 0 & 0.2 & -0.4 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 & -3 & 3 & 5 \\ 4 & 3 & 2 & 10 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} & 0 & \frac{-38}{135} \\ \frac{4}{5} & \frac{3}{5} & 0 & \frac{8}{5} \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

7.

(a) 符合距離夠遠或觀察物厚度相對距離小，選 (c)

(b)

Affine model:

解出 $\hat{D} = \hat{A}\hat{X} = (\hat{A}Q)(Q^{-1}\hat{X}) = \hat{A}'\hat{X}'$

Q 有多組解，所以 more than one solution

orthographic: $M = \begin{pmatrix} r_1^T & t_1 \\ r_2^T & t_2 \\ 0 & 1 \end{pmatrix}$

where $r_1^T r_1 = r_2^T r_2 = 1$, $r_1^T r_2 = 0$

想找 affine transform Q，使 $\hat{A}$ 變 orthographic

$\hat{A}_i Q Q^T A_i^T = I \quad i = 1, 2 \dots m$

each $i = 1 \dots, m$ let

$\hat{A}_i = \begin{pmatrix} a_{i1}^T \\ a_{i2}^T \end{pmatrix}$

$$\hat{A}_i Q Q^T \hat{A}_i^T = I$$ 有 constraints

$$a_{i1}^T Q Q^T a_{i2} = 0$$

$$a_{i1}^T Q Q^T a_{i1} = 1$$

$$a_{i2}^T Q Q^T a_{i2} = 1$$

let $L = QQ^T$ for $i = 1 \dots m$

$$a_{i1}^T L a_{i2} = 0$$

$$a_{i1}^T L a_{i1} = 1$$

$$a_{i2}^T L a_{i2} = 1$$

由 Cholesky decomposition 找 $\hat{Q}$

$$\hat{L} = \hat{Q} \hat{Q}^T$$

而 Cholesky decomposition 為唯一解

故 orthographic projection 只有唯一解

有唯一解可還原 euclidean reconstructio

#

另個觀點：

orthographic camera model

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

其 $\alpha_x, \alpha_y = 1$, $s = 0$，經過轉換角度和平行線會保留，故可還原euclidean reconstruction

───

我

8. x_i 和 x_i' 為 camera C 和 C' 上的 correspondence

X_i 和 X_i' 是 CCS 中座標，u_i, u_i' 是 pixel 座標

1.

C 和 X_i 連線，C' 和 X_i' 連線，兩線若不相交，取此兩條線之公垂線中點當作找到的 3D 座標 $\mathbf{X}_i$，並 minimize 3D error

$$\varepsilon_i = d(u_i, P\mathbf{X}_i)^2 + d(u_i', P'\mathbf{X}_i)^2$$

2.

每個 camera P 會產生兩條式子，故 P, P' 共可生 4 條式子來解 $\mathbf{X}_i$ 3 個自由度的問題

$$u_i = P\mathbf{X}_i \Rightarrow \begin{bmatrix} v_i p^{3T} - p^{2T} \\ u_i p^{3T} - p^{1T} \end{bmatrix} \mathbf{X}_i = 0$$

$$u_i' = P'\mathbf{X}_i \Rightarrow \begin{bmatrix} v_i' p'^{3T} - p'^{2T} \\ u_i' p'^{3T} - p'^{1T} \end{bmatrix} \mathbf{X}_i = 0$$

$$A = \begin{bmatrix} v_i p^{3T} - p^{2T} \\ u_i p^{3T} - p^{1T} \\ v_i' p'^{3T} - p'^{2T} \\ u_i' p'^{3T} - p'^{1T} \end{bmatrix}$$

，對 A 作 SVD 來解 $\underbrace{\ }_{A\mathbf{X}_i = 0 \text{ 的}}$ optimize 問題