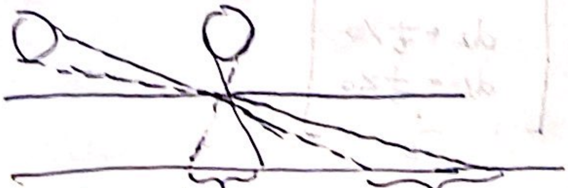
透视 projection

丢失：深度、角度、长度

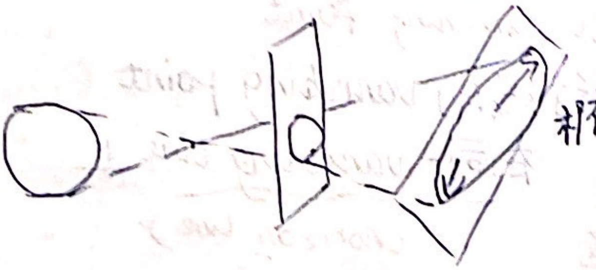
直线不会变曲线

平行线 vanishing ~~line~~ point

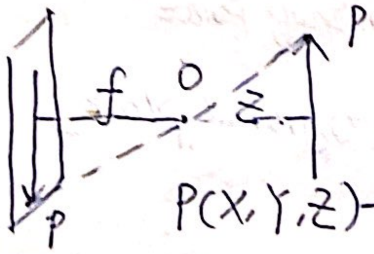
平行面 vanishing ~~point~~ line.



越向两边，像越失真，物体越拉长



椭圆



像机到image plane距离为f

$P(X,Y,Z) \to p(f\frac{X}{Z}, f\frac{Y}{Z})$

$$p = K\hat{p} \to \text{normalized image plane}$$

$$\hat{p} = [R \ \ t]P$$

$$p = \underbrace{\begin{pmatrix} \frac{1}{dx} & s & x_0 \\ 0 & \frac{1}{dy} & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{K} \underbrace{[R \ \ t]P}_{\text{world}\to\text{camera}}$$

camera space → image plane

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix}$$

$\alpha = \frac{1}{dx}f$

$\beta = \frac{1}{dy}f$

(normalized) image plane → pixel

x_0, y_0 : principal offset point

θ : skew angle

$\alpha(f_x)$, $\beta(f_y)$ focal length

$\hat{p}$ 单位：m　　p：单位 pixel

$\alpha = kf$, $\beta = lf$　　unit aspect ratio $\Rightarrow \alpha = \beta = f$

k, l：拉伸的系数

K：自由度 5

$[R \ \ t]$ 自由度 $3+3=6$

$$B \xrightarrow{R,t} A \quad {}^{A}P = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} {}^{B}P$$

$$R \stackrel{\text{def}}{=} \begin{pmatrix} {}^{A}i_B, & {}^{A}j_B, & {}^{A}k_B \end{pmatrix} = \begin{pmatrix} i_A \cdot i_B & i_A \cdot j_B & i_A \cdot k_B \\ j_A \cdot i_B & j_A \cdot j_B & j_A \cdot k_B \\ k_A \cdot i_B & k_A \cdot j_B & k_A \cdot k_B \end{pmatrix}.$$

$$\lambda x_i = \underbrace{K[R \ \ t]}X_i = PX_i$$

$$\Rightarrow \underbrace{x_i \times PX_i = 0}_{\text{提供2个方程}}$$

at least 6 pairs $\Rightarrow P$

物理世界 单位：m，pixel world 单位：(pixel)

Pixel width (dx) m, height (dy) m.

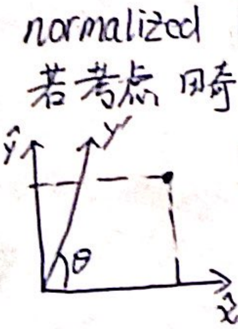
$\alpha = \frac{1}{dx}f$, $\beta = \frac{1}{dy}f$

物理坐标 (x, y)

normalized image plane 读取 $(\hat{x}, \hat{y})$

若考虑 畸变的实际坐标 (x', y')

$\left.\begin{matrix} \\ \\ \end{matrix}\right\} \Rightarrow x = \alpha x' + x_0$, $y = \beta y' + y_0$



$\hat{x} = x' + y'\cos\theta$

$\hat{y} = y'\sin\theta$

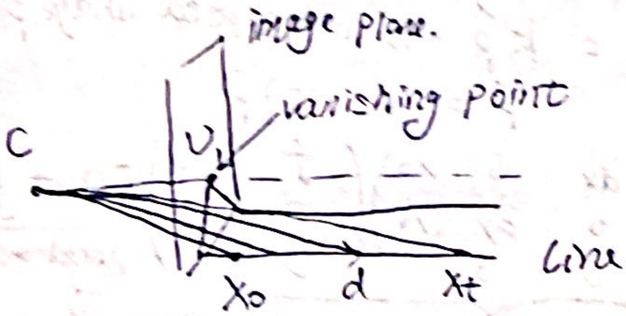
$\Rightarrow y' = \frac{\hat{y}}{\sin\theta}$, $x' = \hat{x} - \hat{y}\cot\theta$

$$\Rightarrow x = \alpha\hat{x} - \alpha\hat{y}\cot\theta + x_0$$

$$y = \beta\frac{\hat{y}}{\sin\theta} + y_0$$

$$\Downarrow$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \hat{x} \\ \hat{y} \\ 1 \end{pmatrix}$$



same direction ⇒ same vanishing point.

$d \Rightarrow v$

$$X_t = X_0 + t\cdot d = \begin{bmatrix} x_0+td_1 \\ y_0+td_2 \\ z_0+td_3 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} d_1+\frac{1}{t}x_0 \\ d_2+\frac{1}{t}y_0 \\ d_3+\frac{1}{t}z_0 \\ \frac{1}{t} \end{bmatrix}$$

$t \to \infty \quad X_\infty = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ 0 \end{pmatrix}$ $\quad v = PX_\infty = K[R\ t]X_\infty$

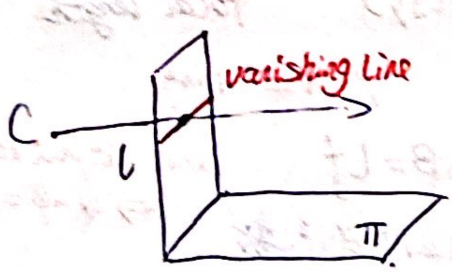
finite vanishing point

$\boxed{\vec{CV} \parallel d}$

同一平面，各平行线的 vanishing point 在同一 vanishing line 上 (horizon line)

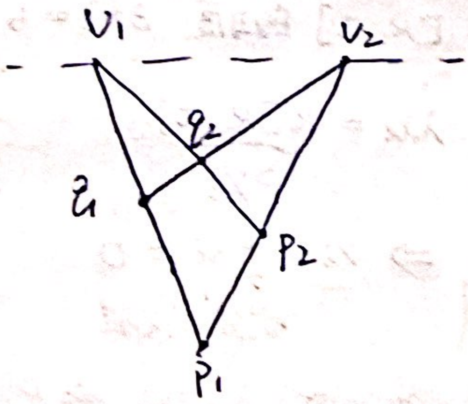
不平行的plane ⇒ 不同的vanishing line

若 $d \parallel$ 像平面 ⇒ infinite vanishing point



$l \parallel \pi$

C和 l 等高（相对于 π 作为ground plane）



~~步骤~~

- p_1q_1 直线：$p_1 \times q_1 \Rightarrow l_1$
- p_2q_2 直线：$p_2 \times q_2 \Rightarrow l_2$
- → 直线交点 $l_1 \times l_2 = v$

$p_1 = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ 2D的齐次坐标、

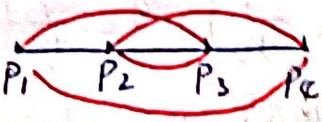
$l = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \Rightarrow ax+by+c=0$

$P_i = \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix}$

$$\frac{\|P_1-P_3\|\|P_2-P_4\|}{\|P_2-P_3\|\cdot\|P_1-P_4\|}$$

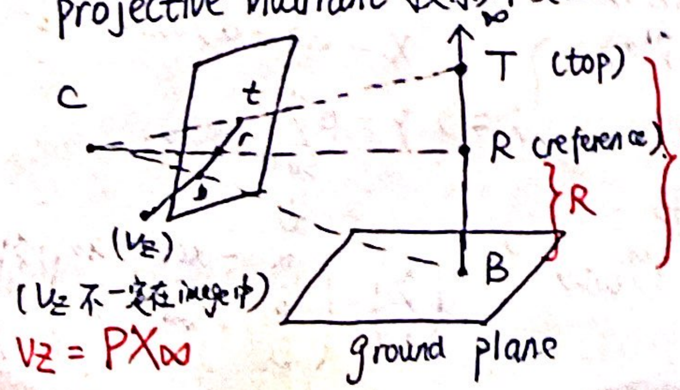
measuring height

projective invariant 投影不变量 → cross ratio



$$\text{ratio} = \frac{\|\cancel{\infty}-R\|\cdot\|T-B\|}{\|\cancel{\infty}-B\|\cdot\|T-R\|} = \frac{H}{H-R}$$

$$= \frac{\|v_z-r\|\cdot\|t-b\|}{\|v_z-b\|\cdot\|t-r\|}$$



$v_z = PX_\infty$

$$P = K[R \quad t] = [P_1 \quad P_2 \quad P_3 \quad P_4]$$

$P_1 = P(1,0,0,0)^T$, $P_2 = P(0,1,0,0)^T$, $P_3 = P(0,0,1,0)^T$

P_i: x, y, z direction 的 vanishing point

$P_4 = P(0,0,0,1)^T \Rightarrow$ 原点在 image 上的位置

$$\lambda_i v_i = K[R \quad t]\begin{bmatrix} e_i \\ 0 \end{bmatrix} = KRe_i \Rightarrow e_i = \lambda_i R^T K^{-1} v_i$$

$$e_i^T e_j = 0 \Rightarrow v_i^T K^{-T} \underline{RR^T} K^{-1} v_j = 0$$

$$\Rightarrow v_i^T K^{-T} K^{-1} v_j = 0$$

假设 $\alpha = \beta$, $\theta = 90°$, 则 K 共有 f, u_0, v_0 3个未知数 $\Rightarrow$ 3 vanishing point 求 K

1° 3 finite vanishing point $\Rightarrow$ 求出 K (f, u_0, v_0)

2° 2 finite vanishing point $\Rightarrow$

3° 1 finite vanishing point $\Rightarrow$ f 无法求出，(u_0, v_0) 为该 vanishing point

K 得知后可求 $R = [r_1 \quad r_2 \quad r_3]$

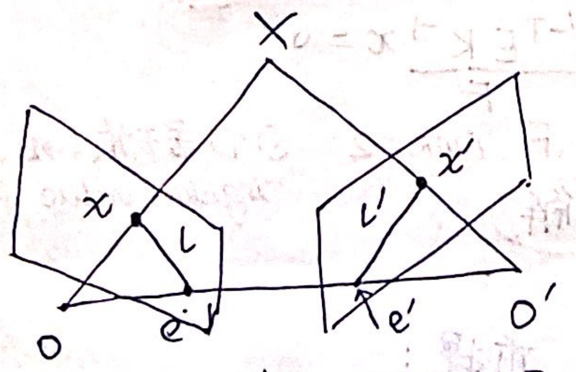
$$\lambda_i v_i = K[R \quad t]\begin{bmatrix} e_i \\ 0 \end{bmatrix} \Rightarrow \lambda_i K^{-1} v_i = r_i, \quad \|r_i\| = 1$$

stereo

structure from motion (SFM)

motion: given set of correspondence point $\Rightarrow$ Camera params

epipolar geometry



OO': base line

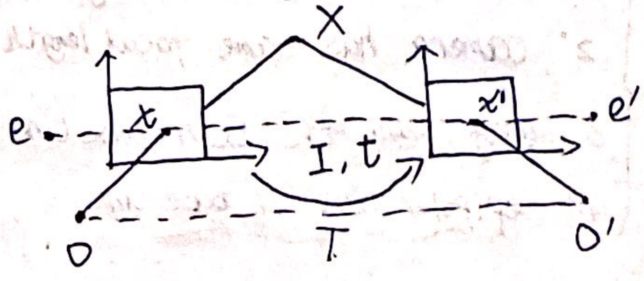
epipolar plane: 过 base line 的所有平面

epipolar: e、e'

epipolar line: l、l'

x 的对应点必在 l' 上 : 面上找 $\Rightarrow$ 线上找

$\Downarrow$ 2成像面平行



无 epipolar (epipolar 在无穷远)

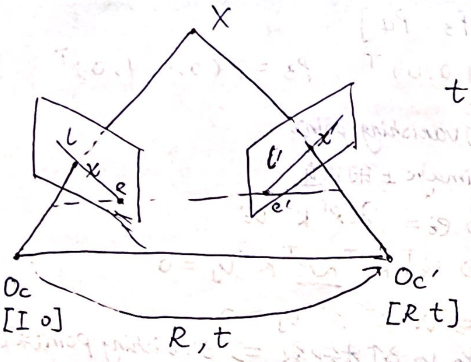
x, x' 等高

$$E = [t_\times]R = [t_\times] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix}$$

$\vec{t} = (T, 0, 0)$

$$x'^T E x = 0$$

$$(u' \quad v' \quad 1)\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{pmatrix}\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = 0 \Rightarrow vT - v'T = 0 \Rightarrow v = v'$$



$$X' = RX + t$$

$$t \times X' = t \times RX + \underbrace{t \times t}_{0}$$

$$= t \times RX$$

$$X' \cdot (t \times X') = 0$$

$$X' \cdot (t \times RX) = 0$$

$$X'^T [t\times] R X = 0$$

$$[t\times] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

设 $X = (x, 1)^T \Rightarrow \left.\begin{array}{l} x = [I \quad 0] X = x \\ x' = [R \quad t] X = Rx + t \end{array}\right\}$

$$\Rightarrow x'^T \underbrace{[t\times] R}_{E} x = 0$$

$[t\times]: 3$, $R: 3$

$x'^T E x = 0$ $rank(E) = 2$ ⇒ 将1个归一化：$\boxed{5}$个自由元

∴ 已知 $x \Rightarrow l' = [t\times] R x = Ex$

已知 $x' \Rightarrow x'^T (E^T x') = 0 \Rightarrow l = E^T x'$ $null(E) = e$

$\Rightarrow rank(E) = 2$

l, l' : epipolar line $\Rightarrow Ee = 0$, $E^T e' = 0$

epipolar 无 epipolar line.

$\hat{x}'^T E \hat{x} = 0$ （真正写法 有^）

若 C_1, C_2 内参已知：

$$x' = K' \hat{x}', \quad x = K \hat{x}$$

$$\Rightarrow \hat{x}' = K'^{-1} x', \quad \hat{x} = K^{-1} x$$

$$\Rightarrow (K'^{-1} x')^T E (K^{-1} x) = 0$$

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

$$(x', y', 1) F \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

$(f_{11} \cdots)$

$Ax = 0$

$rank(A) = 9 \Rightarrow$ 只有0解

$\Rightarrow rank(A) = 8 \Rightarrow$ 8点法

$$x'^T \underbrace{K'^{-T} E K^{-1}}_{F} x = 0$$

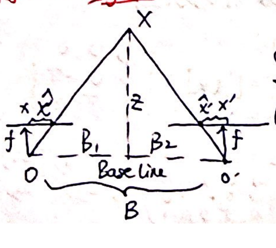
保证 E、F $rank = 2$：SVD 舍去最小的 singular value.

将1个归一化：$\boxed{8}$个自由元 ⇒ 8点法求解

$F \begin{pmatrix} 1 & f_{12} & \\ f_{21} & f_{22} & \\ f_{22} & f_{32} & \end{pmatrix} \rightarrow \alpha c_1 + \beta c_2$

depth from disparity

像机到 image plane 距离为f



$$\left\{\begin{array}{l} \frac{x}{f} = \frac{B_1}{z} \\ \frac{-x'}{f} = \frac{B_2}{z} \end{array}\right.$$

$$\underbrace{x - x'}_{\text{disparity}} = \frac{B \cdot f}{z}$$

disparity $\propto \frac{1}{z}$

前提：

1° Camera center same height
2° camera has same focal length
3° image plane 平行，且 // base line
4° epipolar line // base line

homography 单应性

⇒ 多视角重建 multi-view stereo MVS ⇒ depth ⇒ 点云

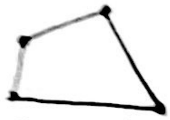
2D 情况

1. □ $\xrightarrow{R,t}$ ◇ $R:\theta$ 1, $t:(x,y)$ 2

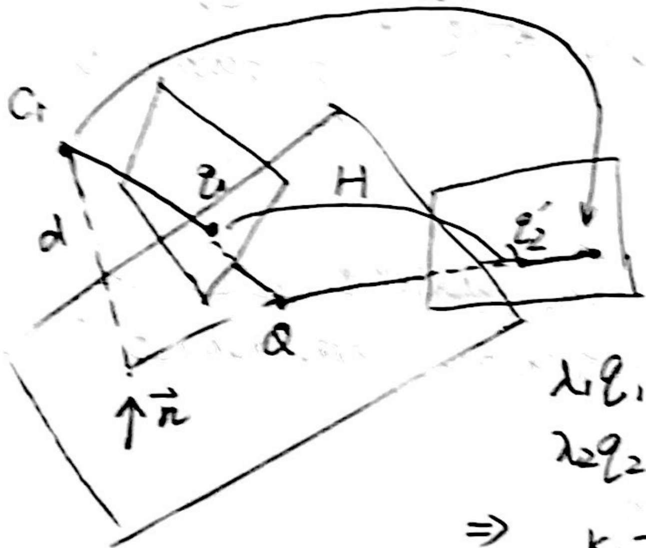
2. affine 仿射 平行关系不变

□ ⟶ ▱ 3点 确定关系 (x,y) 6

$x' = Mx + t$

3. projective (homography) □ ⟶ 

4点确定关系(8个自由度)



Q在C系下 Q_1

C' Q_2

$Q_2 = RQ_1 + t$

$\vec{n}\cdot\vec{Q_1} + d = 0 \Rightarrow 1 = -\frac{n^TQ_1}{d}$

$Q_2 = RQ_1 + t\cdot 1 = (R - t\frac{n^T}{d})Q_1$

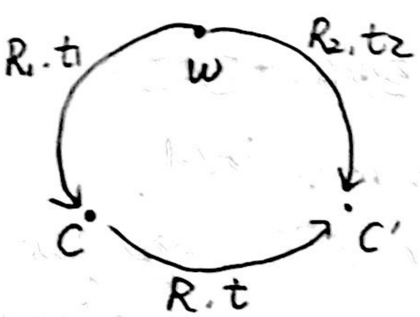
$\lambda_1 q_1 = K_1[I\ 0]\begin{bmatrix} Q_1 \\ 1 \end{bmatrix} = K_1Q_1$

$\lambda_2 q_2 = K_2[I\ 0]\begin{bmatrix} Q_2 \\ 1 \end{bmatrix} = K_2Q_2$

$\Rightarrow K_2^{-1}q_2 = \lambda(R - t\frac{n^T}{d})K_1^{-1}q_1$

$q_2 = \lambda \underbrace{K_2(R - t\frac{n^T}{d})K_1^{-1}}_{H} q_1$

考虑外参:



$$\underbrace{\begin{bmatrix} R_2 & t_2 \\ 0 & 1 \end{bmatrix}^{-1}}_{C'\to w}\underbrace{\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}}_{C\to C'}\underbrace{\begin{bmatrix} R_1 & t_1 \\ 0 & 1 \end{bmatrix}}_{w\to C}\underbrace{P}_{w下} = P$$

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_2 & t_2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} R_1 & t_1 \\ 0 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} R_2 & t_2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} R_1^{-1} & -R_1^{-1}t_1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} R_2R_1^{-1} & -R_2R_1^{-1}t_1 + t_2 \\ 0 & 1 \end{bmatrix}$$

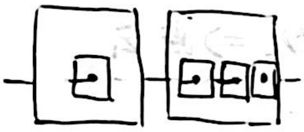
$R = R_2R_1^{-1}$

$t = -R_2R_1^{-1}t_1 + t_2$

$\Rightarrow H = K_2(R_2\underbrace{R_1^{-1}}_{R_1^T} - (-R_2R_1^{-1}t_1 + t_2)\frac{n^T}{d})K^{-1}$

$= K_2R_2(I - \frac{1}{d}(R_1^{-1}t_1 + R_2^{-1}t_2)n^TR_1)R_1^TK^{-1}$

stereo matching



图像特征 1° Pixel $\longrightarrow$ 2° local feature $\longrightarrow$ 3° global feature

$\Downarrow$ 信息不够

 感受野 reception feature

 ⇒各向异性

failure

1° 无纹理/弱纹理 textureless

2° 重复纹理 repetition

3° 动态 dynamic

4° 遮挡 occlusion

马氏距离

Mahalanobias distance.

$$d^2(x,y) = \sum \alpha_i (x_i - y_i)^2 = \sum M_{ii}(x_i - y_i)^2$$

↳对角矩阵

$$x^T M y = \sum M_{ij}\, x_i^T y_j \geq 0 \Rightarrow M: \text{半正定} \begin{cases} 1.\ \text{超参} \\ 2.\ \text{data driven} \end{cases}$$

M_{ij} 的相关性 ⇒ attention

distance F、L_1、L_2、马氏

↳ 对比核直方图，限制sparse的个数：L_1正则化

VQ-VE vector qantification 量化 ⇒ more robust.

Fitting a homography

$$\lambda x_i' = H x_i \Rightarrow x_i' \times H x_i = 0$$

$$H = \begin{pmatrix} h_1^T \\ h_2^T \\ h_3^T \end{pmatrix} \qquad Hx = \begin{pmatrix} h_1^T x_i \\ h_2^T x_i \\ h_3^T x_i \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \times \begin{pmatrix} h_1^T x_i \\ h_2^T x_i \\ h_3^T x_i \end{pmatrix} = 0$$

(H) h：自由度 8 ⇒ 至少4对匹配点

$$\begin{pmatrix} 0 & -x_i^T & y' x_i^T \\ x_i^T & 0 & -x' x_i^T \\ -y' x_i^T & x' x_i^T & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = 0$$

$$Ah = 0$$

find $h \Rightarrow \min |Ah|$

let h: min A^TA's eigenvalue
$\Uparrow$
eigenvector

$A^TAh = \lambda_{min} h$.

$h^TA^TAh = \lambda \|h\| = \lambda_{min}$

$\|Ah\| = \lambda_{min}$.

MVS Net

视锥  对深度做slice.

匹配: basic: pixel层面
deep learning ⇒ feature层面
MVS Net: 不同视角的图像, 相同网络提feature
⇓
核心: 计算volume.
$\boxed{\text{cost volume}}$

假设每个view depth已知 ⇒ warp到 reference view下

$\xRightarrow{\text{n-view}}$ n-feature volume V_i (匹配特征)

$\xRightarrow{\text{找深度}}$ s.t. V_i匹配到 reference.

cost volume: 特征 $\xRightarrow{\text{卷积}}$ 回归 g.t. 深度.

1. 离散化 深度　　2. depth求正

MVS Net.	Nerf.
先得到depth ⇒ 点云	一个Network: 唯一的场景 (SDF)
10个场景train出 ⇒ 可用于第11个	

Combine: MVS 得到depth, 作为prior约束NeRF

Limitation: cost volume 3维
high resolution ⇒ Memory

Ransac 随机采样一致. RANdom SAmple Consensus

eg. Homography

1. random select a group of matches
2. compute H of selected matches
3. Find inlier to the ~~the~~ H
4. 若inlier数目超threshold
   对所有inlier做least square.

Pros: 1. simple, general.
2. work well

Cons: 1. params
2. low ~~inlier~~ inlier rate ⇒ work ×
3. initialization ×

NeRF.

$\sigma(x)$: density prob: hit at x

$T(x)$ transmittance, prob: $0 \to t$ not hit any particle

$$T(t+\delta t) = T(t)(1-\delta t\,\sigma(t))$$

$$\frac{T(t+\delta t)-T(t)}{\delta t} = -T(t)\sigma(t)$$

$\delta(t) \to 0$ $$\frac{T'(t)}{T(t)} = -\sigma(t)$$

$$\int_a^b \frac{T'(t)}{T(t)}dt = -\int_a^b \sigma(t)dt.$$

$$\log T(t)\Big|_a^b = -\int_a^b \sigma(t)dt$$

$$T(a\to b) = e^{-\int_a^b \sigma(t)dt}$$

$a \to b \to c$

$$T(a\to c) = e^{-\int_a^b \sigma(t)dt \cdot \int_b^c \sigma(t)dt}$$

$$= T(a\to b)\cdot T(b\to c)$$

$$C(a\to b) = \int_a^b \underline{T(a\to t)\sigma(t)c(t)}dt + \boxed{T(b)}C_{bg}$$

$a \to t$ 没hit，在t处hit到了color为$c(t)$的particle

Piecewise constant 逐段一致.

$T(t)=T(0\to t)=T(0\to t_n)\cdot T(t_n\to t)$

$$C(t_{N+1}) = \sum_{n=1}^{N}\int_{t_n}^{t_{n+1}} T(t)\sigma(t)c(t)dt$$

constant

$$= \sum_{n=1}^{N}\int_{t_n}^{t_{n+1}} T(t_n)T(t_n\to t)\sigma_n c_n dt$$

$$= \sum_{n=1}^{N} T(t_n)\int_{t_n}^{t_{n+1}} T(t_n\to t)\sigma_n c_n dt$$

$$\hookrightarrow = \sigma_n c_n \int_{t_n}^{t_{n+1}} \exp\left(-\int_{t_n}^{t} \sigma(\tau)d\tau\right)$$

constant

$$= \int_{t_n}^{t_{n+1}} \exp\left(-\sigma(t_n)\int_{t_n}^{t} d\tau\right)dt$$

$$= \int_{t_n}^{t_{n+1}} \exp(-\sigma(t_n)(t-n))dt$$

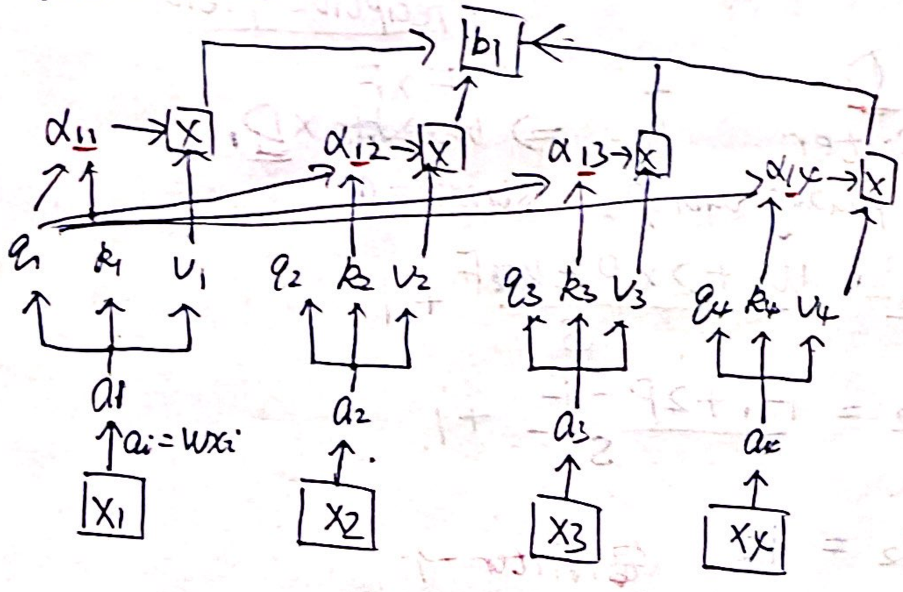
$$= \frac{1}{\sigma_n}\cdot(1-\exp(-\sigma_n(t_{n+1}-t_n)))$$

$$= \sum_{n=1}^{N} T(t_n)(1-\exp(-\sigma_n(t_{n+1}-t_n)))c_n.$$

SDF. signed distance field.

$$SDF(p) = sign(p)\cdot \min_{q\in S}\|p-q\|.$$

attention



better capture long-range dependencies within sequence.

self attention

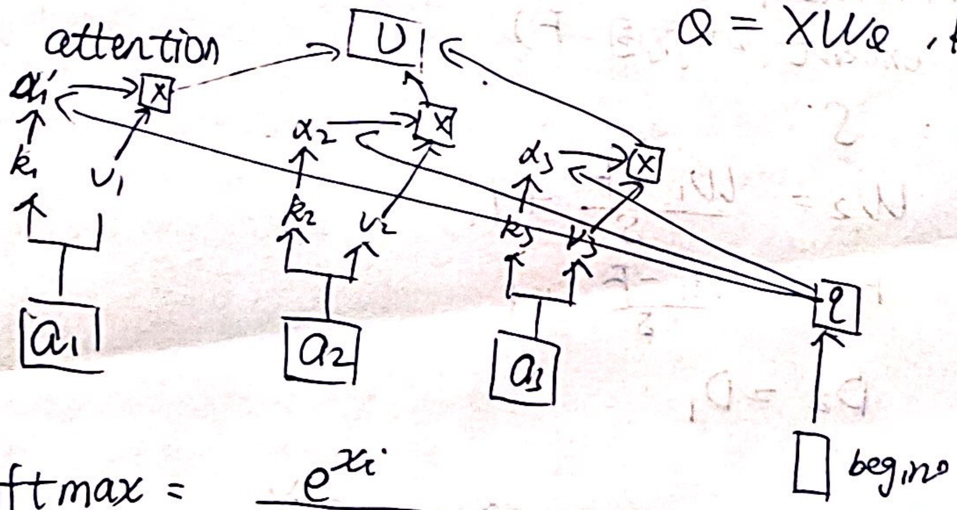
$$\alpha_{i,j} = \frac{q_i \cdot k_j}{\sqrt{d_k}}$$

output: weighted sum of V

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$Q = XW_Q,\ K = XW_K,\ V = XW_V$$

Cross attention



$$\text{softmax} = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

adjust focus dynamically

transformer可起到降维的作用（局部压缩

CNN

reciptive field

image: $W_1 \times H_2 \times D$ I $F \times F$

kernel: 共 K 个 filter   ker $\Rightarrow \cancel{W_2 \times H_2} \times D_1$

zero padding with P.   stride = S

$I * K \Rightarrow$

$$W_2 = \frac{W_1 + 2 \times P - \cancel{W}F}{S} + 1.$$

$$H_2 = \frac{H_1 + 2P - F}{S} + 1.$$

$$D_2 = K.$$

每个filter一个

每个kernel: $w^T x + b \rightarrow$ 单独的数

Pooling: 无参数，不改变通道数.

input $W_1 \times H_1 \times D_1$.

Spatial extent (空间范围 F)

stride: S

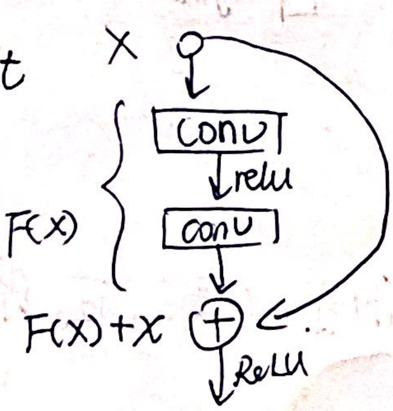
output:

$$W_2 = \frac{W_1 - F}{S} + 1.$$

$$H_2 = \frac{H_1 - F}{S} + 1$$

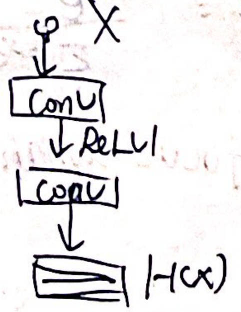
$$D_2 = D_1$$

ResNet



$\boxed{F(x) = H(x) - x}$

origin:



skip connection

不同深度可提feature

⇓

ResNext

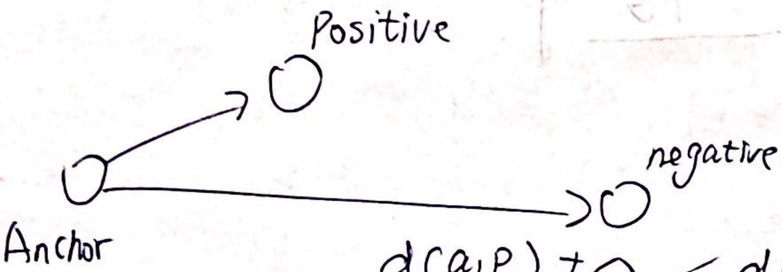
GoogLeNet

deepernet 扩大深度

⇒ computational efficiency

triplet loss

$$L = max(d(a,p) - d(a,n) + \Delta, 0)$$



$$d(a,p) + \Delta < d(a,n)$$

Δ margin: s.t. a距n尽可能大
a距p ——小、

fine-grained 增加显著性

catagory level：区分人、狗、猫

fine-grame image classification
细密度~~三~~区分：eg. 人脸识别
区分狗的品种

可引入 bias 作为先验.

先用 cross-entrpy
后用 triplet loss

image sampling  不可导
⇒ bilinear interacton ⇒ 连续.

depth estimator
用几何特征 Geometry feature 做约束

仿 masked-auto-encoder

图像信息 reductant ⇒ mask 75% 以上 feature

1°~7° 数据增强

1. auto encoder.

NLP { Bert bidirectional encoder ~~represt~~ representation from transform
Mask中间，周围推出
GPT given 前，predict 后.

2° coloring $\boxed{rgb} \longrightarrow \boxed{grey}$

Fo.

3° rotation

4. 图像切割，打乱

/ given $\boxed{①}$, predict 位置

5. contrastive learning 对比学习 (self-supervised)

$\boxed{I_1} \to \boxed{I_{11}}, \boxed{I_{12}}$ $\boxed{I_2} \to \boxed{I_{21}}, \boxed{I_{22}}$

$$||I_{11} - I_{12}|| < ||I_{11} - I_{21}||$$

6. supervised learning ⇒ knowledge

interaction ⇔ feedback

7. scale / flip vertically / horizontally

Monodepth ⇒ multi-task learning

匹配 ⇐ 难点 弱/无纹理，

重复纹理，动态，遮挡

occulsion.

简化MLP:

① random connection ② local + global.

③ remove connect，减少node，depth，local connection

⇒ convolution.

④ knowledge distillation 知识蒸馏.

$I^{teacher}_{1000层}$ $I^{student}_{500\ layer}$. $||f_{cnn500} - f_{cnn1000}||$.