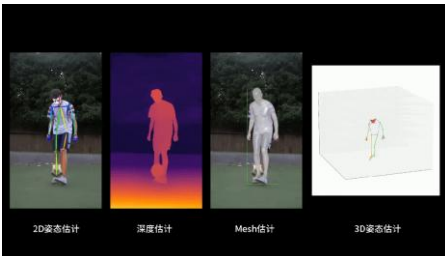




计算机视觉

第10章 三维重建

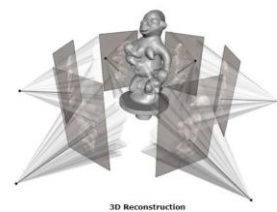
陈飞: chenfei314@fzu.edu.cn



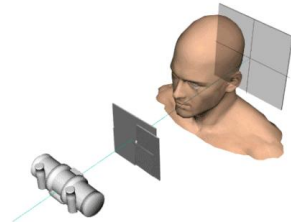
本章内容



- ❑ 三维重建概述
- ❑ 摄像机几何
- ❑ 摄像机标定
- ❑ 双视图与三角化
- ❑ 双目立体视觉
- ❑ 多视图几何



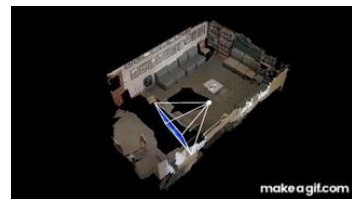
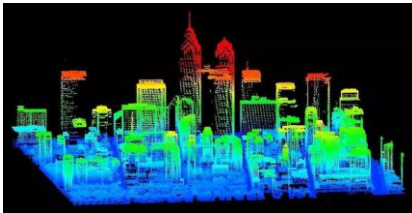
3D Reconstruction



三维重建概述

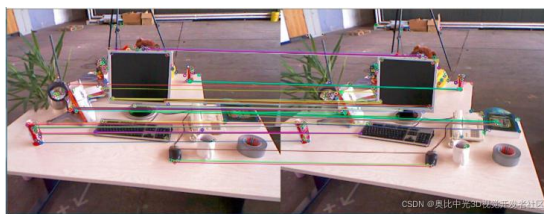
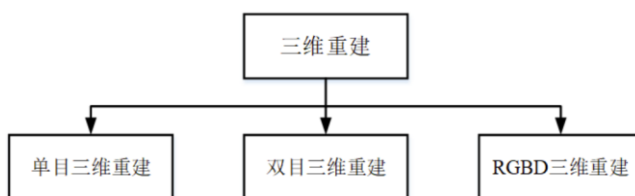


- 三维重建是指对三维物体建立**适合计算机表示和处理的数学模型**，是在计算机环境下对其进行处理、操作和分析其性质的基础，也是在计算机中建立表达客观世界的虚拟现实的关键技术。
- 物体三维重建是**计算机辅助几何设计(CAGD)**、**计算机图形学(CG)**、**计算机动画**、**计算机视觉**、**医学图像处理**、**科学计算和虚拟现实**、**数字媒体创作**等领域的共性科学问题和核心技术。



3

三维重建概述



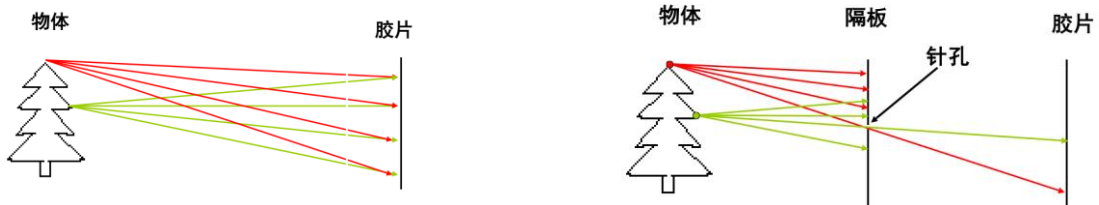
4

摄像机几何



• 小孔成像原理

当将胶片直接放置在物体前方时，3D物体上的同一点会在胶片的多个位置产生成像；



光通过一个小孔后，会在后面的屏幕上形成一个倒立的实像。

5

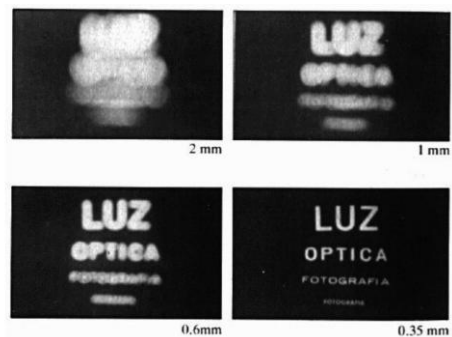
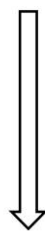
小孔成像原理



- 在物体和胶片之间放置一个带有针孔的隔板时，假设针孔大小只允许穿过一条光线，那么3D物体上的同一点只能有一条光线穿过小孔并在胶片上成像；

此处的小孔相当于就是光圈，光圈的尺寸越大，会有更多的光线穿过小孔，成像会越亮，但是同一个3D点会有多条光线穿过小孔并在胶片上的多个位置（相近）产生成像，导致物体成像模糊。光圈越小，物体成像越清晰，但是亮度也越暗；

缩小
光圈

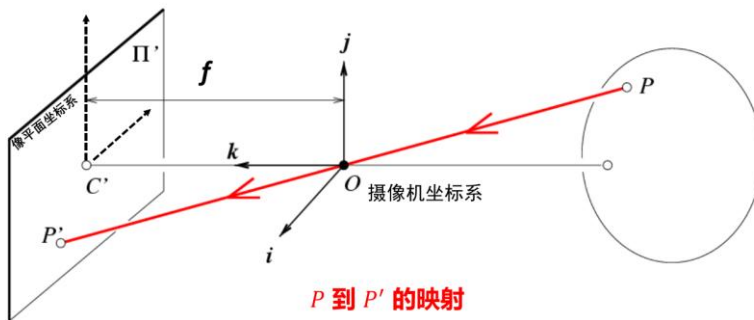


6

针孔相机数学模型



- 针孔相机数学模型



坐标系	维度
世界坐标系	3D
相机坐标系	3D
图像坐标系	2D
像素坐标系	2D

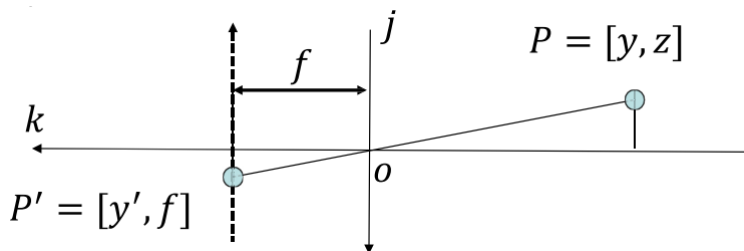
描述如何将三维世界中的点 (3D) 投影到二维图像平面 (2D) 的理想化几何模型。

7

针孔相机数学模型



- P为空间3D点, P'为成像后的像平面的2D点



$$\frac{y'}{f} = \frac{y}{z}$$

$$y' = f \frac{y}{z}$$

$$\frac{x'}{f} = \frac{x}{z}$$

$$x' = f \frac{x}{z}$$

将上面的三维示意图投影到二维平面 (k-j平面, 垂直于i轴) 后, 会产生一对相似三角形:

同理, 若将将上面的三维示意图投影到二维平面 (k-i平面, 垂直于j轴) 后, 会产生另一对相似三角形:

13

摄像机坐标系 - 像素坐标系



- 3D摄像机坐标系 - 2D CCD坐标系
- 2D CCD坐标系 - 2D像素坐标系

$$\begin{cases} x' = f \frac{x}{z} \\ y' = f \frac{y}{z} \end{cases} \quad \begin{cases} u = fk \frac{x}{z} + c_x \\ v = fl \frac{y}{z} + c_y \end{cases}$$

CCD坐标系和像素坐标系存在两个不同之处:

坐标系原点不一致: CCD坐标系的远点在图像中心, 像素坐标系的原点在左上角, c_x 和 c_y 就是原点的平移量; 度量单位不一致: CCD坐标系的单位为m, 像素坐标系的单位为pixel, k, l 分别为 x, y 方向的度量单位变换量, 单位为pixel/m.

14

摄像机坐标系 - 像素坐标系



将 fk 和 fl 表示为 α 和 β 可得:

$$\begin{cases} u = \alpha \frac{x}{z} + c_x \\ v = \beta \frac{y}{z} + c_y \end{cases}$$

齐次坐标变换:

$$P' = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha x + c_x z \\ \beta y + c_y z \\ z \end{bmatrix} = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

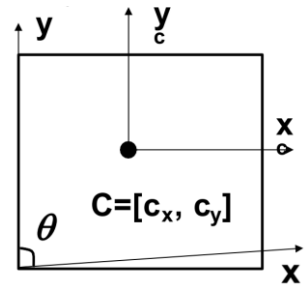
15

摄像机坐标系 - 像素坐标系



- 摄像机偏斜

$$P' = \begin{bmatrix} \alpha & -\alpha \cot\theta & c_x & 0 \\ 0 & \frac{\beta}{\sin\theta} & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = MP = K [I \quad 0] P$$



内参数矩阵K

$$K = \begin{bmatrix} \alpha & -\alpha \cot\theta & c_x \\ 0 & \frac{\beta}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

其中，**K**是摄像机内参数矩阵，内参数矩阵决定了摄像机坐标系下3D空间点到2D图像点的映射；**K**有个**5**自由度DOF；

16

像素坐标系-世界坐标系



- 投影矩阵

像素平面上的点的
齐次坐标

维度: 3×1

$$P' = K [I \quad 0] P = K [I \quad 0] \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} P_w = K [R \quad T] P_w = M P_w$$

$$K = \begin{bmatrix} \alpha & -\alpha \cot\theta & c_x \\ 0 & \frac{\beta}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

维度: 3×3

三维点在摄像机坐标系下的齐次坐标

维度: 4×1

三维点在世界坐标系下的齐次坐标

维度: 4×1

投影矩阵

维度: 3×4

17

像素坐标系-世界坐标系



- 投影矩阵有11个自由度，其中包含5个摄像机内参数+6个摄像机外参数。

$$P' = K[I \ 0]P = K[I \ 0] \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} P_w = K \begin{bmatrix} R & T \end{bmatrix} P_w = MP_w$$

完整的摄像机模型!

内参矩阵 外参矩阵

投影矩阵

错切
旋转 放缩 → 线性变换

平移

$$\begin{bmatrix} a_0 & a_1 & b_1 \\ a_2 & a_3 & b_2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$P' = MP_w = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} P_w = \begin{bmatrix} m_1 P_w \\ m_2 P_w \\ m_3 P_w \end{bmatrix} \quad M = \begin{bmatrix} A & b \\ v & 1 \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

透视

$$\rightarrow \left(\frac{m_1 P_w}{m_3 P_w}, \frac{m_2 P_w}{m_3 P_w} \right)$$

18

投影矩阵性质定理



Faugeras定理:

在给定足够多的视图和相应的摄像机参数的情况下，我们可以唯一地重建出场景的三维结构。

$$M = K[R \ T] = [KR \ KT] = [A \ b]$$

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ a_3^T \end{bmatrix}$$

令 $M = (A \ b)$ 为 3×4 的矩阵， $a_i^T (i = 1, 2, 3)$ 表示由矩阵 A 的行

- M 是透视投影矩阵的一个充分必要条件是 $\text{Det}(A) \neq 0$
- M 是零倾斜透视投影矩阵的一个充分必要条件是 $\text{Det}(A) \neq 0$ 且 $(a_1 \times a_3) \cdot (a_2 \times a_3) = 0$
- M 是零倾斜且宽高比为1的透视投影矩阵的一个充分必要条件是 $\text{Det}(A) \neq 0$ 且

$$\begin{cases} (a_1 \times a_3) \cdot (a_2 \times a_3) = 0 \\ (a_1 \times a_3) \cdot (a_1 \times a_3) = (a_2 \times a_3) \cdot (a_2 \times a_3) \end{cases}$$

19

最小二乘解



线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2q}x_q = y_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q = y_p \end{cases} \iff \mathbf{Ax} = \mathbf{y}$$

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$$

当矩阵 A 列满秩时:

- $p < q$ 欠定方程组, 方程个数少于未知量个数, 有多解;
- $p = q$ 方程个数等于未知量个数时, 有唯一解;
- $p > q$ 超定方程组, 方程个数多于未知量个数, 无解。(除非 \mathbf{y} 可以由 A 的列向量线性表示)

20

最小二乘解



当矩阵 A 列满秩时: $p > q$ 超定方程组, 方程个数多于未知量个数, 无解。

$$E(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{i=1}^p (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{iq}x_q - y_i)^2 = \|\mathbf{Ax} - \mathbf{y}\|^2$$

线性方程组的最小二乘解 $\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{Ax} - \mathbf{y}\|^2$

求解方法1: $\mathbf{x}^* = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$

求解方法2: 1. 矩阵 A 进行奇异值分解 $A = \mathbf{U} \mathbf{D} \mathbf{V}^T$

2. 令 $\tilde{\mathbf{y}} = \mathbf{U}^T \mathbf{y}$, 计算 \mathbf{b} , 其中 $b_i = \tilde{y}_i / d_i$, d_i 为对角矩阵 D 的第 i 个元素

3. $\mathbf{x}^* = \mathbf{V} \mathbf{b}$

求解方法3: 利用牛顿法及其改进或梯度下降法迭代求解

21

线性方程组



线性方程组

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_q \end{pmatrix}$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2q}x_q = 0 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q = 0 \end{cases} \iff Ax = 0$$

当矩阵 A 列满秩时:

- $p < q$ 欠定方程组, 方程个数少于未知量个数, 有多解
- $p = q$ 方程个数等于未知量个数时, 只有零解;
- $p > q$ 超定方程组, 方程个数多于未知量个数, 除零解外无解。

22

最小二乘解



当矩阵 A 列满秩时: $p > q$ 超定方程组, 方程个数多于未知量个数, 除零解外无解。

$$E(x) \stackrel{\text{def}}{=} \sum_{i=1}^p (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{iq}x_q)^2 = \|Ax\|^2$$

注意: 如果 x 是方程组的解, 对于任意 $\lambda \neq 0$, λx 也是该方程的解;

齐次线性方程组的最小二乘解

$$\begin{aligned} x^* &= \arg \min_x \|Ax\| \\ \text{s.t. } \|x\| &= 1 \end{aligned}$$



求解方法: 1. 矩阵 A 进行奇异值分解 $A = UDV^T$
2. x^* 为 V 矩阵的最后一列 (最小特征值对应的右特征向量)

23

非线性方程组的最小二乘解



一般系统, 有 p 个方程, q 个变量

$$\begin{cases} f_1(x_1, x_2, \dots, x_q) = 0 \\ f_2(x_1, x_2, \dots, x_q) = 0 \\ \vdots \\ f_p(x_1, x_2, \dots, x_q) = 0 \end{cases} \iff f(\mathbf{x}) = \mathbf{0}$$

当 $p > q$ 时的最小二乘解:

$$E(\mathbf{x}) \triangleq \sum_{i=1}^p f_i^2(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|^2$$

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{f}(\mathbf{x})\|^2$$

- 牛顿法与 列文伯格-马夸尔特法 (L-M方法)

- 从初始解开始迭代, 若初始解与实际相距较远, 可能会很慢
- 估计解可能是初始解的函数 (由于局部最小值)
- 牛顿法需要计算一阶导数矩阵 J (雅可比矩阵), 二阶导数矩阵 H (海塞矩阵)
- L-M算法不用计算 H

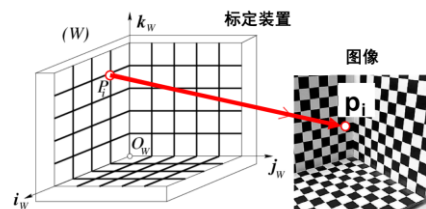
24

摄像机标定



- 摄像机标定: 求解摄像机内、外参数矩阵 $K [R T]$;
- 因为摄像机内外参数矩阵描述了三维世界到二维像素的映射关系;

$$p = MP = K[R T]P$$



- 世界坐标系中 P_1, \dots, P_n 位置已知
- 图像中 p_1, \dots, p_n 位置已知

25

摄像机标定



- 投影矩阵M求解

$$p_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \frac{m_1 P_i}{m_3 P_i} \\ \frac{m_2 P_i}{m_3 P_i} \end{bmatrix}$$

像素

$$M = K[R \ T] = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

11个未知量!

最少6对点对应!

$$u_i = \frac{m_1 P_i}{m_3 P_i} \rightarrow u_i(m_3 P_i) = m_1 P_i \rightarrow m_1 P_i - u_i(m_3 P_i) = 0$$

$$v_i = \frac{m_2 P_i}{m_3 P_i} \rightarrow v_i(m_3 P_i) = m_2 P_i \rightarrow m_2 P_i - v_i(m_3 P_i) = 0$$

实际操作中使用多于六对点来获得更加鲁棒的结果。

26

摄像机标定



- 摄像机内外参数共同构成了投影矩阵，投影矩阵共有11个未知量；
- 每对点可以列出两个方程，因此，最少需要6对对应点；
- 实际操作中通常使用多于6对点来获得更加鲁棒的结果；

$$\begin{cases} -u_1(m_3 P_1) + m_1 P_1 = 0 \\ -v_1(m_3 P_1) + m_2 P_1 = 0 \\ \vdots \\ -u_n(m_3 P_n) + m_1 P_n = 0 \\ -v_n(m_3 P_n) + m_2 P_n = 0 \end{cases} \rightarrow \begin{matrix} \text{已知} & \text{未知} \\ \boxed{P\bar{m} = 0} \end{matrix}$$

$$P \stackrel{\text{def}}{=} \begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \vdots & \vdots & \vdots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix}_{2n \times 12}$$

$$\bar{m} \stackrel{\text{def}}{=} \begin{pmatrix} m_1^T \\ m_2^T \\ m_3^T \end{pmatrix}_{12 \times 1}$$

方程个数 $2n$ 个，且 $n > 6$ ；未知参数11个；这是一个超定齐次线性方程组；

27



投影矩阵M求解

$$P\vec{m} = 0 \quad \begin{array}{l} \text{最小二乘解} \\ \iff \end{array} \quad \begin{array}{l} \min_m \|P\vec{m}\| \\ \text{s. t. } \|\vec{m}\|=1 \end{array}$$

奇异值分解!!!

$U_{2n \times 12} D_{12 \times 12} V_{12 \times 12}^T$

求解方法: 1. 矩阵 P 进行奇异值分解 $P = UDV^T$
 2. x^* 为 V 矩阵的最后一列 (最小特征值对应的右特征向量)

结论: \vec{m} 为 V 矩阵的最后一列, 且 $\|\vec{m}\| = 1$

28

提取摄像机内参数



$$p = MP = K[R \ T]P$$

内参数

$$K = \begin{bmatrix} \alpha & -\alpha \cot \theta & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

外参数

$$R = \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix} \quad T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

$$M = K[R \ T] = [KR \ KT] = [A \ b]$$

29

提取摄像机内参数



考虑投影矩阵M求解时，假定 $\|m\| = 1$ ，因此在M前面乘了一个 ρ

$$\rho M = \begin{pmatrix} \alpha r_1^T - \alpha \cot \theta r_2^T + u_0 r_3^T & \alpha t_x - \alpha \cot \theta t_y + u_0 t_z \\ \frac{\beta}{\sin \theta} r_2^T + v_0 r_3^T & \frac{\beta}{\sin \theta} t_y + v_0 t_z \\ r_3^T & t_z \end{pmatrix}$$

$$\rho A = \rho \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \end{pmatrix} = \begin{pmatrix} \alpha r_1^T - \alpha \cot \theta r_2^T + u_0 r_3^T \\ \frac{\beta}{\sin \theta} r_2^T + v_0 r_3^T \\ r_3^T \end{pmatrix} = KR$$

30

摄像机标定结果



内参数

$$\rho = \frac{\pm 1}{|a_3|} \quad \begin{aligned} u_0 &= \rho^2 (a_1 \cdot a_3) \\ v_0 &= \rho^2 (a_2 \cdot a_3) \end{aligned}$$

$$\cos \theta = -\frac{(a_1 \times a_3) \cdot (a_2 \times a_3)}{|a_1 \times a_3| \cdot |a_2 \times a_3|}$$

$$\alpha = \rho^2 |a_1 \times a_3| \sin \theta$$

$$\beta = \rho^2 |a_2 \times a_3| \sin \theta$$

外参数

$$r_1 = \frac{(a_2 \times a_3)}{|a_2 \times a_3|}$$

$$r_3 = \frac{\pm a_3}{|a_3|}$$

$$r_2 = r_3 \times r_1$$

$$T = \rho K^{-1} b$$

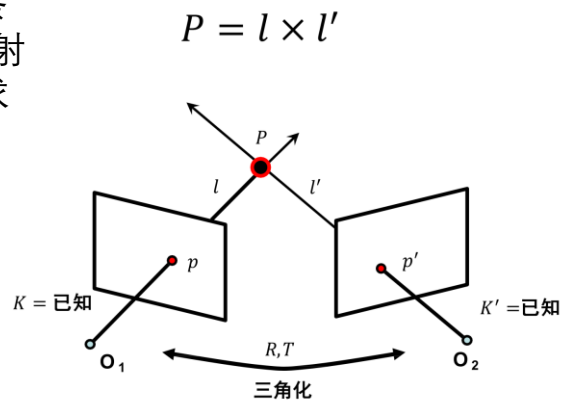
31

双视图与三角化



只要已知直线 l 和 l' ，则三维点 P 就是两条直线的交点。因此，只要将直线 l 和 l' 映射到同一坐标系下，就可以通过 $P = l \times l'$ 求出点 P 的三维坐标。

此时，问题演变成了：已知二维像素坐标 p 和 p' ， K 和 K' ，以及两视图之间的变换矩阵 R 、 T ，求解 P 点的三维坐标？



32

线性解



$$\begin{cases} p = MP = K[I \ 0]P \\ p' = M'P = K'[R \ T]P \end{cases} \Rightarrow \begin{cases} u = \frac{m_1 P}{m_3 P} \rightarrow m_1 P - u(m_3 P) = 0 \\ v = \frac{m_2 P}{m_3 P} \rightarrow m_2 P - v(m_3 P) = 0 \\ u' = \frac{m'_1 P}{m_3 P} \rightarrow m'_1 P - u'(m_3 P) = 0 \\ v' = \frac{m'_2 P}{m_3 P} \rightarrow m'_2 P - v'(m_3 P) = 0 \end{cases} \Rightarrow AP = 0$$

$$A = \begin{bmatrix} um_3 - m_1 \\ vm_3 - m_2 \\ u'm'_3 - m'_1 \\ v'm'_3 - m'_2 \end{bmatrix}$$

➤ 方程数4个
 ➤ 未知参数3个

超定齐次线性方程组

最小二乘解：1. 矩阵 A 进行奇异值分解 $A = UDV^T$
 2. P 为 V 矩阵的最后一列

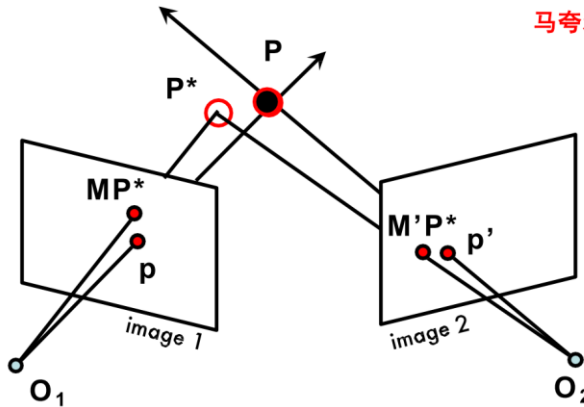
33



非线性解

- 寻找 P 最小化 $d(p, MP) + d(p', M'P)$

求解：牛顿法 与 列文伯格-
马夸尔特法 (L-M方法)



34

实际情况中的问题定义



- 在实际情况下，由于噪声的存在，两条直线通常不相交；
- 且上述的线性解和非线性解都需要已知 K 和 K' ， R 、 T ；
- 然而，实际情况下，摄像机的内参 K 和 K' 和两视图之间的变换矩阵 R 、 T 通常不可知，那么实际情况就变成了如下的问题：

问题1： 已知 p 和 p' ，摄像机内参数 K 和 K' ，

- 求解：摄像机的 R 、 T 以及 P 点的三维坐标？

问题2： 已知 p 和 p'

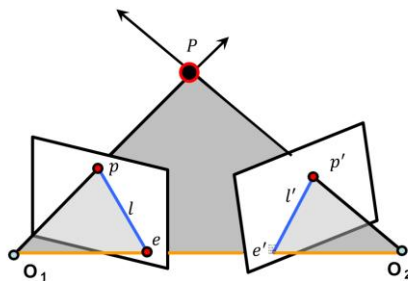
- 求解：摄像机内参数 K 和 K' ，摄像机的 R 、 T 以及 P 点的三维坐标？

35

极几何



- 极几何描述了同一场景或者物体的两个视点图像间的几何关系



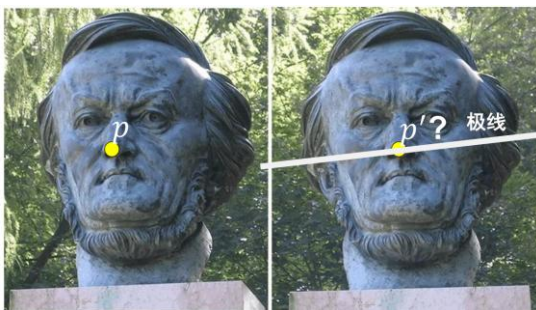
作用：
将搜索范围缩小到对应的极线上。

- 极平面: 过点 P , O_1 与 O_2 的平面
- 基线: O_1 与 O_2 的连线
- 极线: 极平面与成像平面的交线 l_1, l_2
- 极点: 基线与成像平面的交点 e_1, e_2

- 极平面相交与基线
- 极线相交于极点
- p 的对应点在极线 l' 上
- p' 的对应点在极线 l 上

36

极几何



- 极几何可以将对应点搜索范围缩小到对应的极线上；

极点关系：

- $F e_1 = 0$ (右极点与左极线的关系)
- $e_2^T F = 0$ (左极点与右极线的关系)

对极线: $l_2 = F x_1, l_1 = F^T x_2$

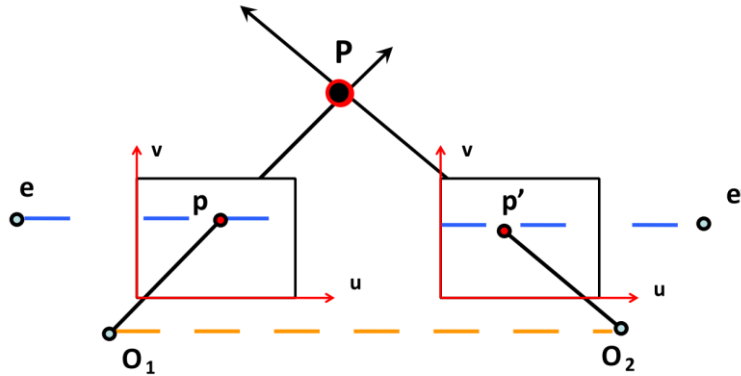
基础矩阵 F 是一个 3×3 的秩为 2 的齐次矩阵, 满足对任意一对对应点 $x_1 \leftrightarrow x_2$:

$$x_2^T F x_1 = 0$$

其中 x_1, x_2 是齐次图像坐标 (维度 3×1)。

37

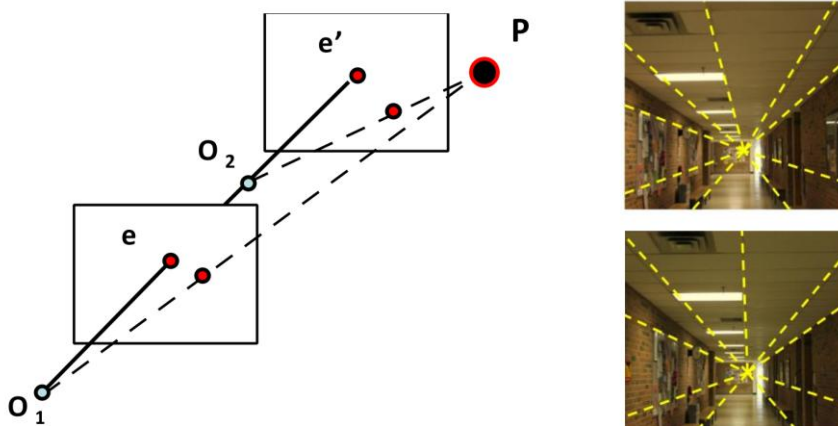
极几何特例——平行视图



- 两个图像平面平行；
- 基线平行于图像平面，极点 e 和 e' 位于无穷远处
- 极线平行于图像坐标系的 u 轴

38

极几何特例——前向平移（无旋转）



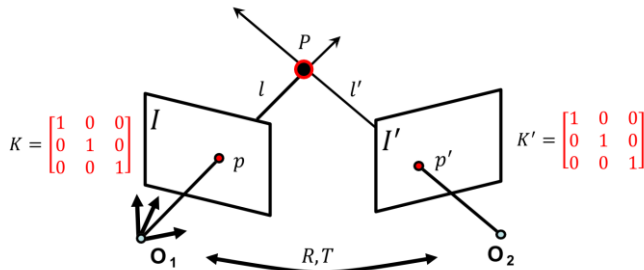
两幅图上极点的位置相同，极点称作展开焦点(focus of expansion)。

39



本质矩阵

- 本质矩阵对规范化摄像机拍摄的两个视点图像间的极几何关系进行代数描述；



图像 I 上的点 p 像素坐标为 (u, v)

图像 I' 上的点 p' 像素坐标为 (u', v')

$K = K'$ 已知且为规范化相机 \Rightarrow 三维点的欧式（非齐次）坐标 = 图像点的齐次坐标

空间点 p 在 O_1 坐标系下的非齐次坐标为 $(u, v, 1)$ 空间点 p' 在 O_2 坐标系下的非齐次坐标为 $(u', v', 1)$

40

基础矩阵



$$F = K'^{-T} [T_{\times}] R K^{-1}$$

基础矩阵 F 为：

- p 对应的极线是 l' ($l' = Fp$)
- p' 对应的极线是 l ($l = F^T p'$)
- $Fe = 0$ 与 $e'^T F = 0$
- F 是奇异的 (秩为2)
- F 7个自由度 (尺度无法确定, $\det(F) = 0$)

$$p'^T F p = 0$$

$$F = K'^{-T} [T_{\times}] R K^{-1}$$

F 是一个 3×3 、秩为 2、行列式为 0 的齐次矩阵，它完整刻画了两幅图像间的极几何关系，将左图像中的点映射为右图像中的极线，并通过 $x_2^T F x_1 = 0$ 约束对应点。

41

基础矩阵估计



- 八点法

$$p'^T F p = 0$$

F有7个自由度,
理论上7点可以解出F,
但计算比较复杂;

$$p'^T F p = 0 \quad p = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad p' = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}$$

$$(u', v', 1) \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = 0$$

单点方程矩阵形式展开:

$$\Rightarrow (uu', vu', u', uv', vv', v', u, v, 1) \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0$$

42

基础矩阵估计



- 选取八个点，列出齐次线性方程:

$$W \begin{bmatrix} u_1 u'_1 & v_1 u'_1 & u'_1 & u_1 v'_1 & v_1 v'_1 & v'_1 & u_1 & v_1 & 1 \\ u_2 u'_2 & v_2 u'_2 & u'_2 & u_2 v'_2 & v_2 v'_2 & v'_2 & u_2 & v_2 & 1 \\ u_3 u'_3 & v_3 u'_3 & u'_3 & u_3 v'_3 & v_3 v'_3 & v'_3 & u_3 & v_3 & 1 \\ u_4 u'_4 & v_4 u'_4 & u'_4 & u_4 v'_4 & v_4 v'_4 & v'_4 & u_4 & v_4 & 1 \\ u_5 u'_5 & v_5 u'_5 & u'_5 & u_5 v'_5 & v_5 v'_5 & v'_5 & u_5 & v_5 & 1 \\ u_6 u'_6 & v_6 u'_6 & u'_6 & u_6 v'_6 & v_6 v'_6 & v'_6 & u_6 & v_6 & 1 \\ u_7 u'_7 & v_7 u'_7 & u'_7 & u_7 v'_7 & v_7 v'_7 & v'_7 & u_7 & v_7 & 1 \\ u_8 u'_8 & v_8 u'_8 & u'_8 & u_8 v'_8 & v_8 v'_8 & v'_8 & u_8 & v_8 & 1 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = 0$$

- 齐次系统

$$Wf = 0$$

 f

43



最小二乘求解:

- 通常 $N > 8$, 为超定方程组, 采用最小二乘求解:

$$\begin{array}{l} \min_f \|Wf\| \\ \text{s.t. } \|f\| = 1 \end{array} \xrightarrow{\text{SVD}} \hat{F}$$

最小二乘解

f 为 W 矩阵最小奇异值的右奇异向量, 且 $\|f\| = 1$

44

满秩矩阵分解:



- 上述经过最小二乘估计得到的通常秩为3, 为满秩矩阵, 但基础矩阵 F 的秩为2。

$$\text{寻找} F \text{ 最小化 } \|F - \hat{F}\|_F \quad \text{s.t. } \det(F) = 0$$

$$\text{SVD}(\hat{F}) = U \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} V^T \Rightarrow F = U \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T$$

45

八点法求解总结



1. 构建 W 矩阵

2. 对 W 矩阵进行奇异值分解求 \hat{F}

f 为 W 矩阵最小奇异值的右奇异向量, 且 $\|f\| = 1$

3. 执行秩2约束 $\rightarrow F$

$$SVD(\hat{F}) = U \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} V^T \Rightarrow F = U \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T$$

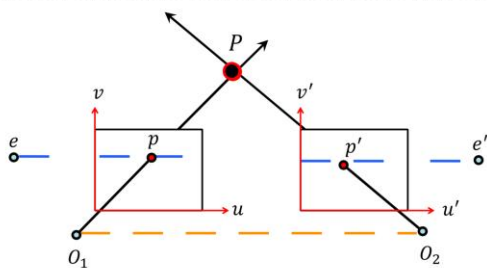
$$f = \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix}$$

46

基于平行视图的双目立体视觉



平行视图



- 两个图像平面平行;
- 基线平行于图像平面, 极点 e 和 e' 位于无穷远处

本质矩阵与基础矩阵

E = 本质矩阵

$$K = K' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

p 点像素坐标为 (u, v)
 p' 点像素坐标为 (u', v')

$$p'^T E p = 0$$

$$E = T \times R = [T_x] R$$

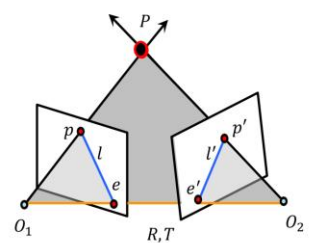
F = 基础矩阵

内参数分别为 K 和 K'

p 点像素坐标为 (u, v)
 p' 点像素坐标为 (u', v')

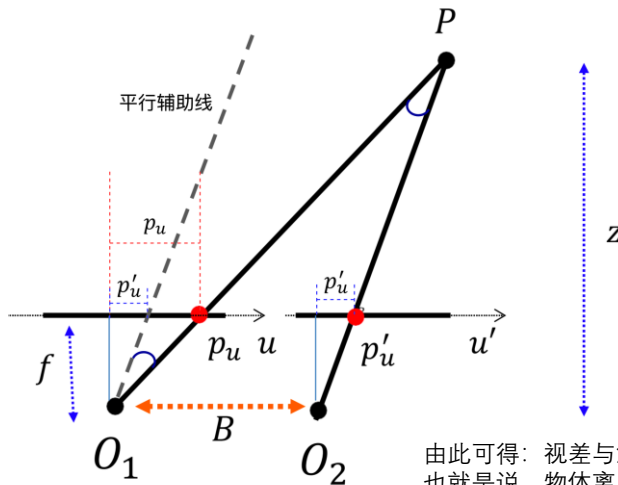
$$p'^T F p = 0$$

$$F = K'^{-T} [T_x] R K^{-1} = K'^{-T} E K^{-1}$$



47

平行视图三角测量



$$\frac{p_u - p'_u}{f} = \frac{B}{Z}$$

$$p_u - p'_u = \frac{B \cdot f}{Z}$$

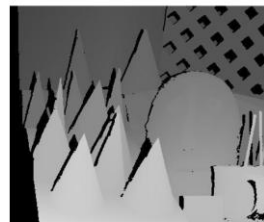
由此可得：视差与深度Z成反比！
也就是说，物体离人眼越远，左右眼观察到的图像越相似；

48

平行视图三角测量



“视差与深度Z成反比”这个结论可以方便我们从视差图中推导出深度图。（视差图颜色越暗，距离双目摄像机越远）



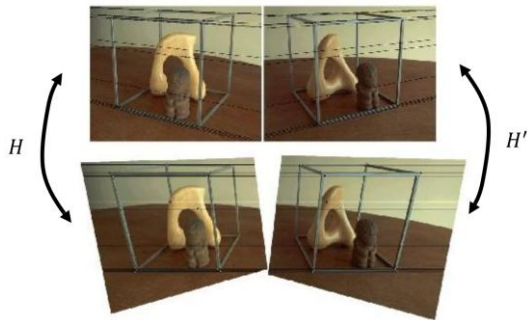
视差图 / 深度图

49

图像校正



- 在平行视图中，可以很方便利用视差获取深度图，但是，实际构建的双目立体视觉系统中，如何保证两个视图是完全平行的呢，这就需要进行图像校正。



步骤:

1. 在两幅图像 I 和 I' 找到一组匹配点 $p_i \leftrightarrow p'_i$ ，不少于8个。
2. 计算基础矩阵 F ，求解两幅图像中的极点 e 和 e'
3. 选择透视变换 H' 将 e' 映射到无穷远点 $(f, 0, 0)$ 。

$$H' = T^{-1}GRT$$
4. 寻找对应的透视变换矩阵 H 使得下式最小

$$\sum_i d(Hp_i, H'p'_i)$$
5. 分别用矩阵 H 和 H' ，对左右两幅图像 I 和 I' 进行重采样。

求解 e

$$l_i = F^T p'_i$$

$$\begin{bmatrix} l_1^T \\ \vdots \\ l_n^T \end{bmatrix} e = 0$$

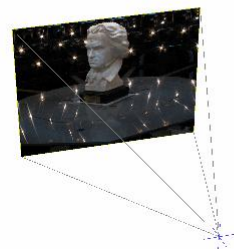
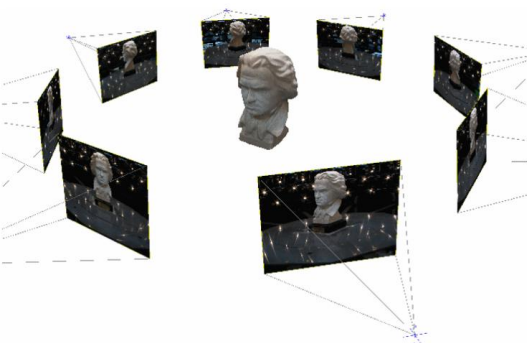
求解 e'

$$l'_i = F p_i$$

$$\begin{bmatrix} l'_1 \\ \vdots \\ l'_n \end{bmatrix} e' = 0$$

50

多视图几何



51

多视图几何



已知: n 个3D点 X_j 在 m 张图像中的对应点的像素坐标 x_{ij} ($i = 1, \dots, m, j = 1 \dots, n$)

$$\text{且 } x_{ij} = M_i X_j \quad i = 1, \dots, m; j = 1 \dots, n$$

其中, M_i 为第 i 张图片对应的摄像机的投影矩阵

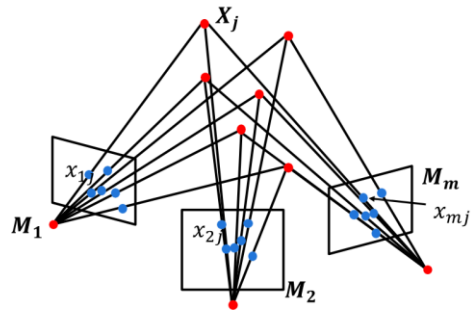
求解:

- m 个摄像机投影矩阵 M_i ($i = 1, \dots, m$);

运动 (motion)

- n 个三维点 X_j ($j = 1, \dots, n$)的坐标。

结构 (structure)



因此, 该类问题也称为“运动恢复结构问题”!

52

欧式结构恢复问题定义



已知:

- n 个三维点 X_j ($j = 1, \dots, n$) 在 m 张图像中的对应点的像素坐标 x_{ij}

- m 张图像对应的摄像机的内参数矩阵 K_i ($i = 1, \dots, m$)

$$\text{且 } x_{ij} = M_i X_j = K_i [R_i \ T_i] X_j \quad i = 1, \dots, \boxed{m}; j = 1 \dots \boxed{n}$$

↑ 图像个数 ↑ 3D点个数

其中, $M_i, K_i, [R_i \ T_i]$ 为第 i 张图片对应的摄像机的投影矩阵、内参数及外参数矩阵

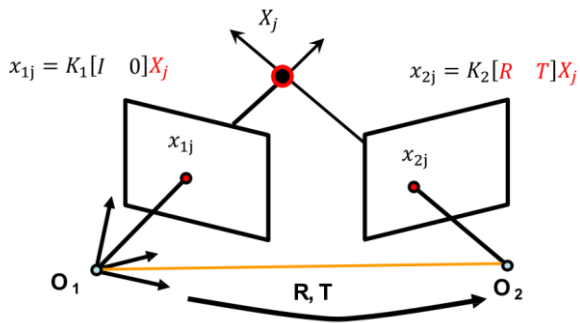
求解:

- n 个三维点 X_j ($j = 1, \dots, n$)的坐标;
- m 个摄像机的外参数 R_i 及 T_i ($i = 1, \dots, m$)

53



两视图欧式结构恢复



求解步骤:

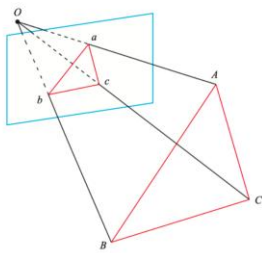
1. 求解基础矩阵 F (归一化八点法)
2. 求解本质矩阵 $E = K_2^T F K_1$
3. 分解本质矩阵 $E \rightarrow R, T$
4. 三角化

54

PnP问题

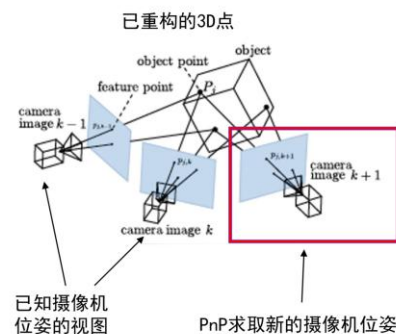


P3P: 通过世界中的3个特征点与图像成像中的3个像点, 获得相机位姿。



- **已知:** 摄像机内参数 K , 像素平面上 a, b, c 点的像素坐标以及其对应的三维点 A, B, C 在世界坐标系中的世界坐标;
- **求解:** 摄像机外参数 R, T

PnP (Perspective-n-Point) 问题: 指通过世界中的 N 个特征点与图像成像中的 N 个像点, 计算出其投影关系, 从而获得相机或物体位姿的问题。



已知摄像机位姿的视图

PnP求取新的摄像机位姿

55

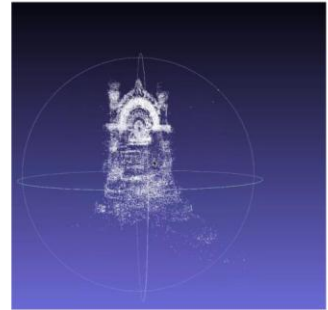
多视图几何



- 输入：多张图片
- 输出：3D点云 (structure)
摄像机位姿 (motion)



输入



输出

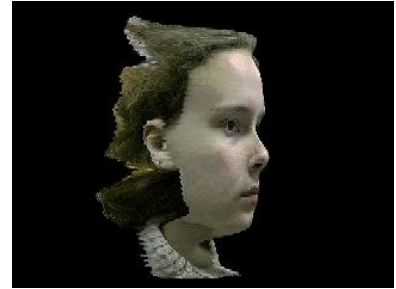
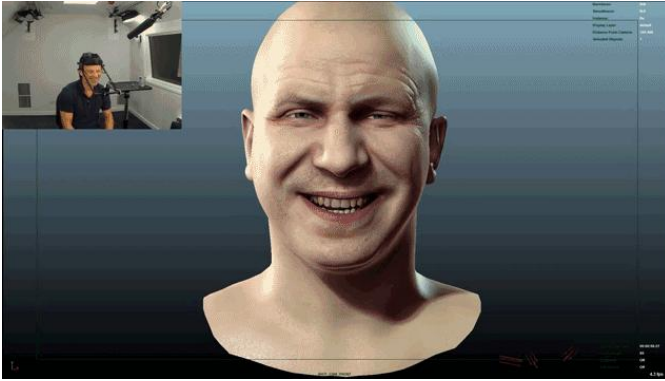
56

多视图几何



57

人脸重建



58

三维重建算法发展

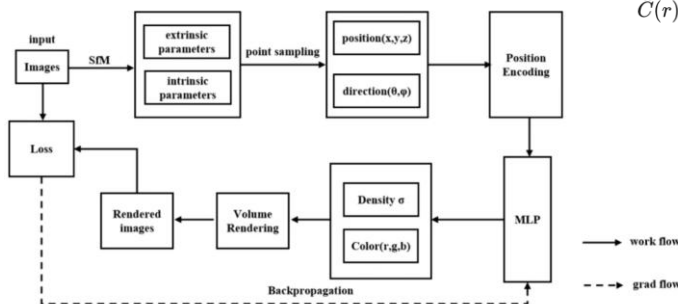


- NeRF: 隐式神经辐射场 (2020-2024)

把场景建模为连续 5D 函数:

$$F_{\Theta} : (x, y, z, \theta, \phi) \rightarrow (c, \sigma)$$

- 通过体渲染积分得到像素颜色:



$$\hat{C}(r) = \sum_{i=1}^N \underbrace{\exp\left(-\sum_{j=1}^{i-1} \sigma_j \delta_j\right)}_{\text{透射 } T_i} \cdot \underbrace{(1 - \exp(-\sigma_i \delta_i))}_{\text{alpha}} \cdot c_i$$

方法	关键创新	解决的问题	代价
Mip-NeRF	锥体追踪+集成位置编码	锯齿、混叠	训练×2
NeRF-W	外观嵌入+可变光照	天气/曝光变化	推理需调 latent
NeRF++	退化模拟器+双层场景	远景退化	参数量+30%
NeRFliX	退化模拟器+视角混合	伪影、噪声	需合成数据预训练
BAD-NeRF	运动模糊物理模型	模糊输入鲁棒	需已知模糊核
UHDNeRF	隐式体+稀疏点云高频	8K 超高清	显存↑

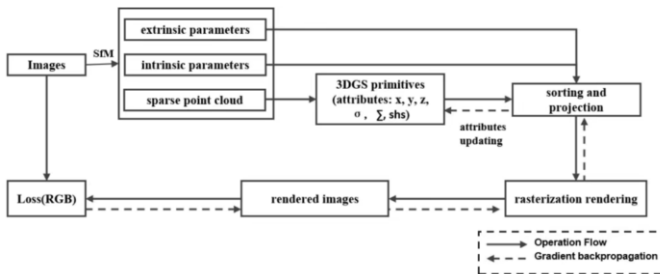
59

3DGS: 三维高斯溅射 (2023-2025)



- 将场景表示为 3D 高斯集合: $\mathcal{G} = \{\mu_i, \Sigma_i, \alpha_i, SH_i\}_{i=1}^M$

投影到图像平面后按深度排序, 做 α -混合
$$C = \sum_{i \in N} c_i \alpha'_i \prod_{j=1}^{i-1} (1 - \alpha_j)$$



方法	技术要点	结果
Mip-Splatting	3D/2D Mip 滤波	抗锯齿, LPIPS110%
Scaffold-GS	锚点生长-剪枝	内存179%, 覆盖1
GaussianPro	渐进传播+深度一致	低纹理 PSNR+1.7 dB
GSDf	高斯+SDF 双分支	几何误差130%
SuperGS	粗到细+梯度引导分裂	4K 超分实时

60



Thank you!

61