

# CCCV 2017讲习班

# 基于图像的大规模场景三维重建

申抒含、崔海楠  
中国科学院自动化研究所  
模式识别国家重点实验室

<http://vision.ia.ac.cn>



模式识别国家重点实验室  
National Laboratory of Pattern Recognition

# 讲者简介



**申抒含**，中国科学院自动化研究所模式识别国家重点实验室副研究员、中科院自动化所特聘青年骨干、中国计算机学会计算机视觉专委会委员。研究方向为三维计算机视觉理论与应用，包括基于图像的大规模场景三维重建、智能机器人三维环境感知、场景三维语义建模等。



**崔海楠**，中国科学院自动化研究所模式识别国家重点实验室助理研究员。研究方向为三维计算机视觉理论与应用，包括基于图像的大规模场景建模、机器人三维环境感知等。



# 提 纲

## 1、简介（申抒含，20分钟）

- 图像三维重建基本流程
- 多视几何基础

## 2、稀疏重建（崔海楠，40分钟）

- 增量式稀疏重建
- 全局式稀疏重建
- 混合式稀疏重建

## 3、稠密重建（申抒含，40分钟）

- 基于体素的稠密重建
- 基于特征点扩散的稠密重建
- 基于深度图融合的稠密重建

## 4、资源简介（申抒含，20分钟）

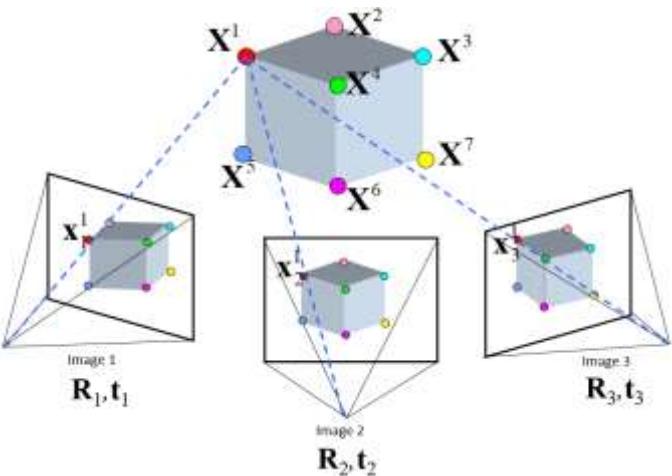
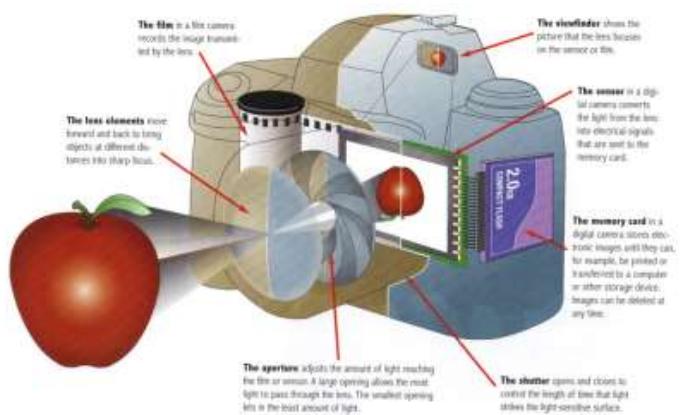
- 三维重建开源与商业软件
- 三维重建评测数据库



# 参考文献

1. **Hainan Cui**, Xiang Gao, **Shuhan Shen**, Zhanyi Hu. HSfM: Hybrid Structure-from-Motion. *CVPR* 2017.
2. Miao Yu, **Shuhan Shen**, Zhanyi Hu. Dynamic Graph Cuts in Parallel, *IEEE Trans. on Image Processing*, 26(8): 3775–3788, 2017.
3. Miao Yu, **Shuhan Shen**, Zhanyi Hu. Dynamic Parallel and Distributed Graph Cuts. *IEEE Trans. on Image Processing*, 25(12): 5511–5525, 2016.
4. Johannes L. Schönberger, Jan-Michael Frahm. Structure-from-Motion Revisited, *CVPR* 2016.
5. Johannes L. Schönberger, Enliang Zheng, Marc Pollefeys, Jan-Michael Frahm. Pixelwise View Selection for Unstructured Multi-View Stereo, *ECCV* 2016.
6. Onur Ozyesil, Amit Singer. Robust Camera Location Estimation by Convex Programming. *CVPR* 2015.
7. **Hainan Cui**, **Shuhan Shen**, Wei Gao, Zhanyi Hu. Efficient Large-Scale Structure from Motion by Fusing Auxiliary Imaging Information. *IEEE Trans. on Image Processing*, 24(11): 3561–3573, 2015.
8. **Hainan Cui**, **Shuhan Shen**, Wei Gao, Zhanyi Hu. Fusion of Auxiliary Imaging Information for Robust, Scalable and Fast 3D Reconstruction. *ACCV* 2014.
9. **Shuhan Shen**, Zhanyi Hu. How to Select Good Neighboring Images in Depth-Map Merging Based 3D Modeling. *IEEE Trans. on Image Processing*, 23(1): 308–318, 2014.
10. **Shuhan Shen**. Accurate Multiple View 3D Reconstruction Using Patch Based Stereo for Large-Scale Scenes. *IEEE Trans. on Image Processing*, 22(5): 1901–1914, 2013.
11. Yasutaka Furukawa, Jean Ponce. Accurate, Dense, and Robust Multiview Stereopsis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(8): 1362–1376, 2010.
12. Richard Hartley, Andrew Zisserman. Multiple View Geometry in Computer Vision. *Cambridge University Press*, NY, USA, 2003.

# 1. 简介





# 提 纲

## 1、简介（申抒含，20分钟）

- 图像三维重建基本流程
- 多视几何基础

## 2、稀疏重建（崔海楠，40分钟）

- 增量式稀疏重建
- 全局式稀疏重建
- 混合式稀疏重建

## 3、稠密重建（申抒含，40分钟）

- 基于体素的稠密重建
- 基于特征点扩散的稠密重建
- 基于深度图融合的稠密重建

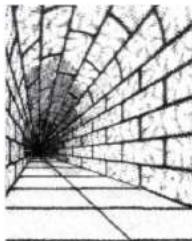
## 4、资源简介（申抒含，20分钟）

- 三维重建开源与商业软件
- 三维重建评测数据库

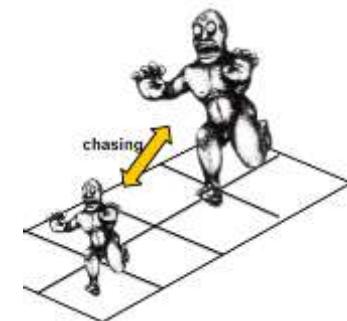
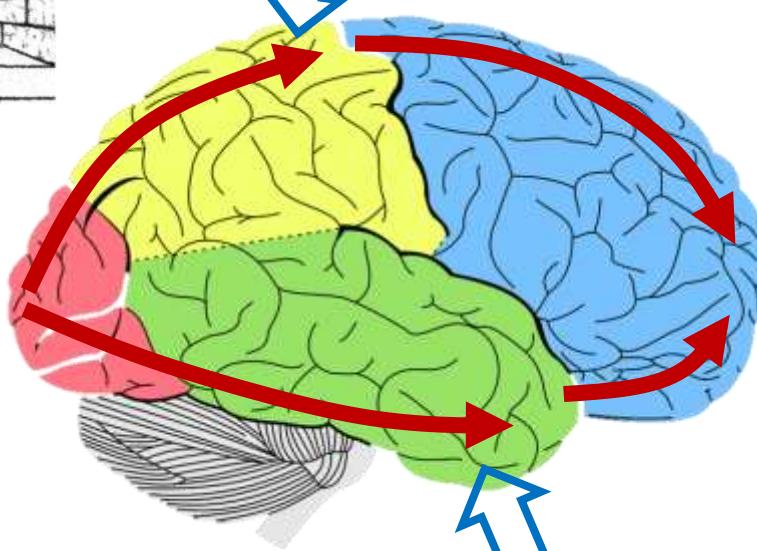
# 几何视觉与物体视觉

背部通道进行**空间视觉**, 目前

主要采用基于**几何**的方法



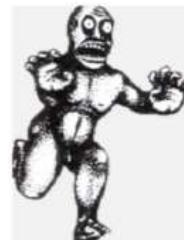
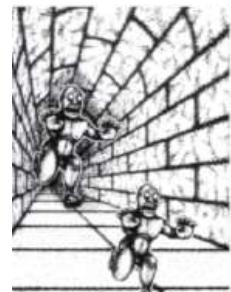
V1



前额皮质通过信息  
融合实现**高级认知**

腹部通道进行**物体视觉**, 目前

主要采用基于**学习**的方法



# 几何视觉的核心问题



无人机



无人车



服务机器人

- 智能机器人工在三维环境中，除了对二维图像的理解，还需要对三维环境的感知；
- 场景的**三维结构**和相机的六自由度**空间位姿**是智能机器人感知、决策、动作的基础信息。

# 图像三维重建基本流程

多视角  
图像

图像特征  
提取匹配

稀疏重建  
(SfM)

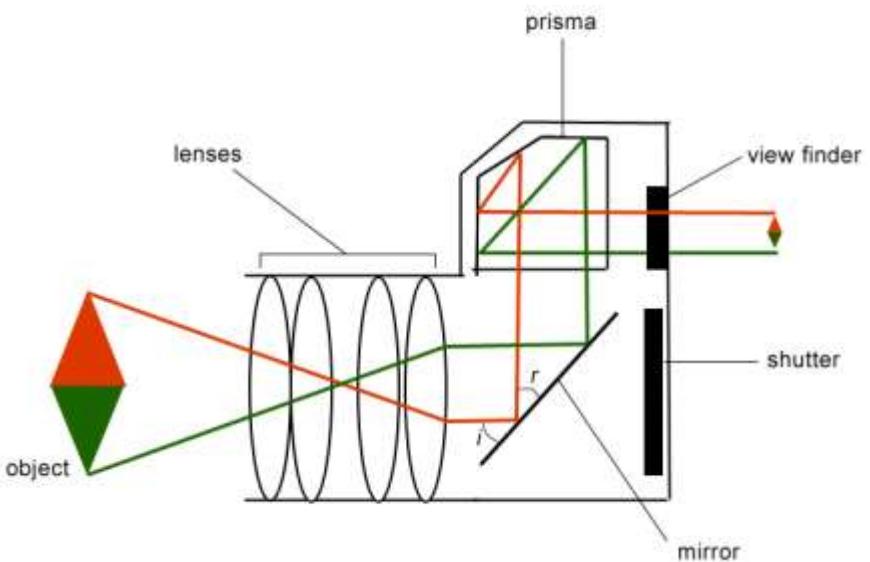
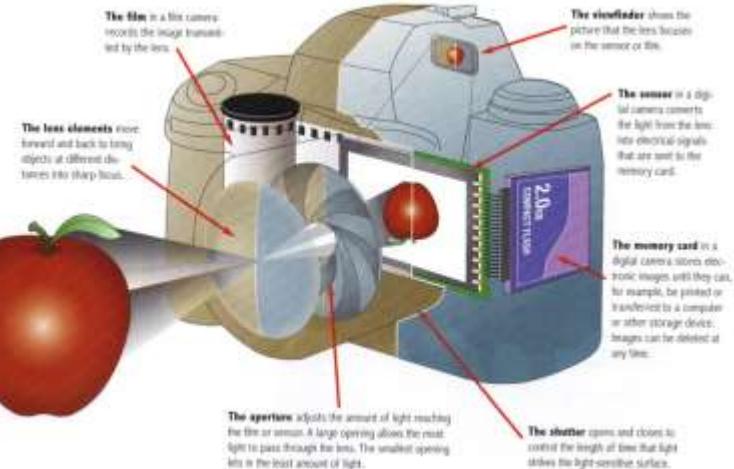
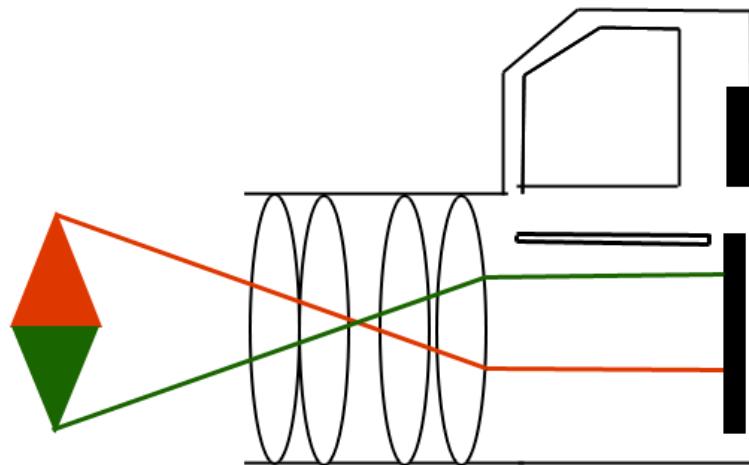
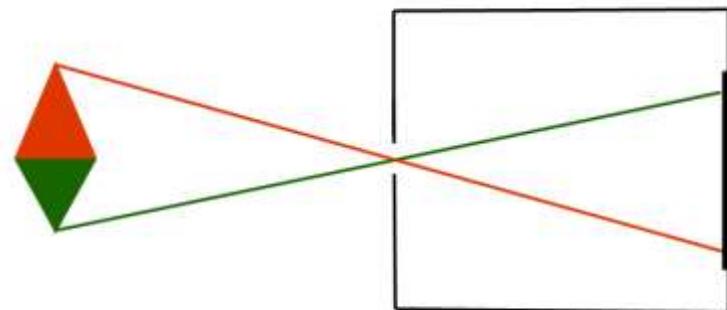
稠密重建  
(MVS)

点云  
模型化

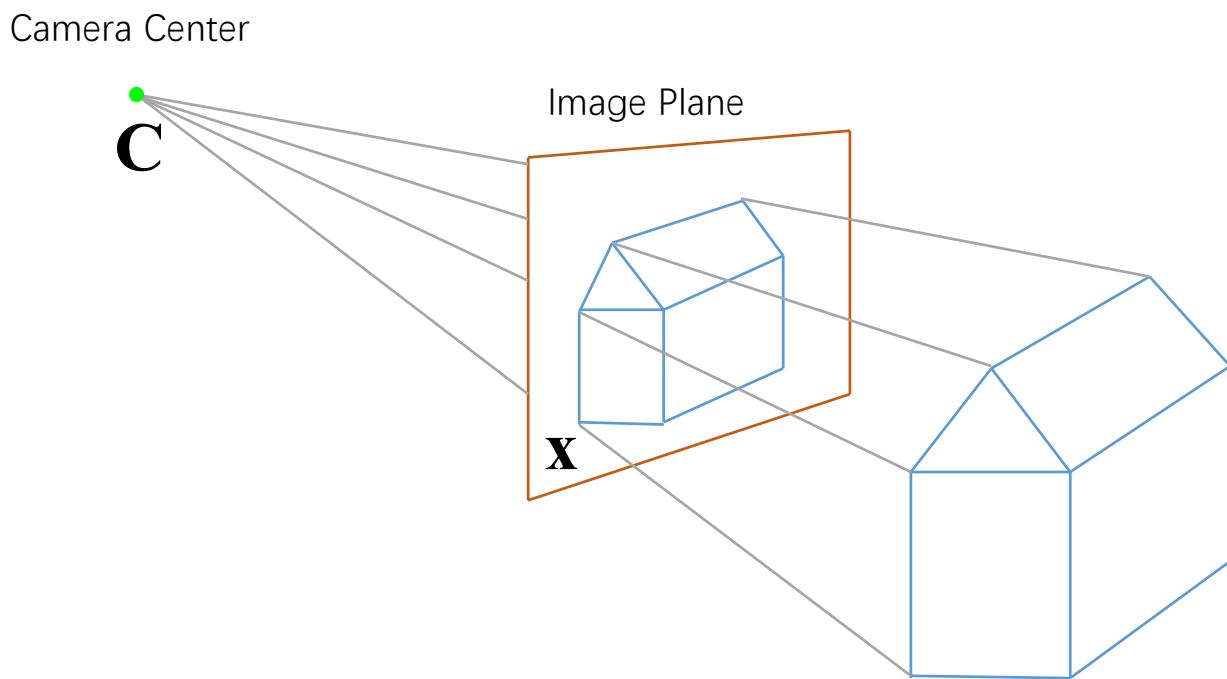
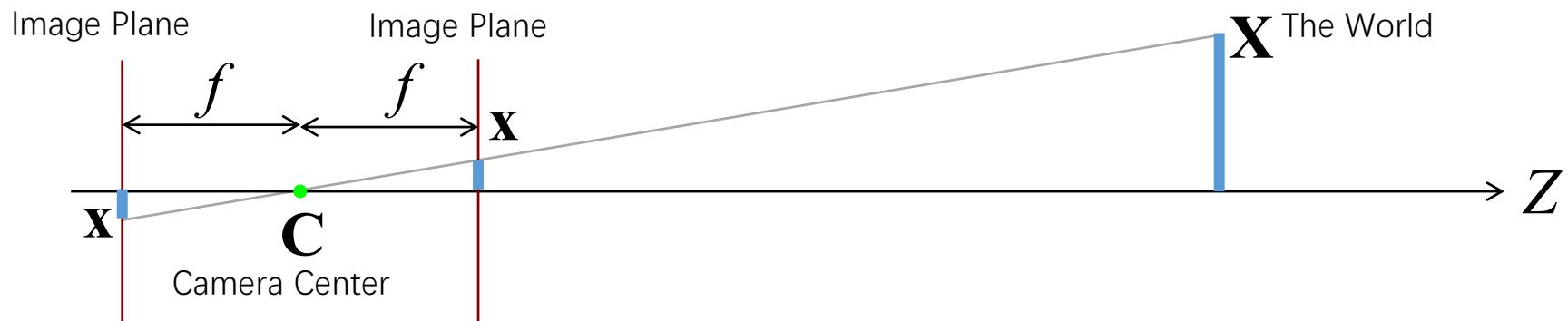
三维  
模型



# 小孔成像与透镜系统



# 小孔相机模型



$$\lambda \begin{bmatrix} x \\ y \\ f \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

# 欧氏空间与射影空间

无穷远元素



图片来源：视觉中国 www.vcg.com

# 欧氏空间与射影空间

齐次坐标是射影空间的坐标表达方式

- 非齐次坐标到齐次坐标的转换：

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- 齐次坐标到非其次坐标的转换：

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

# 欧氏空间与射影空间

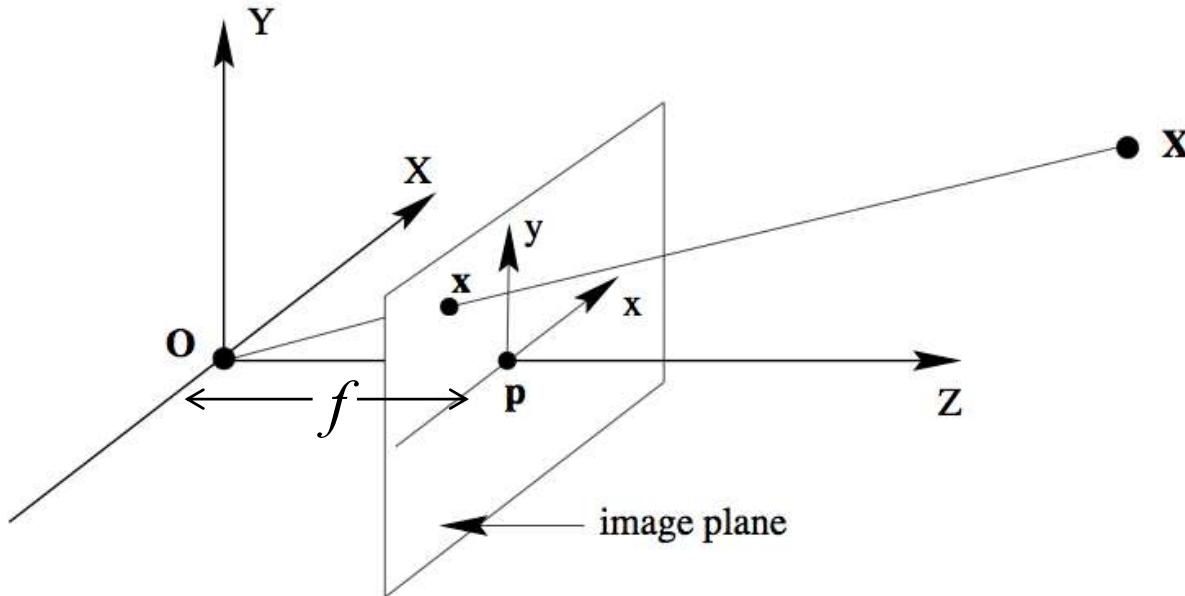
- 齐次坐标在相差一个尺度时等价：

$$\lambda \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- 无穷远点的齐次坐标：

$$\begin{bmatrix} x \\ y \\ z \\ 0 \end{bmatrix}$$

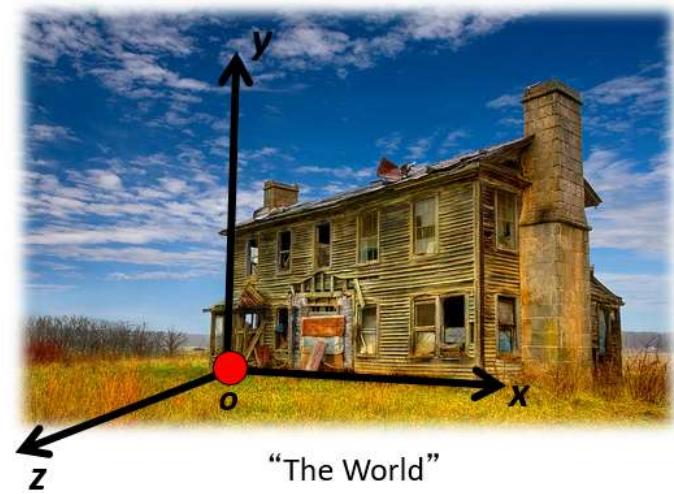
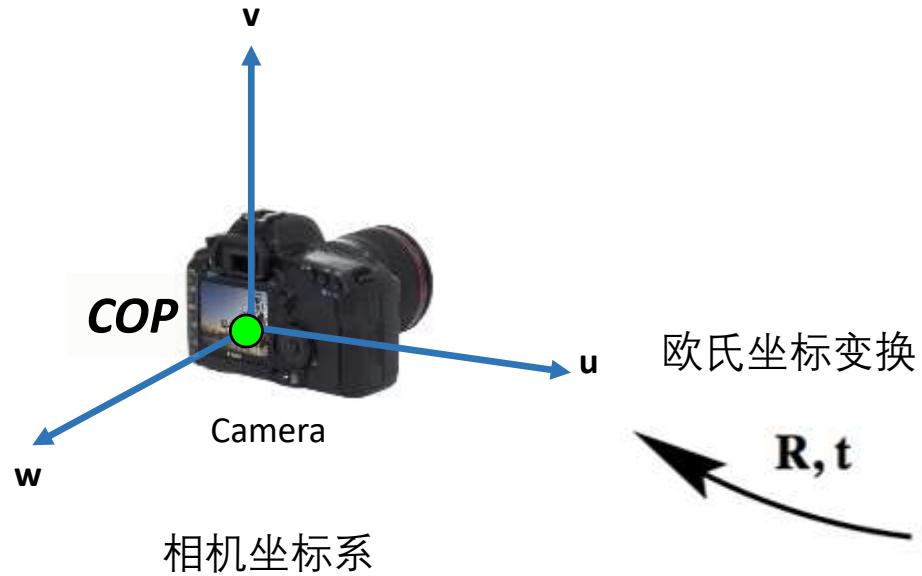
# 小孔相机模型



$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u \\ 0 & f & v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Up to a scale 焦距 主点

# 小孔相机模型



$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

# 小孔相机模型

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u \\ 0 & f & v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

图像点

相机内参数

相机硬件

相机旋转 相机平移

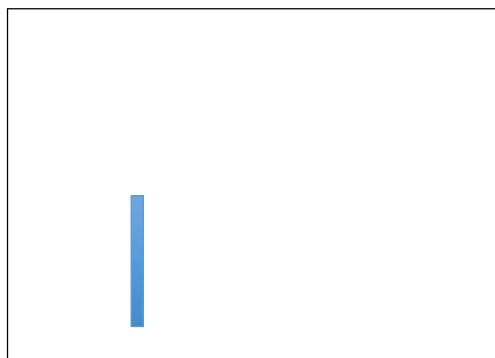
相机位姿

空间点

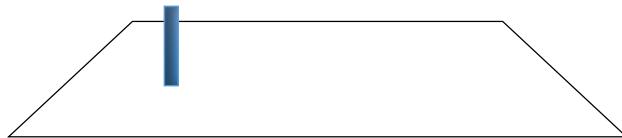
场景结构

# 多视几何

- 通过单幅图像无法重建场景结构



Image



3D world

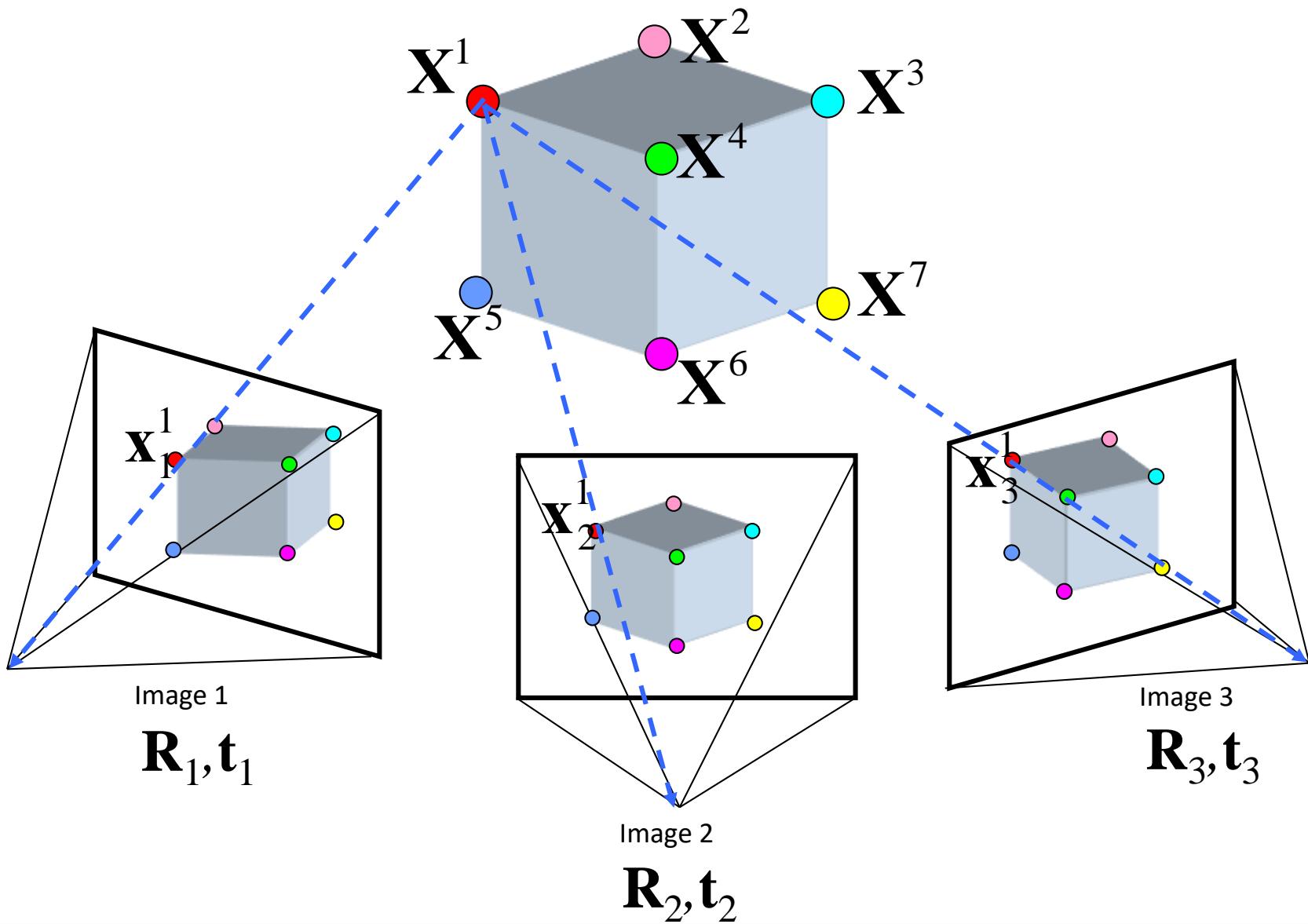


3D world



3D world

# 多视几何



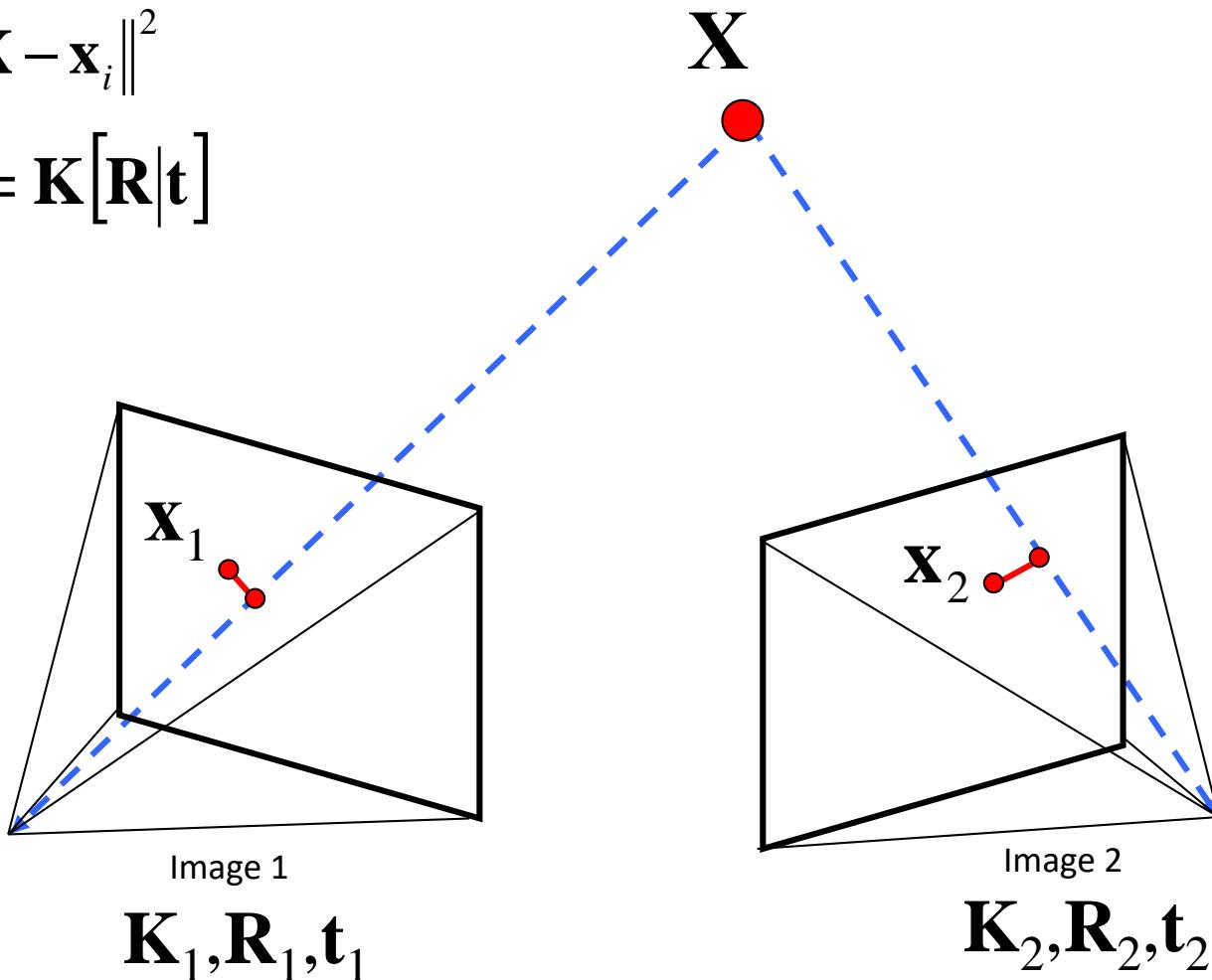
# 多视几何

	Point 1	Point 2	Point 3
Image 1	$\mathbf{x}_1^1 = \mathbf{K}_1 [\mathbf{R}_1   \mathbf{t}_1] \mathbf{X}^1$	$\mathbf{x}_1^2 = \mathbf{K}_1 [\mathbf{R}_1   \mathbf{t}_1] \mathbf{X}^2$	
Image 2	$\mathbf{x}_2^1 = \mathbf{K}_2 [\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^1$	$\mathbf{x}_2^2 = \mathbf{K}_2 [\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^2$	$\mathbf{x}_2^3 = \mathbf{K}_2 [\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^3$
Image 3	$\mathbf{x}_3^1 = \mathbf{K}_3 [\mathbf{R}_3   \mathbf{t}_3] \mathbf{X}^1$		$\mathbf{x}_3^3 = \mathbf{K}_3 [\mathbf{R}_3   \mathbf{t}_3] \mathbf{X}^3$
	⋮	⋮	

# 多视几何

- 求解思路：最小化重投影误差平方和

$$\min \sum_i \| \mathbf{P}_i \mathbf{X} - \mathbf{x}_i \|^2$$
$$\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}]$$



# 多视几何

- 重投影误差最小化问题的求解

$$\min \sum_i \|\mathbf{P}_i \mathbf{X}_i - \mathbf{x}_i\|^2$$



$$\min \sum_i \sum_j \left\| \frac{\mathbf{P}_{1:2}^i \mathbf{X}_j}{\mathbf{P}_3^i \mathbf{X}_j} - \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} \right\|^2$$



$$\min \sum_i \sum_j \|f_i(\hat{\mathbf{X}}) - b_{ij}\|^2$$



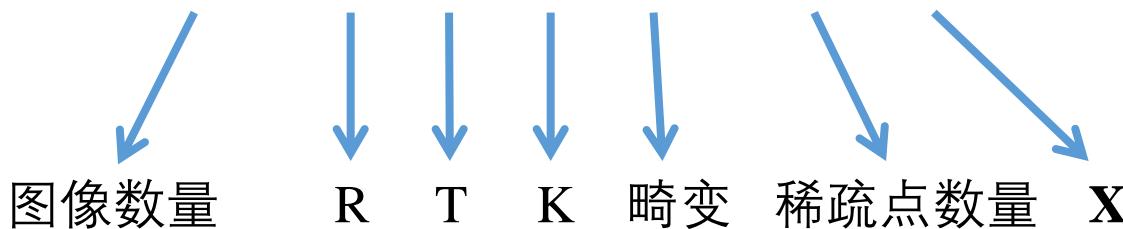
$$\min \|F(\hat{\mathbf{X}}) - \mathbf{b}\|^2 \quad F(\hat{\mathbf{X}}) = \begin{bmatrix} \vdots \\ f_i(\hat{\mathbf{X}}) \\ \vdots \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \vdots \\ b_{ij} \\ \vdots \end{bmatrix}$$

# 多视几何

- 重投影误差最小化问题的求解

$$\min \left\| F(\hat{\mathbf{X}}) - \mathbf{b} \right\|^2$$

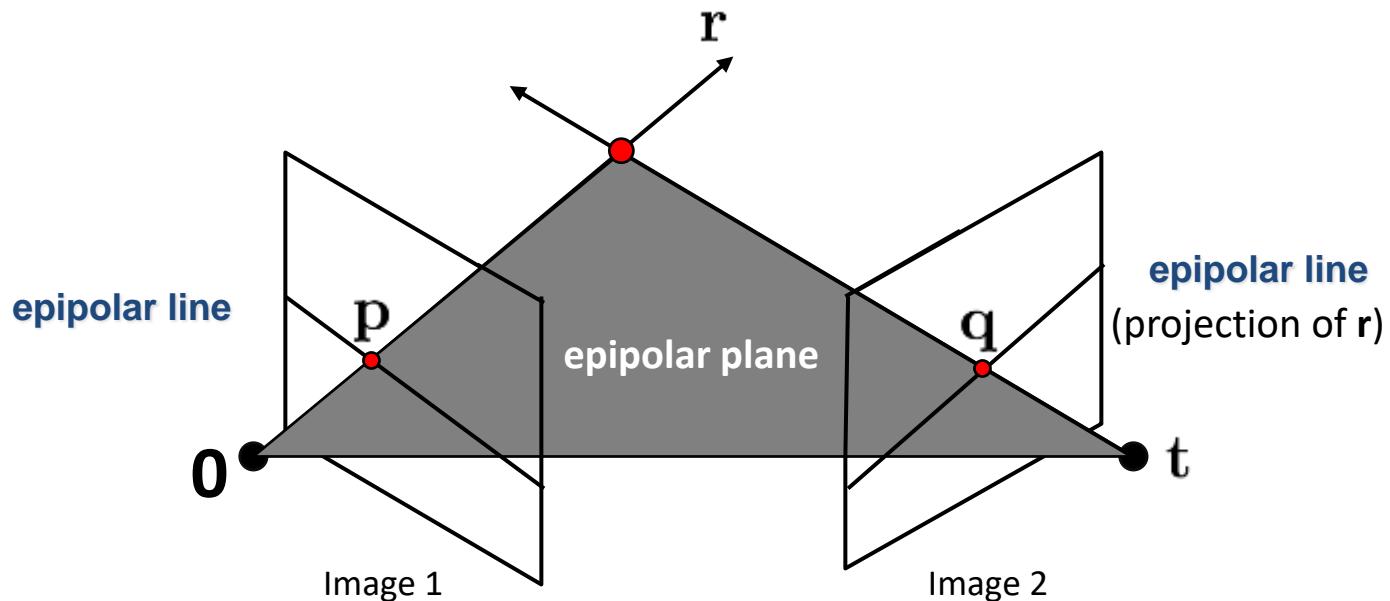
- 重投影误差最小化是一个高维非线性最小二乘问题
- 未知量数量 :  $N \times (3 + 3 + 3 + 2) + M \times 3$



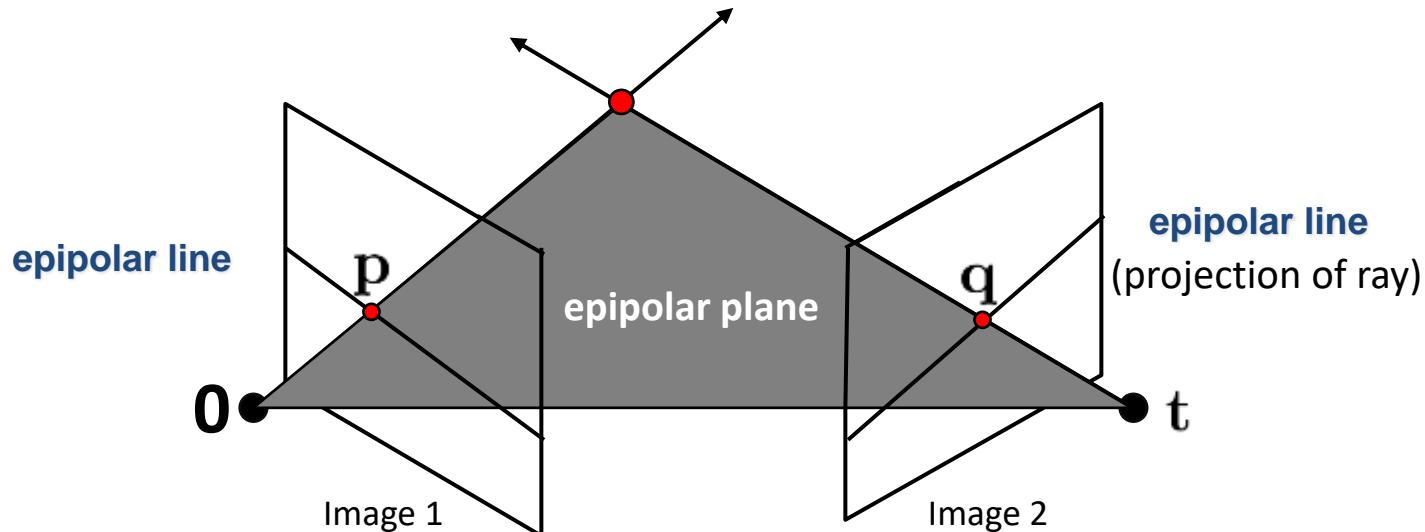
- 求解工具 : **Bundle Adjustment** (捆绑调整)
  - BA一种启发式的阻尼高斯牛顿法，在几何视觉中广泛使用
  - BA有效性的关键在于提供理想的初始值

# 两视图几何

- 初始值求解的基础：两视图几何



# 两视图几何



$\mathbf{K}_1$  : 左相机内参数矩阵

$\mathbf{K}_2$  : 右相机内参数矩阵

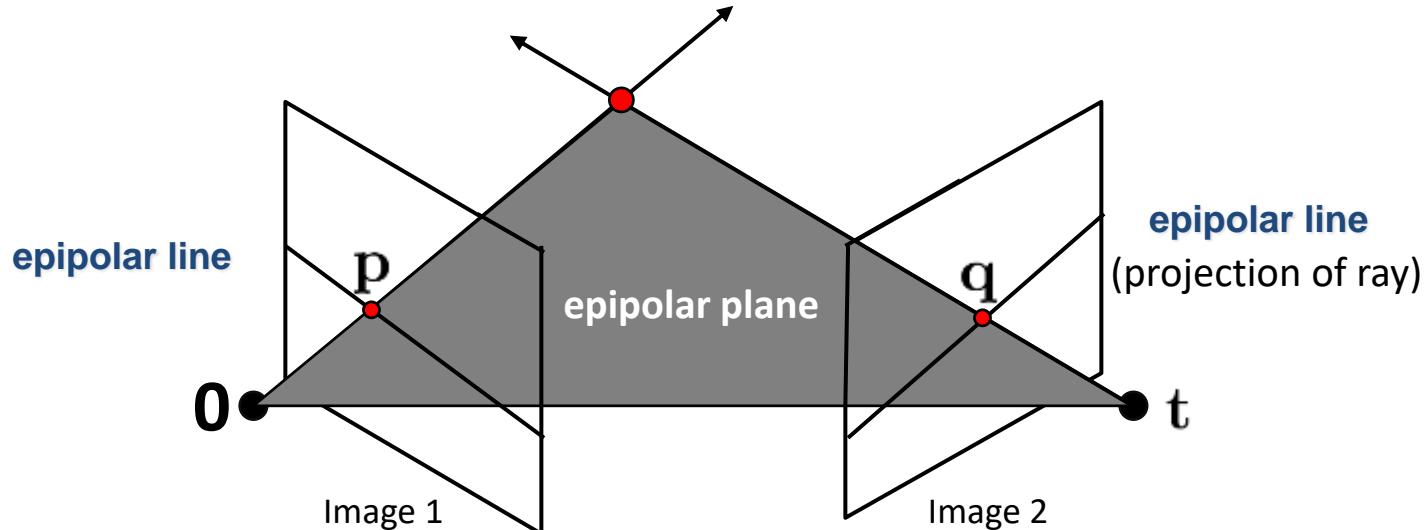
$\mathbf{R}$  : 左右相机间的相对旋转

$\mathbf{t}$  : 左右相机间的平移

$$\underbrace{\mathbf{q}^T \mathbf{K}_2^{-T} \mathbf{R} [\mathbf{t}]_\times \mathbf{K}_1^{-1} \mathbf{p}}_F = 0$$

$\mathbf{F} \leftarrow$  the Fundamental matrix

# 两视图几何



当内参数矩阵  $\mathbf{K}_1$  和  $\mathbf{K}_2$  已知:  $\tilde{\mathbf{p}} = \mathbf{K}_1^{-1}\mathbf{p}$      $\tilde{\mathbf{q}} = \mathbf{K}_2^{-1}\mathbf{q}$

$$\underbrace{\tilde{\mathbf{q}}^T \mathbf{R} [\mathbf{t}]_\times \tilde{\mathbf{p}}}_E = 0 \quad \tilde{\mathbf{q}}^T \mathbf{E} \tilde{\mathbf{p}} = 0$$

$E \leftarrow$  the Essential matrix (本质矩阵)

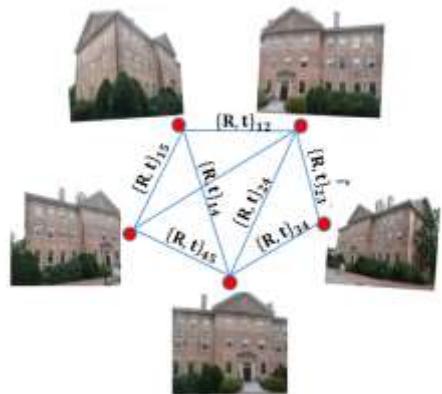


# 两视图几何

1. 基本矩阵F可通过8组对应点线性求解（8点法）、本质矩阵E可通过5组对应点求解（5点法）
2. 通过SVD分解可以从本质矩阵E中分解相机R和t
3. 已知相机K、R、T，可通过三角化求解三维点X
4. 至此，得到两视图重建的初始值，之后通过BA对相机内参、相机位姿和三维结构K、R、T、X进行非线性优化

多视图的初始值计算：More complicated! Next.....

## 2. 稀疏重建





# 提 纲

## 1、简介（申抒含，20分钟）

- 图像三维重建基本流程
- 多视几何基础

## 2、**稀疏重建**（崔海楠，40分钟）

- 增量式稀疏重建
- 全局式稀疏重建
- 混合式稀疏重建

## 3、稠密重建（申抒含，40分钟）

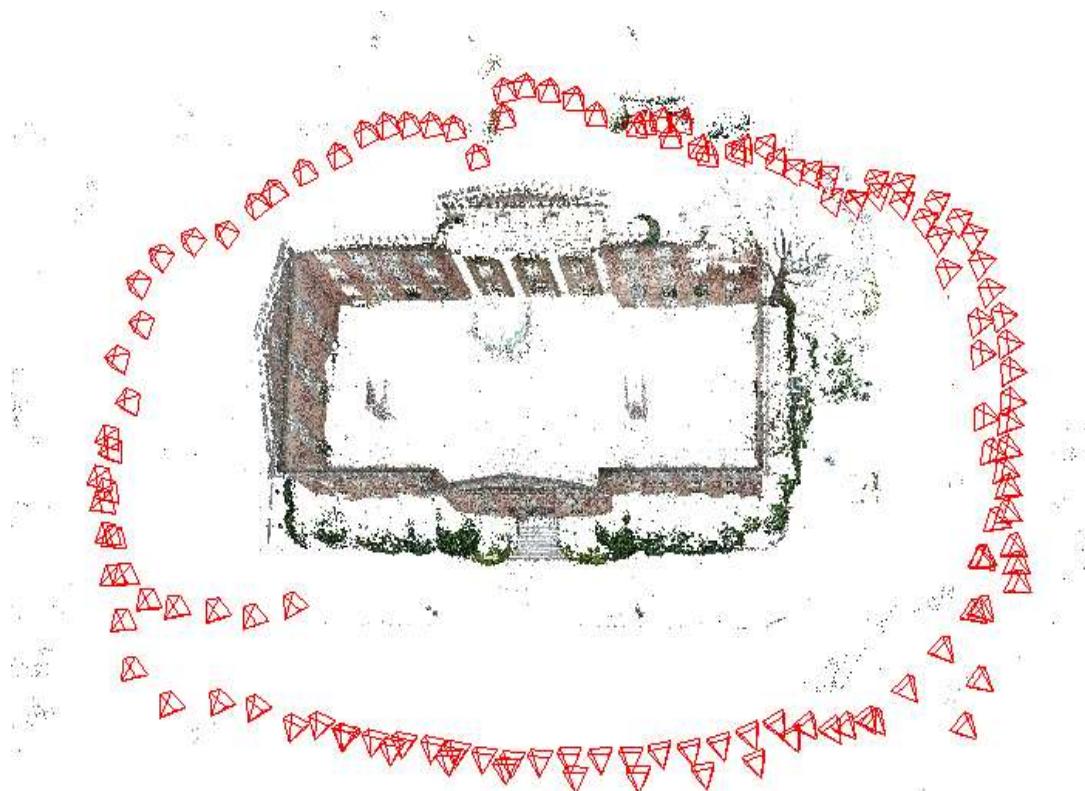
- 基于体素的稠密重建
- 基于特征点扩散的稠密重建
- 基于深度图融合的稠密重建

## 4、资源简介（申抒含，20分钟）

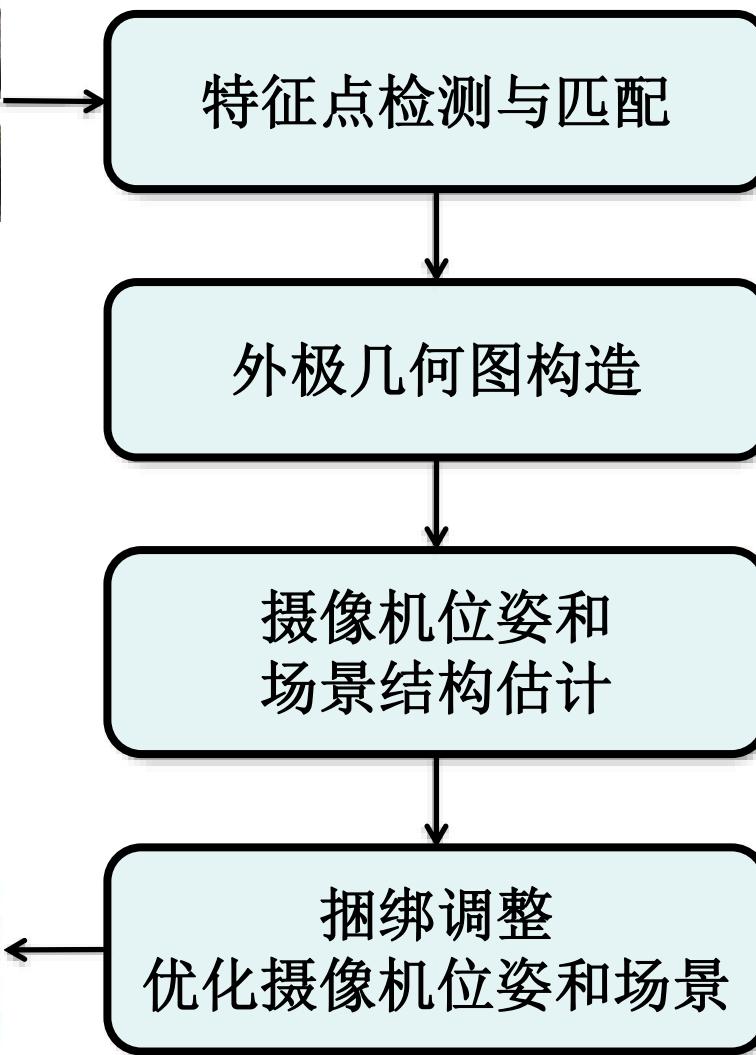
- 三维重建开源与商业软件
- 三维重建评测数据库

# 从运动恢复结构 (SfM)

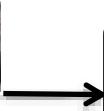
通过相机运动恢复场景的结构(Structure-from-Motion)：



# 从运动恢复结构 (SfM)



# 从运动恢复结构 (SfM)



特征点检测与匹配



外极几何图构造

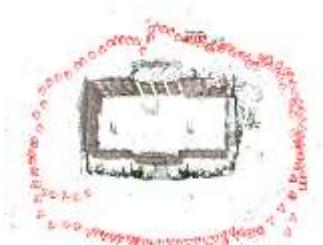


摄像机位姿和  
场景结构估计

- 增量式  
● 全局式  
● 混合式

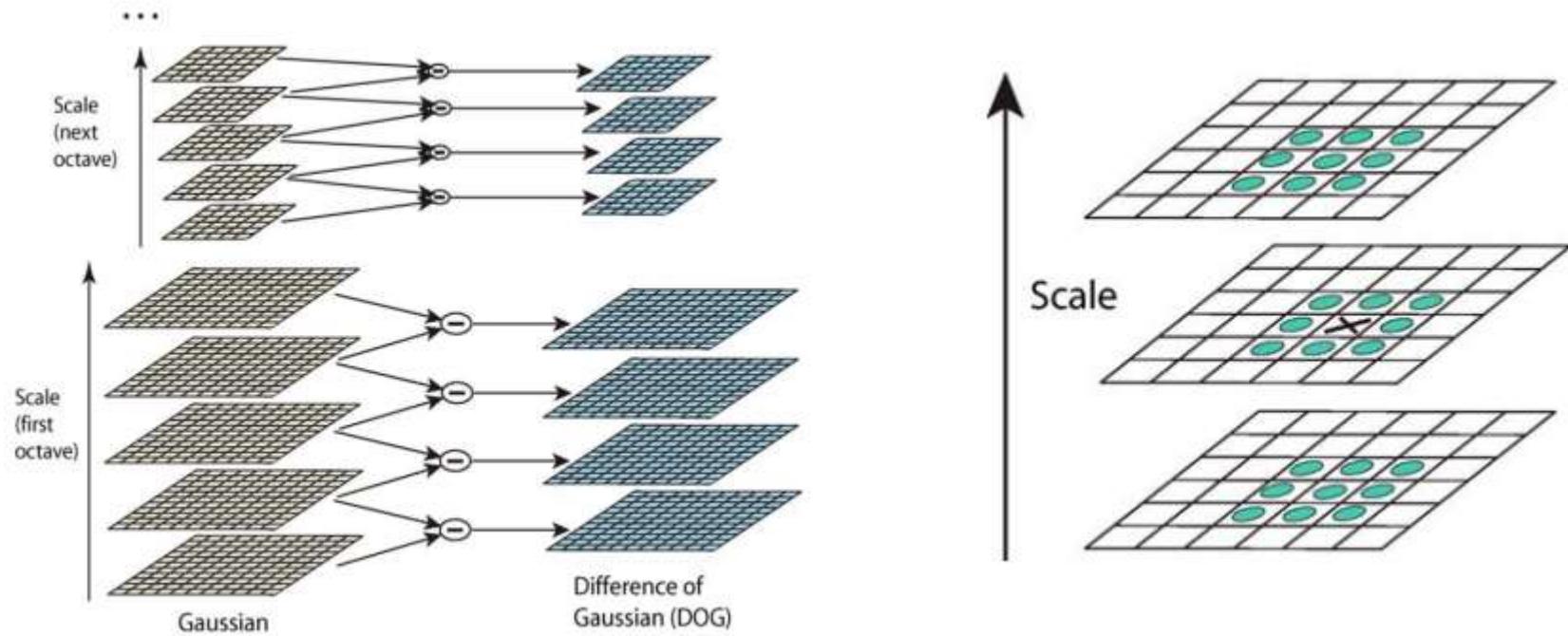


捆绑调整  
优化摄像机位姿和场景



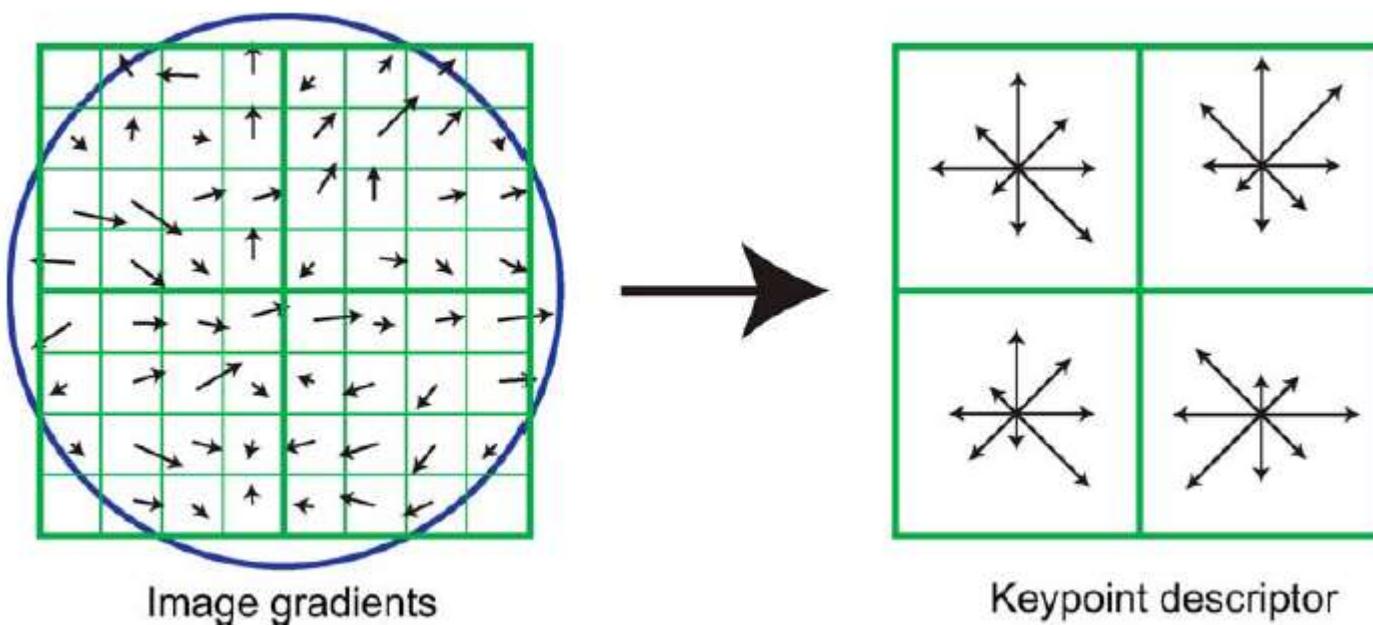
# 特征点检测 (SIFT)

尺度空间极大值检测，特征点定位



# 特征点描述 (SIFT)

主方向求取以及特征点描述 (128维向量)



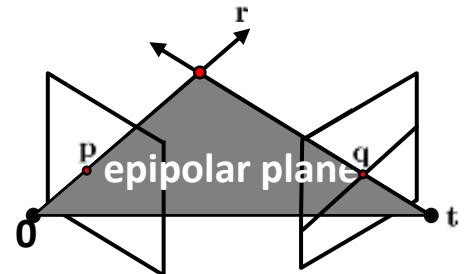
# 特征点匹配

- (1) L2范数距离最小；
- (2) 比值约束：最近邻/次近邻小于一定阈值



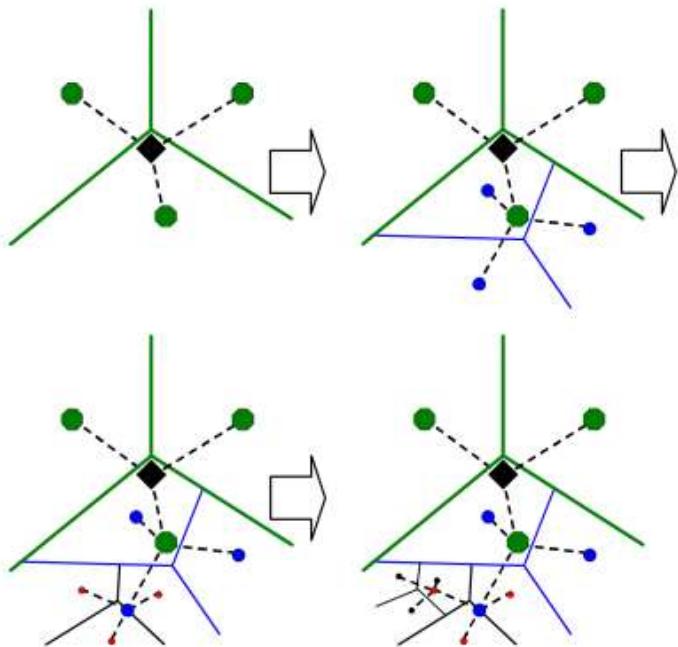
# 特征点匹配

外极几何约束（匹配特征点在对应的外极线附近）



# 特征点匹配难点-效率

- 候选图像匹配对 ( $N^2$ )

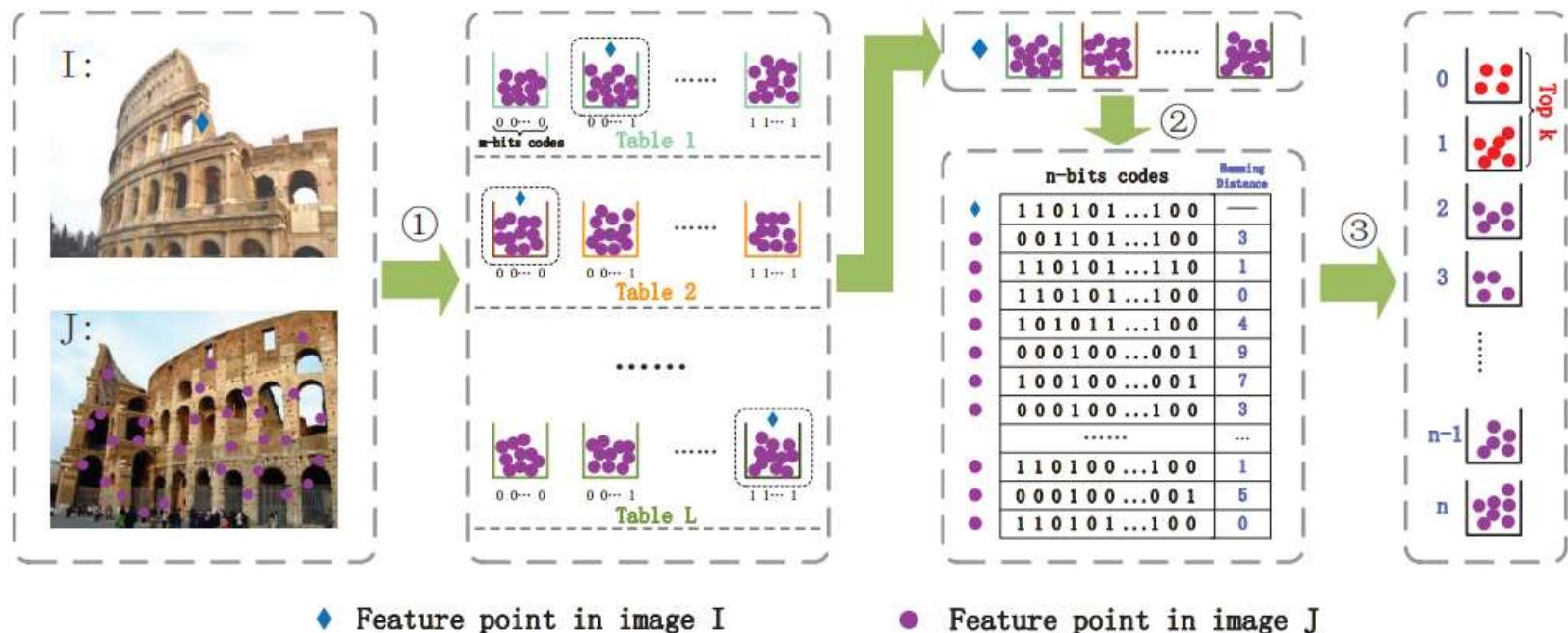


a.聚类->建树->匹配 (K-means 层级树)

b.  $O(N^2)$  降到了  $O(kN)$

# 特征点匹配难点-效率

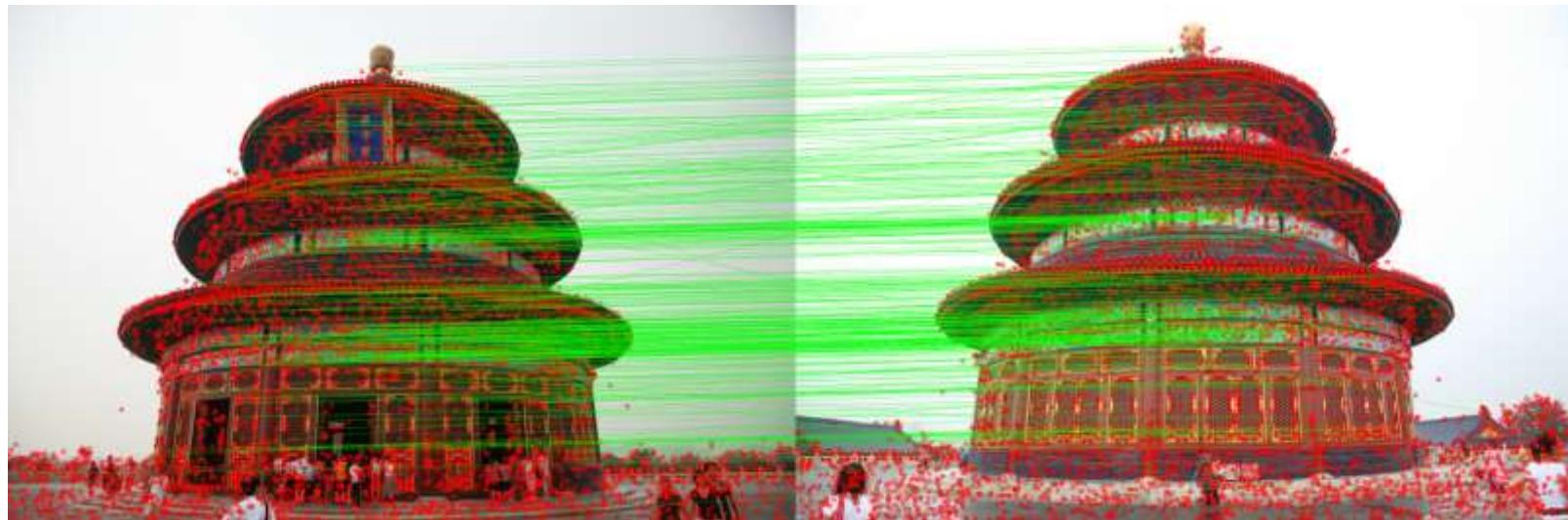
- 每对图像之间特征点匹配 ( $N^2$ )



Cheng J, et al. "Fast and accurate image matching with cascade hashing for 3d reconstruction." CVPR2014

# 特征点匹配难点-误匹配

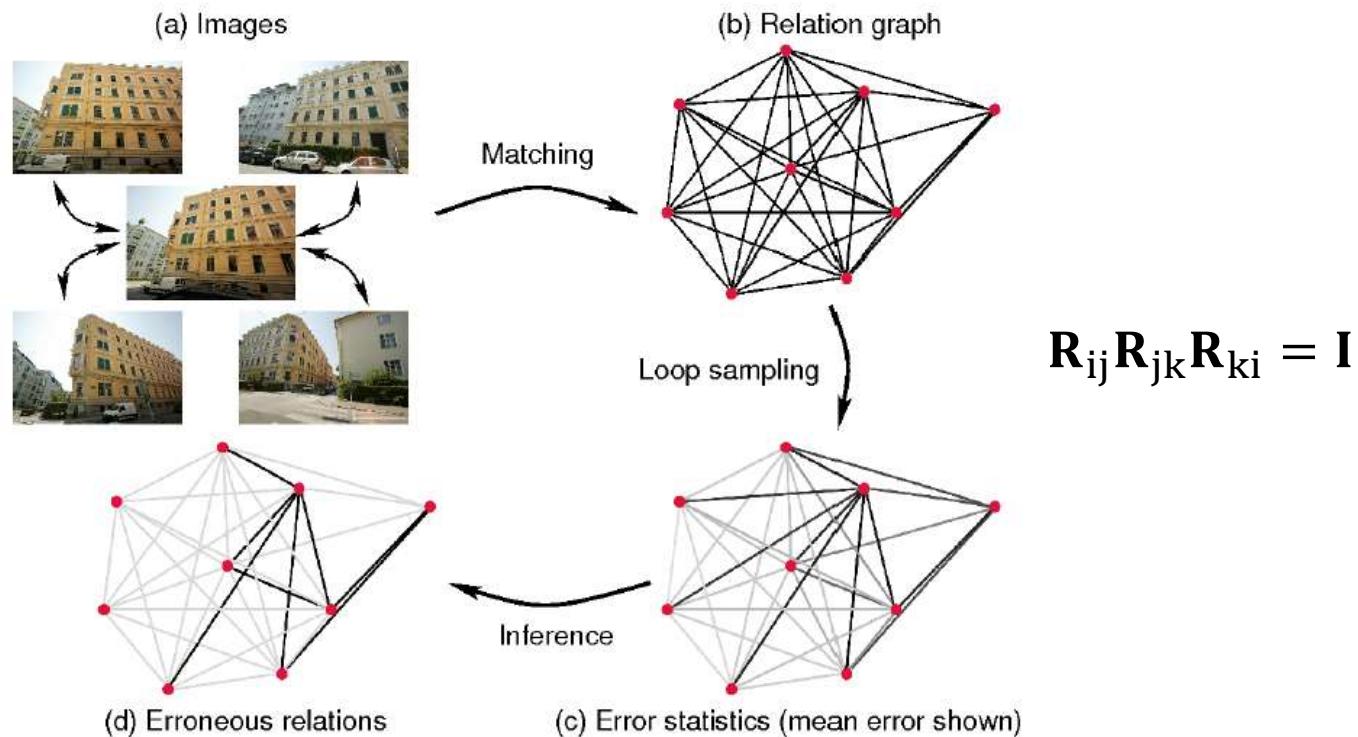
- 错误的外极几何关系



天坛前面-背面特征点匹配示例

# 特征点匹配难点及解决方案

- 错误的外极几何关系



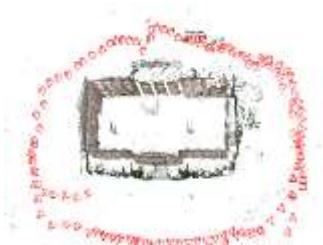
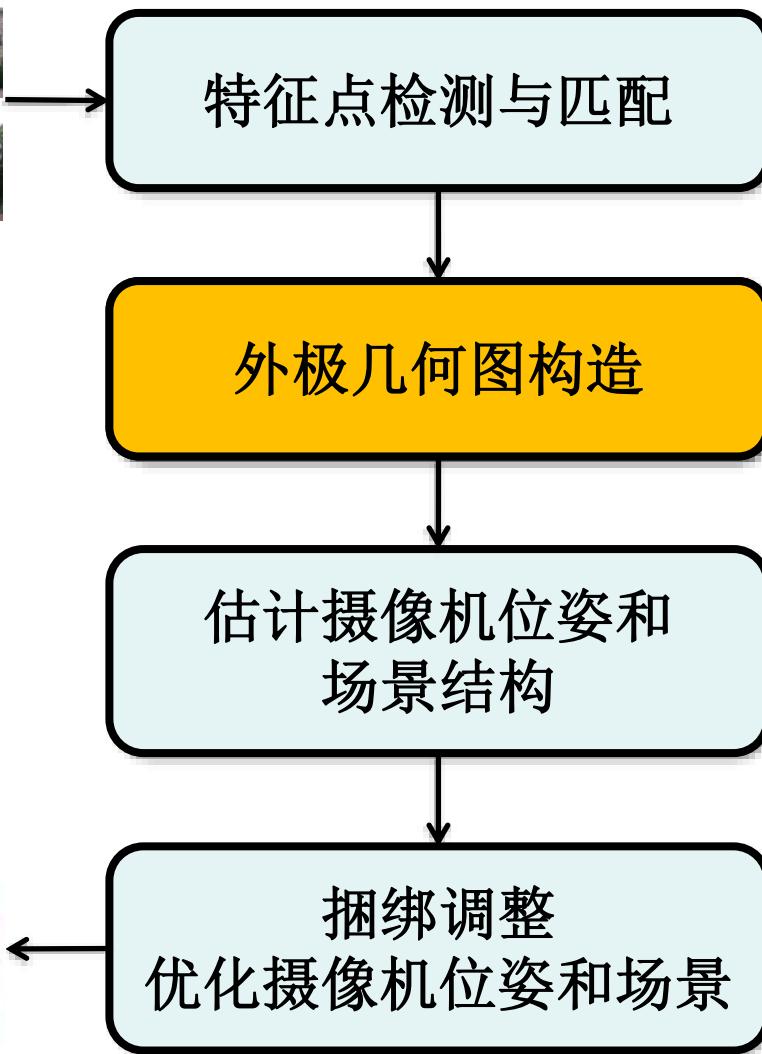


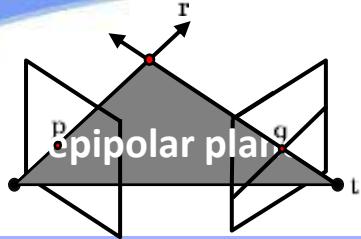
# 特征点匹配难点及解决方案

- 其他解决方案

- 1) Roberts, Richard, et al. "Structure from motion for scenes with large duplicate structures." CVPR 2011. (基于丢失匹配与时间戳进行对极几何图采样)
- 2) Wilson, Kyle, and Noah Snavely. "Network principles for sfm: Disambiguating repeated structures with local context." CVPR 2013 (基于track的可视性判断误匹配)
- 3) Wu,C. "Towards linear-time incremental structure from motion." 3DV2013 (基于大尺度特征点快速寻找候选匹配对)
- 4) T. Shen, S. Zhu, T. Fang, R. Zhang, and L. Quan. "Graph-based consistent matching for structure-from-motion." ECCV 2016 (基于语义树和团体检测减少候选匹配对)
- 5) Qingan Yan, et al. "Distinguishing the Indistinguishable: Exploring Structural Ambiguities via Geodesic Context." CVPR 2017 (聚类得到代表图像)

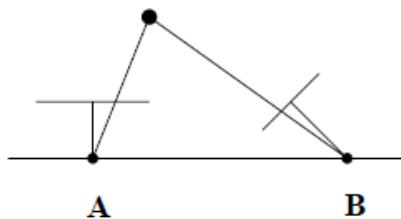
# 从运动恢复结构 (SfM)





# 本质矩阵分解

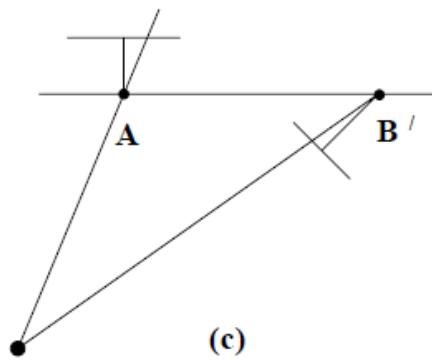
$$E = [t_{ij}] \times R_{ij}$$



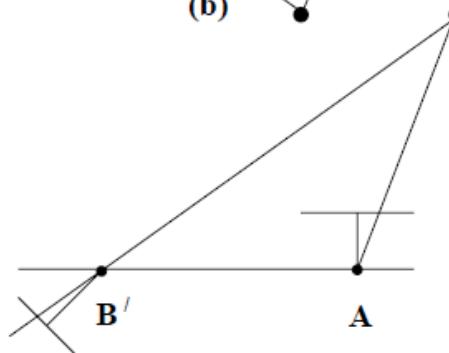
(a)



(b)



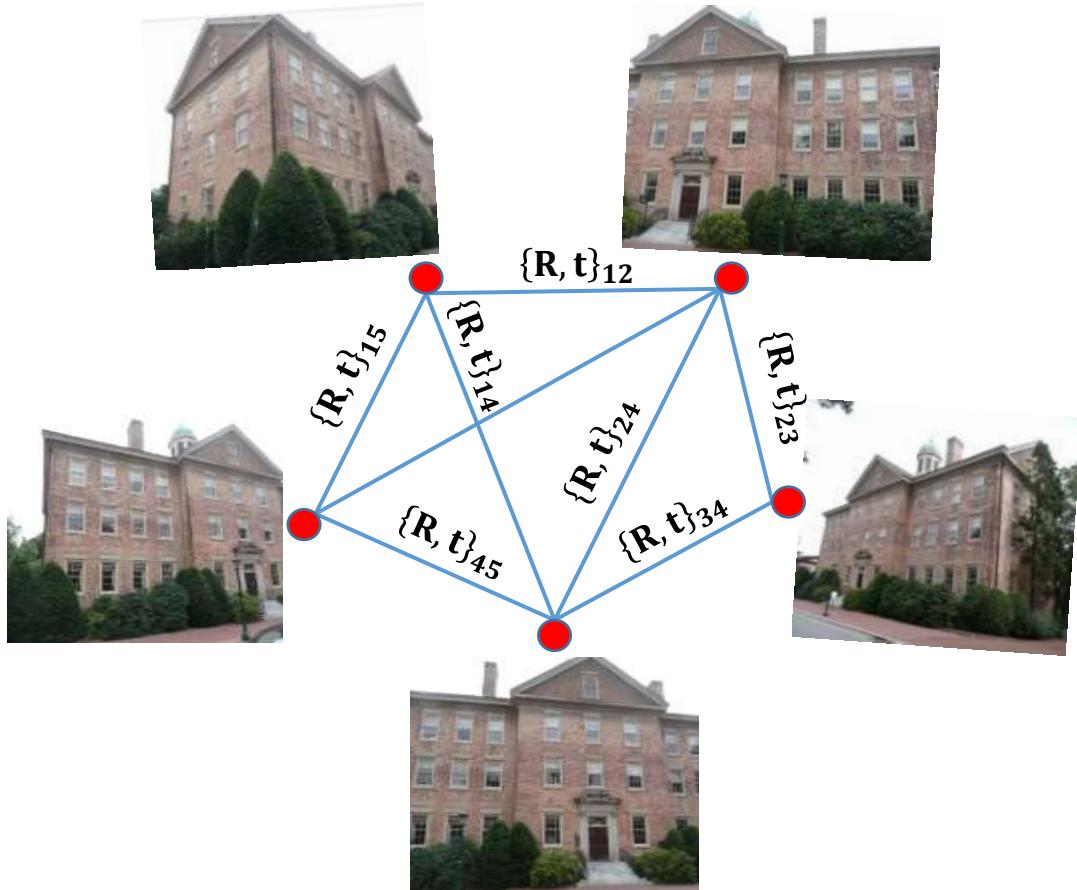
(c)



(d)

三角化选取内点最多的一组配置，作为两幅图像之间的外极几何关系

# 外极几何图

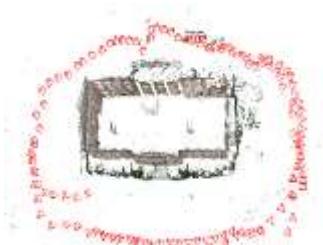
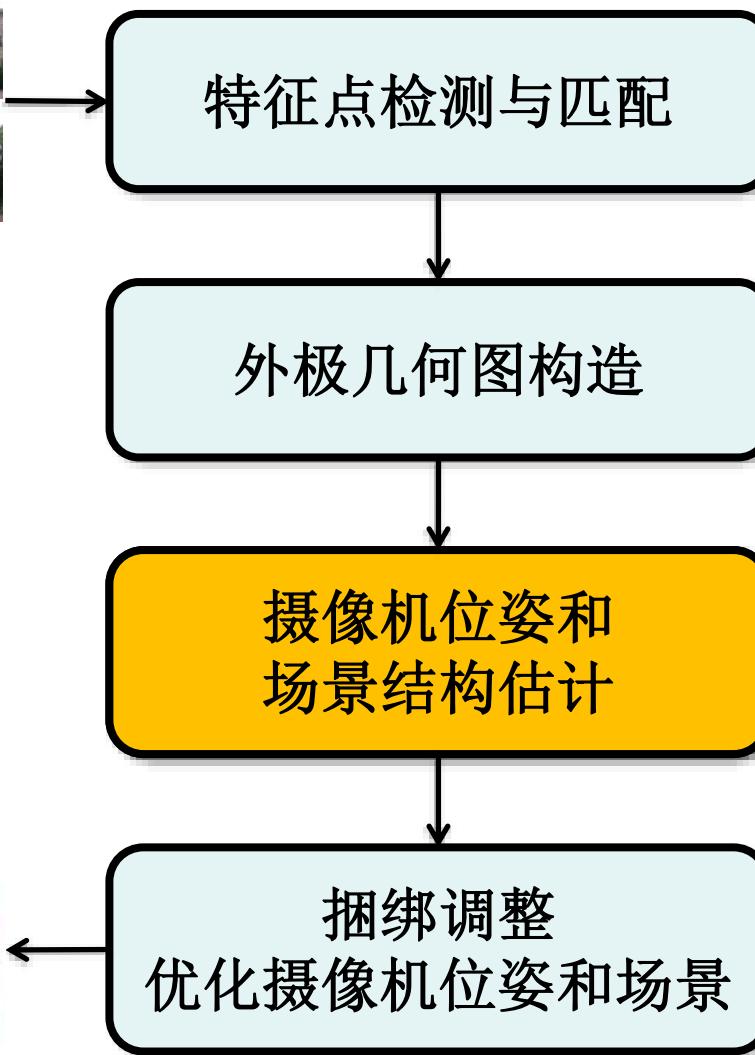


- (1) 图中每个顶点表示一幅图像，每一条边连接有公共可见区域的两张图像
- (2) 每一条边上表达了对应图像对之间的外极几何关系，其中：

$$R_{ij} = R_j R_i^T$$

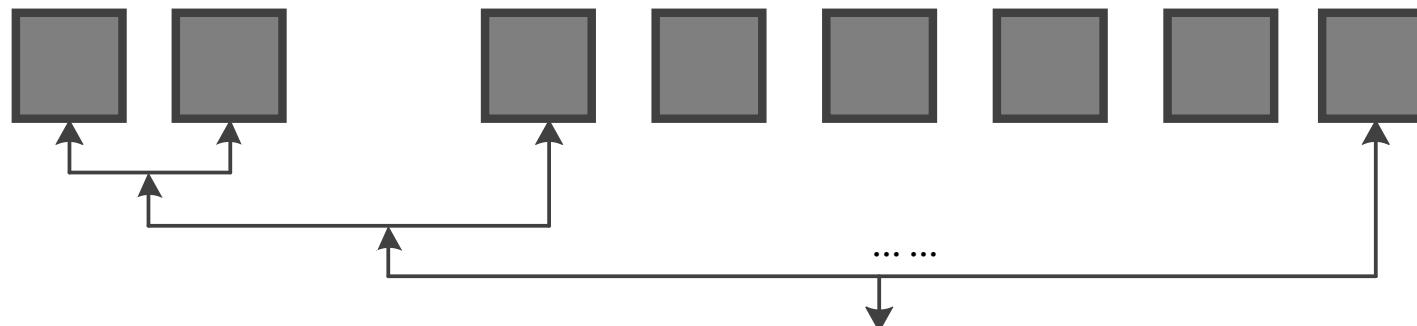
$$\lambda_{ij} t_{ij} = R_j (C_i - C_j)$$

# 从运动恢复结构 (SfM)

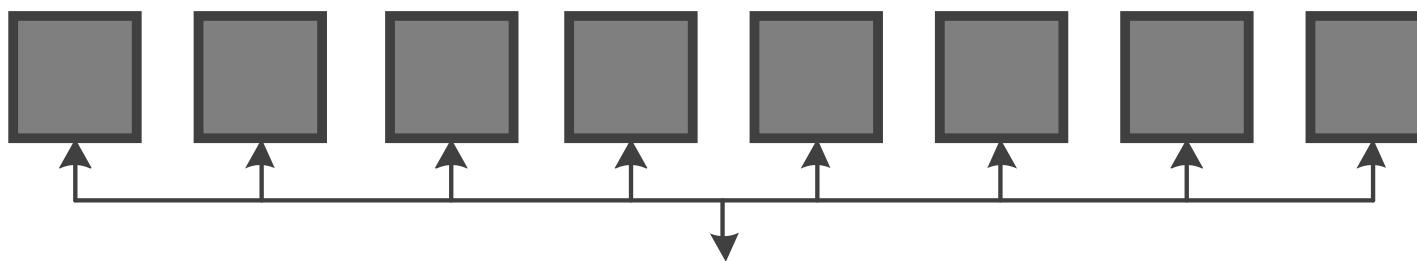


# 估计摄像机位姿和场景结构

- 增量式 (Incremental)

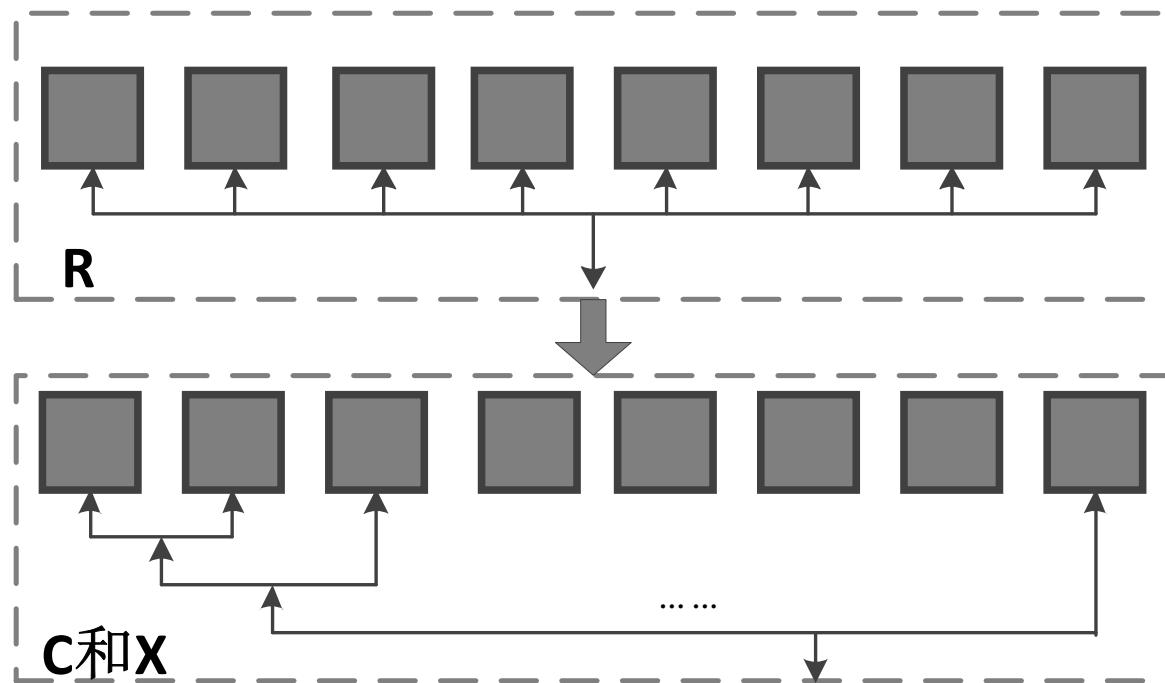


- 全局式 (Global)

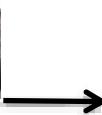


# 估计摄像机位姿和场景结构

- 混合式 (Hybrid)



# 从运动恢复结构 (SfM)



特征点检测与匹配



外极几何图构造



摄像机位姿和  
场景结构估计

● 增量式

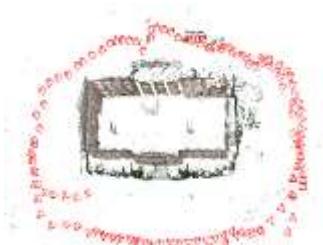


● 全局式

● 混合式

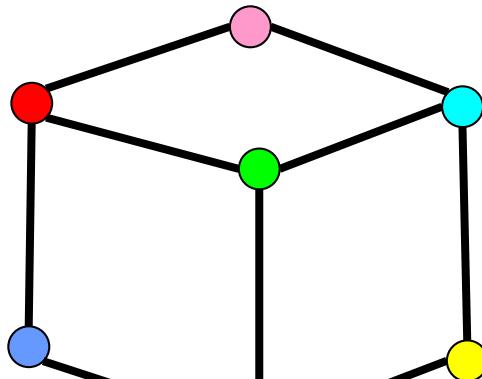


捆绑调整  
优化摄像机位姿和场景

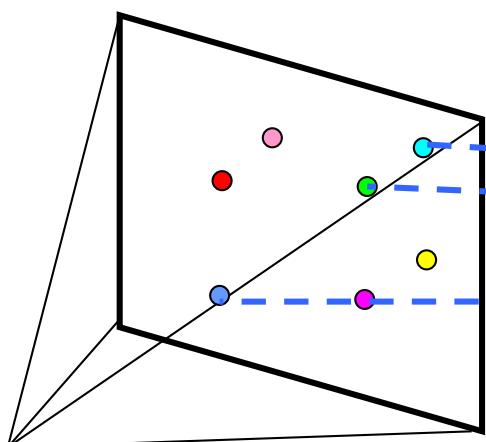


# 增量式

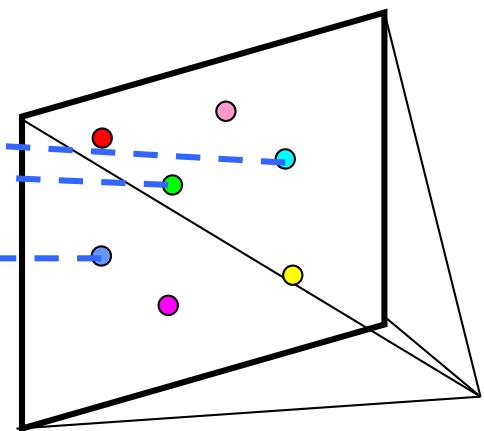
1. 基线足够宽
2. 匹配点足够多



$$\min_{P, X} \|x - \pi(P, X)\|$$

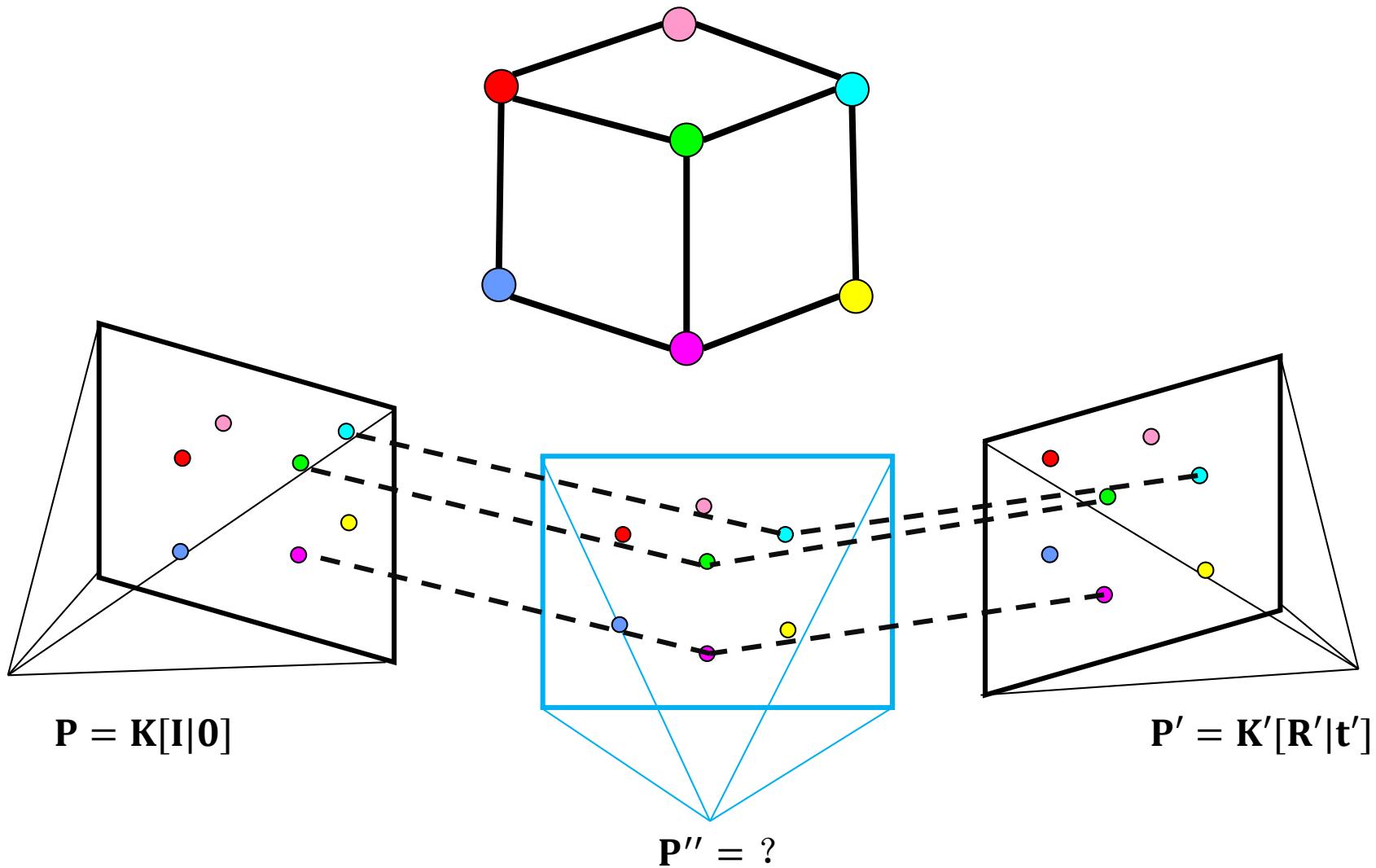


$$P = K[I|0]$$

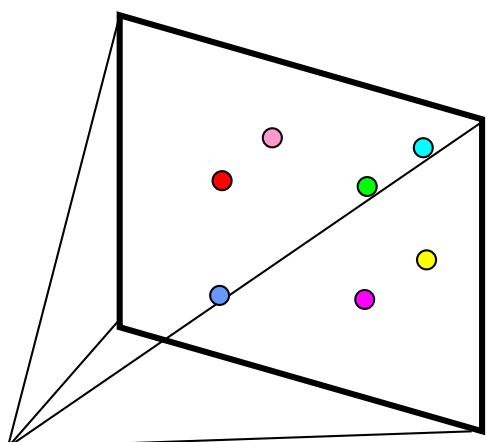


$$P' = K'[R'|t']$$

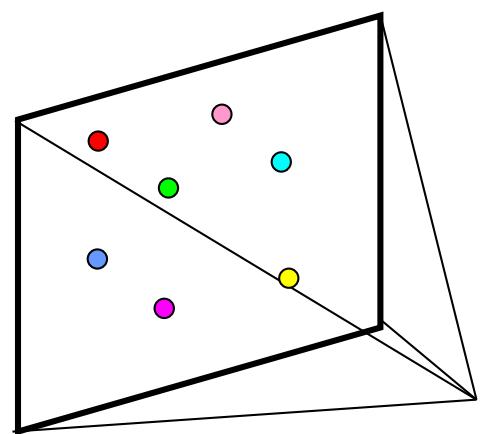
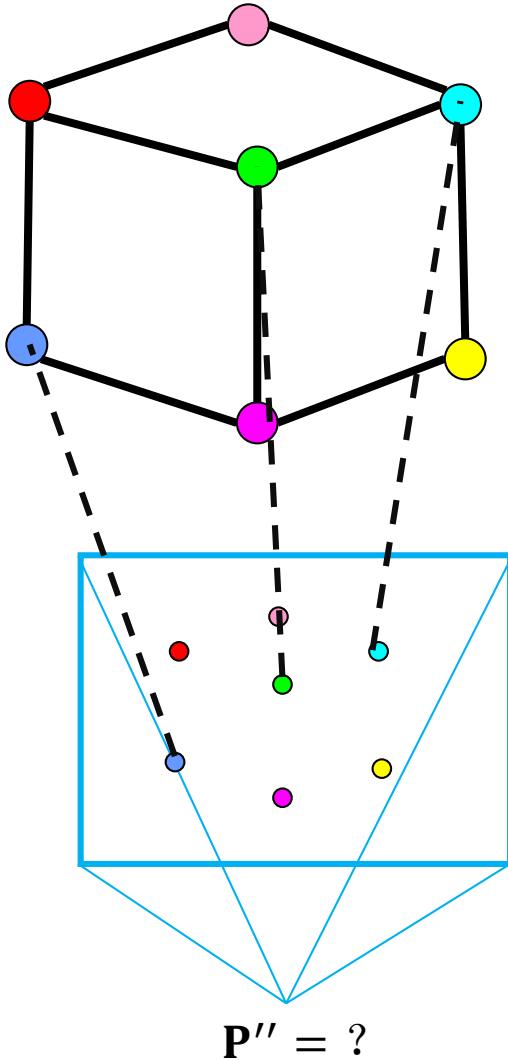
# 增量式



# 增量式

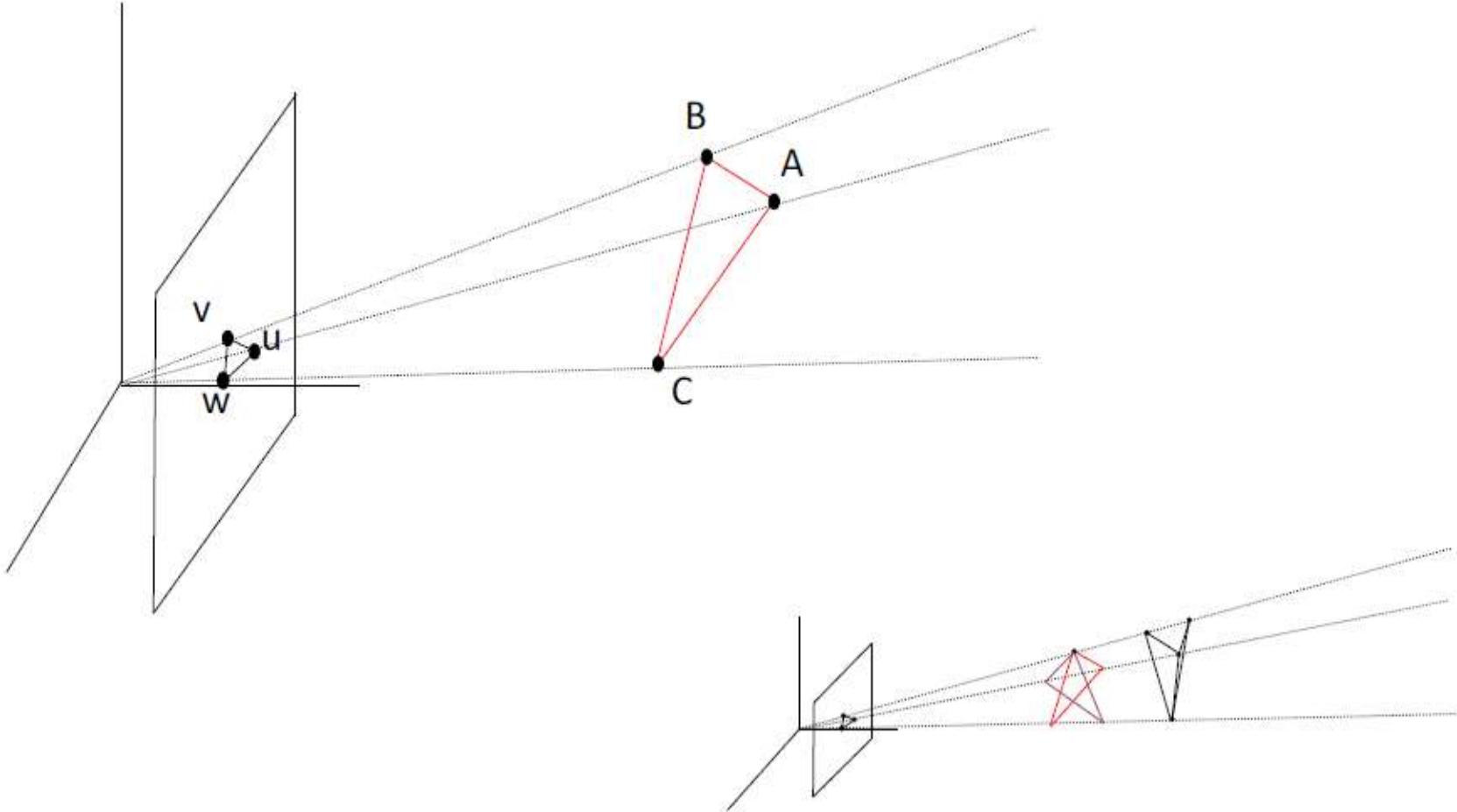


$$\mathbf{P} = \mathbf{K}[\mathbf{I}|0]$$

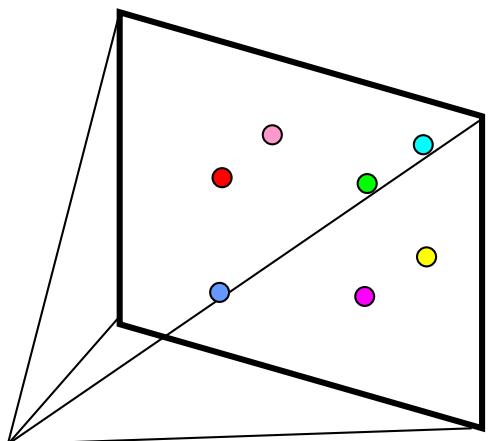


$$\mathbf{P}' = \mathbf{K}'[\mathbf{R}'|\mathbf{t}']$$

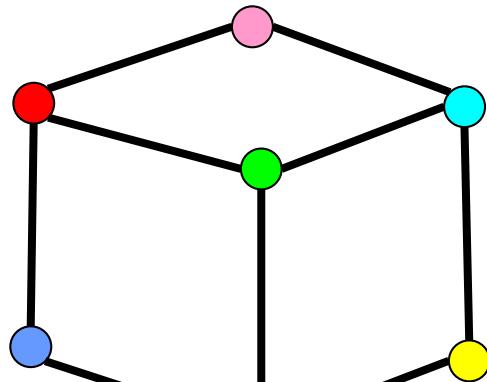
# P3P问题



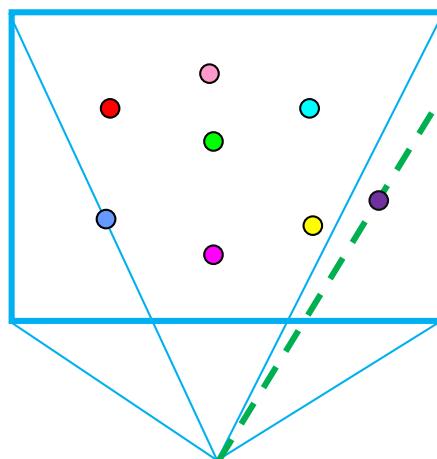
# 三角化



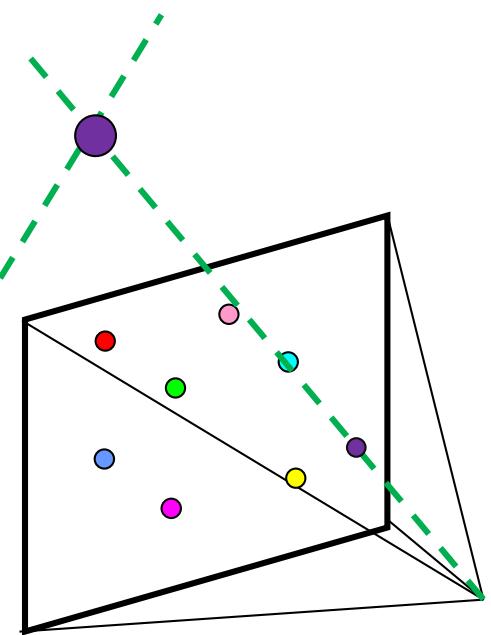
$$\mathbf{P} = \mathbf{K}[\mathbf{I}|0]$$



$$\min_X \|x - \pi(P, X)\|$$



$$\mathbf{P}'' = \mathbf{K}''[\mathbf{R}''|\mathbf{t}'']$$



$$\mathbf{P}' = \mathbf{K}'[\mathbf{R}'|\mathbf{t}']$$

# 捆绑调整

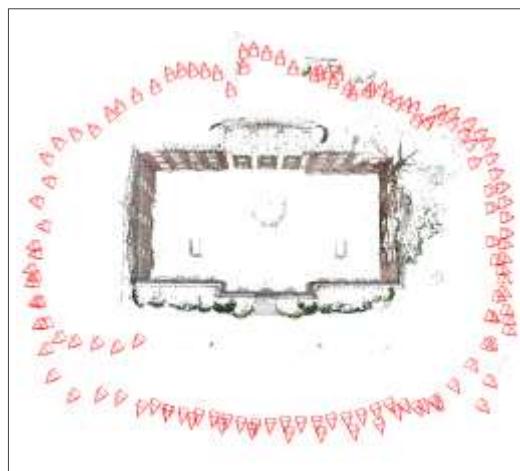
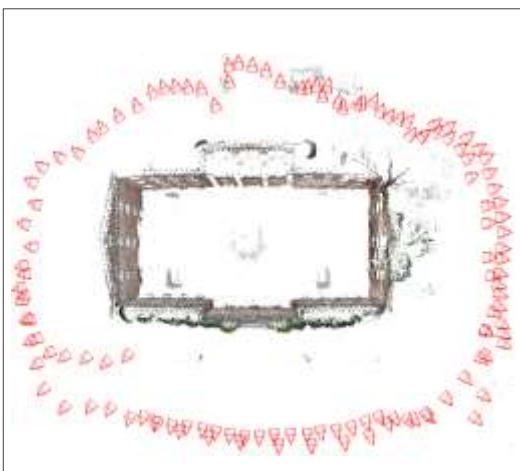
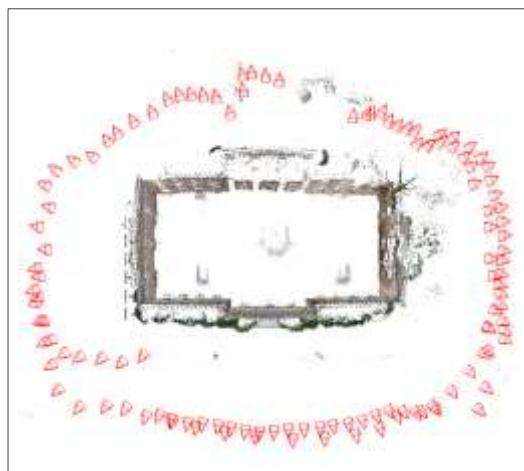
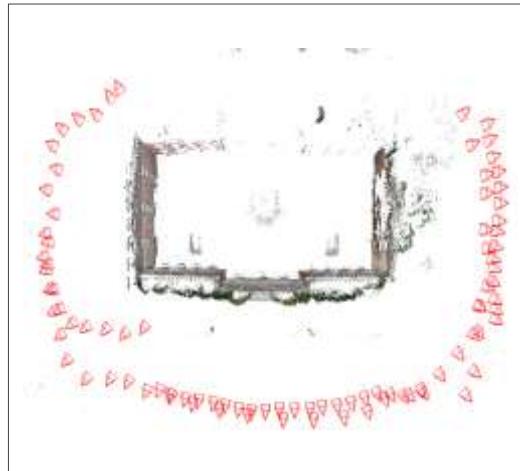
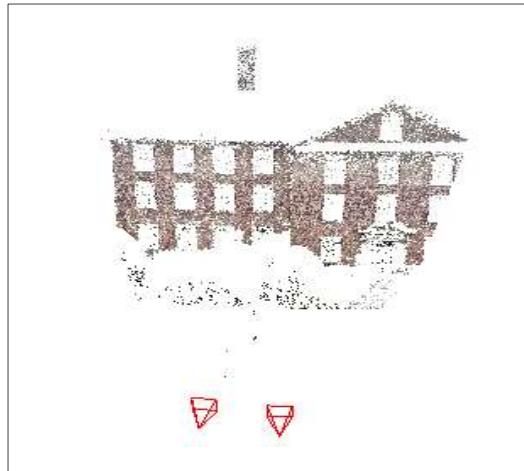
$$\min_{P, X} \|x - \pi(P, X)\|$$

$$\begin{aligned}
 \frac{\partial \mathbf{X}}{\partial \mathbf{P}} &= \left( \begin{array}{ccccccccc}
 \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} & \mathbf{B}_{12} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{0} & \mathbf{A}_{13} & \mathbf{B}_{13} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
 \mathbf{A}_{21} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{21} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{22} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{0} & \mathbf{A}_{23} & \mathbf{0} & \mathbf{B}_{23} & \mathbf{0} & \mathbf{0} \\
 \mathbf{A}_{31} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{31} & \mathbf{0} \\
 \mathbf{0} & \mathbf{A}_{32} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{32} & \mathbf{0} \\
 \mathbf{0} & \mathbf{0} & \mathbf{A}_{33} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{33} & \mathbf{0} \\
 \mathbf{A}_{41} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{41} \\
 \mathbf{0} & \mathbf{A}_{42} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{42} \\
 \mathbf{0} & \mathbf{0} & \mathbf{A}_{43} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{43}
 \end{array} \right) \rightarrow \left( \begin{array}{ccc|cccc}
 \mathbf{U}_1 & \mathbf{0} & \mathbf{0} & \mathbf{W}_{11} & \mathbf{W}_{21} & \mathbf{W}_{31} & \mathbf{W}_{41} \\
 \mathbf{0} & \mathbf{U}_2 & \mathbf{0} & \mathbf{W}_{12} & \mathbf{W}_{22} & \mathbf{W}_{32} & \mathbf{W}_{42} \\
 \mathbf{0} & \mathbf{0} & \mathbf{U}_3 & \mathbf{W}_{13} & \mathbf{W}_{23} & \mathbf{W}_{33} & \mathbf{W}_{43} \\ \hline
 \mathbf{W}_{11}^T & \mathbf{W}_{12}^T & \mathbf{W}_{13}^T & \mathbf{V}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
 \mathbf{W}_{21}^T & \mathbf{W}_{22}^T & \mathbf{W}_{23}^T & \mathbf{0} & \mathbf{V}_2 & \mathbf{0} & \mathbf{0} \\
 \mathbf{W}_{31}^T & \mathbf{W}_{32}^T & \mathbf{W}_{33}^T & \mathbf{0} & \mathbf{0} & \mathbf{V}_3 & \mathbf{0} \\
 \mathbf{W}_{41}^T & \mathbf{W}_{42}^T & \mathbf{W}_{43}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{V}_4
 \end{array} \right) \begin{pmatrix} \delta_{\mathbf{a}_1} \\ \delta_{\mathbf{a}_2} \\ \delta_{\mathbf{a}_3} \\ \delta_{\mathbf{b}_1} \\ \delta_{\mathbf{b}_2} \\ \delta_{\mathbf{b}_3} \\ \delta_{\mathbf{b}_4} \end{pmatrix} = \begin{pmatrix} \epsilon_{\mathbf{a}_1} \\ \epsilon_{\mathbf{a}_2} \\ \epsilon_{\mathbf{a}_3} \\ \epsilon_{\mathbf{b}_1} \\ \epsilon_{\mathbf{b}_2} \\ \epsilon_{\mathbf{b}_3} \\ \epsilon_{\mathbf{b}_4} \end{pmatrix} \\
 (\mathbf{J}^T \Sigma_{\mathbf{X}}^{-1} \mathbf{J} + \mu \mathbf{I}) \delta = \mathbf{J}^T \Sigma_{\mathbf{X}}^{-1} \epsilon
 \end{aligned}$$

$\downarrow$   
 $\left( \begin{array}{cc} \mathbf{U}^* - \mathbf{W} \mathbf{V}^{*-1} \mathbf{W}^T & \mathbf{0} \\ \mathbf{W}^T & \mathbf{V}^* \end{array} \right) \begin{pmatrix} \delta_{\mathbf{a}} \\ \delta_{\mathbf{b}} \end{pmatrix} = \begin{pmatrix} \epsilon_{\mathbf{a}} - \mathbf{W} \mathbf{V}^{*-1} \epsilon_{\mathbf{b}} \\ \epsilon_{\mathbf{b}} \end{pmatrix}$



# 增量式重建示例





# 增量式的优势与不足

系统对于特征匹配以及外极几何关系的外点比较鲁棒，重建场景精度高：

1. 标定过程中通过RANSAC不断地过滤外点
2. 捆绑调整不断地优化场景结构

主要不足：

1. 敏感：初始图像对选择及摄像机的添加顺序
2. 场景漂移：大场景重建时的累积误差
3. 效率不足：反复的捆绑调整需要大量的计算时间

# 增量式重建示例

	openMVG	VSFM	Theia	COLMAP
Building				
Campus				

openMVG : ACCV2012

VSFM : 3DV2013

Theia : <http://theia-sfm.org> ( 2015 )

COLMAP : CVPR2016

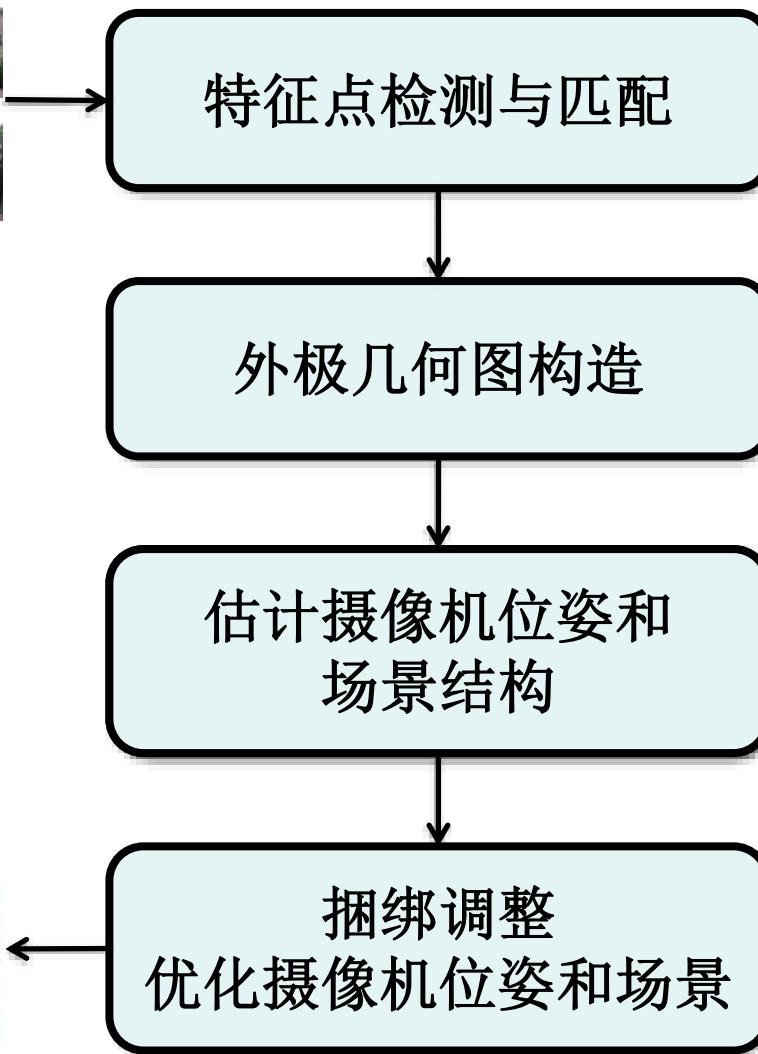




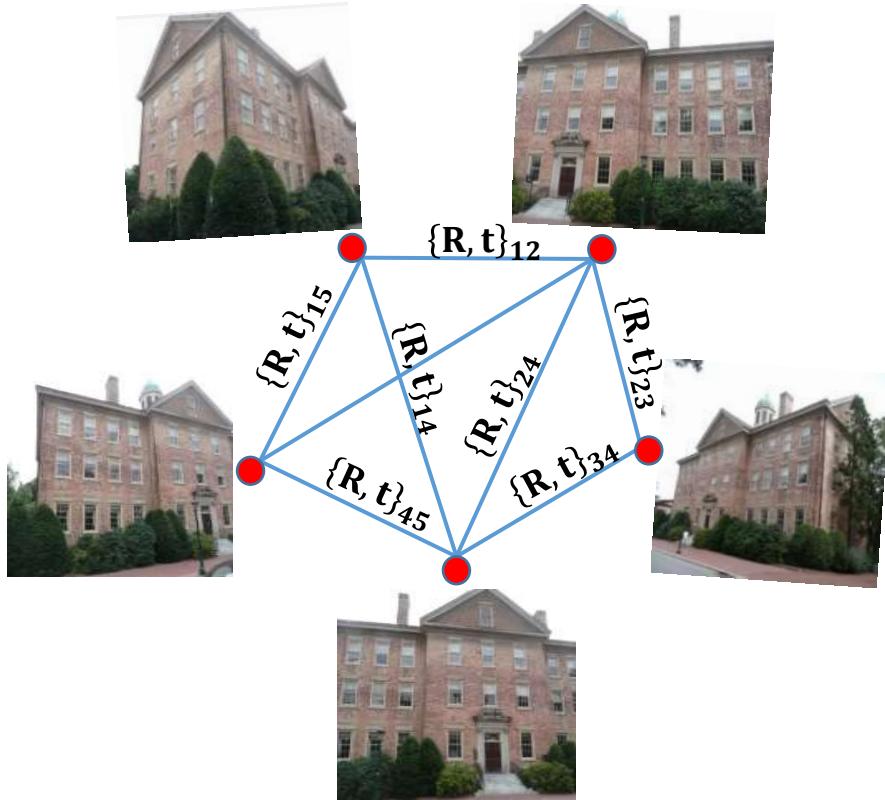
# 相关文献

- 1 ) Bundler : Noah Snavely, et al. “Modeling the World from Internet Photo Collections.” IJCV 2007.
- 2 ) openMVG : Moulon Pierre, et al. “Adaptive Structure from Motion with a contrario model estimation.” ACCV 2012.
- 3 ) VSFM: Wu Changchang. “Towards linear-time incremental structure from motion.” 3DV 2013.
- 4 ) Theia: Chris Sweeney. “Theia Multiview Geometry Library: Tutorial & Reference.” 2015. <http://theia-sfm.org>
- 5 ) Colmap: Schonberger, Johannes Lutz, et al. “Structure-from-Motion Revisited.” CVPR 2016.
- 6 ) BSfM: Hainan Cui, et al. “Batched Incremental Structure-from-Motion.” 3DV 2017.

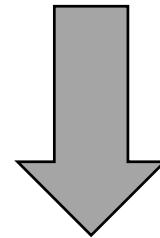
# 从运动恢复结构 (SfM)



# 全局式



$$R_{ij} = R_j R_i^T$$
$$\lambda_{ij} t_{ij} = R_j(C_i - C_j)$$

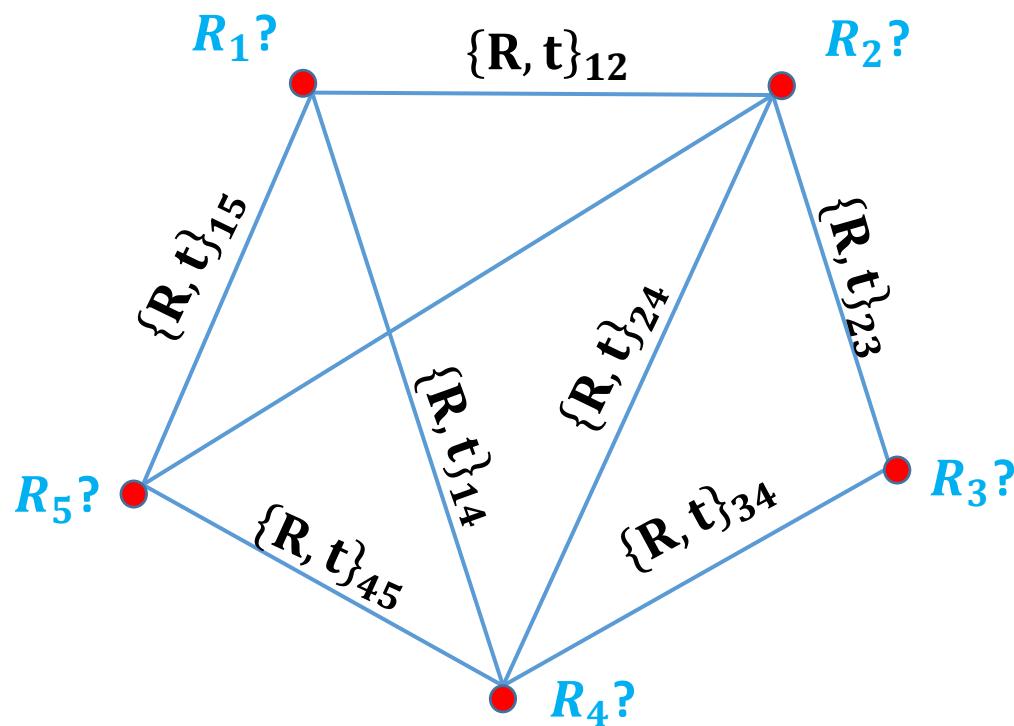


1. 估计所有摄像机的旋转矩阵
2. 估计所有摄像机的位置
3. 三角化初始场景点

# 全局式估计摄像机旋转矩阵

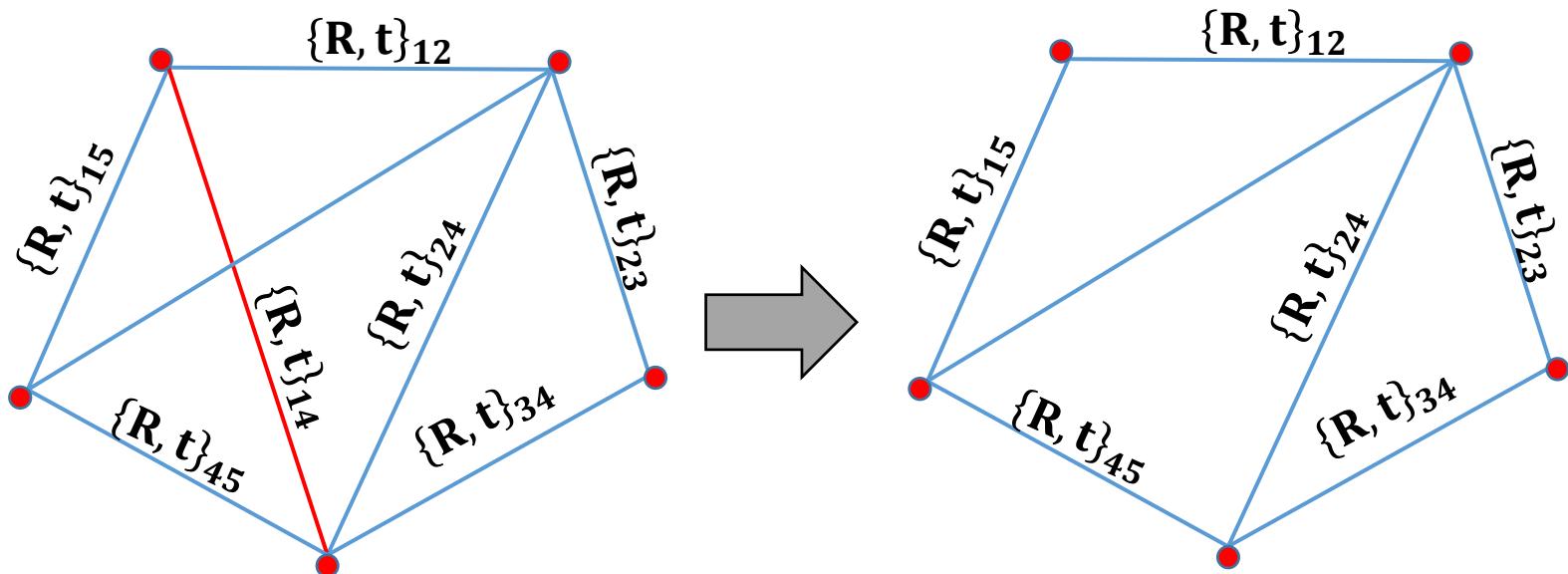
$$R_{ij} = R_j R_i^T \quad \longrightarrow \quad \min_R \|R_{ij} - R_j R_i^T\|$$

$$w_{ij} = w_j - w_i \quad \longrightarrow \quad AX = b$$



# 过滤外极几何边

$$\|R_{ij} - R_j R_i^T\| \geq \epsilon$$

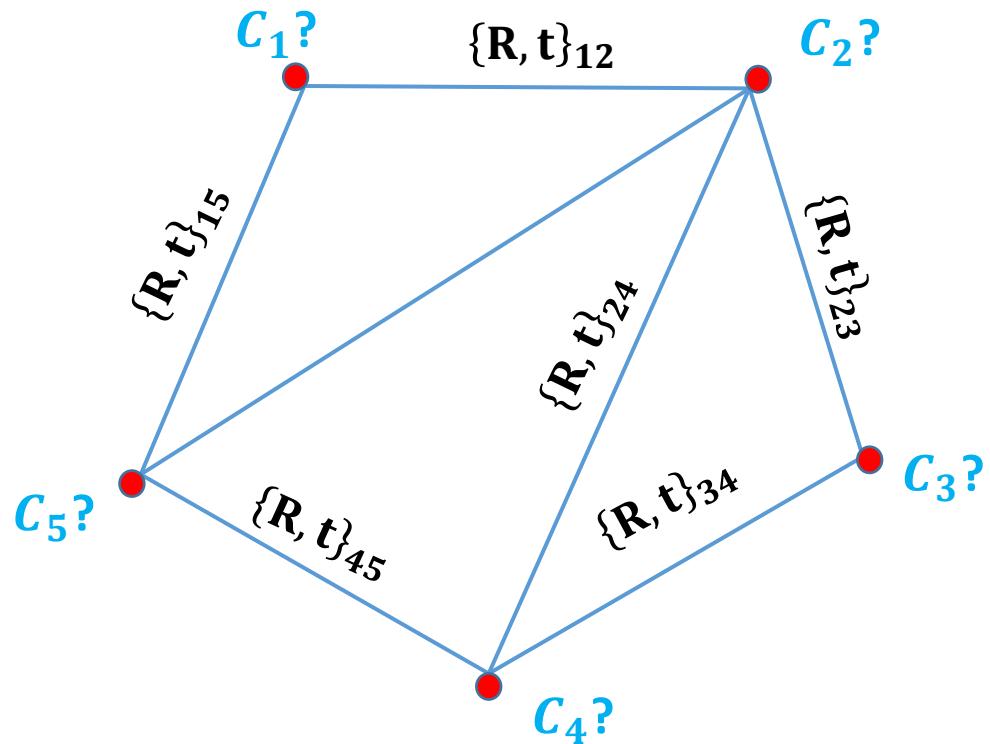


# 全局式估计摄像机位置

$$\begin{aligned}\lambda_{ij} t_{ij} &= R_j(C_i - C_j) \\ \mu_{ij} &= R_j^T t_{ij}\end{aligned}$$

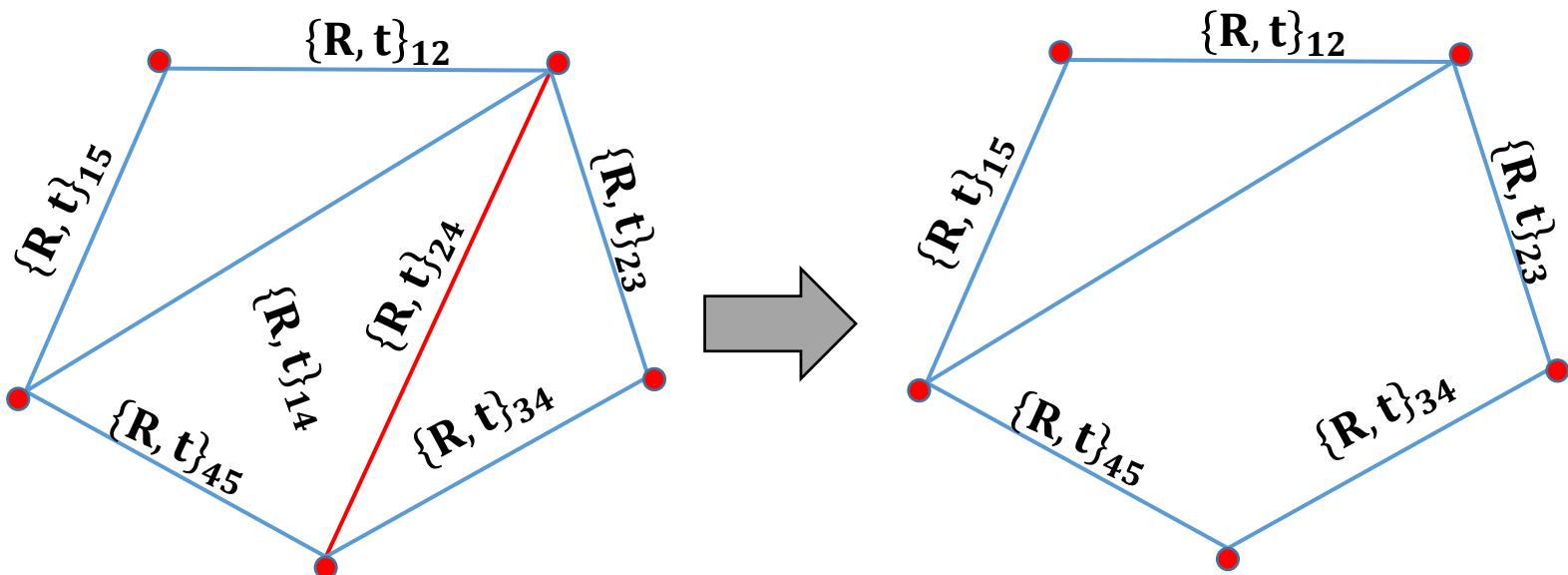


$$\min_C \left\| \mu_{ij} - \frac{C_i - C_j}{\|C_i - C_j\|} \right\|^2$$

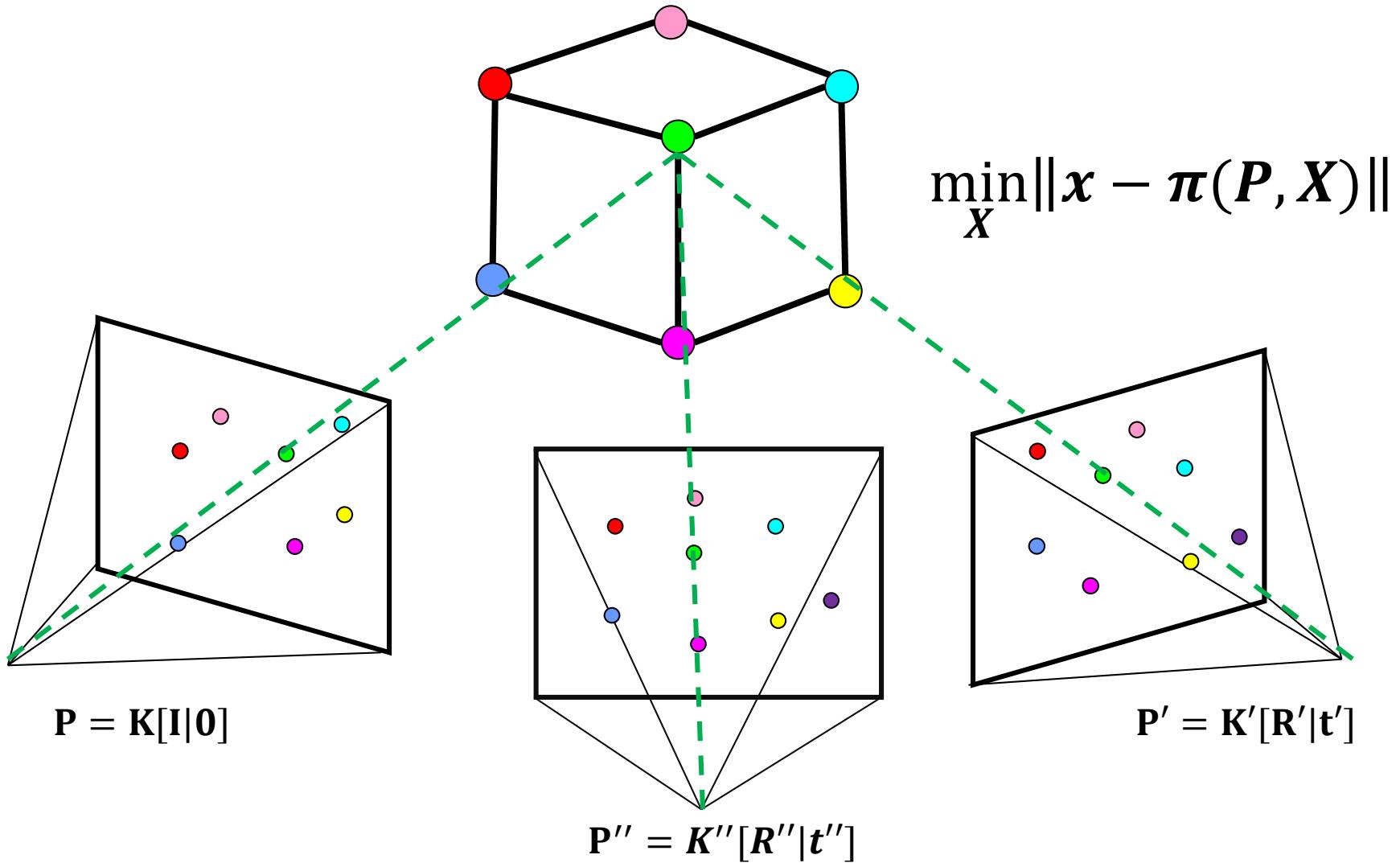


# 过滤外极几何边

$$\left\| \mu_{ij} - \frac{c_i - c_j}{\|c_i - c_j\|} \right\| \leq \epsilon$$



# 三角化初始场景点





# 全局式的优势与不足

主要优势：

1. 将误差均匀分布在外极几何图上，没有误差累积
2. 不需要考虑初始图像和图像添加顺序的问题
3. 仅执行一次捆绑调整，重建效率高

主要不足：

1. 鲁棒性不足：旋转矩阵求解时L1范数对外点相对鲁棒；而摄像机位置求解时相对平移关系对匹配外点比较敏感。
2. 场景完整性：过滤外极几何边，可能会丢失部分图像



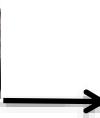
# 相关文献

- 1) Nianjuan Jiang, et al. "A Global Linear Method for Camera Pose Registration." ICCV 2013. (三视图线性求解)
- 2) P. Moulon, et al. "Global Fusion of Relative Motions for Robust, Accurate and Scalable structure from motion." ICCV 2013. (三视图约束)
- 3) Kyle Wilson, et al. "Robust Global Translations with 1DSfM." ECCV 2014. (外极几何边过滤)
- 4) O. Ozyesil, et al. "Robust Camera Location Estimation by Convex Programming." CVPR 2015 . (引入平行刚体概念)
- 5) C. Sweeney, et al. "Optimizing the Viewing Graph for Structure-from-Motion." ICCV 2015. (利用track优化外极几何关系)
- 6) T. Goldstein, et al. "Shapefit and Shapekick for Robust, Scalable Structure from Motion." ECCV 2016.

# 增量式 vs. 全局式

	鲁棒性	精度	效率	场景完整性
增量式				
全局式				

# 从运动恢复结构 (SfM)



特征点检测与匹配



外极几何图构造

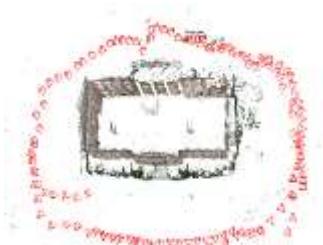


估计摄像机位姿和  
场景结构

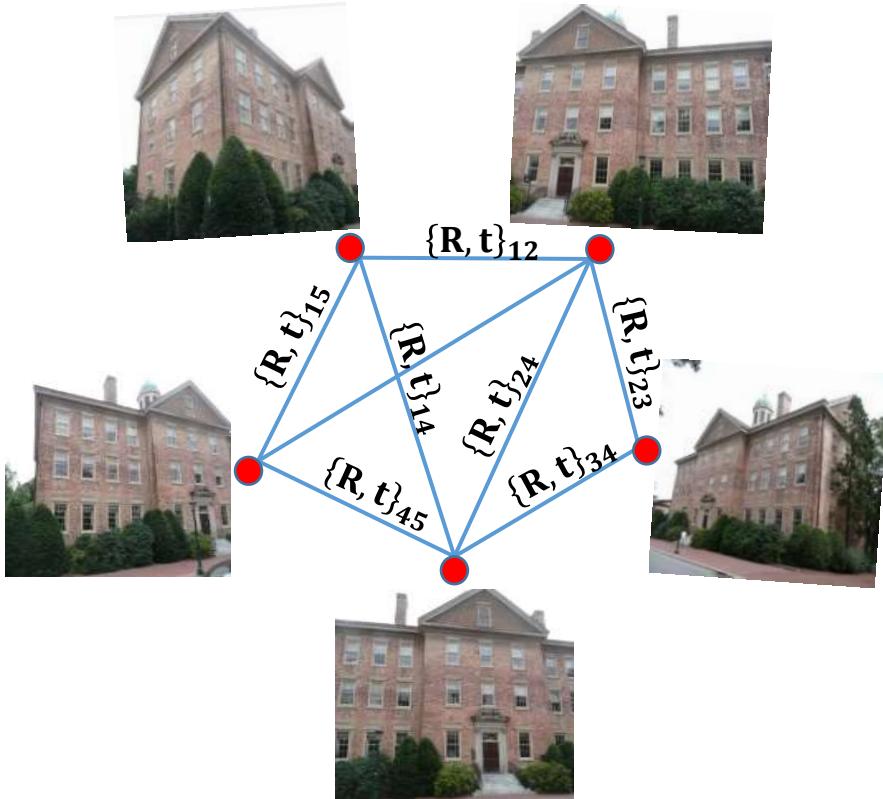
- 增量式
- 全局式
- 混合式



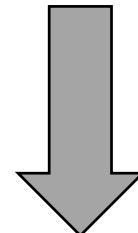
捆绑调整  
优化摄像机位姿和场景



# 混合式



$$R_{ij} = R_j R_i^T$$
$$\lambda_{ij} t_{ij} = R_j(C_i - C_j)$$



1. 全局式估计摄像机旋转矩阵
2. 增量式估计摄像机位置
3. 三角化初始场景点



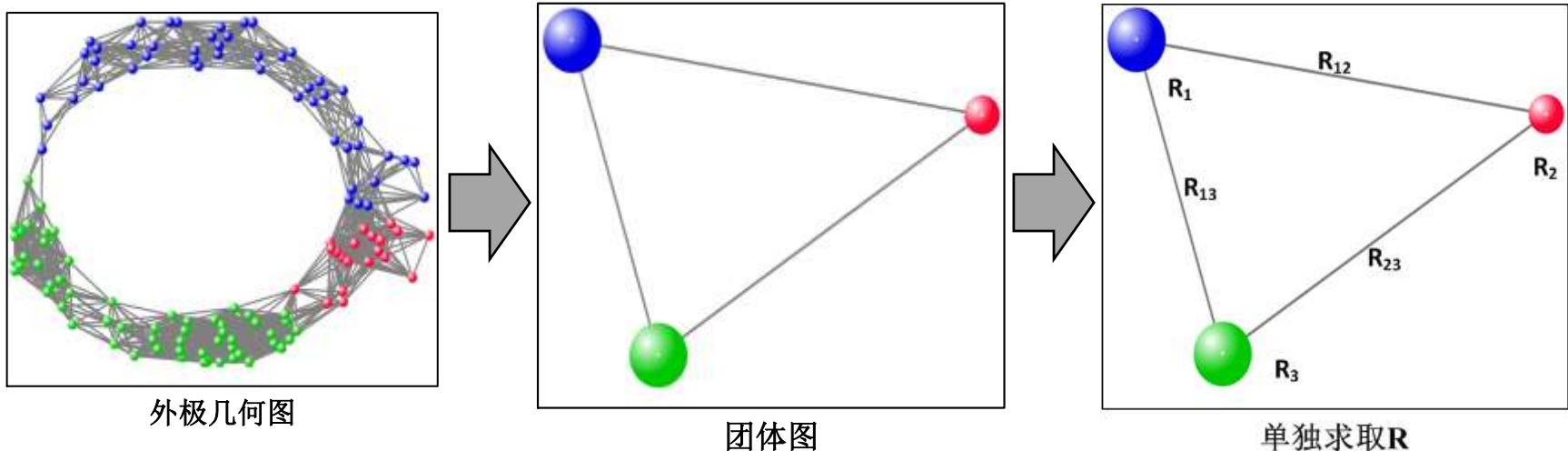
# 全局式估计摄像机旋转矩阵

全局式求取摄像机旋转矩阵精度，与以下两种因素有关：

1. 图像之间的连接紧密程度
2. 外极几何关系的估计精度

$$\frac{\lambda_2(\mathbf{L})}{n} > \frac{\Delta(\theta_{ij})}{\mu(\theta_{max})}$$

# 基于团体的全局式 $R$ 求取



A. Clauset, M. E. Newman, and C. Moore. Finding community structure in very large networks. Physical Review E, 70(6 Pt 2):264–277, 2005. 3, 4



# 基于团体的全局式 $R$ 求取



Community\_1

Community\_2

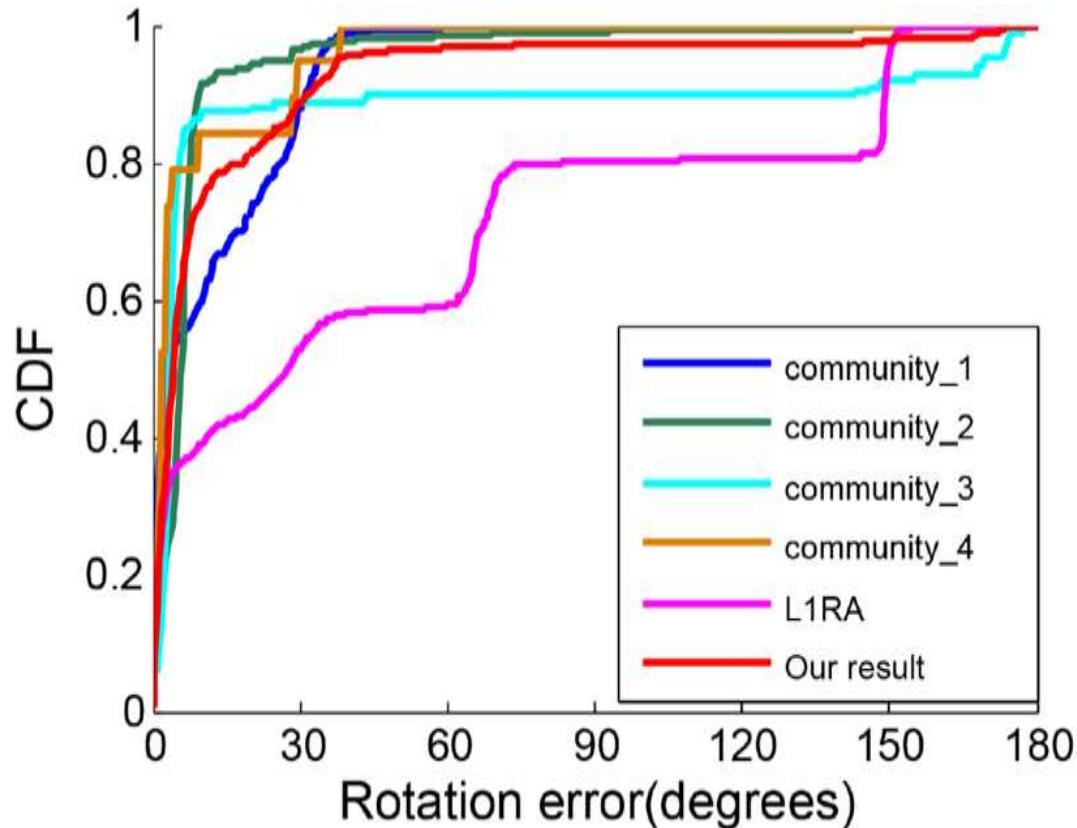


Community\_3

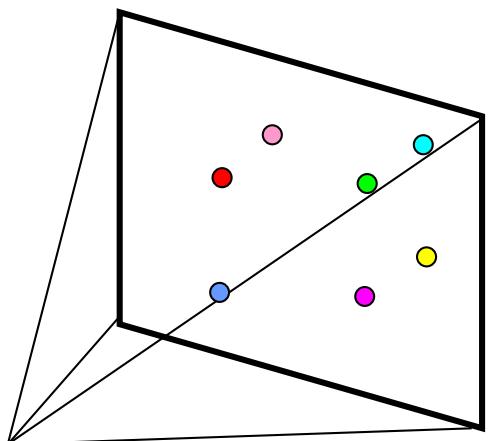
Community\_4



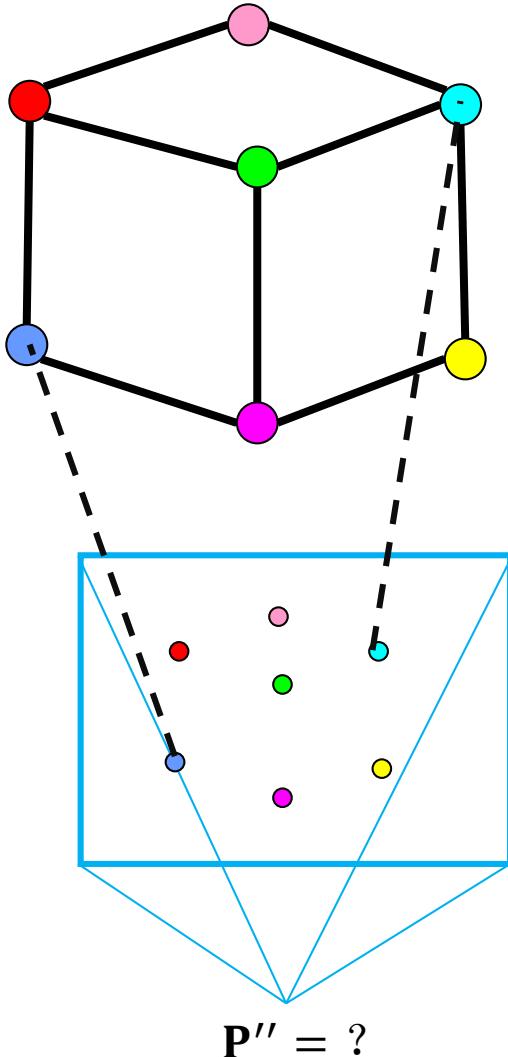
# 基于团体的全局式 $R$ 求取



# 增量式摄像机位置求取

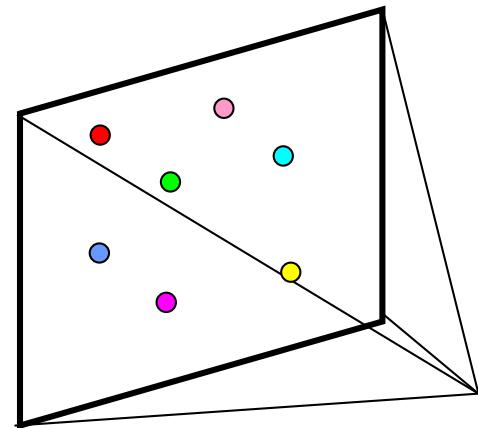


$$P = K[I|0]$$



$$\lambda_{ij}x_{ij} = K_i R_i (X_i - C_j)$$

$$\lambda_{ij} R_i^T K_i^{-1} x_{ij} = X_i - C_j$$



$$P' = K'[R'|t']$$



# 增量式摄像机位置求取

已知摄像机旋转矩阵，两点法求取摄像机位置的优势：

- 1) RANSAC最小配置减少，更大概率随机到正确的三维点与二维点对应；（精度）
- 2) 已知R情况下，可以一定程度上过滤一部分错误的匹配特征点；（鲁棒性）
- 3) 在每次增量过程中，可以同时计算得到更多摄像机的位置，从而大幅度的减少了捆绑调整的次数，加速系统重建。（效率）



# 混合式的优势和不足

主要优势：

1. 求解旋转矩阵时误差均摊到了外极几何图中
2. 全局旋转矩阵的求解缓解了后续增量式的误差累积问题
3. 求解摄像机位置时，两点法求取更加鲁棒，所需的捆绑次数更少

主要不足：

1. 计算摄像机位置仍需要重复迭代的捆绑调整，效率需要进一步改善

# 增量式 vs. 全局式 vs. 混合式

	鲁棒性	精度	效率	场景完整性
增量式				
全局式				
混合式				



# 总结

稀疏重建通过摄像机运动恢复场景结构，当前主流方法主要区分在于摄像机初始位姿计算的模式。

- 1) 增量式鲁棒，场景结构准确，但效率不足；全局式仅一次捆绑调整，效率高，但鲁棒性不足，易受到匹配外点的影响；
- 2) 混合式继承了增量和全局两种模式的优点：不仅提高了旋转矩阵的求取精度，而且在保持鲁棒性的前提下，提高了增量式重建的效率。



# 问题？

稀疏重建中亟待解决的问题：

1) 鲁棒性

图像特征点（学习得到的点或者带语义信息的点）

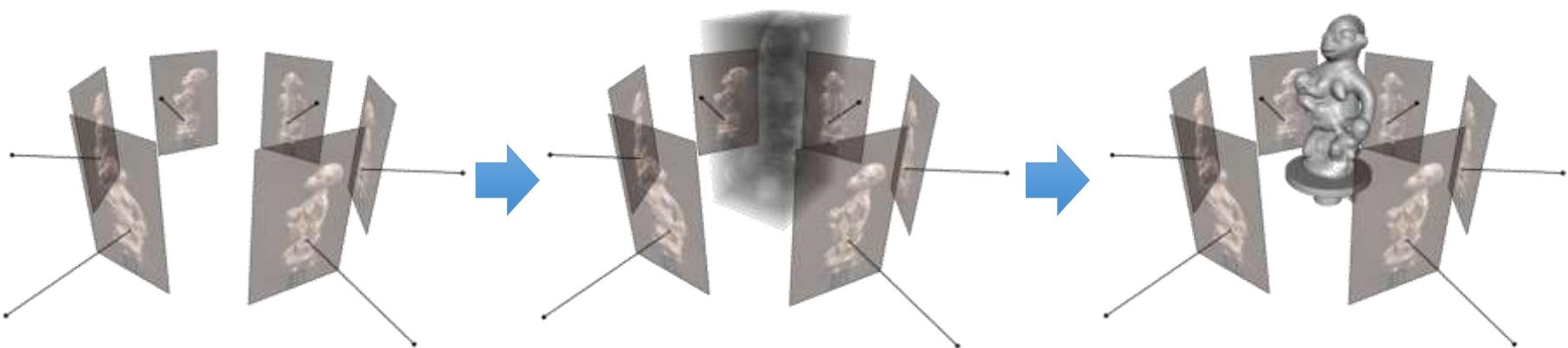
2) 模型的完整性

地面图像与天空图像之间的匹配

3) 效率

更快速的图像匹配策略（保证召回率）；分布式SfM

### 3. 稠密重建





# 提 纲

## 1、简介（申抒含，20分钟）

- 图像三维重建基本流程
- 多视几何基础

## 2、稀疏重建（崔海楠，40分钟）

- 增量式稀疏重建
- 全局式稀疏重建
- 混合式稀疏重建

## 3、稠密重建（申抒含，40分钟）

- 基于体素的稠密重建
- 基于特征点扩散的稠密重建
- 基于深度图融合的稠密重建

## 4、资源简介（申抒含，20分钟）

- 三维重建开源与商业软件
- 三维重建评测数据库

# 稠密重建(Multiple View Stereo)

多视角  
图像

图像特征  
提取匹配

稀疏重建  
(SfM)

稠密重建  
(MVS)

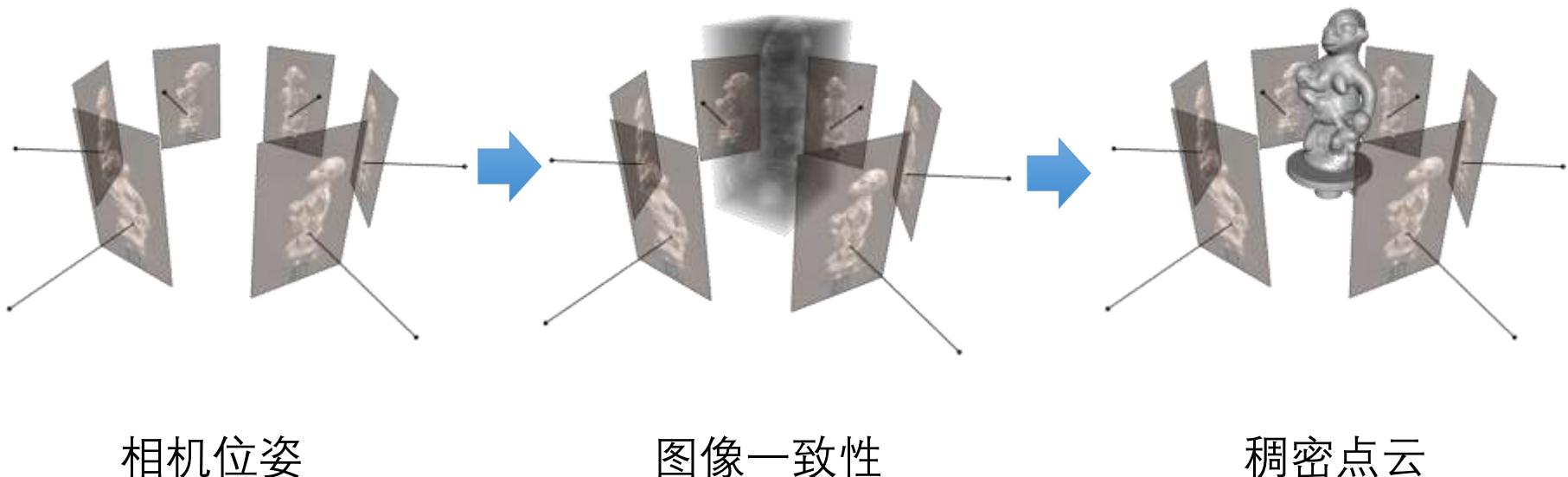
点云  
模型化

三维  
模型

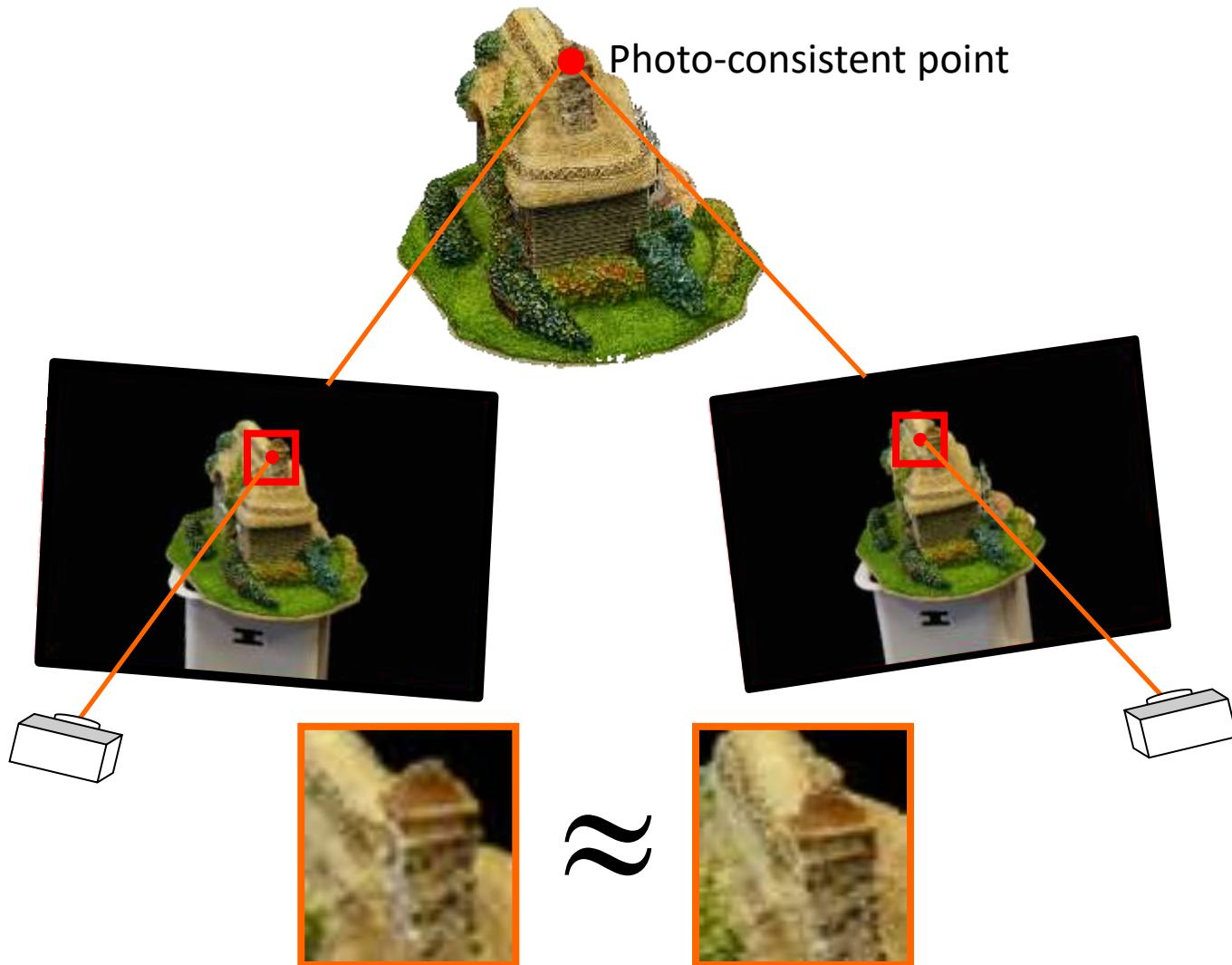


# MVS的基本思路

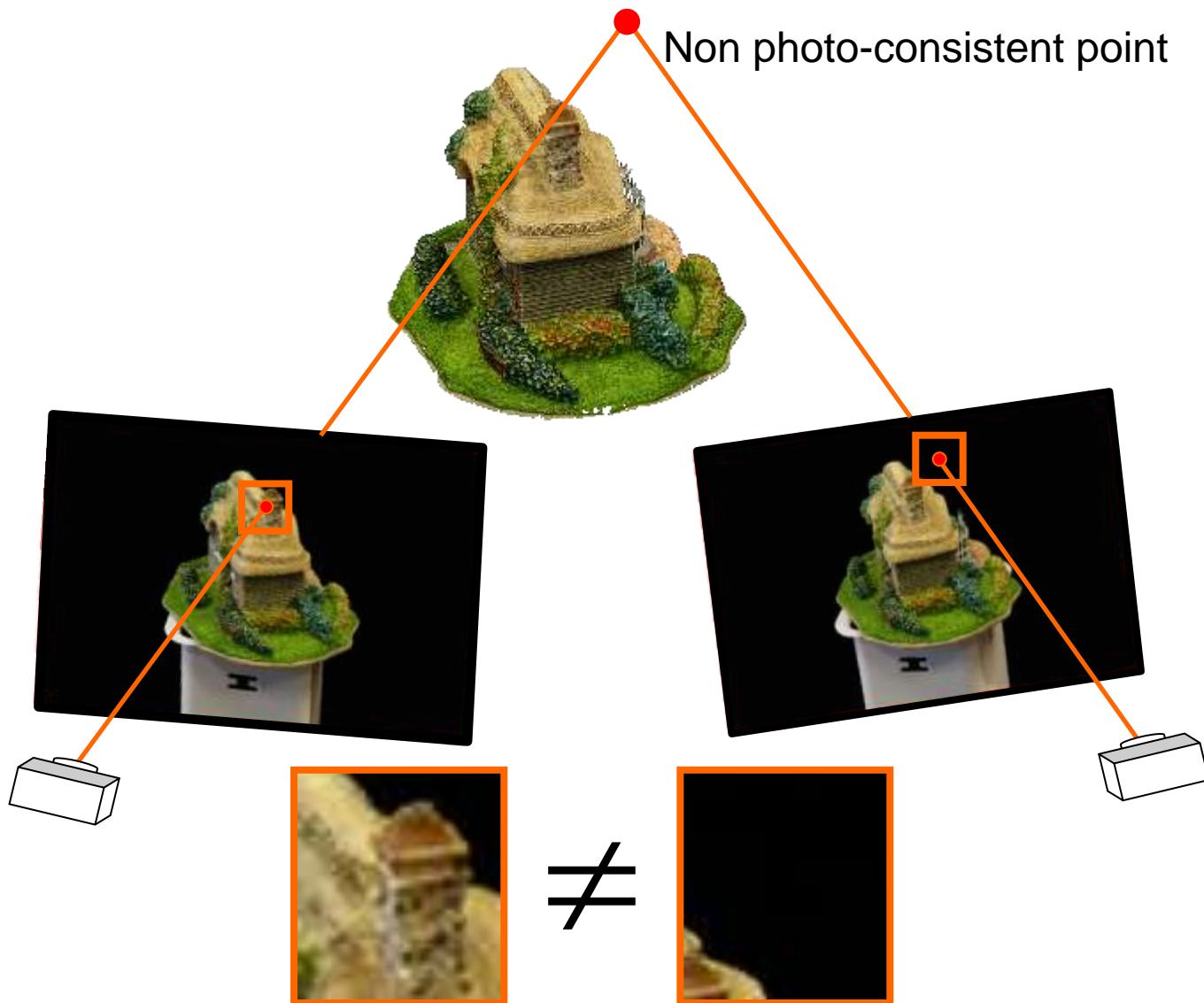
MVS的基本思路：寻找空间中具有图像一致性(Photo-consistency)的点



# Two-view Photo-consistency



# Two-view Photo-consistency



# Two-view Photo-consistency

两视图图像一致性(Photo-consistency)的度量：

SSD (Sum of Squared Differences) :  $\rho_{SSD}(f, g) = \|f - g\|^2$

SAD (Sum of Absolute Differences) :  $\rho_{SAD}(f, g) = \|f - g\|_1$

NCC (Normalized Cross Correlation) :  $\rho_{NCC}(f, g) = \frac{(f - \bar{f}) \cdot (g - \bar{g})}{\delta_f \delta_g}$

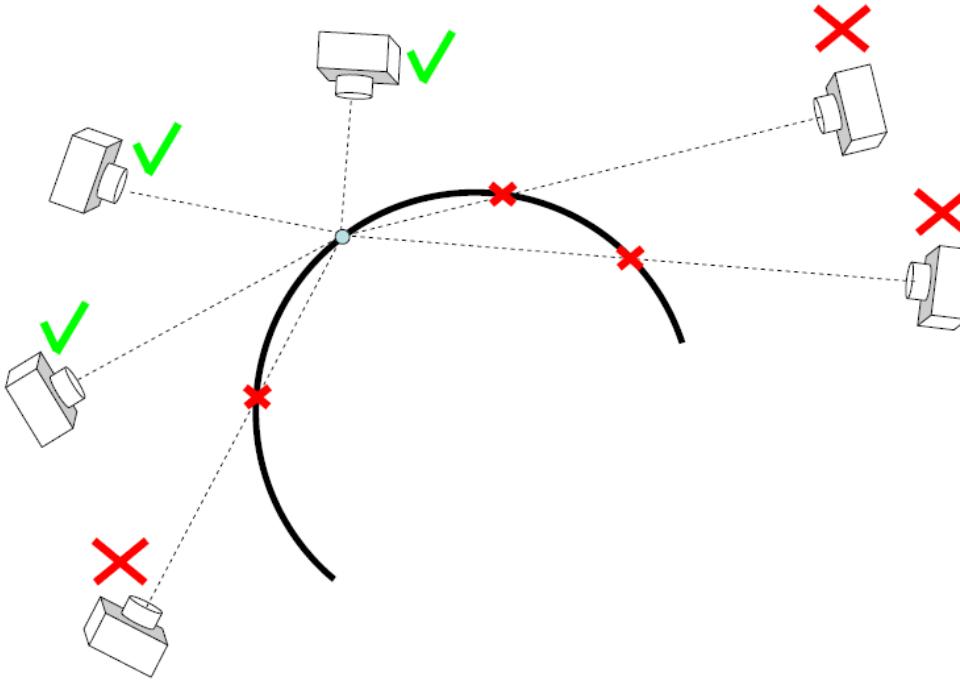
$f =$



$g =$

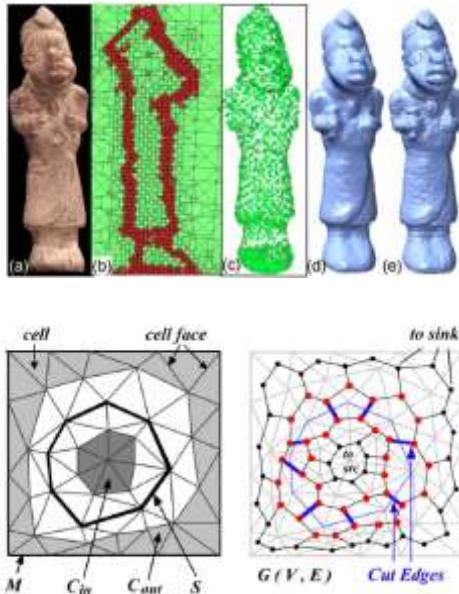


# Multi-view Photo-consistency

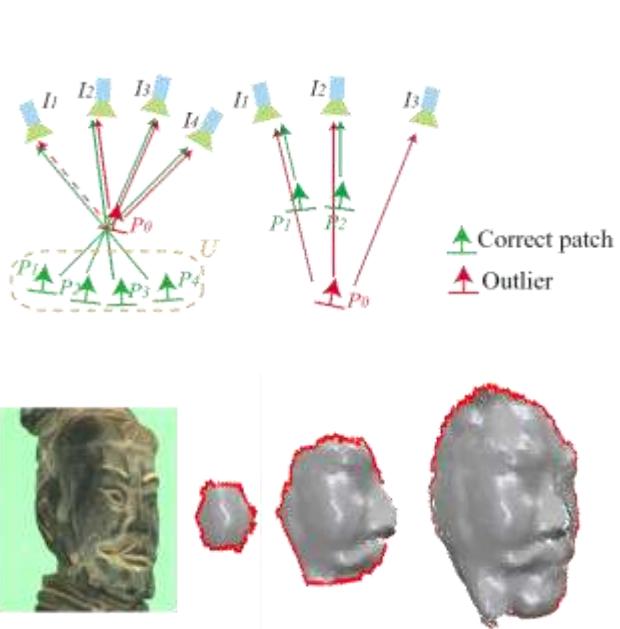


- 多视图图像一致性需要考虑相机的可视性问题
- 相机可视性需要场景结构、场景结构需要相机可视性
- This is a chicken-and-egg problem

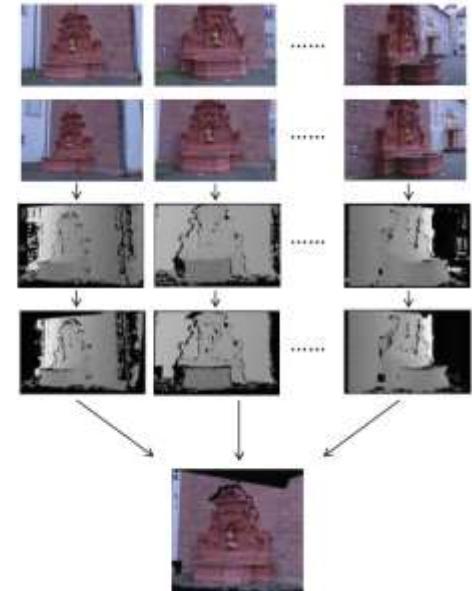
# MVS算法的分类



基于体素的方法  
Voxel based MVS



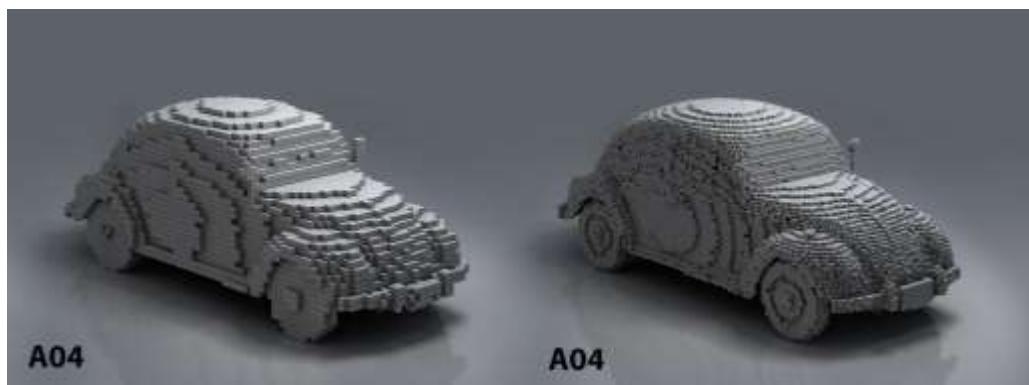
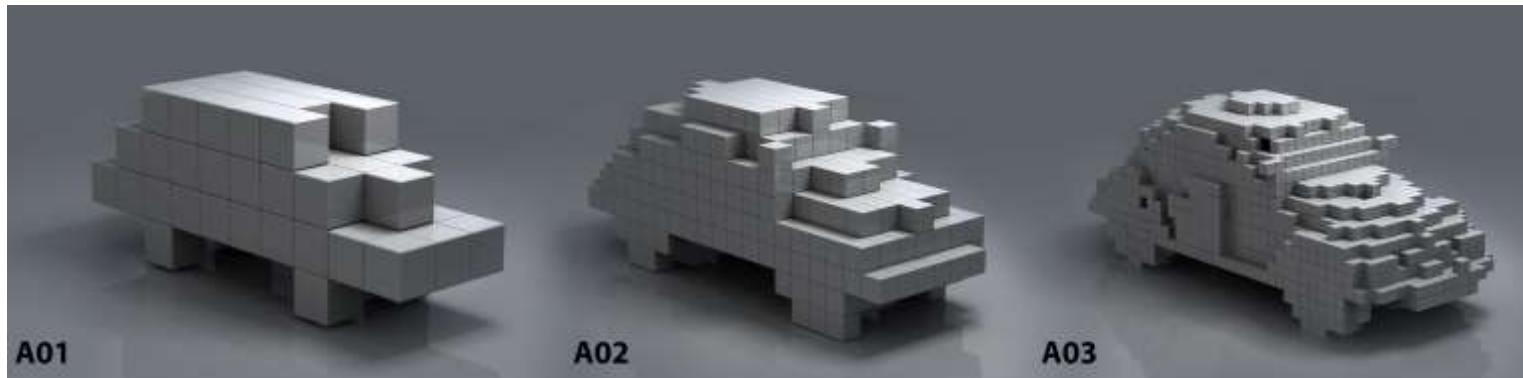
基于点云扩散的方法  
Feature point growing  
based MVS



基于深度图融合的方法  
Depth-map merging  
based MVS

# 基于体素的MVS

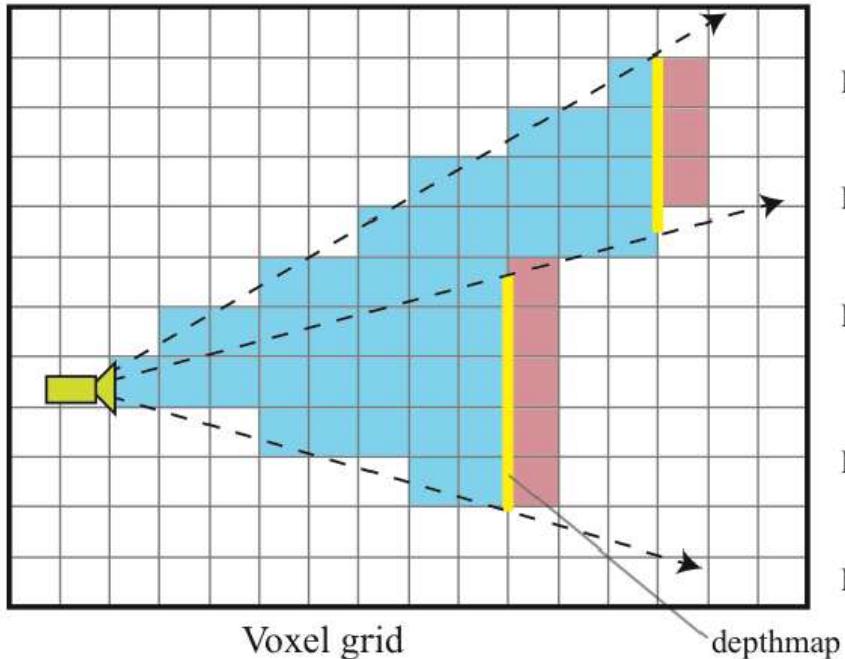
3D空间的Voxel表达：



MVS等价为一个3D空间Voxel的标记(Labeling)问题

# 基于体素的MVS

离散空间的Labeling是一个典型的MRF(Markov Random Field)优化问题：



For  $\Phi(k_v) = \begin{cases} \text{high} & \text{if } interior \\ 0 & \text{if } exterior \end{cases}$

For  $\Phi(k_v) = \begin{cases} 0 & \text{if } interior \\ \text{high} & \text{if } exterior \end{cases}$

For  $\Phi(k_v) = \begin{cases} 0 & \text{if } interior \\ 0 & \text{if } exterior \end{cases}$

For  $\Psi(k_v, k_w) = \begin{cases} 0 & \text{if } k_v = k_w \\ \text{low} & \text{otherwise} \end{cases}$

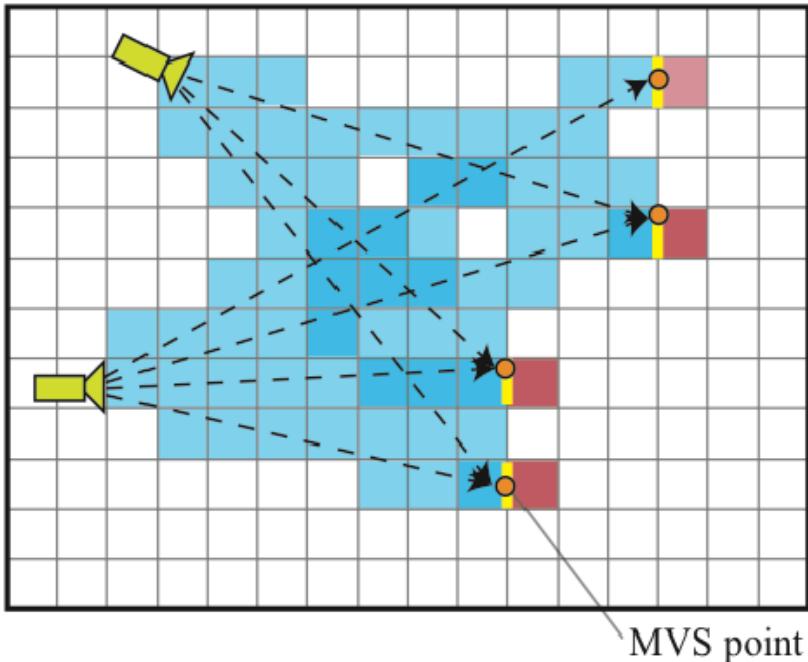
For  $\Psi(k_v, k_w) = \begin{cases} 0 & \text{if } k_v = k_w \\ \text{high} & \text{otherwise} \end{cases}$

$$E(\{k_v\}) = \sum_v \Phi(k_v) + \sum_{(v,w) \in N} \Psi(k_v, k_w)$$

↓      ↓  
 unary term      pairwise term

# 基于体素的MVS

离散空间的Labeling是一个典型的MRF(Markov Random Field)优化问题：



$$\Psi(k_v, k_w) = \begin{cases} 0 & \text{if } k_v = k_w \\ \rho(\frac{k_v + k_w}{2}) & \text{otherwise} \end{cases}$$

photo-consistency

$$\Phi(k_v) = \begin{cases} 0 & \text{if interior} \\ \alpha & \text{if exterior} \end{cases}$$

ballooning force

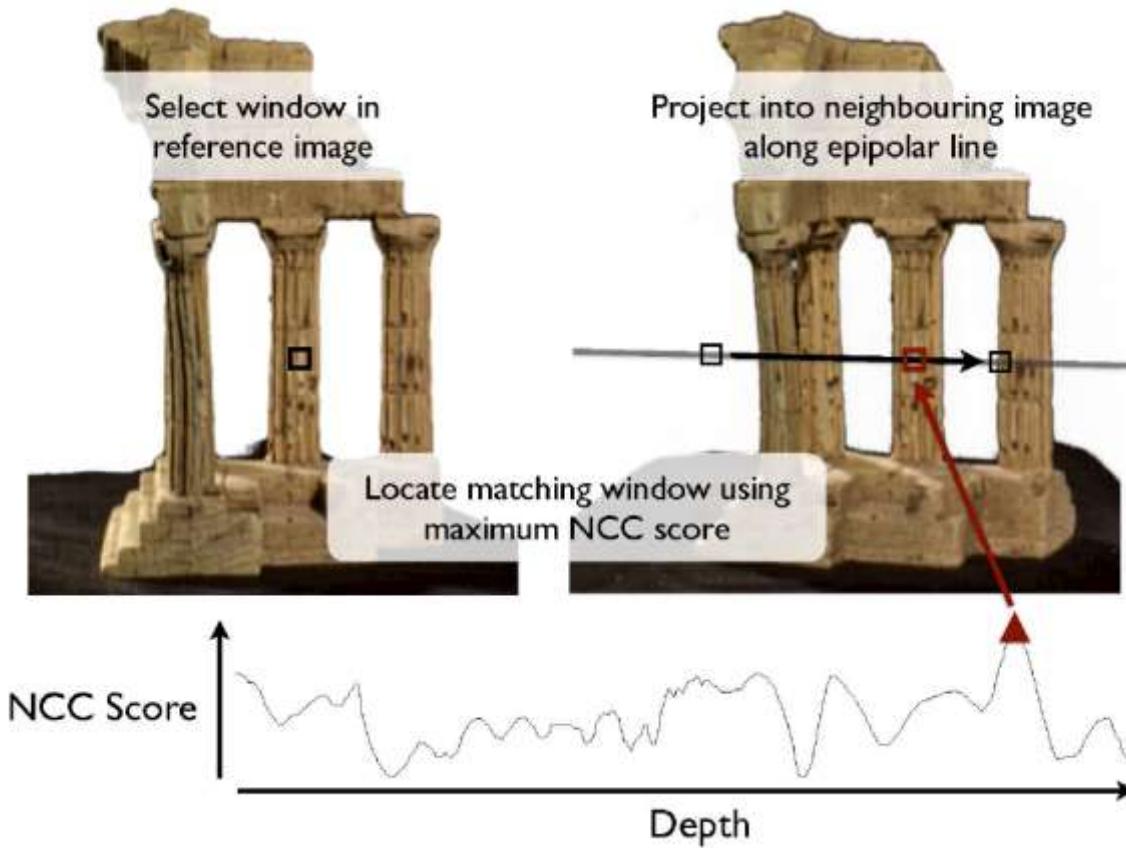
$$E(\{k_v\}) = \sum_v \Phi(k_v) + \sum_{(v,w) \in N} \Psi(k_v, k_w)$$

unary term

pairwise term

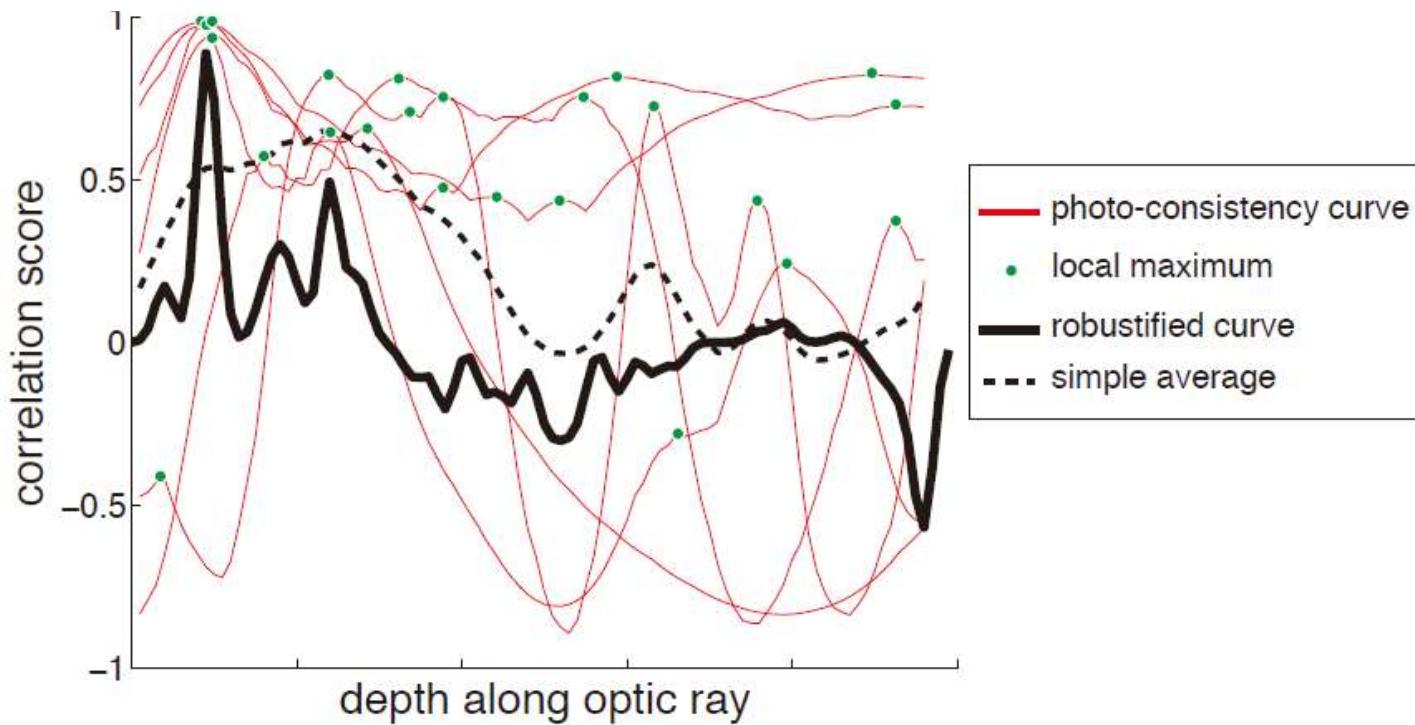
# 基于体素的MVS

两视图一致性的计算：Winner-Takes-All



# 基于体素的MVS

多视图一致性的计算：Robust Voting（寻找局部极值集中的点）



# 基于体素的MVS

MRF优化问题的求解：Graph-cuts

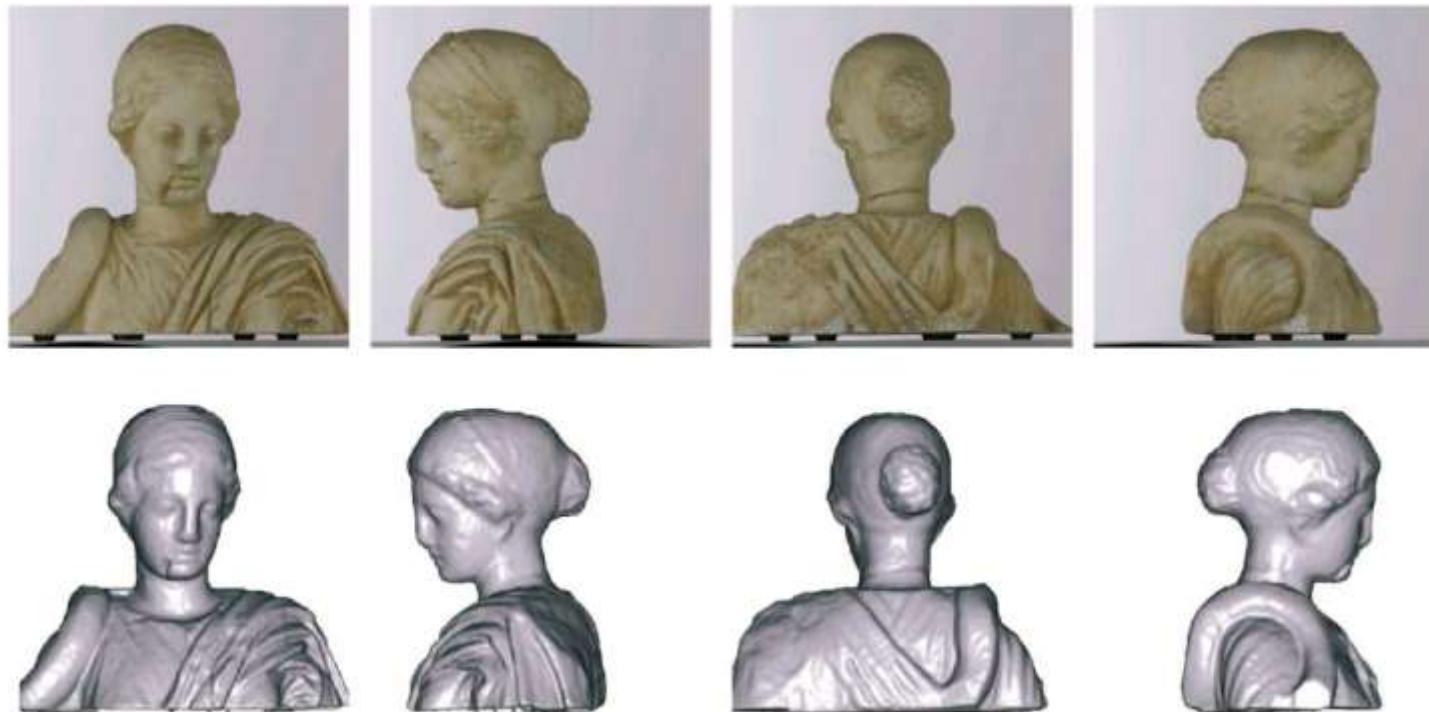
$$E(\{k_v\}) = \sum_v \Phi(k_v) + \sum_{(v,w) \in N} \Psi(k_v, k_w)$$

↑  
unary term                              ↓  
pairwise term

- 只有两个label (interior and exterior)
- $\Psi(k_v, k_w)$  为sub-modular :  $\Psi(\text{interior}, \text{interior}) + \Psi(\text{exterior}, \text{exterior}) \leq \Psi(\text{interior}, \text{exterior}) + \Psi(\text{exterior}, \text{interior})$
- Graph-cuts可以有效求解

# 基于体素的MVS

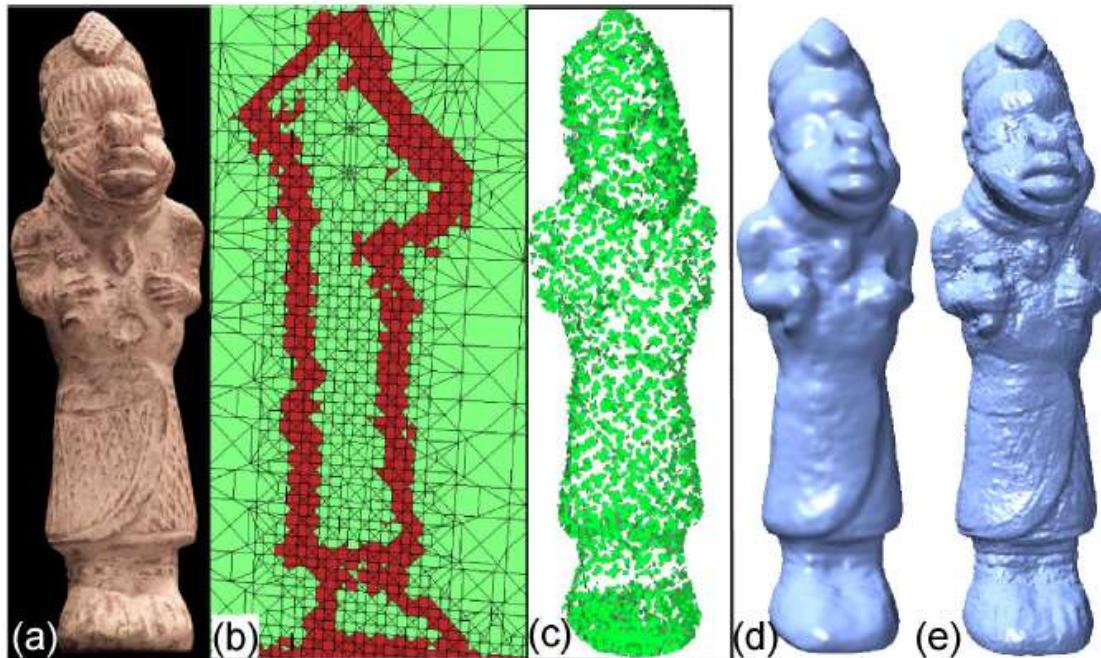
基于体素的MVS重建结果：



200×200×200 voxels

# 基于体素的MVS

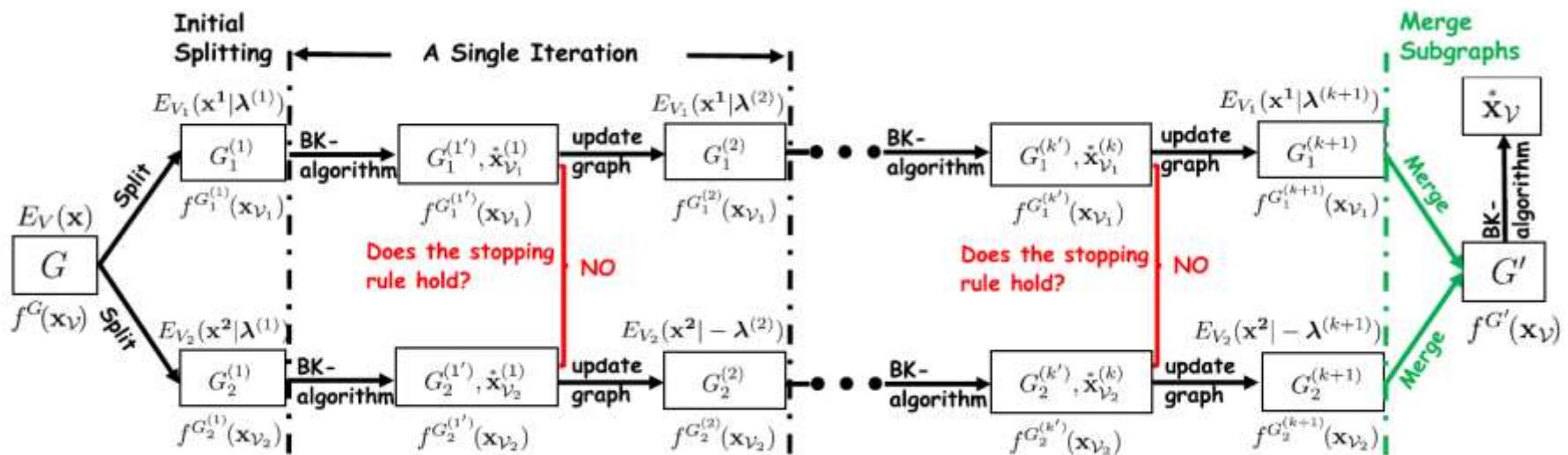
如何处理大场景？（如何处理海量高分辨率Voxel）



自适应多分辨率网格（物体表面高分辨率、其他区域低分辨率）

# 基于体素的MVS

如何处理大场景？（如何处理海量高分辨率Voxel）



并行、分布式Graph-cuts算法



# 基于体素的MVS

基于体素的MVS算法优缺点：

Pros :

- 生成规则点云
- 易于提取Mesh (Marching cube algorithm)

Cons :

- 精度取决于voxel粒度
- 难以处理大场景

# 基于特征点扩散的MVS

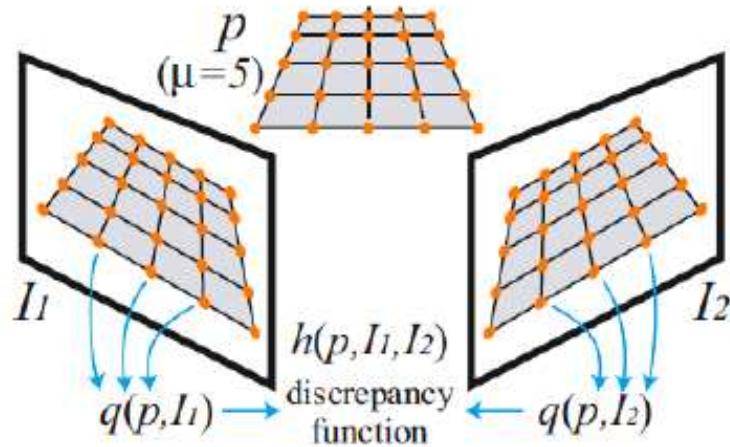
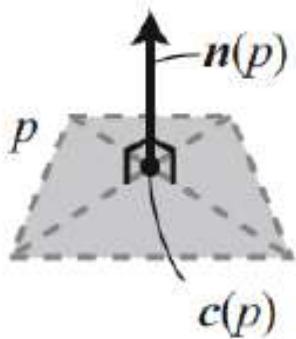


## Patch-based Multi-view Stereo (PMVS)

Open source: <http://www.di.ens.fr/pmvs/>

# 基于特征点扩散的MVS

3D点的patch表达：

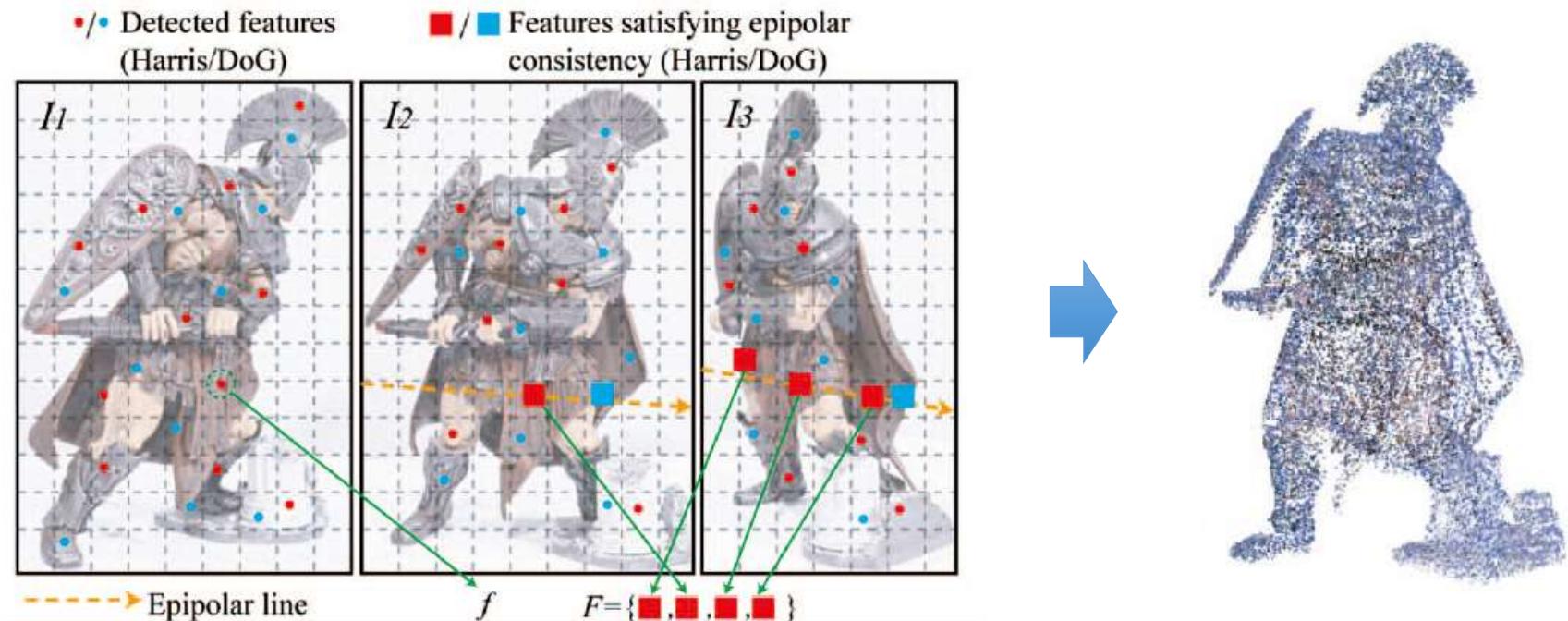


中心 $c(p)$ , 法向 $n(p)$

patch在图像上的投影

# 基于特征点扩散的MVS

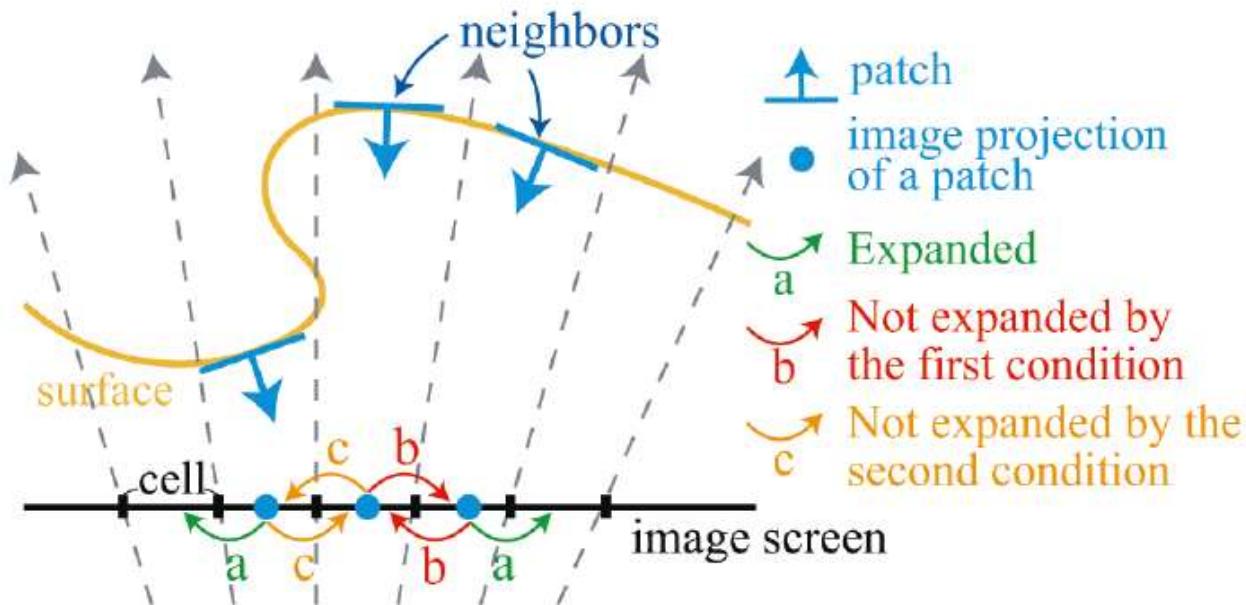
Step 1: 生成初始点云



检测、匹配并三角化图像中的Harris与DoG点

# 基于特征点扩散的MVS

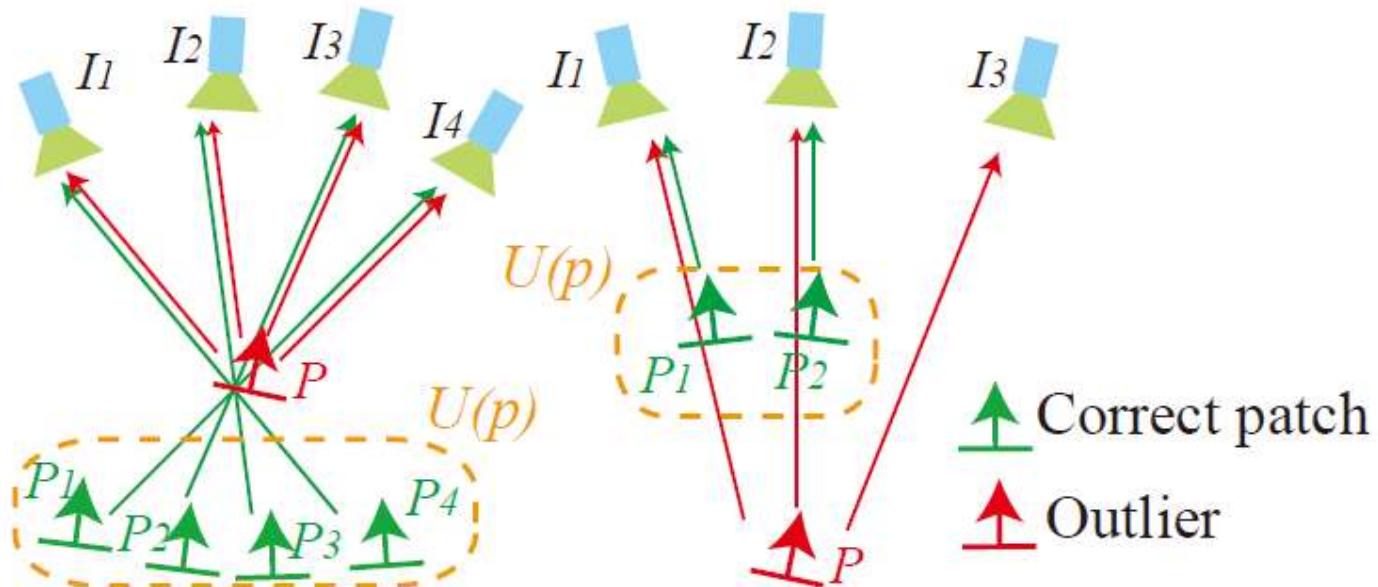
## Step 2: 点云扩散 (Expansion)



3D点投影到图像，并向投影点周边的空白区域扩散

# 基于特征点扩散的MVS

Step 3: 点云过滤 (Filtering)



去除深度值不一致且一致性较低的点

# 基于特征点扩散的MVS

基于特征点扩散的MVS重建结果：





# 基于特征点扩散的MVS

基于特征点扩散的MVS算法优缺点：

Pros :

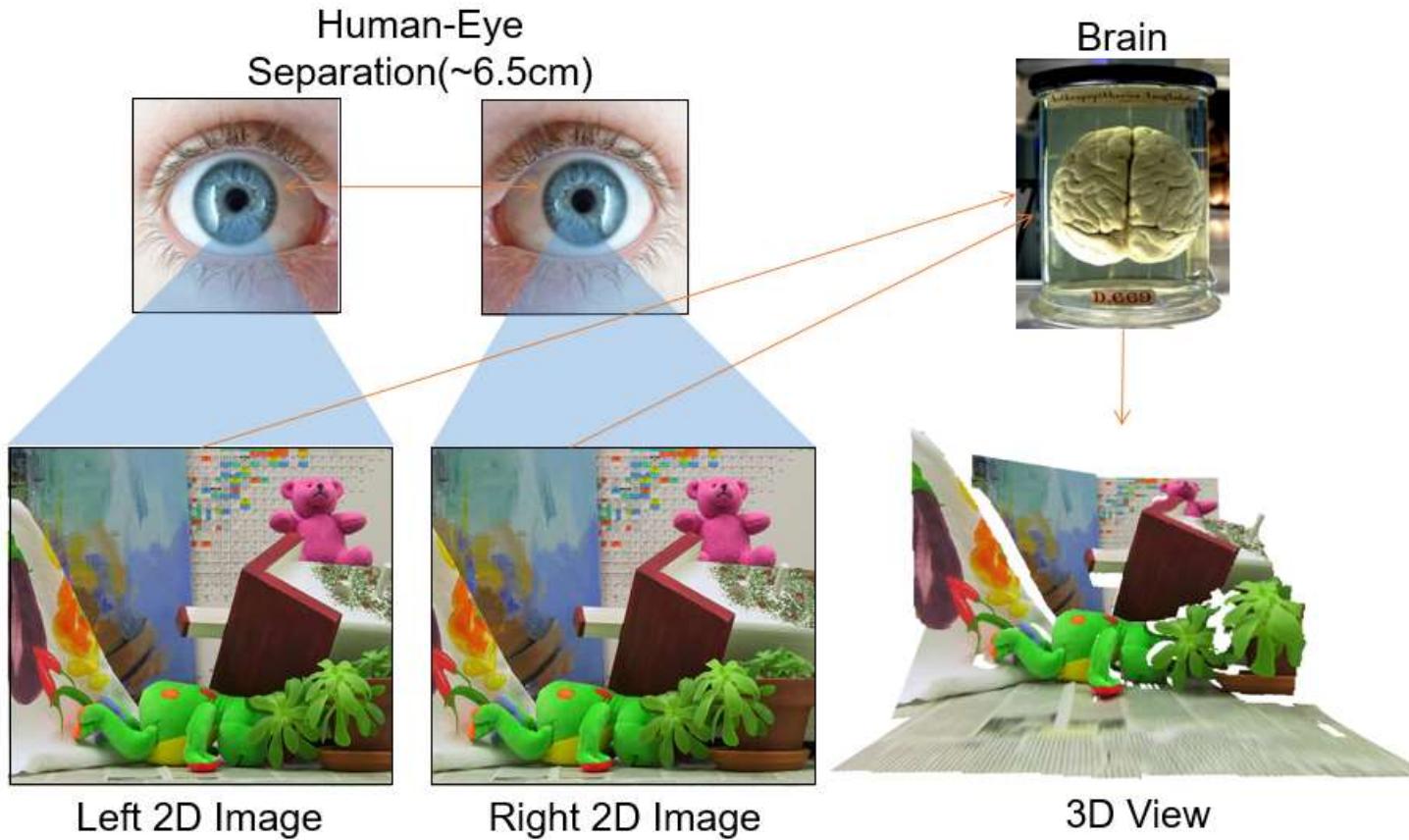
- 点云精度较高
- 点云分布均匀

Cons :

- 弱纹理区域造成扩散空洞
- 需要一次读入所有图像 (改进 : Clustering Views for PMVS, CMVS)

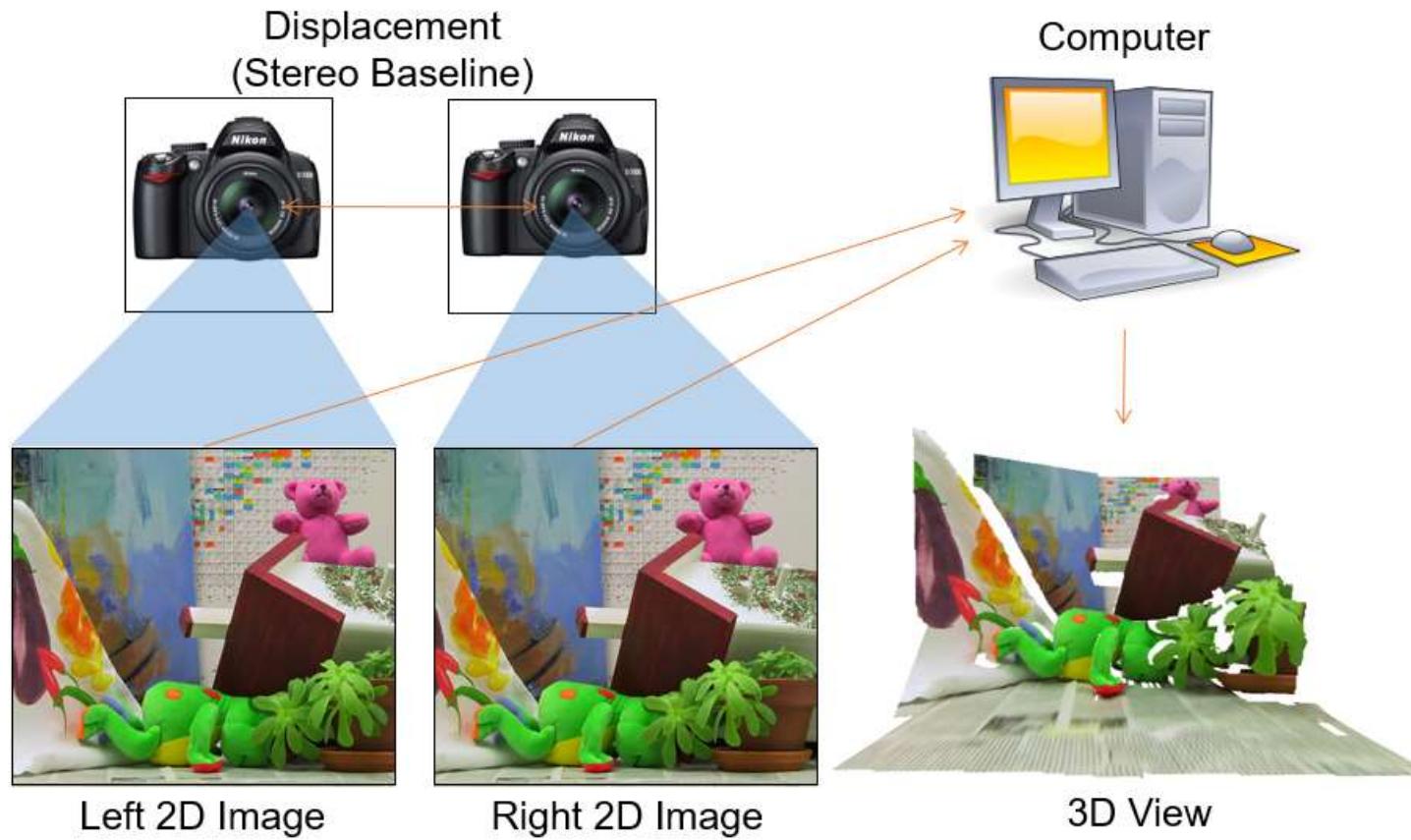
# 基于深度图融合的MVS

立体视觉与深度图：



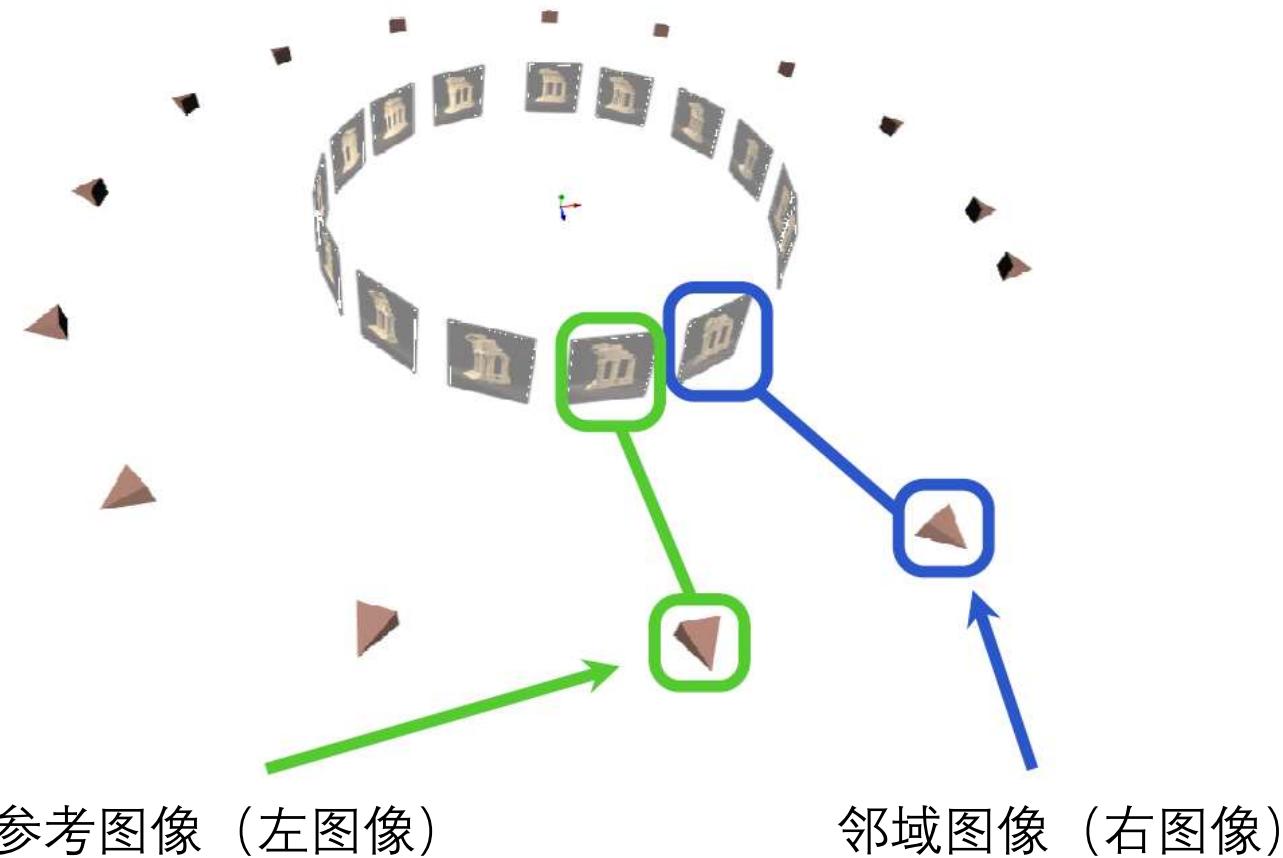
# 基于深度图融合的MVS

立体视觉与深度图：



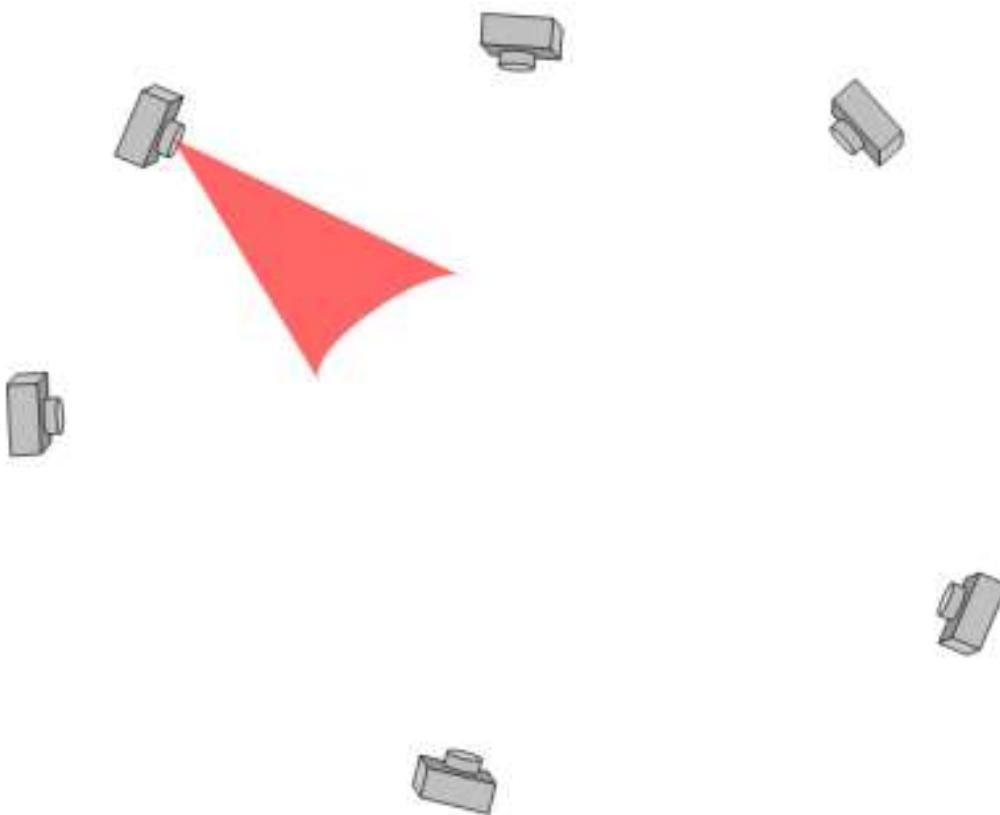
# 基于深度图融合的MVS

Step 1: 为每一幅图像选择邻域图像构成立体图像对(组)



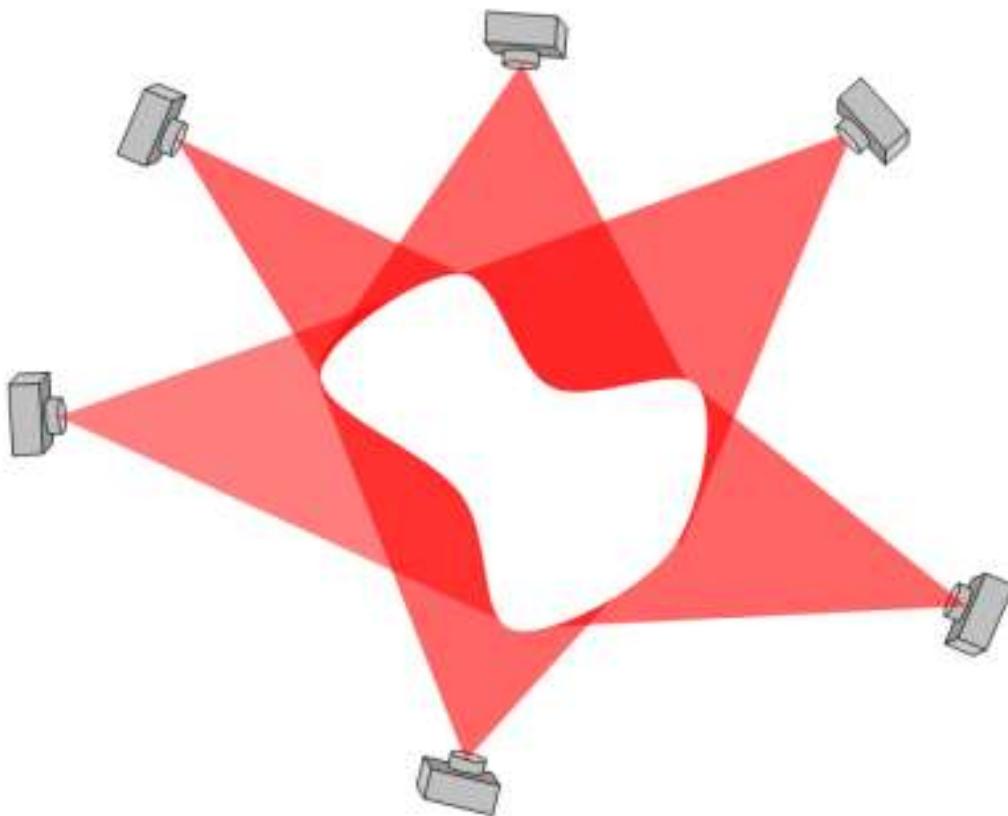
# 基于深度图融合的MVS

Step 2: 计算每一幅图像的深度图



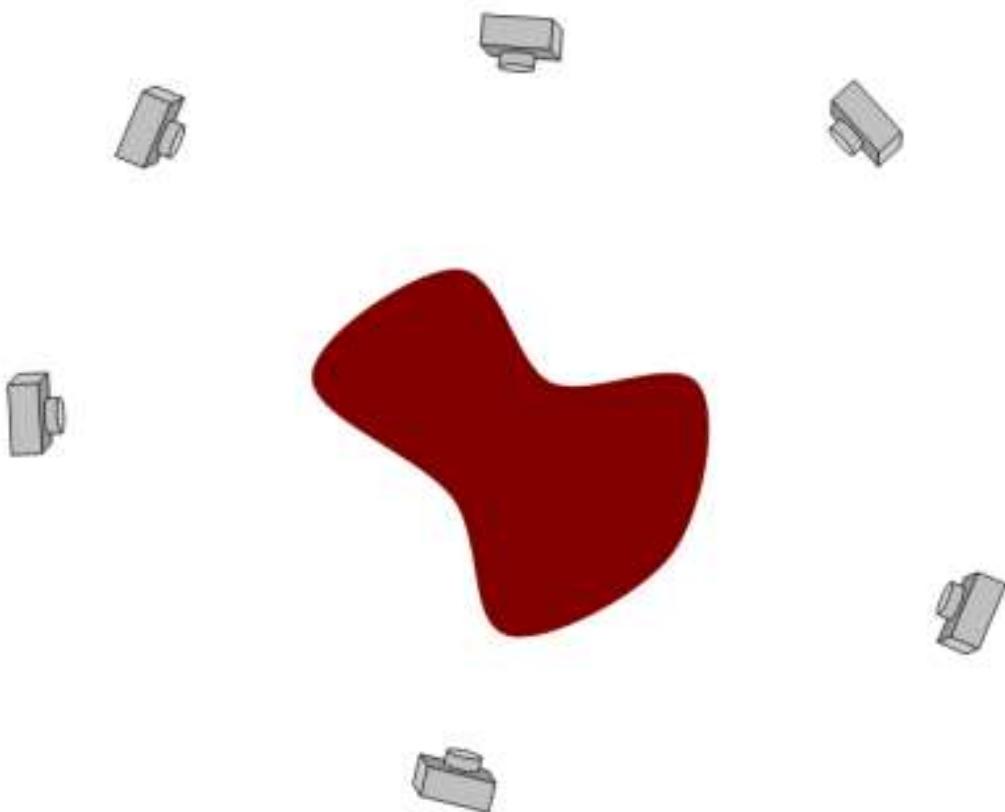
# 基于深度图融合的MVS

Step 3: 深度图融合



# 基于深度图融合的MVS

Step 4: 抽取物体表面





# 基于深度图融合的MVS

基本流程：

Step 1: 为每一幅图像选择邻域图像构成立体图像对(组)

Step 2: 计算每一幅图像的深度图

Step 3: 深度图融合

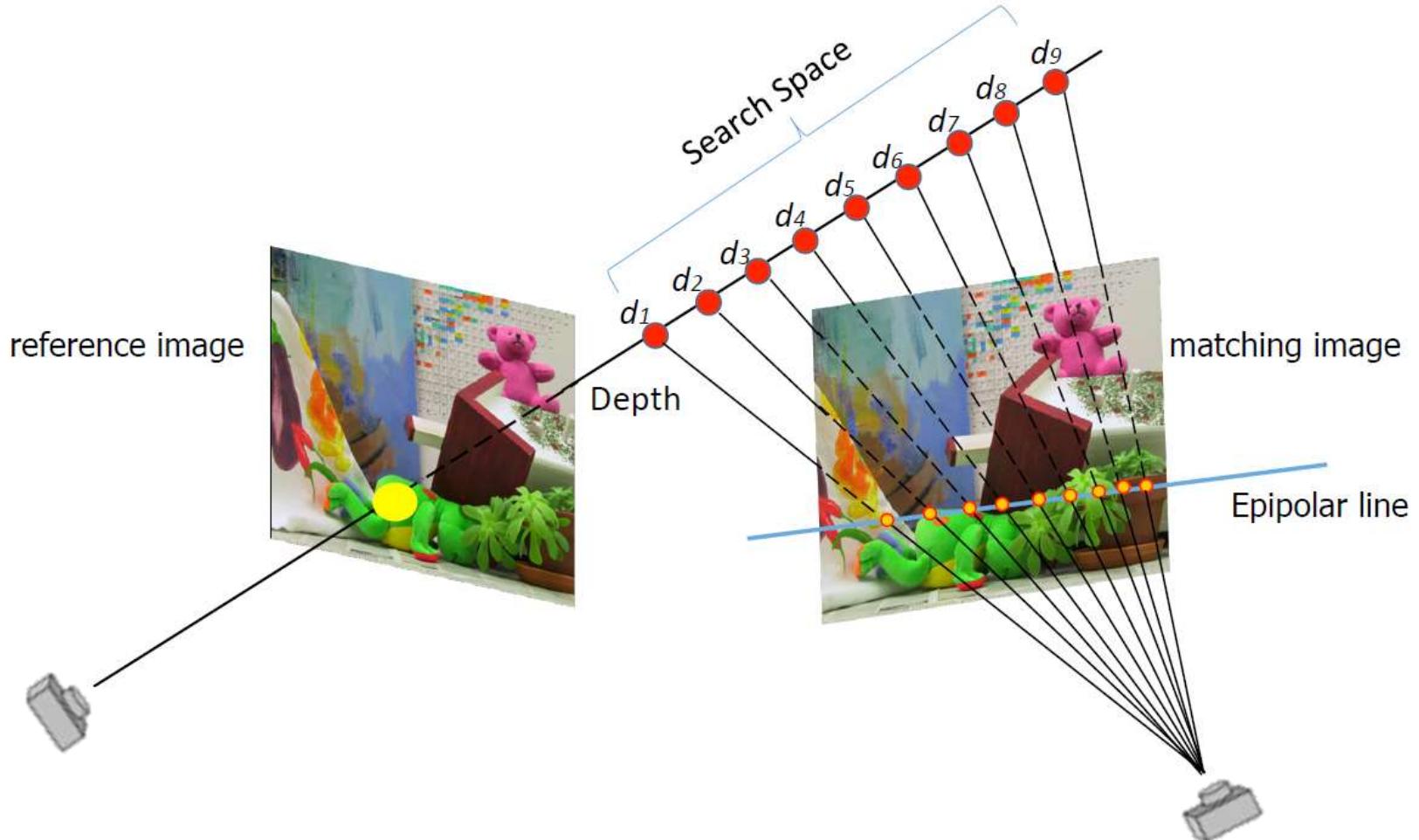
Step 4: 抽取物体表面

关键步骤1：如何  
选择邻域图像组？

关键步骤2：如何  
计算深度图？

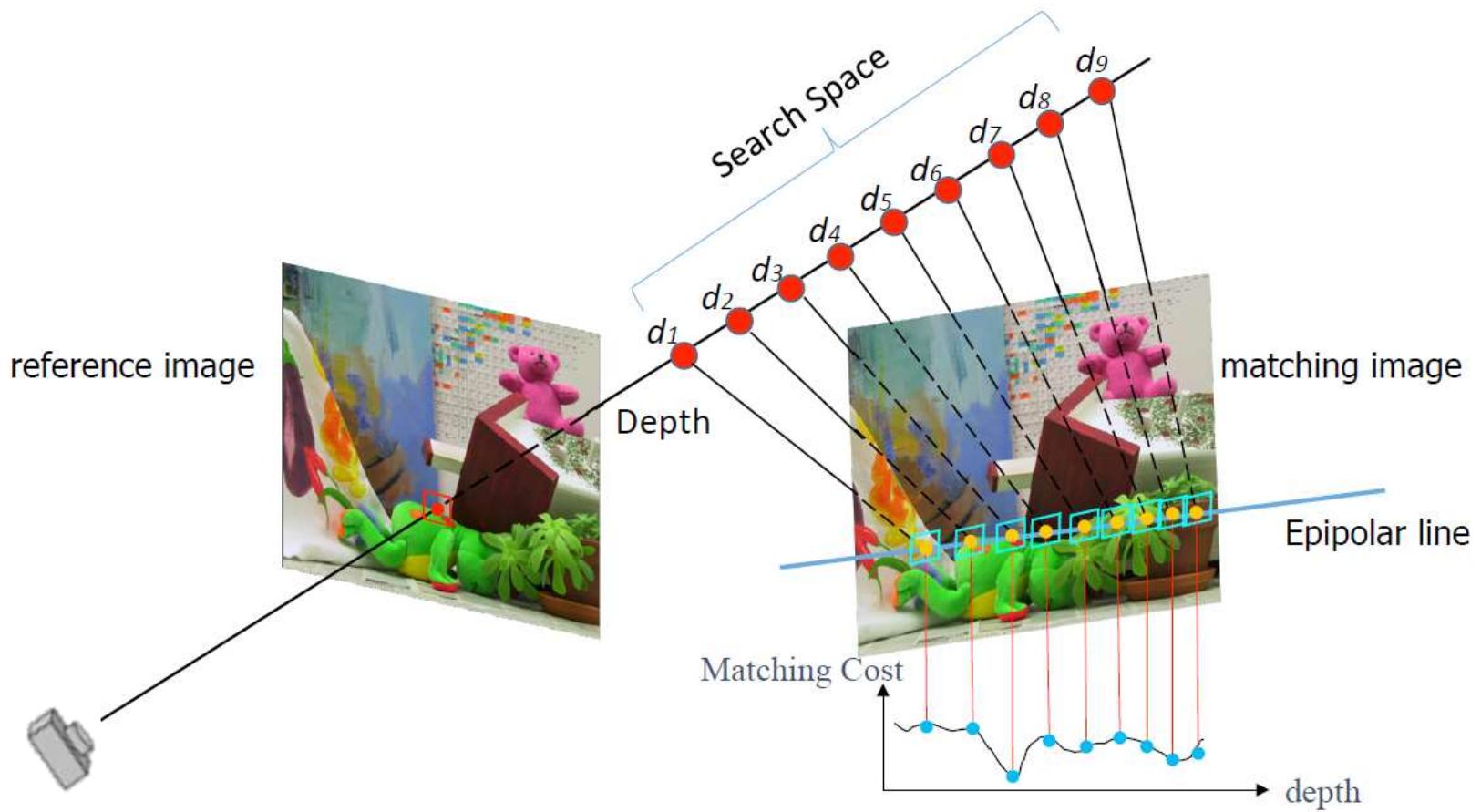
# 基于深度图融合的MVS

每一个像素上深度值的计算：



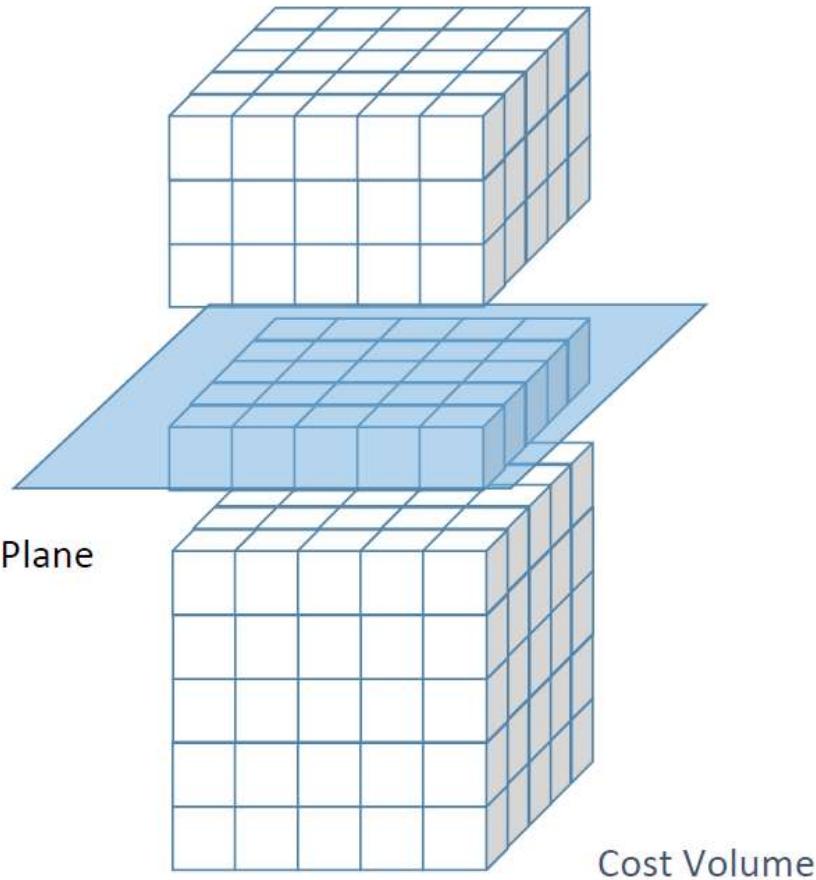
# 基于深度图融合的MVS

每一个像素上深度值的计算：



# 基于深度图融合的MVS

聚合(Cost aggregation)：

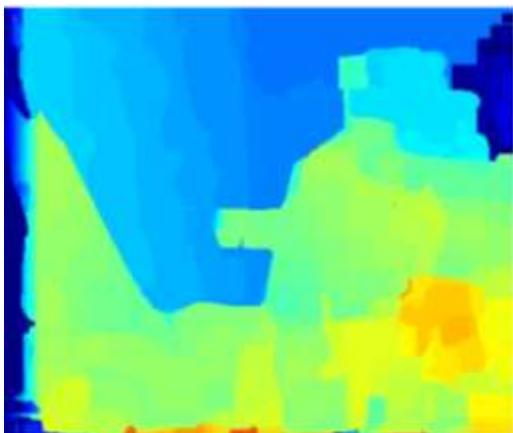


# 基于深度图融合的MVS

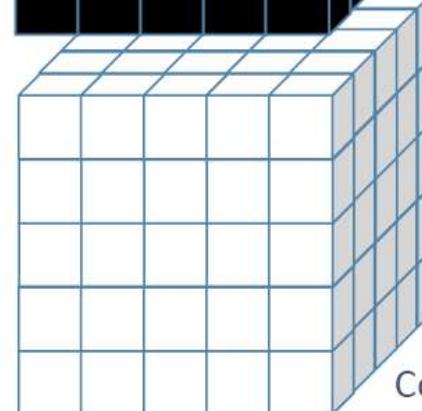
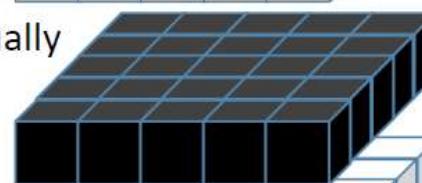
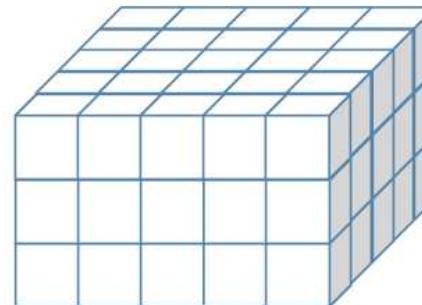
聚合(Cost aggregation) : SAD



Treat neighbors equally



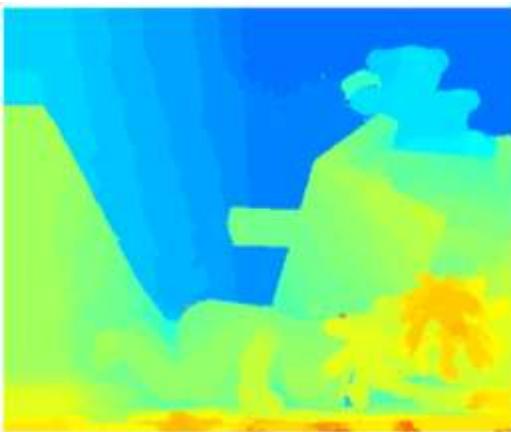
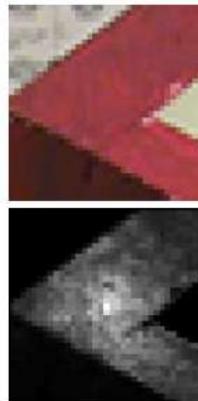
Sum of Absolute  
Differences (SAD)



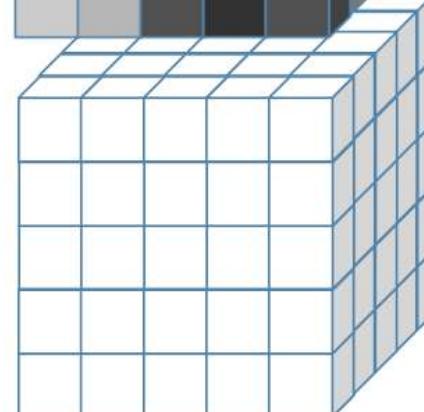
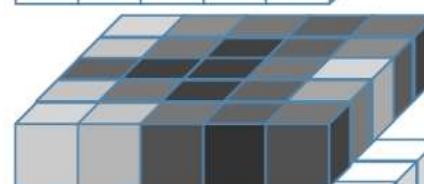
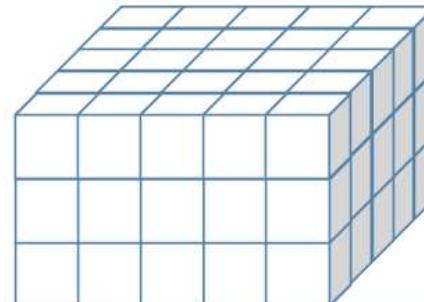
Cost Volume

# 基于深度图融合的MVS

聚合(Cost aggregation) : Adaptive weight

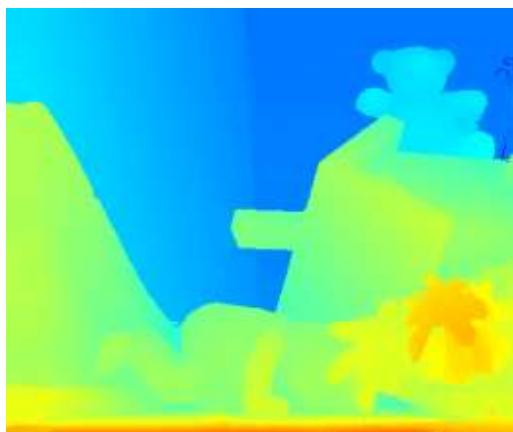


**Adaptive weight**  
• Color differences  
• Spatial distances

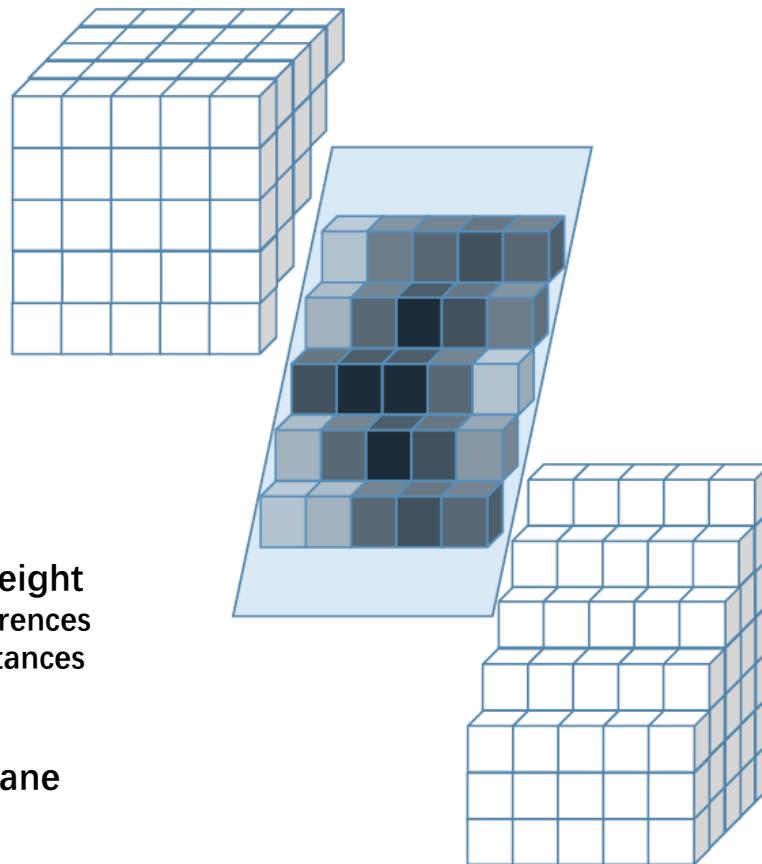


# 基于深度图融合的MVS

聚合(Cost aggregation) : Adaptive weight + Oriented plane

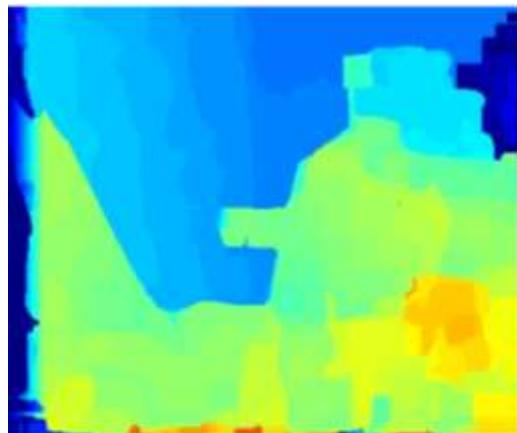


Adaptive weight  
• Color differences  
• Spatial distances  
+  
Oriented plane

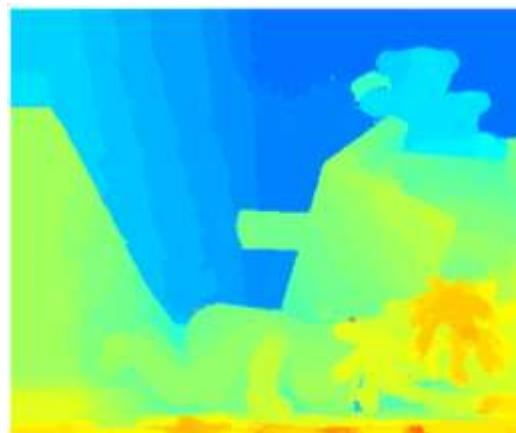


# 基于深度图融合的MVS

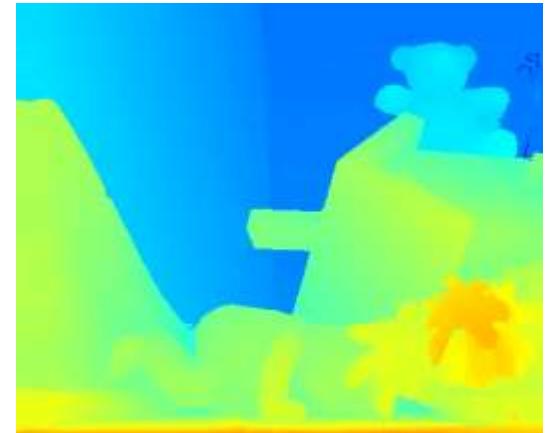
不同聚合(Cost aggregation)方法的比较：



SAD



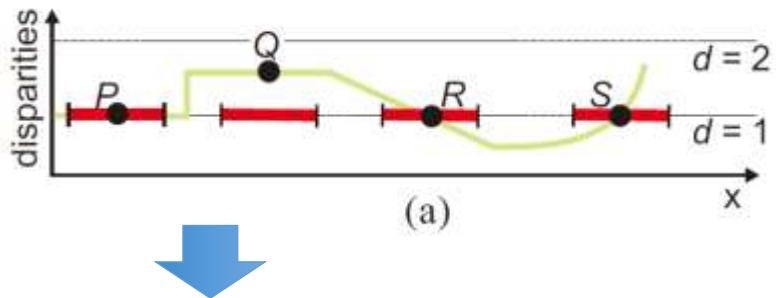
Adaptive weight



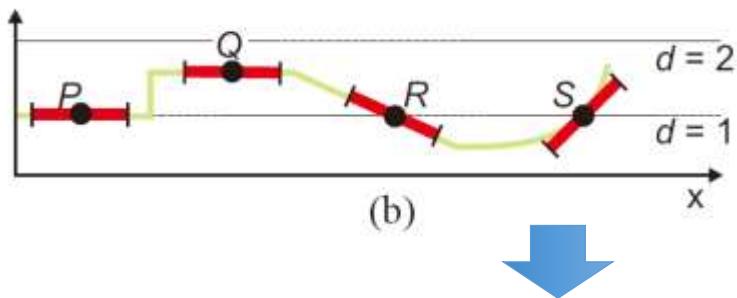
Adaptive weight  
+  
Oriented plane

# 基于深度图融合的MVS

空间平面方向的估计：PathMatch



(a)



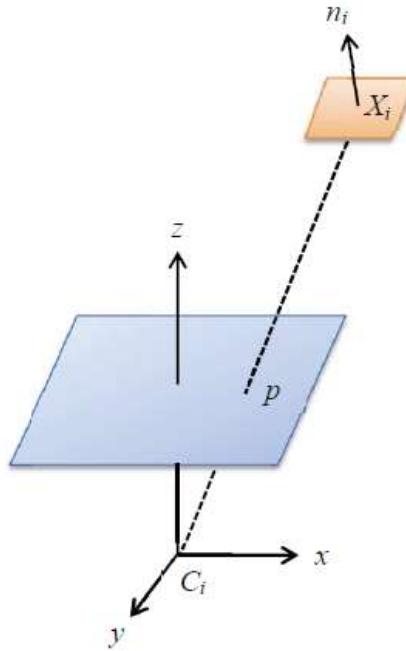
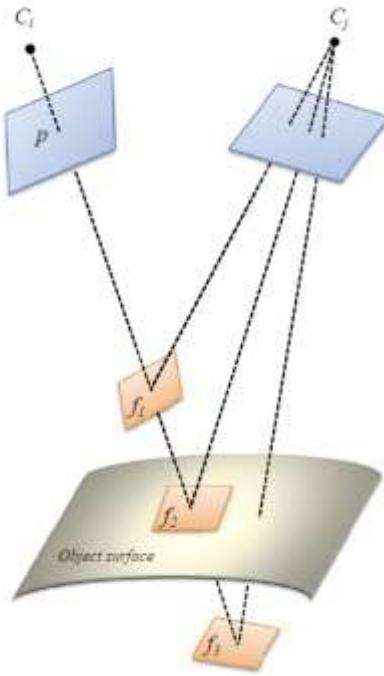
(b)



C. Barnes, et. al. Patchmatch: A randomized correspondence algorithm for structural image editing. *SIGGRAPH* 2009.  
M. Bleyer, et. al. Patchmatch stereo - stereo matching with slanted support windows. *BMVC* 2011.

# 基于深度图融合的MVS

空间平面方向的估计：PathMatch

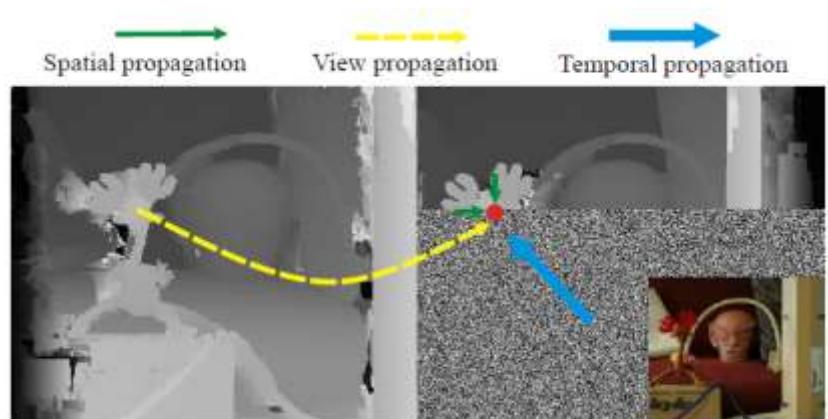


相机 $C_i$ 坐标系下空间面片(Patch)的表达： $d$  (1自由度),  $n$  (2个自由度)

# 基于深度图融合的MVS

两视图PathMatch Stereo :

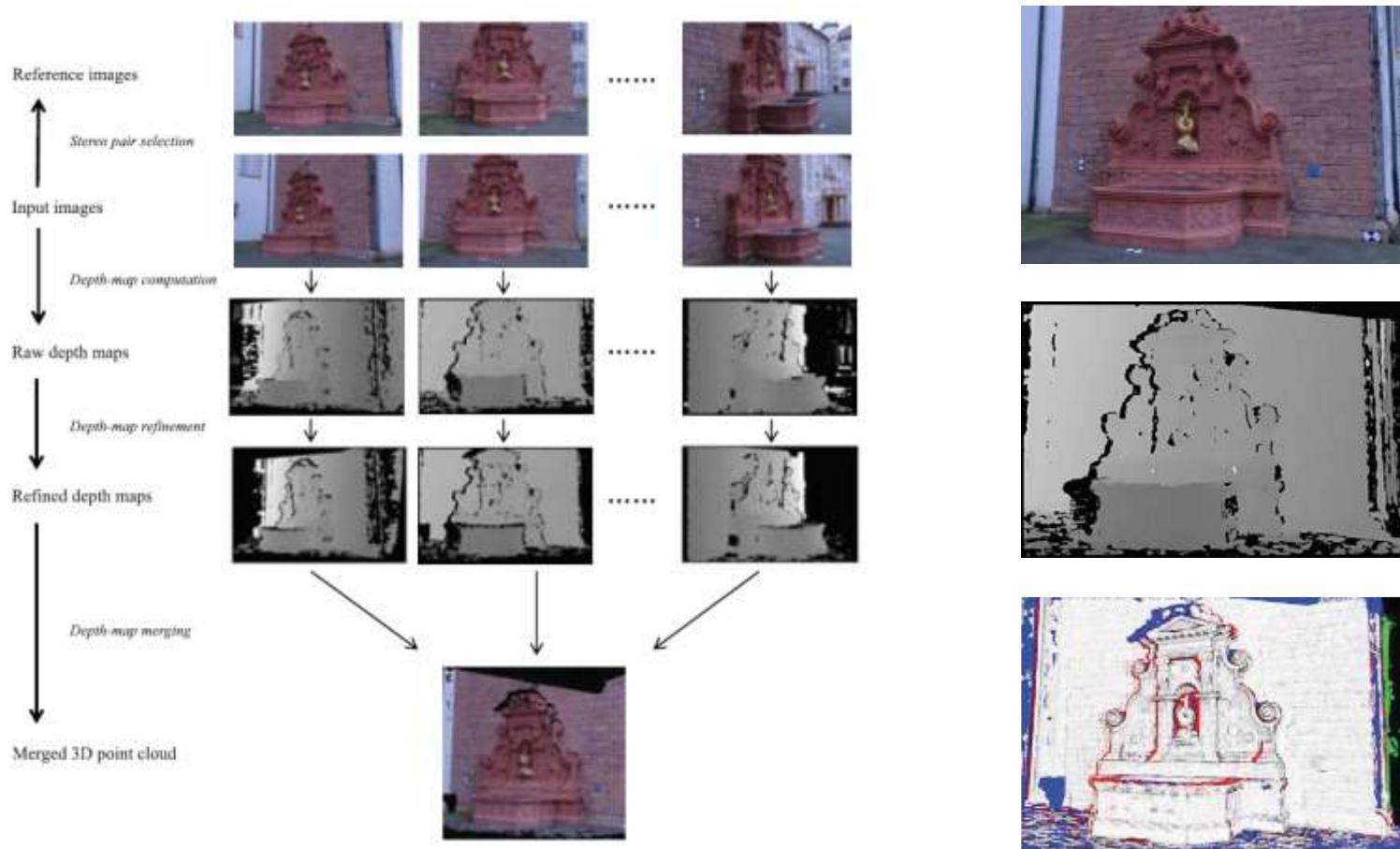
1. 随机生成每个像素点的深度和法向
2. 从左上向右下传播：
  - 1) 检测邻域点的深度和法向是否更好
  - 2) 检测自身随机扰动后的点是否更好
  - 3) 检测立体图像对对应点是否更好
  - 4) 检测前后帧同一位置点是否够更好
3. 从右下向左上再传播一次



Based on the law of large numbers !

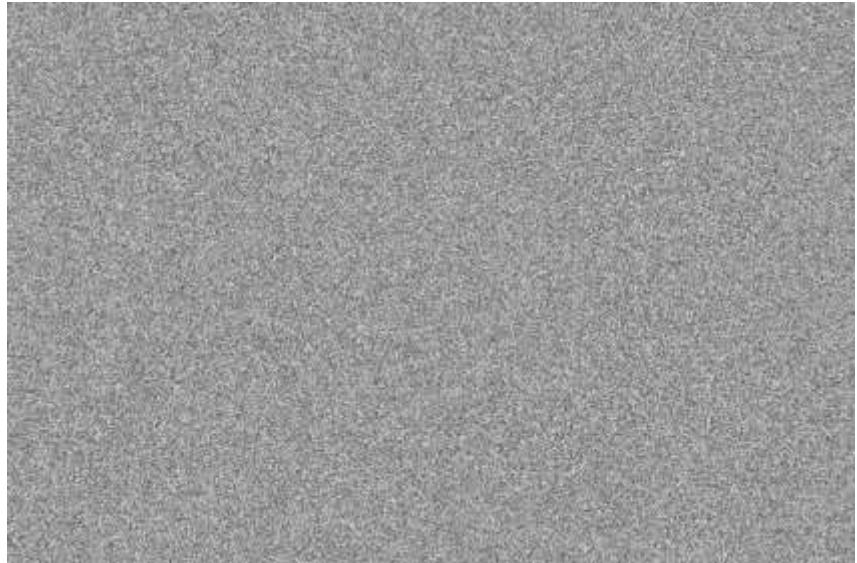
# 基于深度图融合的MVS

多视图PathMatch MVS：



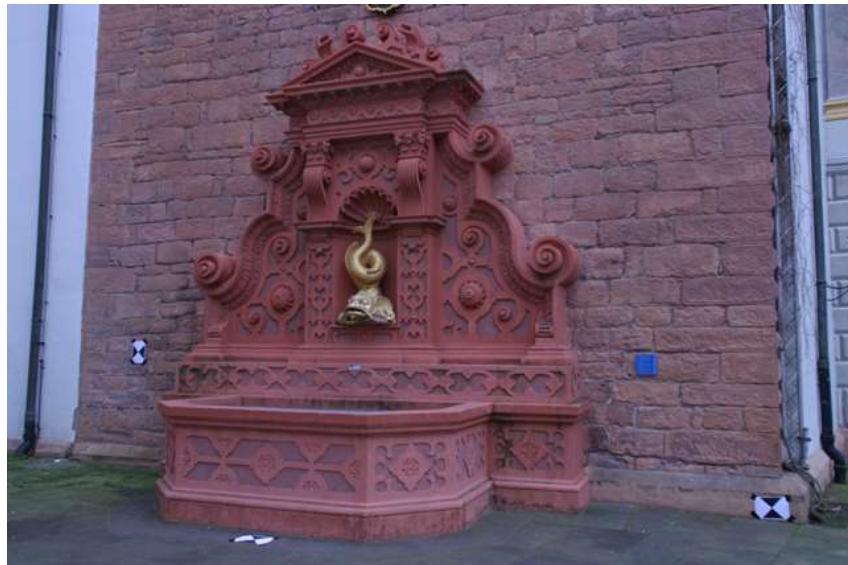
# 基于深度图融合的MVS

多视图PathMatch MVS：



# 基于深度图融合的MVS

多视图PathMatch MVS：



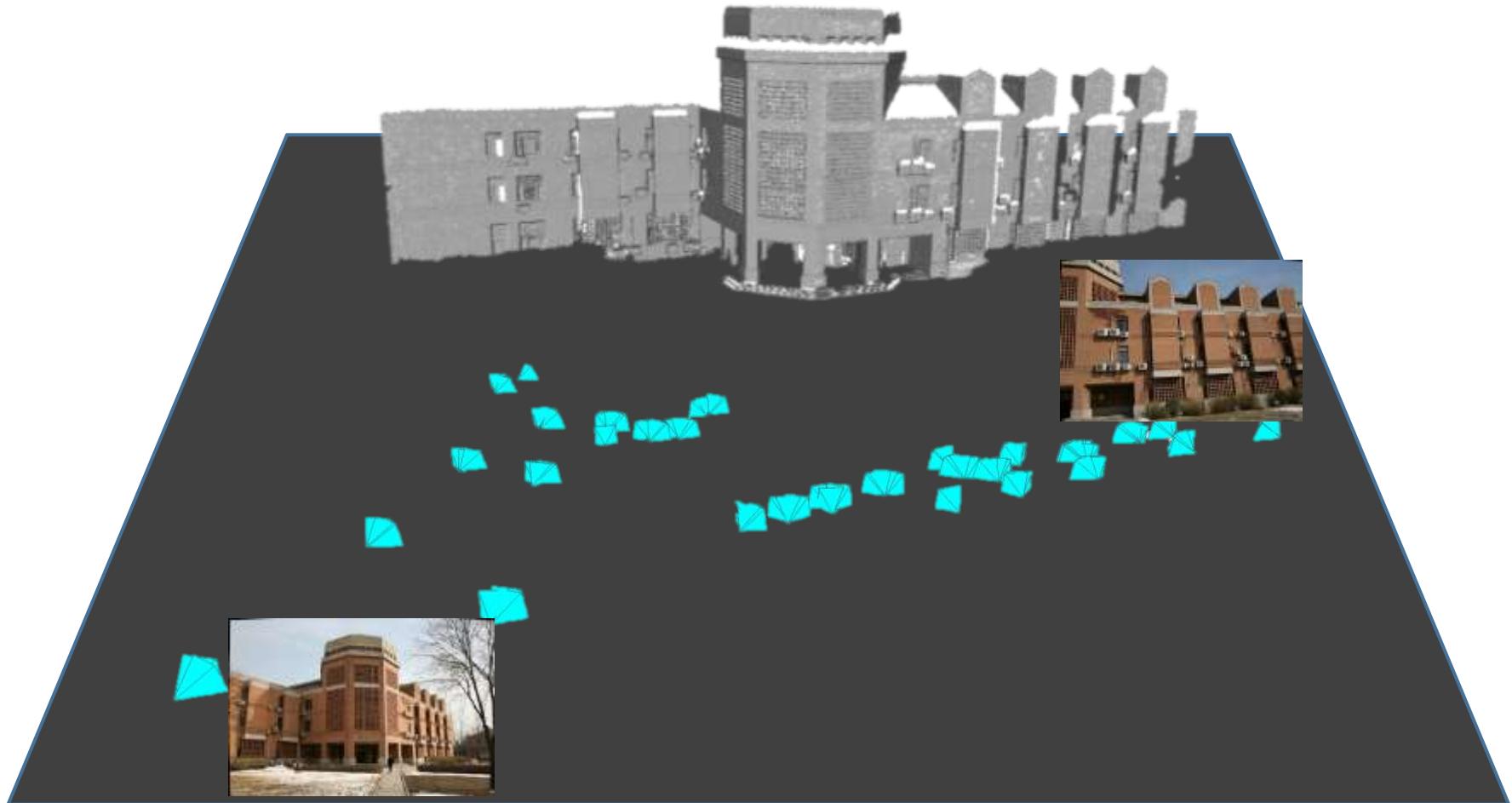
# 基于深度图融合的MVS

多视图PathMatch MVS：

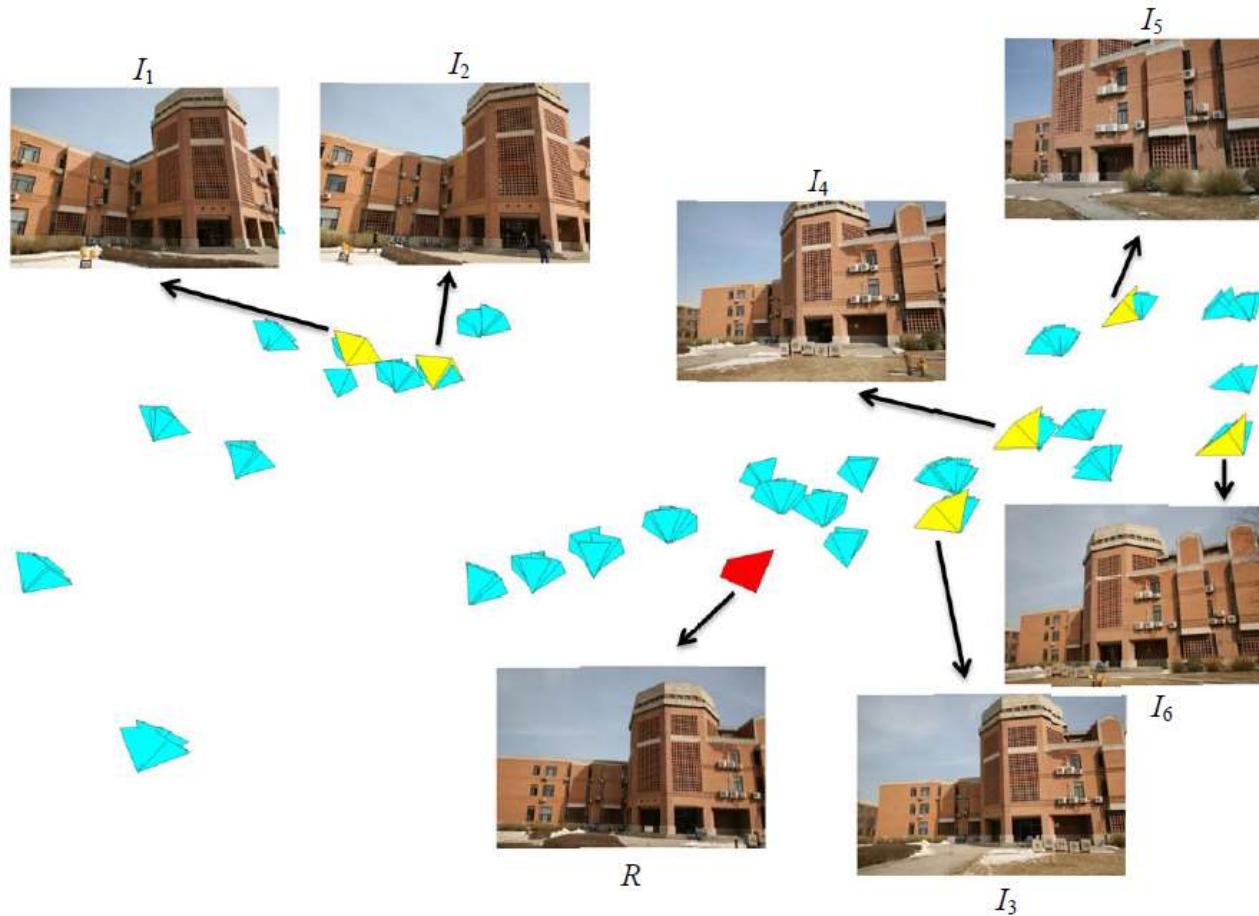


# 基于深度图融合的MVS

多视图PathMatch MVS中邻域图像组的选择：



# 基于深度图融合的MVS



邻域图像组选择准则：视线夹角、物距、覆盖度、分散度等；（通过稀疏点云计算）  
 邻域图像组选择：0-1 Knapsack Problem。（NP-hard，通过QEA求解）

# 基于深度图融合的MVS

逐像素点选择邻域图像组：



E. Zheng, et. al. Patchmatch based joint view selection and depthmap estimation, *CVPR* 2014.  
J. L. Schönberger, et. al. Pixelwise View Selection for Unstructured Multi-View Stereo, *ECCV* 2016.

# 基于深度图融合的MVS

逐像素点选择邻域图像组：



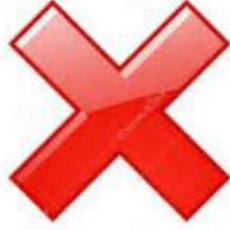
# 基于深度图融合的MVS

逐像素点选择邻域图像组：



# 基于深度图融合的MVS

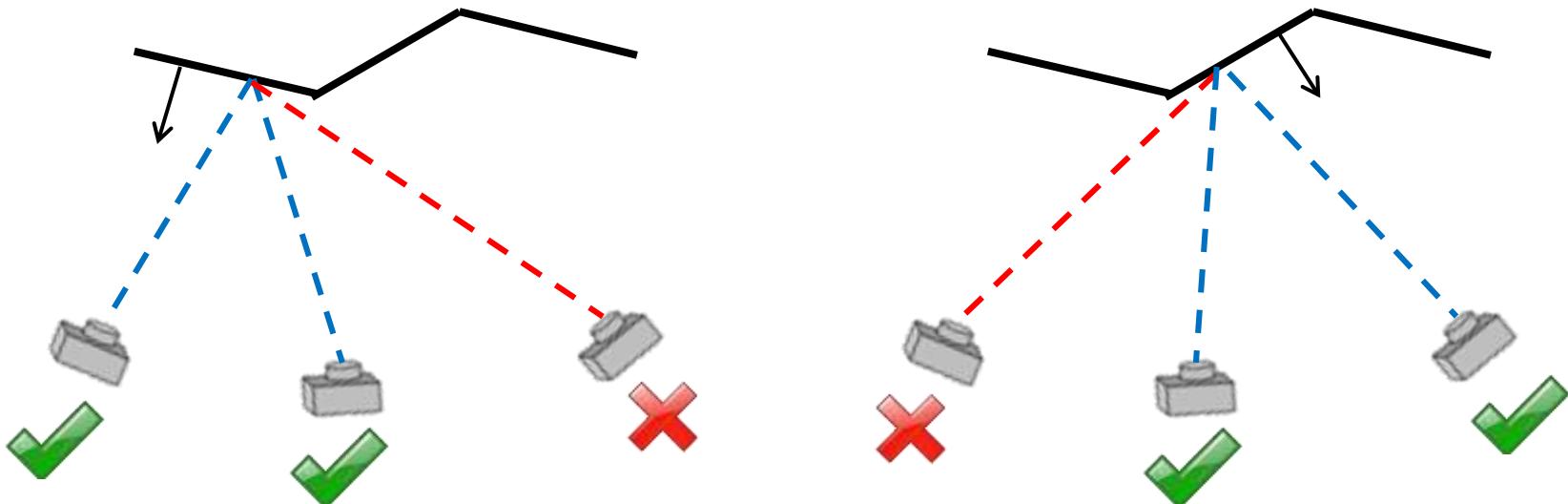
逐像素点选择邻域图像组：



# 基于深度图融合的MVS

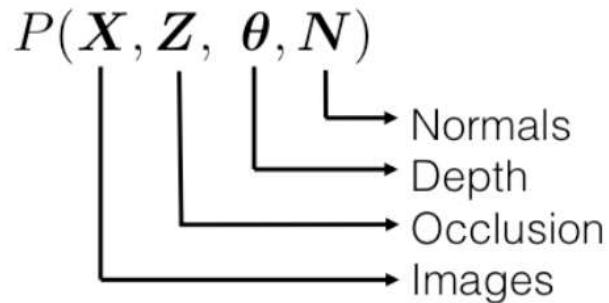
逐像素点选择邻域图像组：

- 像素可视性需要场景结构、场景结构需要像素可视性
- This is a chicken-and-egg problem



# 基于深度图融合的MVS

逐像素点选择邻域图像组：



Generalized Expectation Maximization :

- E-Step  
使用变分推断  $Z$
- M-Step  
使用PatchMatch推断  $\theta, N$

E-Step  
选择可视图像

M-Step  
计算深度和法向



# 基于深度图融合的MVS

基于深度图融合的MVS算法优缺点：

Pros :

- 适用于大场景海量图像(并行计算深度图)
- 得到的点云数量多(冗余)
- 目前的开源和商用软件基本都采用这类方法

Cons :

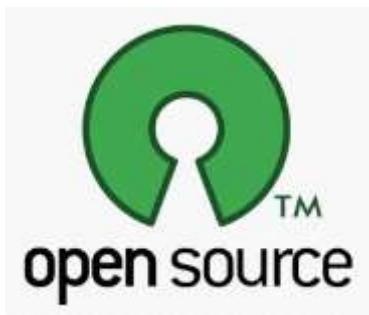
- 很大程度上依赖于邻域图像组的选择



# 稠密重建(MVS)总结

- 稠密重建的目的：
  - 在SfM基础上(相机内外参数已知), 计算每个像素点对应的3D点
- 稠密重建三类基本方法：
  - 基于体素的方法
  - 基于点云扩散的方法
  - 基于深度图融合的方法
- 稠密重建的后续操作：
  - 三角网格建模 (得到mesh模型)
  - LOD建模 (得到标准城市模型)
  - BIM建模 (得到标准建筑物模型)

# 4. 资源简介





# 提 纲

## 1、简介（申抒含，20分钟）

- 图像三维重建基本流程
- 多视几何基础

## 2、稀疏重建（崔海楠，40分钟）

- 增量式稀疏重建
- 全局式稀疏重建
- 混合式稀疏重建

## 3、稠密重建（申抒含，40分钟）

- 基于体素的稠密重建
- 基于特征点扩散的稠密重建
- 基于深度图融合的稠密重建

## 4、资源简介（申抒含，20分钟）

- 三维重建开源与商业软件
- 三维重建评测数据库



# 开源与商业软件

开源SfM软件：

Project	Language	License
<a href="#"><u>Bundler</u></a>	C++	GNU General Public License - contamination
<a href="#"><u>Colmap</u></a>	C++	GNU General Public License - contamination
<a href="#"><u>MAP-Tk</u></a>	C++	BSD 3-Clause license - Permissive
<a href="#"><u>MICMAC</u></a>	C++	CeCILL-B
<a href="#"><u>MVE</u></a>	C++	BSD 3-Clause license + parts under the GPL 3 license
<a href="#"><u>OpenMVG</u></a>	C++	MPL2 - Permissive
<a href="#"><u>OpenSfM</u></a>	Python	Simplified BSD license - Permissive
<a href="#"><u>TheiaSfM</u></a>	C++	New BSD license - Permissive



# 开源与商业软件

开源MVS软件：

Project	Language	License
<a href="#"><b>Colmap</b></a>	C++ CUDA	GNU General Public License - contamination
<a href="#"><b>GPUIma + fusibile</b></a>	C++ CUDA	GNU General Public License - contamination
<a href="#"><b>HPMVS</b></a>	C++	GNU General Public License - contamination
<a href="#"><b>MICMAC</b></a>	C++	CeCILL-B
<a href="#"><b>MVE</b></a>	C++	BSD 3-Clause license + parts under the GPL 3 license
<a href="#"><b>OpenMVS</b></a>	C++ CUDA	AGPL3
<a href="#"><b>PMVS</b></a>	C++ CUDA	GNU General Public License - contamination



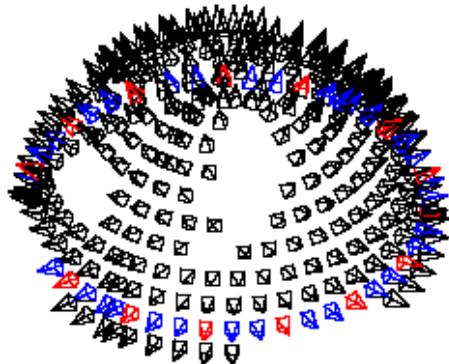
# 开源与商业软件

商业软件 (SfM+MVS+Modeling) :

<b>Software</b>	<b>Company</b>
ContextCapture	Bentley Acute3D
PhotoMesh	SkyLine
StreetFactory	AirBus
PhotoScan	AgiSoft
Pix4DMapper	Pix4D
RealityCapture	RealityCapture

# 评测数据库

小场景数据集：



Middlebury MVS Benchmark



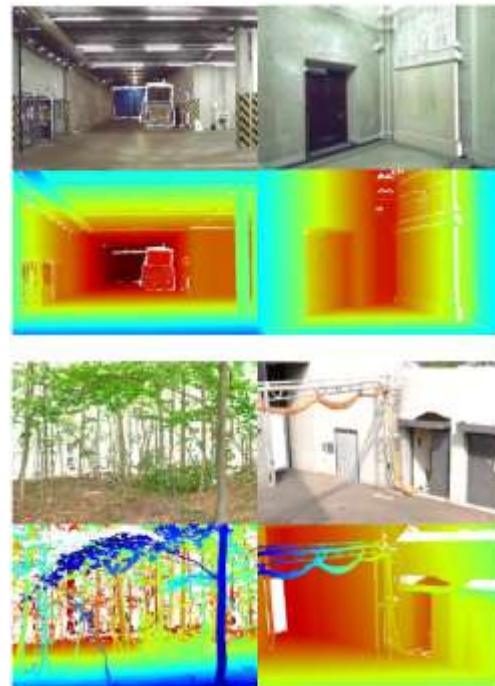
DTU MVS Benchmark

# 评测数据库

大场景数据集：



Tanks and Temples Benchmark



ETH3D Benchmark

# 评测数据库

State-of-the-art :



Intermediate F-score

method	rank ▲	mean	Family	Francis	Horse	Lighthouse	M60	Panther	Playground	Train
COLMAP	<b>2.38</b>	42.14	50.41	22.25	25.63	<b>56.43</b>	44.83	<b>46.97</b>	<b>48.53</b>	<b>42.04</b>
OpenMVG + OpenMVS	2.50	41.71	58.86	<b>32.59</b>	26.25	43.12	44.73	46.85	45.97	35.27
Pix4D	2.50	<b>43.24</b>	<b>64.45</b>	31.91	<b>26.43</b>	54.41	<b>50.58</b>	35.37	47.78	34.96
OpenMVG + MVE	3.75	38.00	49.91	28.19	20.75	43.35	44.51	44.76	36.58	35.95
OpenMVG + SMVS	7.38	30.67	31.93	19.92	15.02	39.38	36.51	41.61	35.89	25.12
MVE	8.25	25.37	48.59	23.84	12.70	5.07	39.62	38.16	5.81	29.19
OpenMVG-G + OpenMVS	8.88	22.86	56.50	29.63	21.69	6.55	39.54	28.48	0.00	0.53
OpenMVG + PMVS	8.88	29.66	41.03	17.70	12.83	36.68	35.93	33.20	31.78	28.10
Theia-I + OpenMVS	9.12	27.93	48.11	19.38	20.66	30.02	30.37	30.79	23.65	20.46
VisualSfM + PMVS	9.62	27.80	38.02	12.93	11.30	41.75	35.47	34.19	35.47	13.26
VisualSfM + OpenMVS	10.00	24.45	49.10	21.38	18.59	25.24	27.02	24.64	16.59	13.07



# 未来发展趋势

## 1. 从三维几何感知到三维语义感知

- Mesh to GIS (CityGML)
- Mesh to BIM (IFC)

## 2. DNN for 3D Reconstruction

- 将语义信息纳入3D Reconstruction过程  
C. Häne, et. al. Dense Semantic 3D Reconstruction. *IEEE PAMI* 2017.
- End-to-End 3D Reconstruction  
H. Su, et. al. A Tutorial on 3D Deep Learning, *CVPR* 2017



# Thanks !

## Q & A