基于图像的三维建模

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
http://vision.ia.ac.cn

三维视觉的核心问题

场景的**三维结构**、场景的**语义信息**、相机的**空间位姿**

几何精确性、场景完整性、语义准确性、高效矢量表达

三维几何视觉的核心问题

三维几何视觉核心问题：场景结构+相机位姿+(相机参数)

途径一：Simultaneous Localization and Mapping (SLAM)

—视频序列

—重建场景稀疏/准稠密/稠密结构与相机位姿（on-line）

—闭环检测+图优化（on-line）

实时处理

途径二：Structure from Motion (SfM)

—多视角图像

—图像完全匹配（off-line）

—重建场景稀疏结构与相机位姿（off-line）

离散图像

三维几何视觉的核心问题

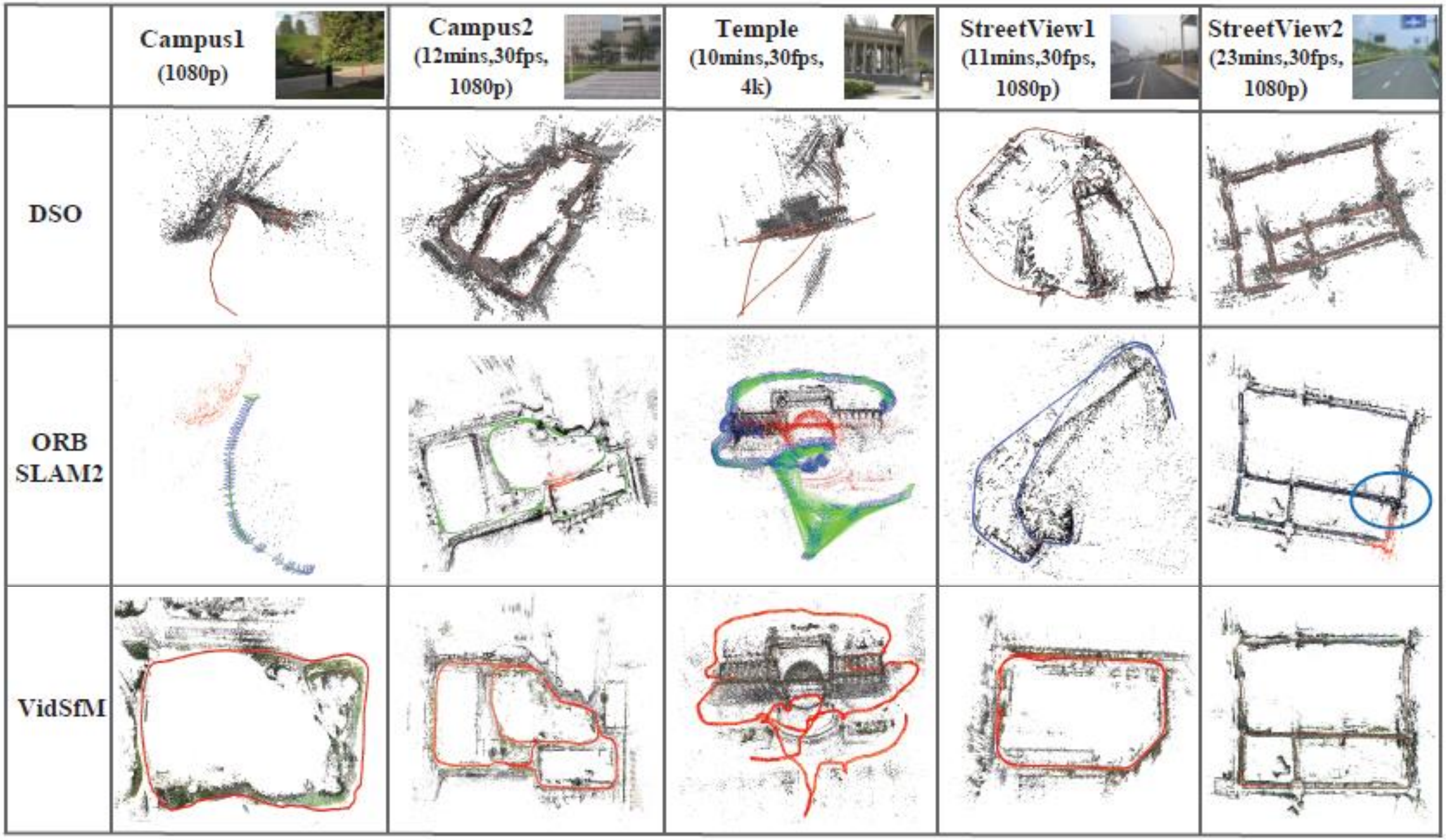
三维几何视觉核心问题：场景结构+相机位姿+(相机参数)

SLAM & Incremental SfM： 错误匹配→误差累计→场景漂移

SfM误差消除策略：

— View graph filtering
— Tracks selection
— Next best view
— Global bundle adjustment

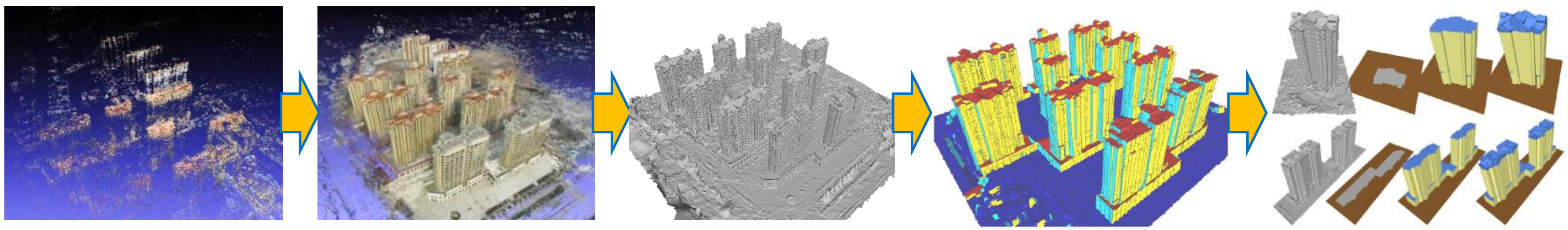
三维几何视觉的核心问题



视频序列+离线计算 （SLAM/VO vs Incremental SfM）

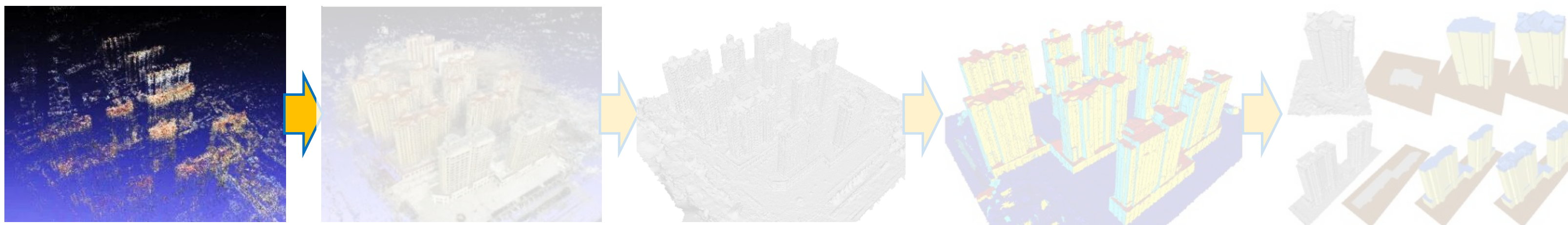
图像三维建模基本流程

稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模



稀疏重建(Structure from Motion)

稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模



输入： 多视角图像

输出： 相机位姿、稀疏点云

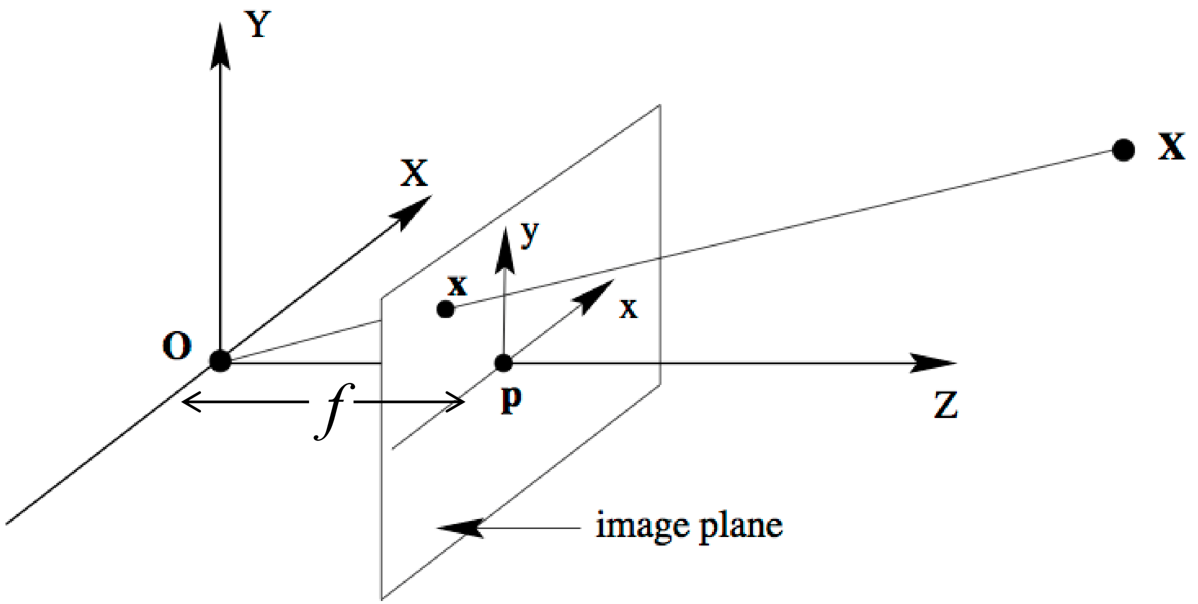
稀疏重建(Structure from Motion)

多视角图像

相机位姿+场景结构

N. Snavely, et. al. IJCV 2008.

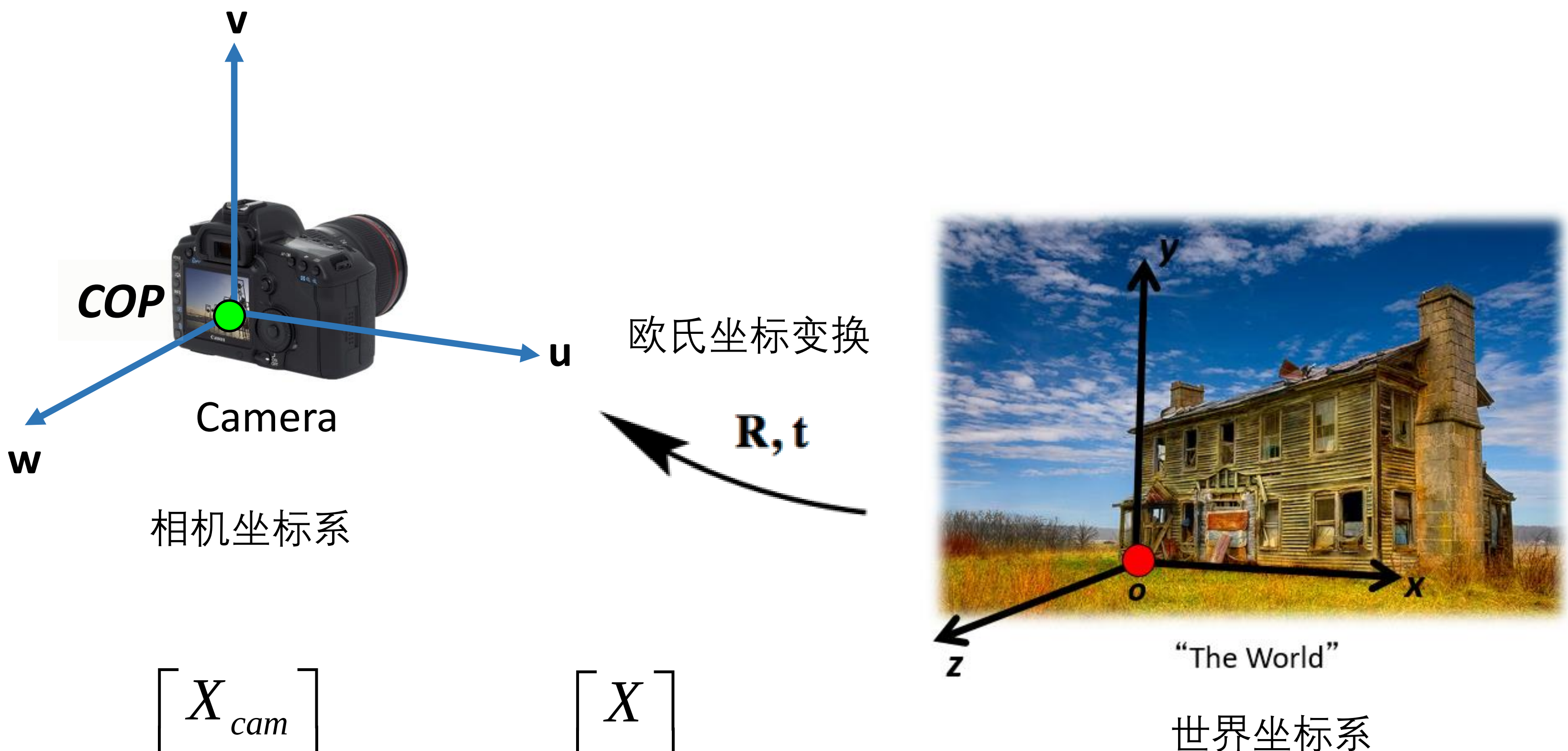
基础：小孔相机模型



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

Up to a scale 焦距　主点

基础：小孔相机模型



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

基础：小孔相机模型

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

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

图像点　相机内参数　相机旋转　相机平移　空间点

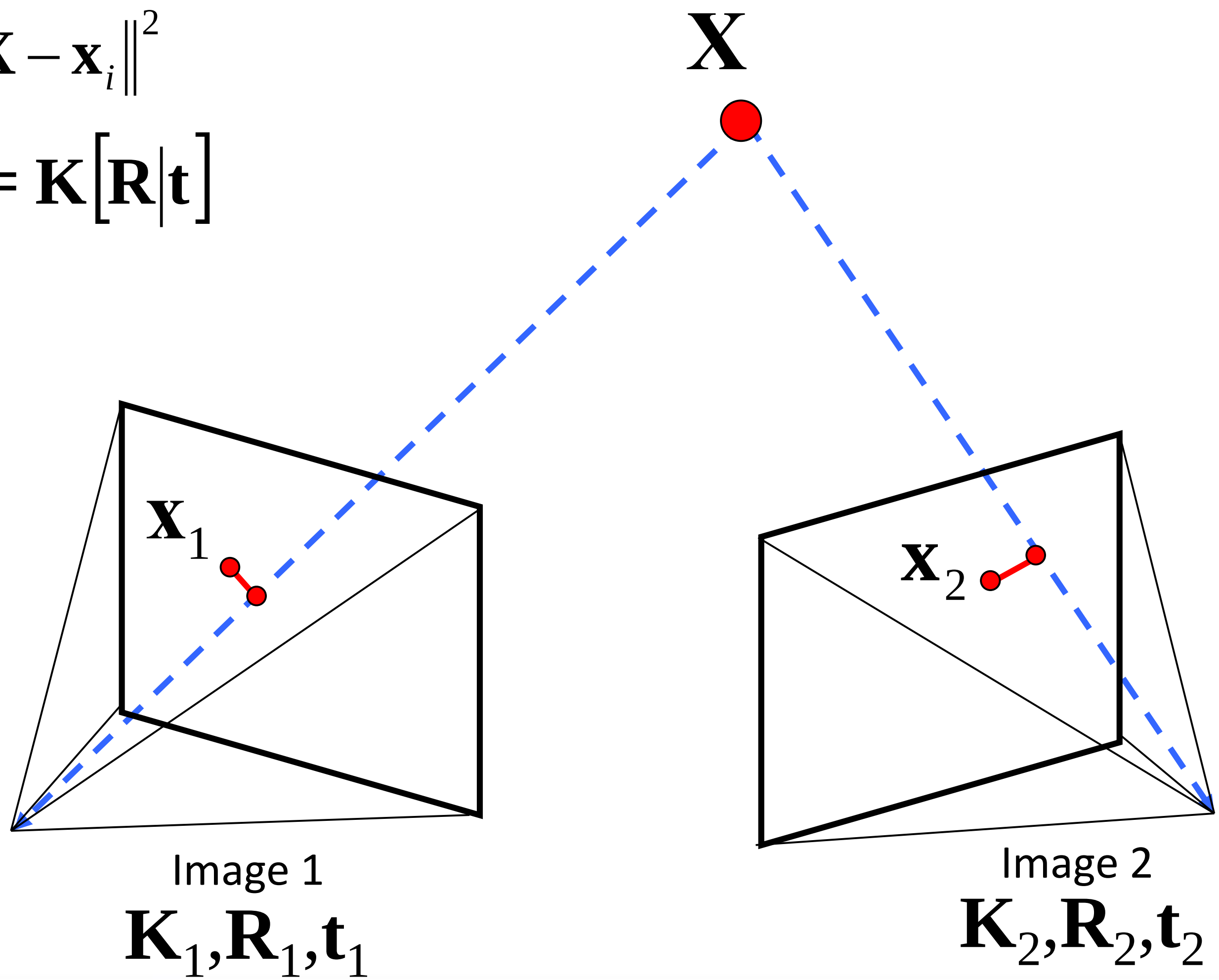
相机硬件　相机位姿　场景结构

基础：重投影误差最小化

- 求解思路：最小化重投影误差平方和

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

$$\mathbf{P} = \mathbf{K}\left[\mathbf{R} \middle| \mathbf{t}\right]$$



基础：重投影误差最小化

- 重投影误差最小化问题的求解

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

$$\Downarrow$$

$$\min \sum_i \sum_j \left\| \frac{\mathrm{P}^i_{1:2}\, \mathbf{X}_j}{\mathrm{P}^i_3\, \mathbf{X}_j} - \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} \right\|^2$$

$$\Downarrow$$

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

$$\Downarrow$$

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

基础：重投影误差最小化

- 重投影误差最小化问题的求解

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

- 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：$N \times (3 + 3 + 3 + 2) + M \times 3$

图像数量　R　T　K　畸变　稀疏点数量　**X**

- 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

稀疏重建(Structure from Motion)

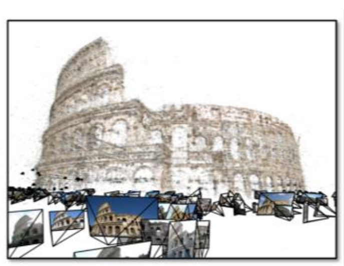
特征点检测

特征点匹配

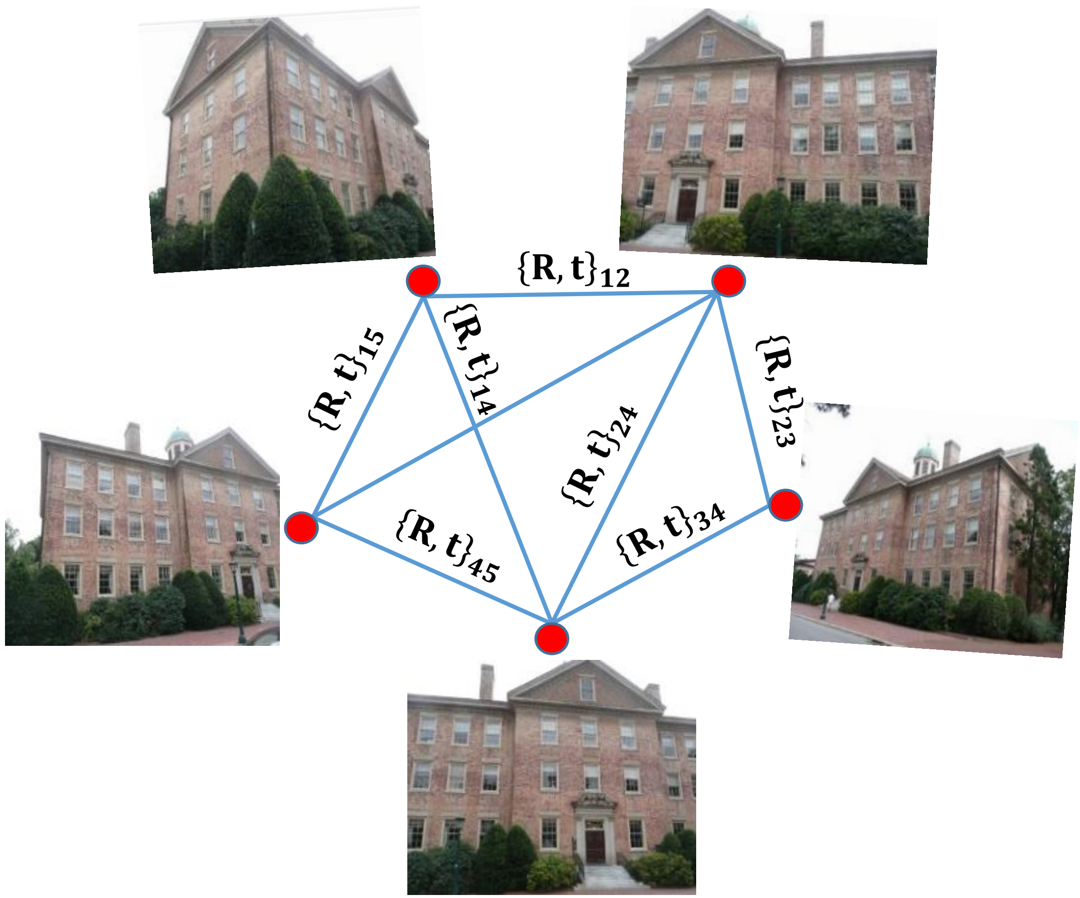
相机位姿(R,T)初始化

1. 增量式
2. 全局式
3. 混合式

捆绑调整
优化相机位姿、相机参数、场景结构



稀疏重建(Structure from Motion)



View Graph

（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

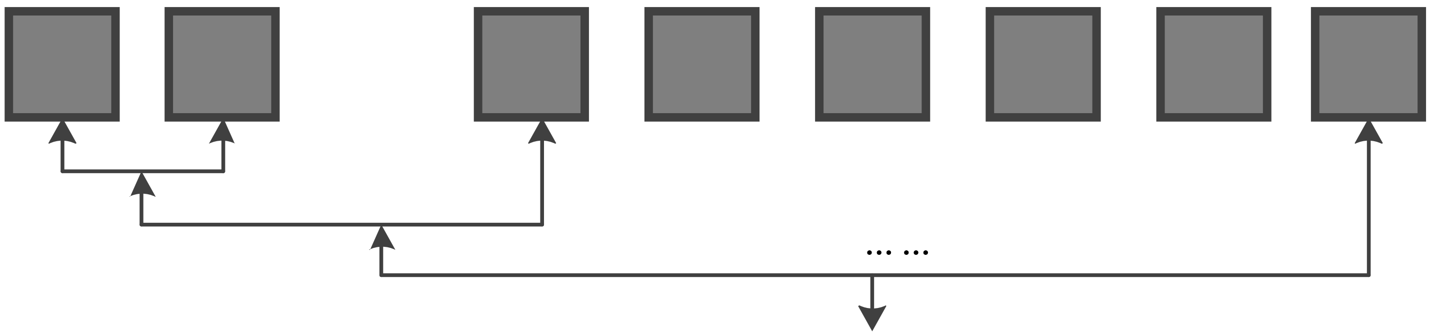
（2）每一条边上表达了对应图像对之间的极几何关系，其中：

$$\mathbf{R}_{ij} = \mathbf{R}_j \mathbf{R}_i^T$$

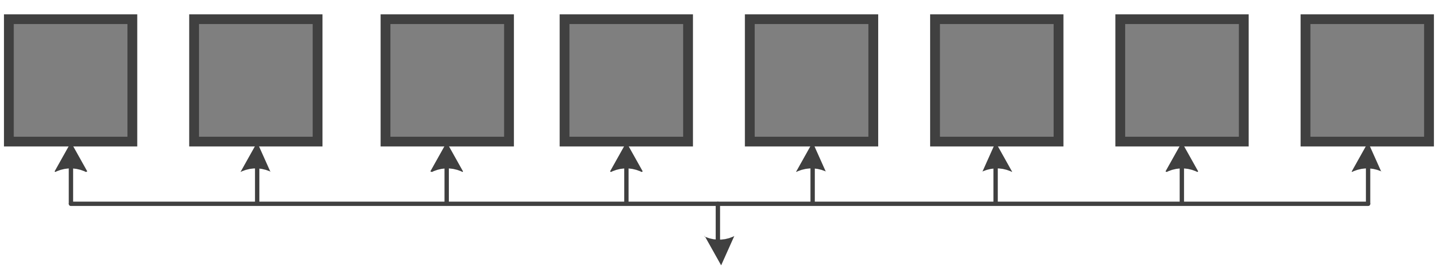
$$\lambda_{ij} \mathbf{t}_{ij} = \mathbf{R}_j (\mathbf{C}_i - \mathbf{C}_j)$$

稀疏重建(Structure from Motion)

- 增量式（Incremental）

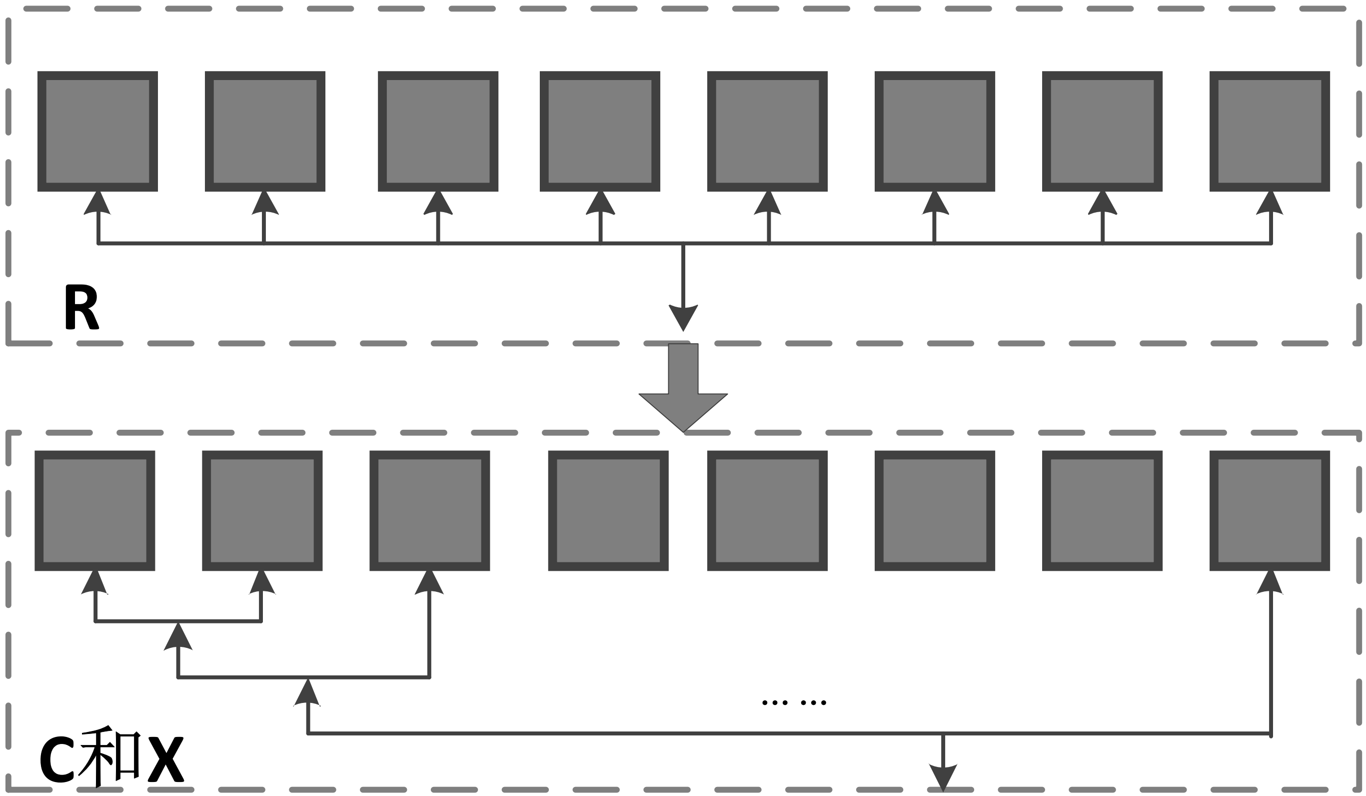


- 全局式（Global）



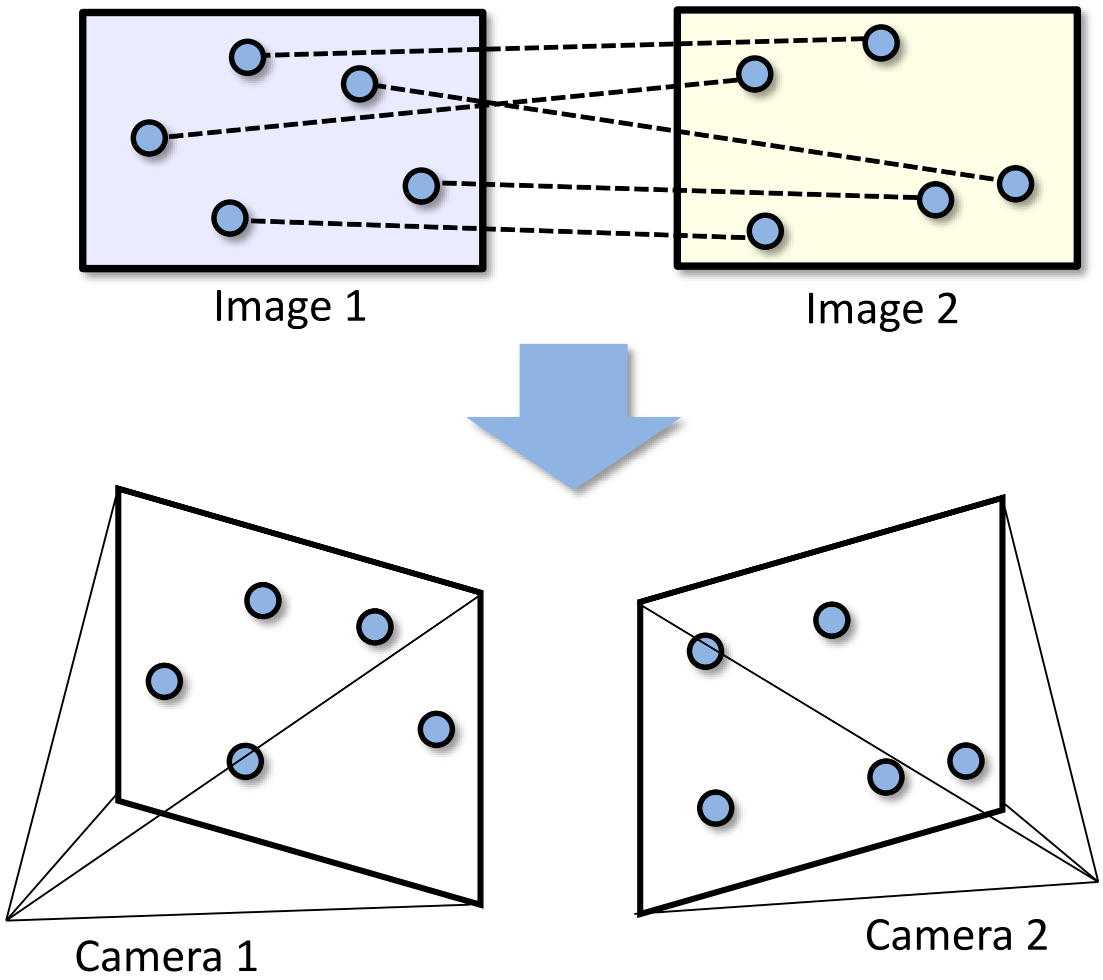
稀疏重建(Structure from Motion)

- 混合式（Hybrid）



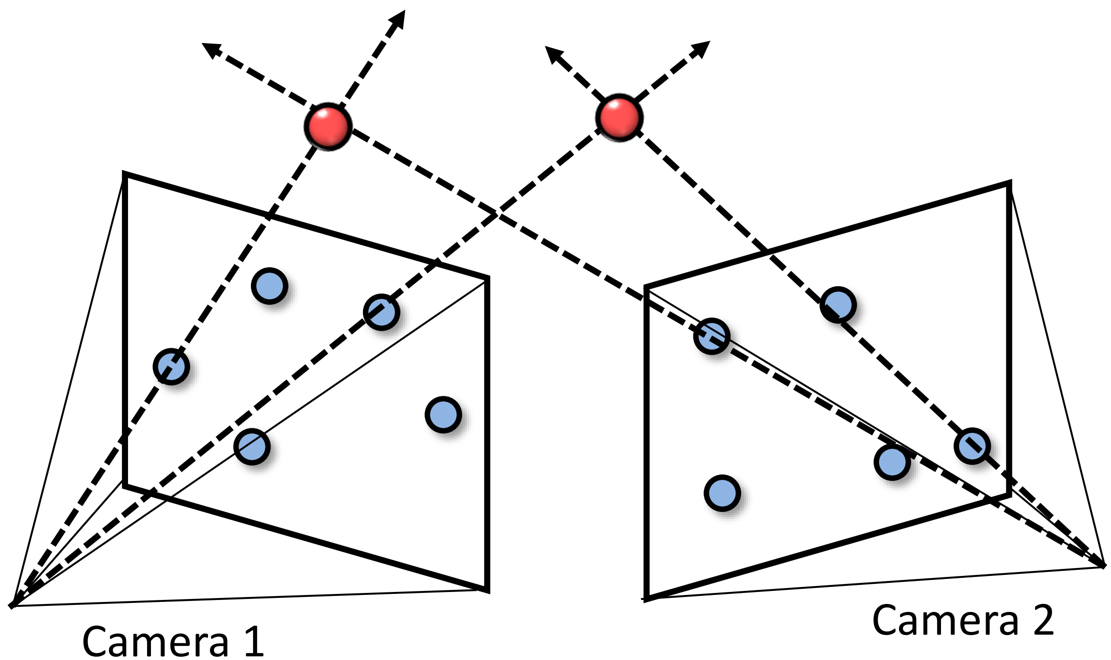
增量式Structure from Motion

- 两视图SfM（8点法、5点法求取两视图相对位姿）



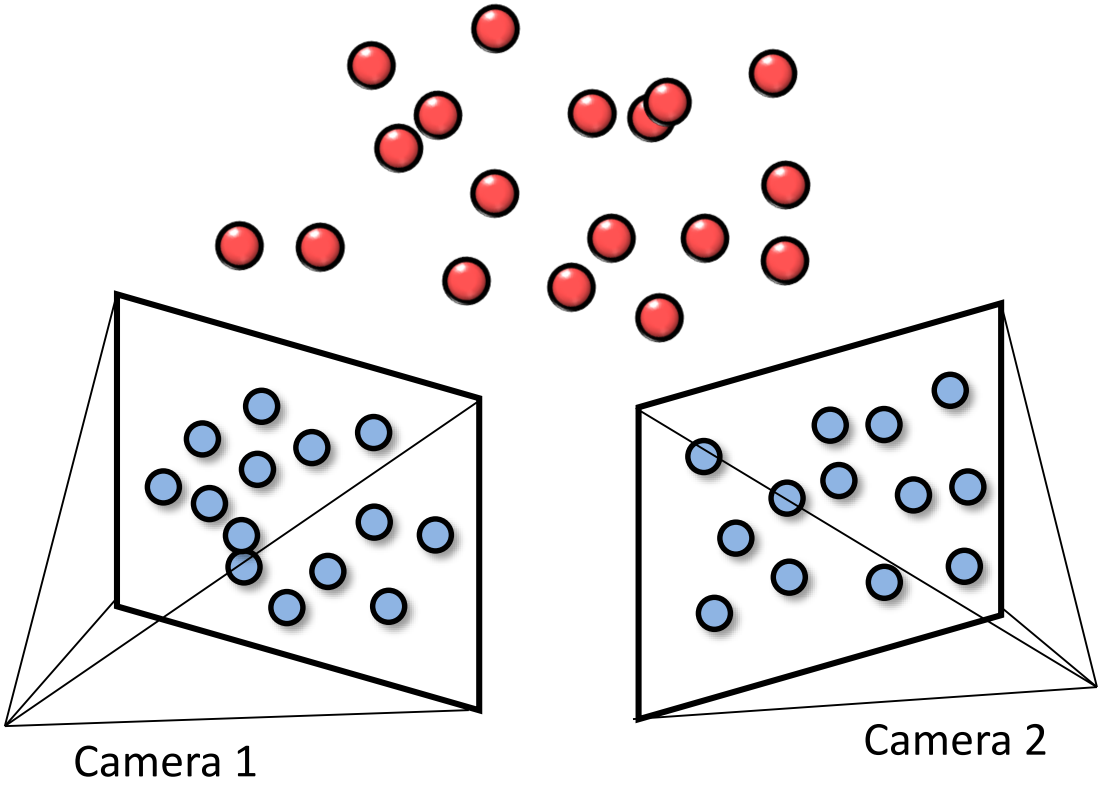
增量式Structure from Motion

- 两视图SfM（两视图三角化）



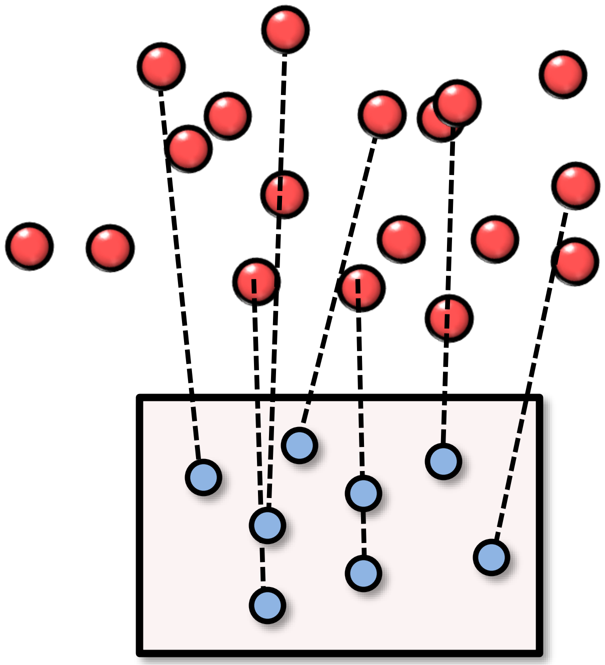
增量式Structure from Motion

- 两视图SfM（两视图三角化）

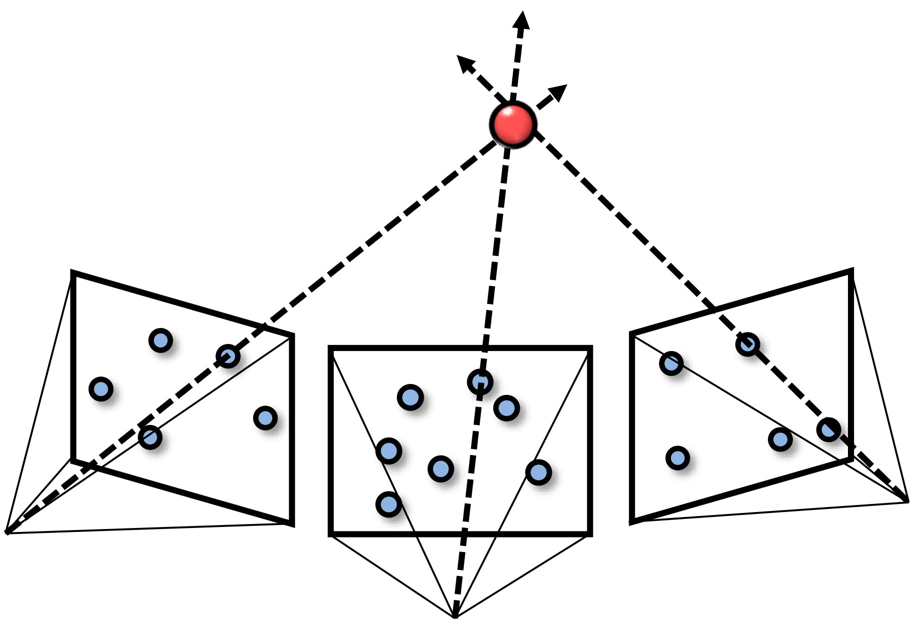


增量式Structure from Motion

- 增量SfM：在初始模型中添加新图像，同时三角化新的3D点



Pose estimation: 2D -> 3D



n-view triangulation

增量式Structure from Motion

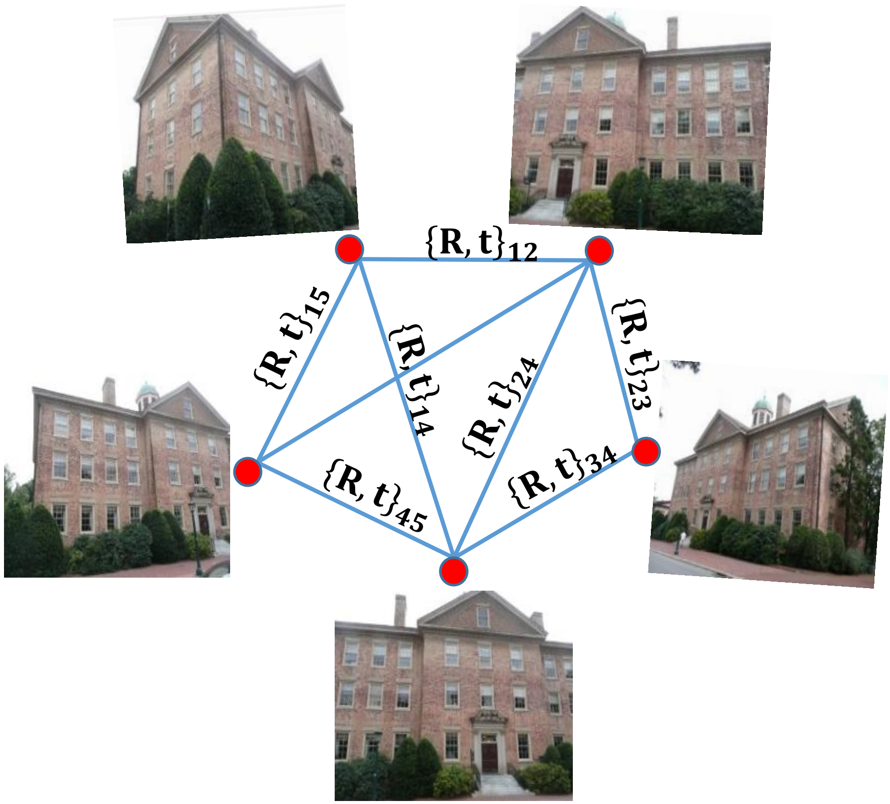
Pros：

- 对匹配外点比较鲁棒，标定过程中通过RANSAC不断地过滤外点
- 重建场景精度高，捆绑调整不断地优化场景结构

Cons：

- 对初始图像对选择及相机的添加顺序敏感
- 大场景重建时的累积误差可能导致场景漂移
- 反复的捆绑调整需要大量的计算时间

全局式Structure from Motion



View Graph

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

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

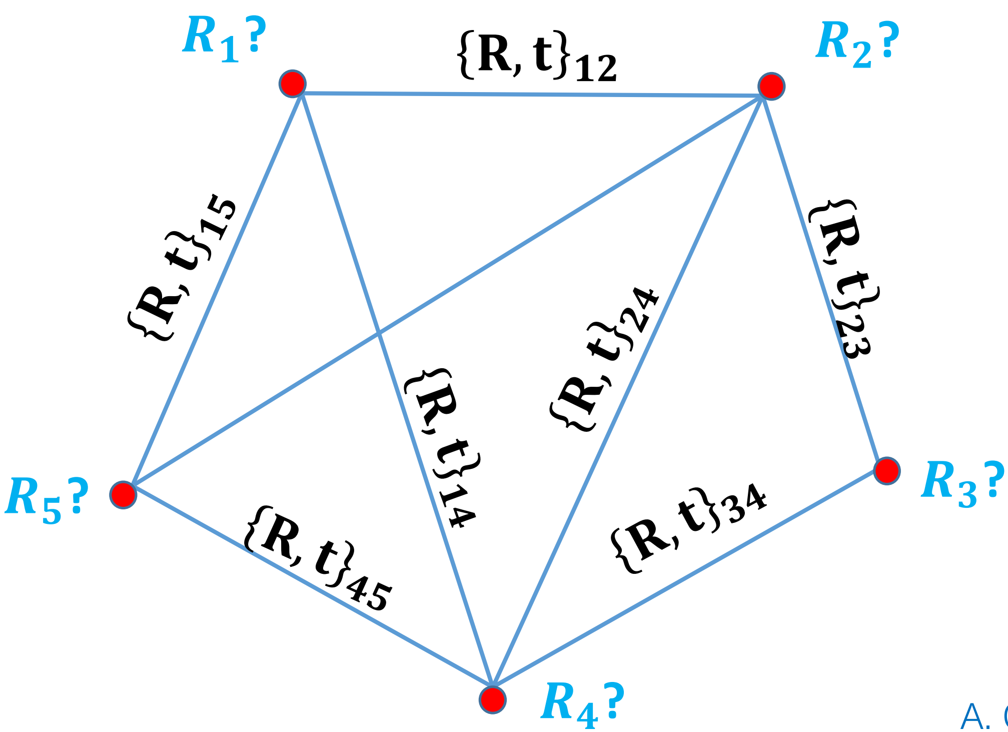
1. 估计所有摄像机的旋转矩阵
2. 估计所有摄像机的位置
3. 三角化初始场景点

全局式Structure from Motion

- 全局式估计摄像机旋转矩阵

$$\mathbf{R_{ij}} = \mathbf{R_j R_i^T} \quad \Rightarrow \quad \min_{R} \left\| \mathbf{R_{ij}} - \mathbf{R_j R_i^T} \right\|$$

Solver: Lie algebra in $\mathfrak{so}(3)$ $\rightarrow$ $\mathcal{L}1$ averaging $\rightarrow$ $\mathcal{L}2$ averaging

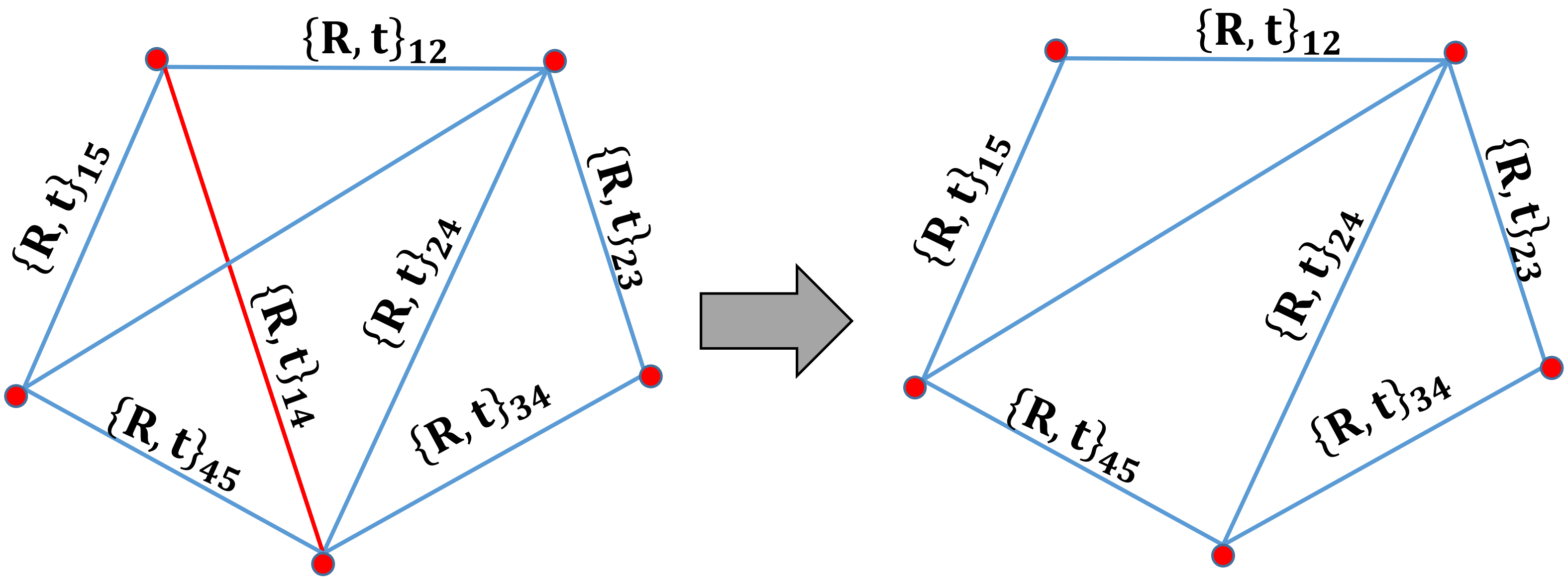


A. Chatterjee, et. al. ICCV 2013.

全局式Structure from Motion

- 过滤不一致边

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

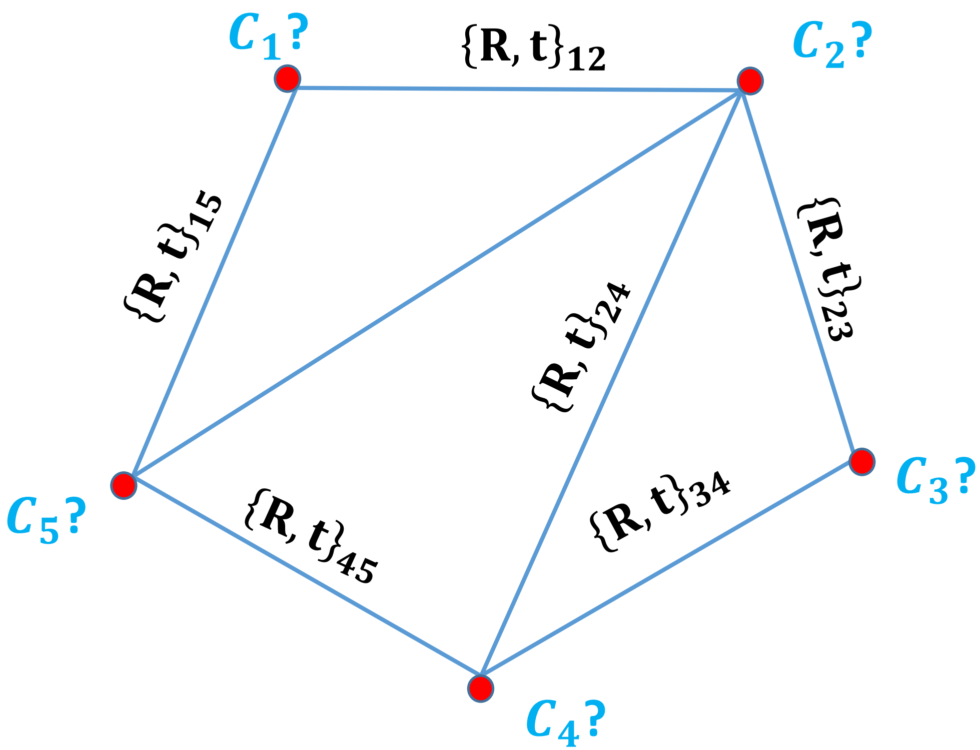


全局式Structure from Motion

- 全局式估计摄像机位置

$$\lambda_{ij}\mathbf{t}_{ij} = \mathbf{R}_j(\mathbf{C}_i - \mathbf{C}_j)$$
$$\mu_{ij} = \mathbf{R}_j^T \mathbf{t}_{ij}$$

$$\Longrightarrow \quad \min_{C} \left\| \mu_{ij} - \frac{C_i - C_j}{\|C_i - C_j\|} \right\|^2$$

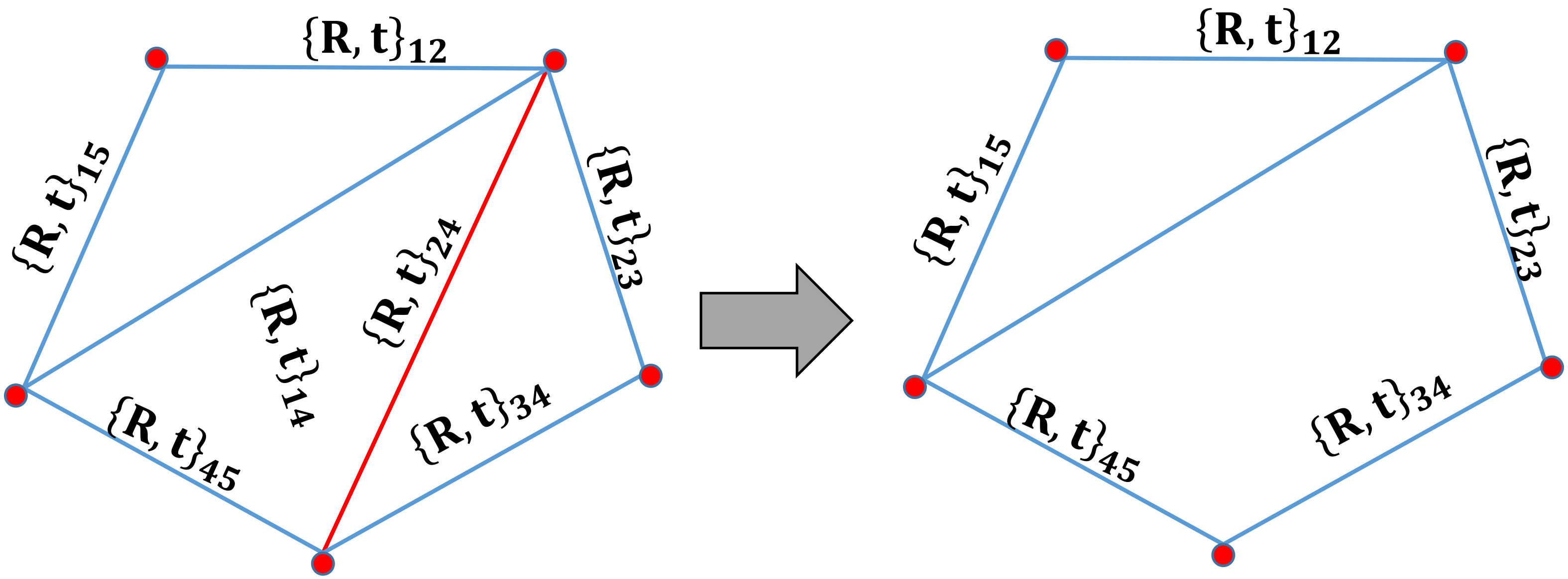


K. Wilson, et. al. ECCV 2014.

全局式Structure from Motion

- 过滤不一致边

$$\left\| \mu_{\mathbf{ij}} - \frac{C_i - C_j}{\left\| C_i - C_j \right\|} \right\| \geq \epsilon$$



全局式Structure from Motion

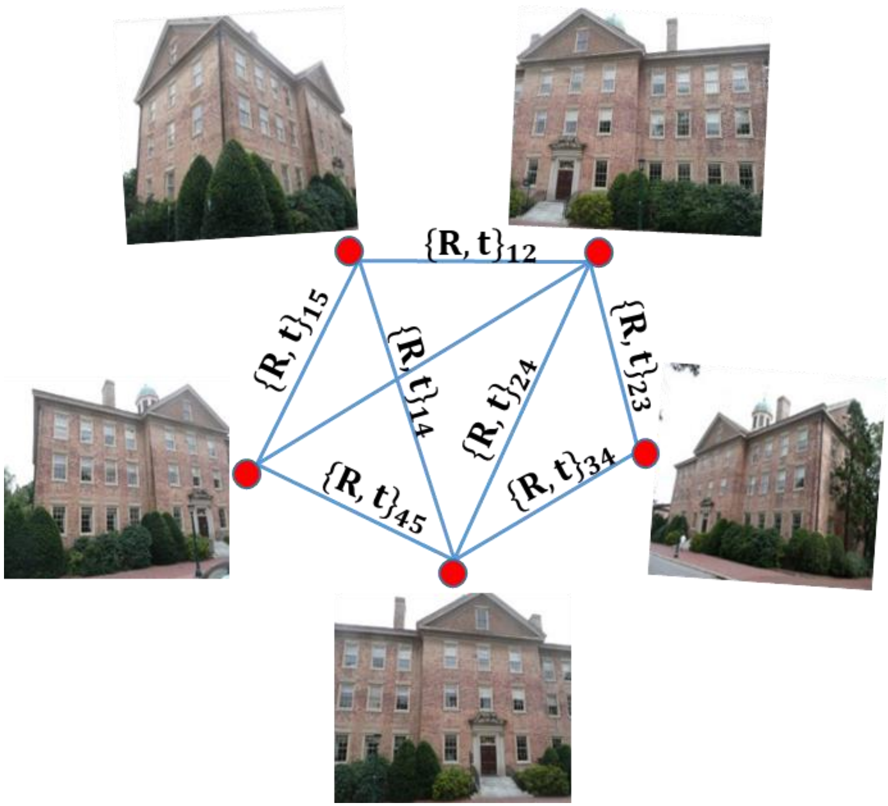
Pros：

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

Cons：

- 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
- 场景完整性：过滤外极几何边可能会丢失部分图像

混合式Structure from Motion



View Graph

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

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

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. 局部捆绑调整

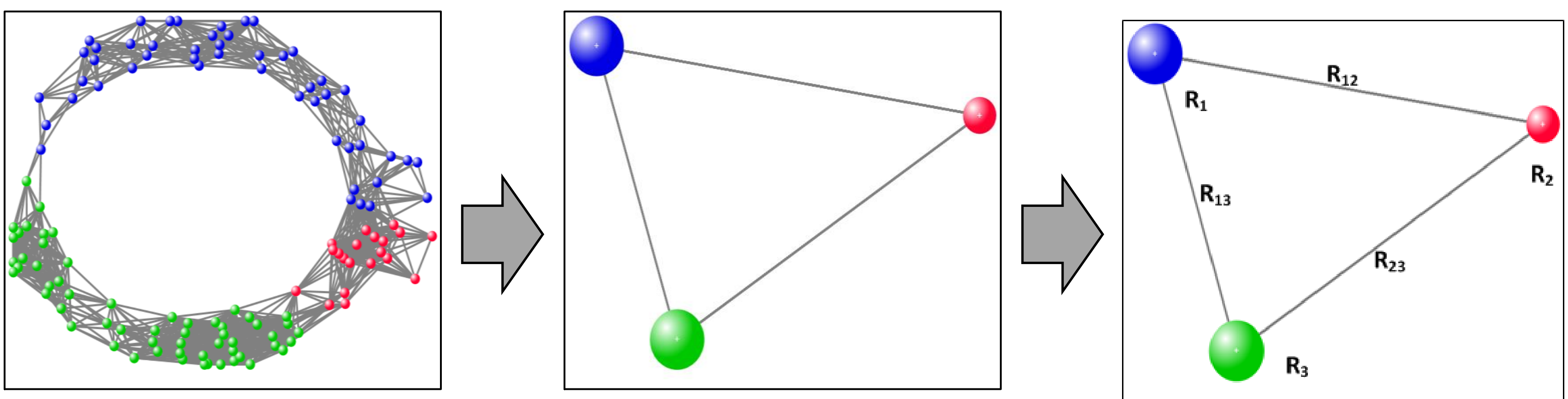
混合式Structure from Motion

全局式求取摄像机旋转矩阵精度的影响因素：

1. 图像之间的连接紧密程度
2. 外极几何关系的估计精度

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

K. Wilson, et al. When is Rotations Averaging Hard. ECCV, 2016.

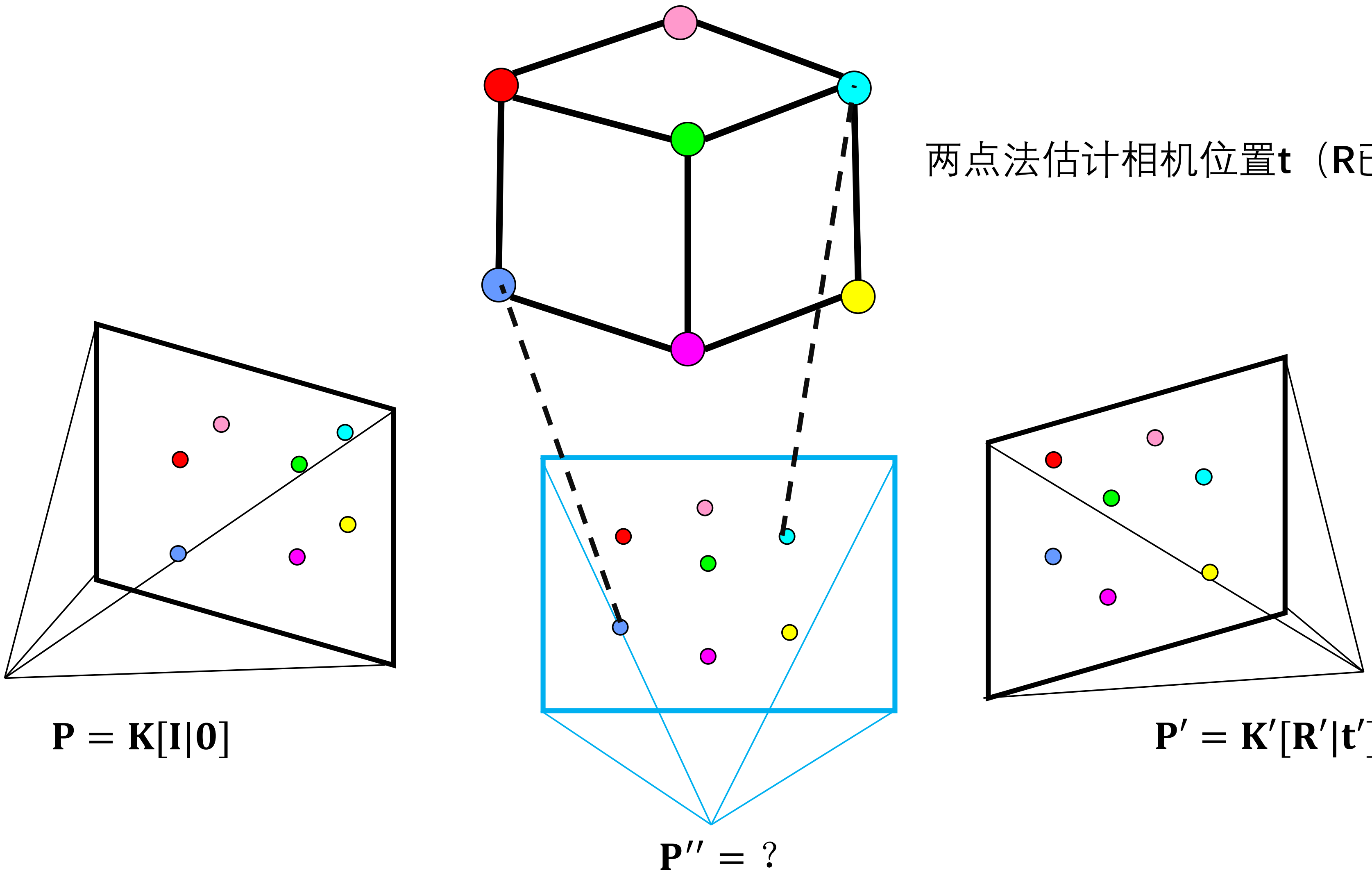


外极几何图　　团体图　　单独求取**R**

A. Clauset, et al. Physical Review E, 2005.
H. Cui, et al. CVPR, 2017.

混合式Structure from Motion

两点法估计相机位置t（R已知）



混合式Structure from Motion

Pros：

- 解旋转矩阵误差均匀分布在外极几何图中
- 全局旋转矩阵的求解缓解了后续增量式的误差累积问题
- 求解摄像机位置时，两点法求取更加鲁棒，捆绑调整次数更少

Cons：

- 对旋转矩阵的精度依赖较高
- 计算相机位置后仍需要重复迭代的捆绑调整

相关文献

增量式SfM：

1. N. Snavely, et al. Modeling the World from Internet Photo Collections. IJCV 2007.
2. P. Moulon, et al. Adaptive Structure from Motion with a contrario model estimation. ACCV 2012.
3. C. Wu. Towards linear-time incremental structure from motion. 3DV 2013.
4. J. Schonberger, et al. Structure-from-Motion Revisited. CVPR 2016.
5. H. Cui, et al. Batched Incremental Structure-from-Motion. 3DV 2017.
6. H. Cui, et al. Progressive Large-Scale Structure-from-Motion with Orthogonal MSTs. 3DV 2018.

全局式SfM：

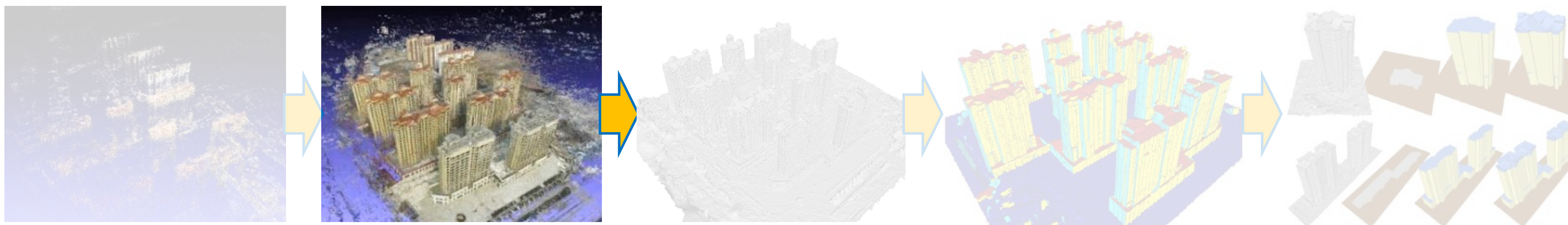
1. N. Jiang, et al. A Global Linear Method for Camera Pose Registration. ICCV 2013.
2. M. Moulon, et al. Global Fusion of Relative Motions for Robust, Accurate and Scalable structure from motion. ICCV 2013.
3. K. Wilson, et al. Robust Global Translations with 1DSfM. ECCV 2014.
4. O. Ozyesil, et al. Robust Camera Location Estimation by Convex Programming. CVPR 2015 .
5. C. Sweeney, et al. Optimizing the Viewing Graph for Structure-from-Motion. ICCV 2015.
6. T. Goldstein, et al. Shapefit and Shapekick for Robust, Scalable Structure from Motion. ECCV 2016.
7. S. Zhu, et al. Very Large-Scale Global SfM by Distributed Motion Averaging. CVPR 2018.

混合式SfM：

1. H. Cui, et al. HSfM: Hybrid Structure-from-Motion. CVPR 2017.
2. S. Zhu, et al. Parallel Structure from Motion from Local Increment to Global Averaging. arXiv:1702.08601.

稠密重建(Multiple View Stereo)

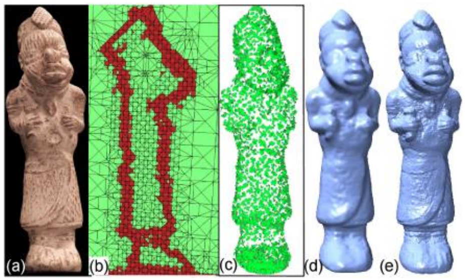
稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模

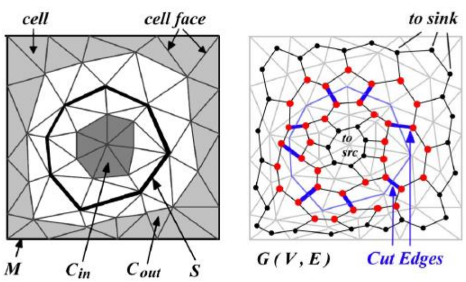


输入：多视角图像、相机位姿

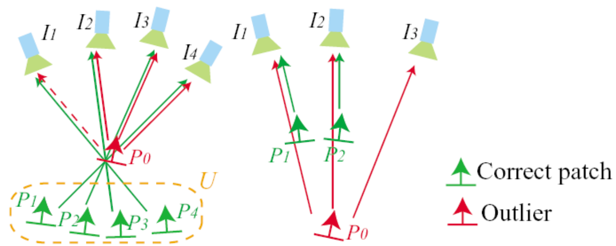
输出：稠密点云

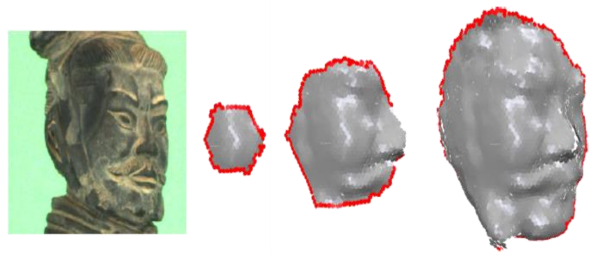
稠密重建(Multiple View Stereo)



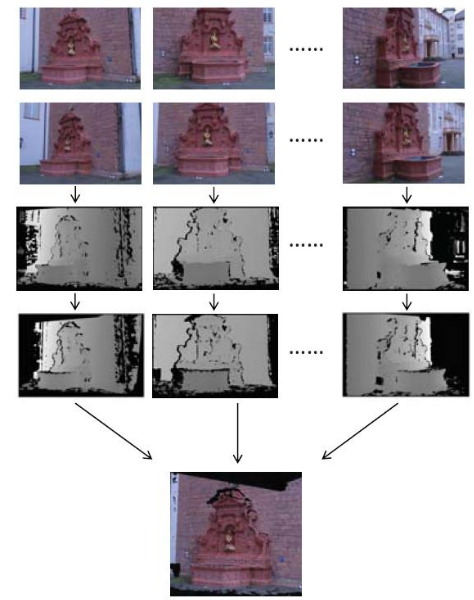


基于体素的方法
Voxel based MVS





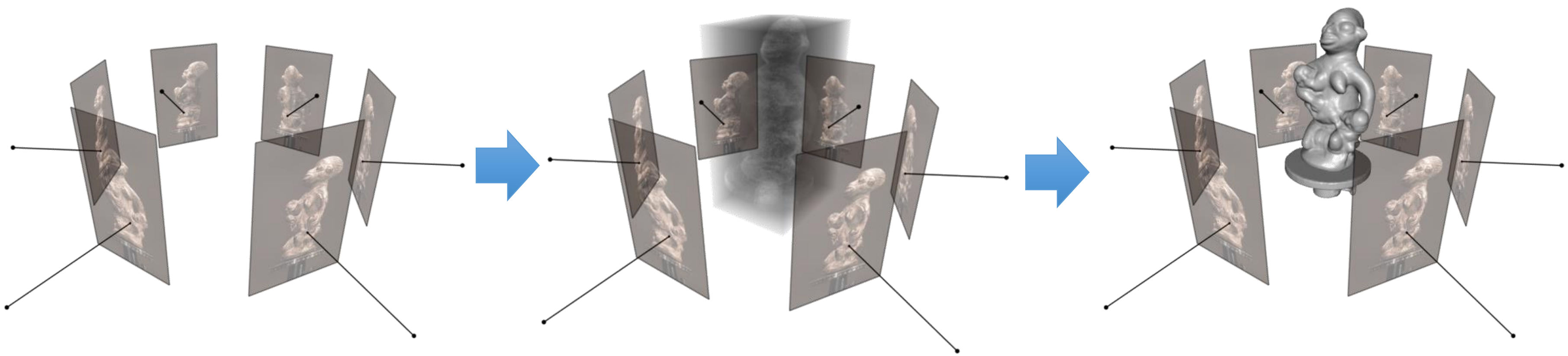
基于点云扩散的方法
Feature point growing based MVS



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

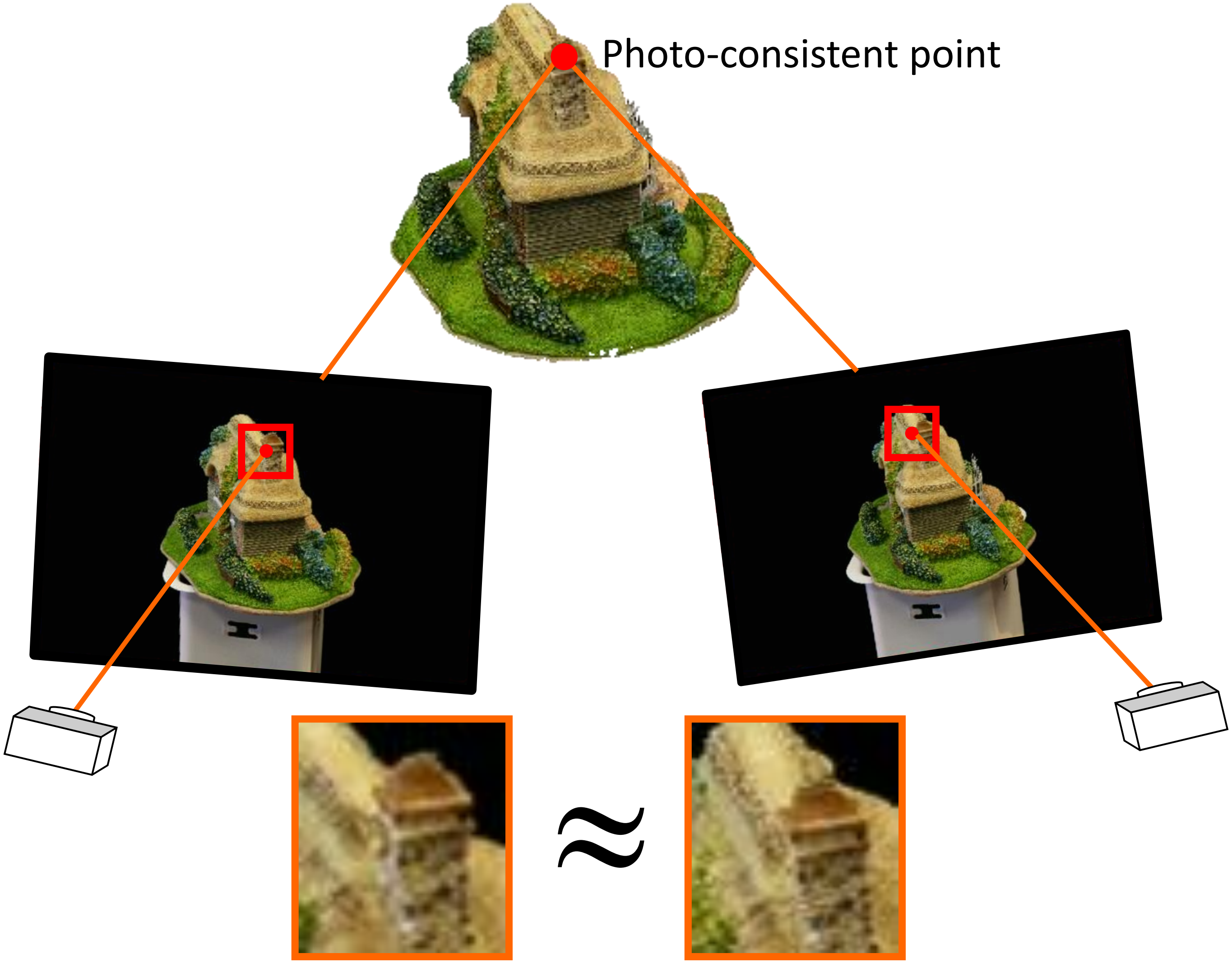
稠密重建(Multiple View Stereo)

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

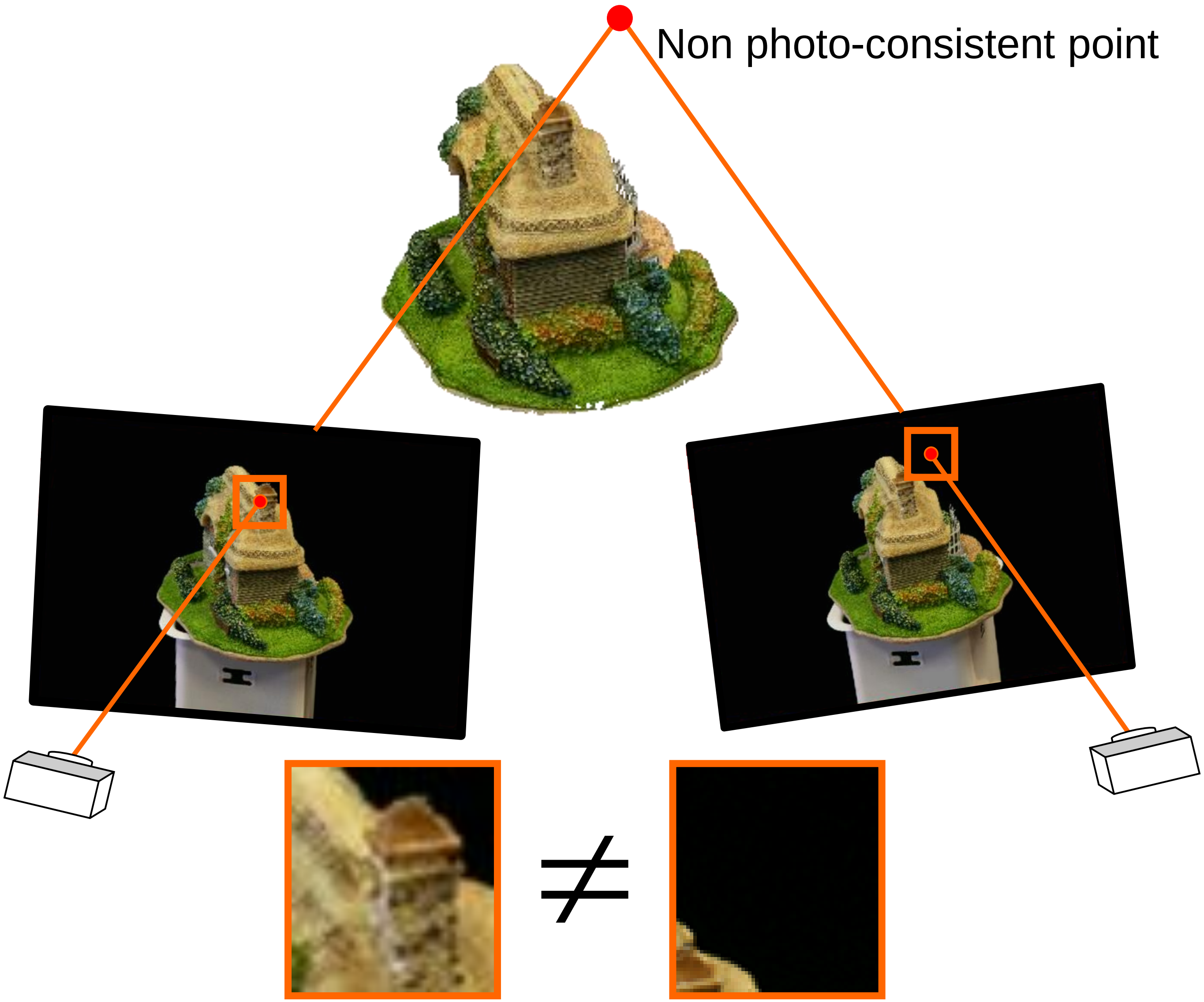


相机位姿　　　　图像一致性　　　　稠密点云

稠密重建(Multiple View Stereo)



稠密重建(Multiple View Stereo)



稠密重建(Multiple View Stereo)

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): $\rho_{NCC}(f,g)=\dfrac{(f-\bar{f})\cdot(g-\bar{g})}{\delta_f\delta_g}$

$f=$  $g=$ 

稠密重建(Multiple View Stereo)

两视图图像一致性(Photo-consistency)的度量:

- Census: $\rho_{census}(f,g) = |census(f) - census(g)|_1$

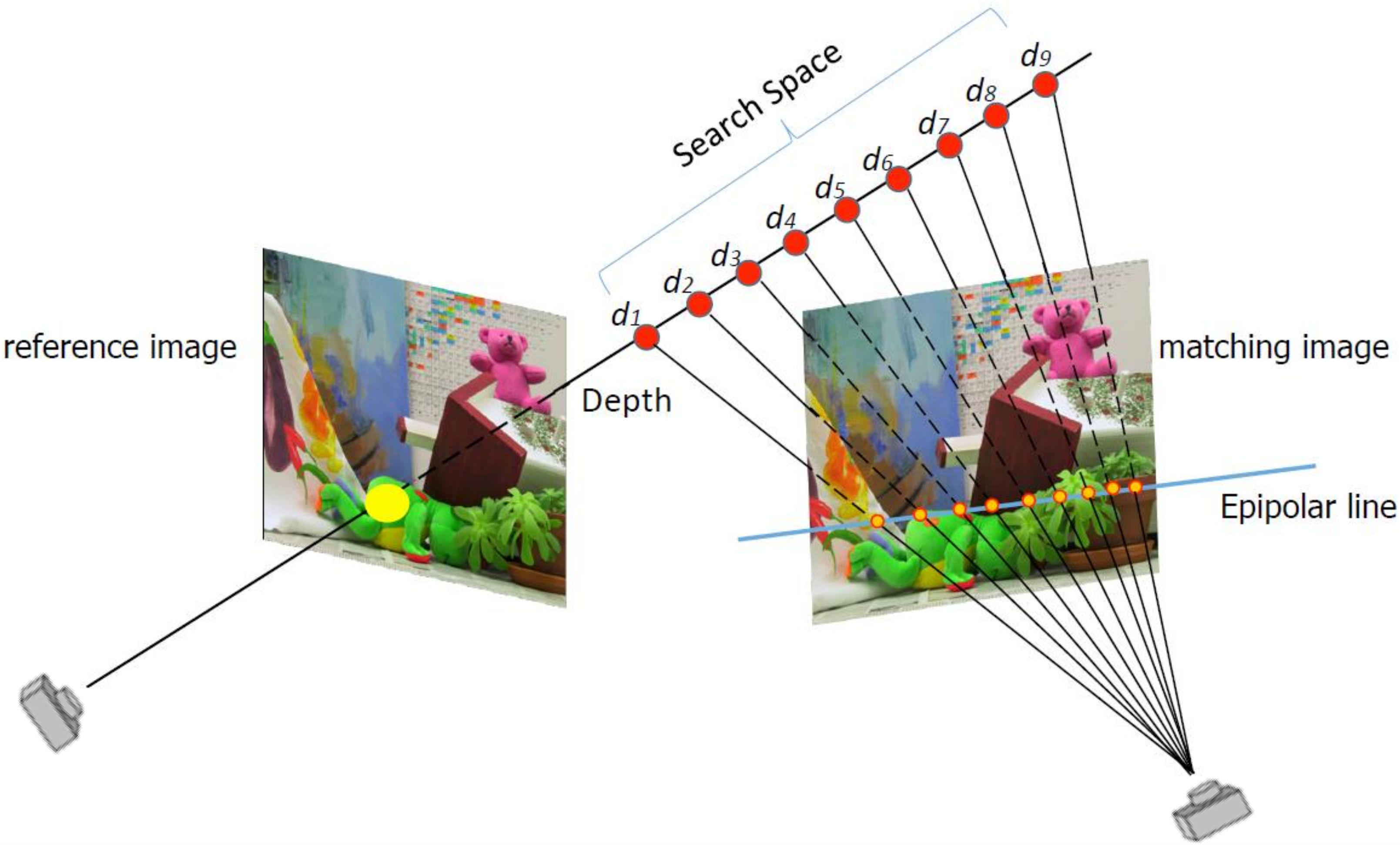
  $census(f) = \otimes_{q\in\Omega}\xi(f(p), f(q))$ $\xi(a,b) = 1\ if\ a < b,\ 0\ otherwise$

- Rank: $\rho_{rank}(f,g) = |rank(f) - rank(g)|$ $rank(f) = \sum_{q\in\Omega}\xi(f(p), f(q))$

$f =$  $g =$ 

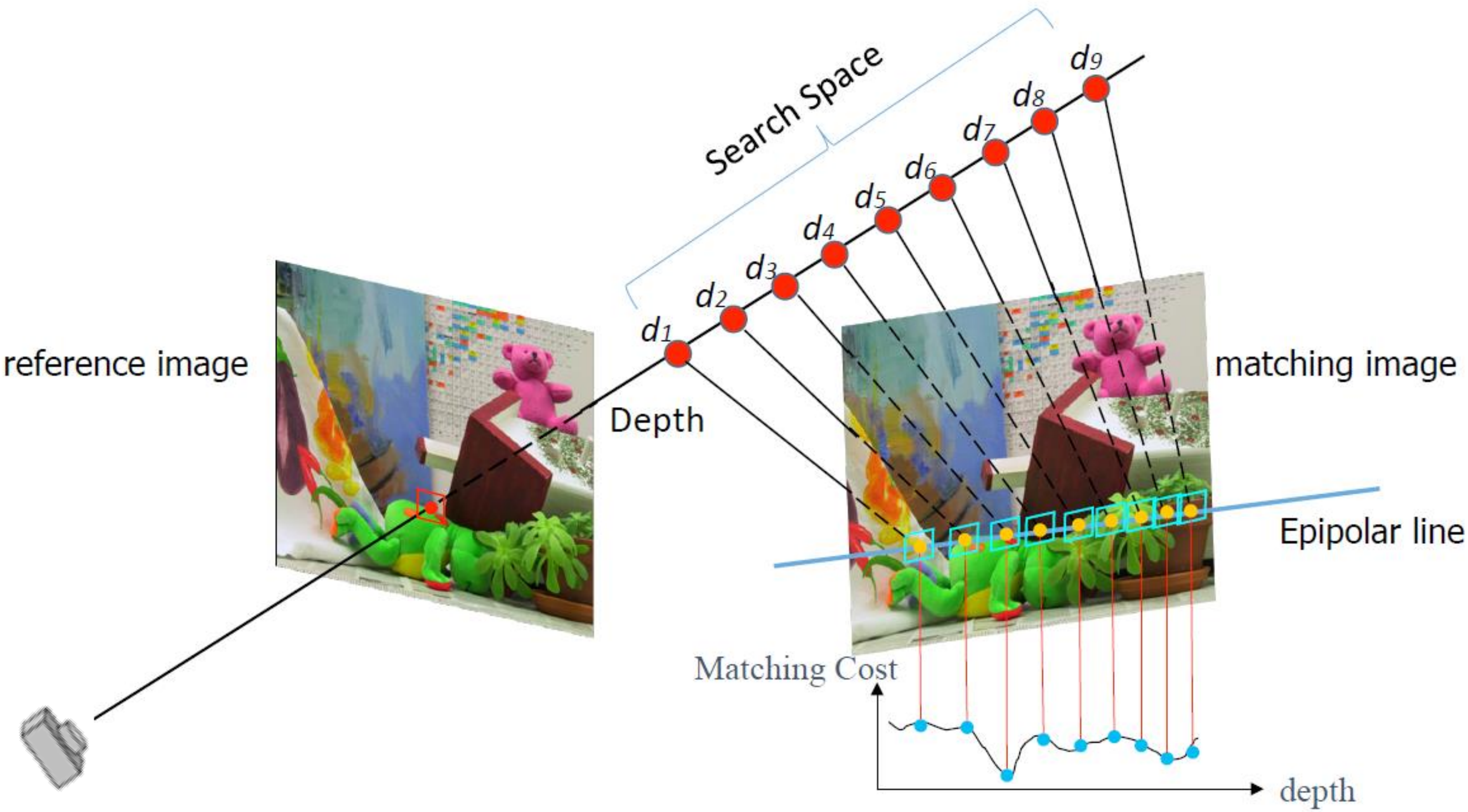
稠密重建(Multiple View Stereo)

每一个像素上深度值的计算：



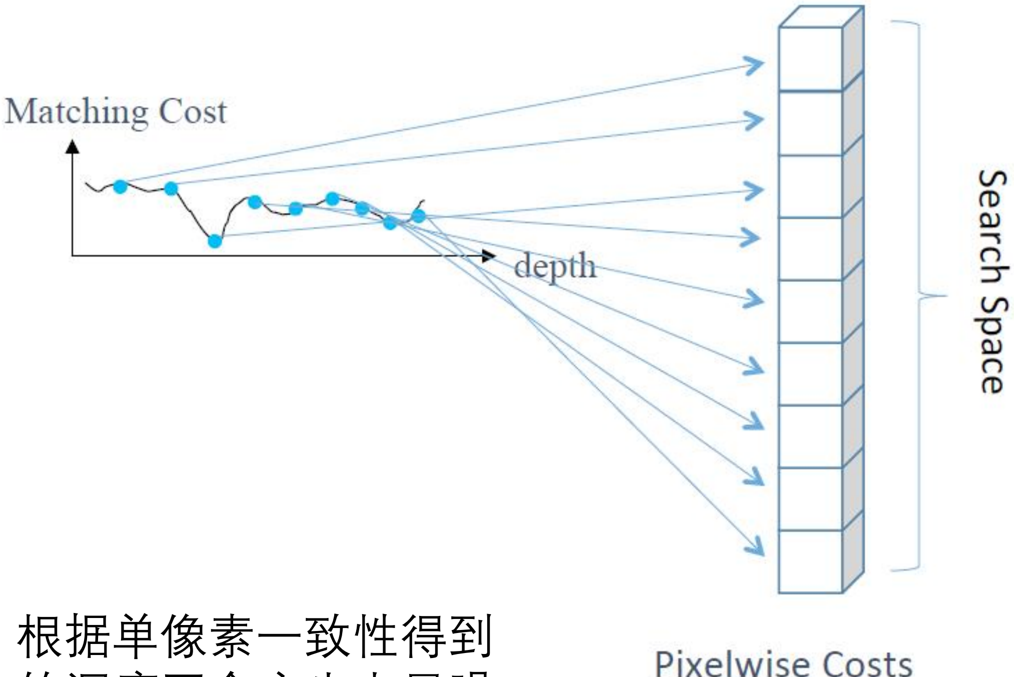
稠密重建(Multiple View Stereo)

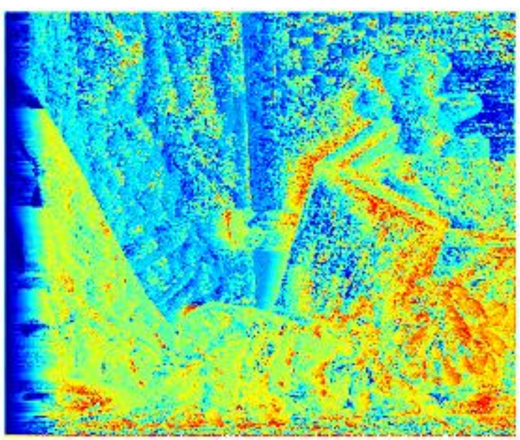
每一个像素上深度值的计算：Winner-Takes-All



稠密重建(Multiple View Stereo)

每一个像素上深度值的计算：Winner-Takes-All





根据单像素一致性得到的深度图会产生大量噪声和外点。

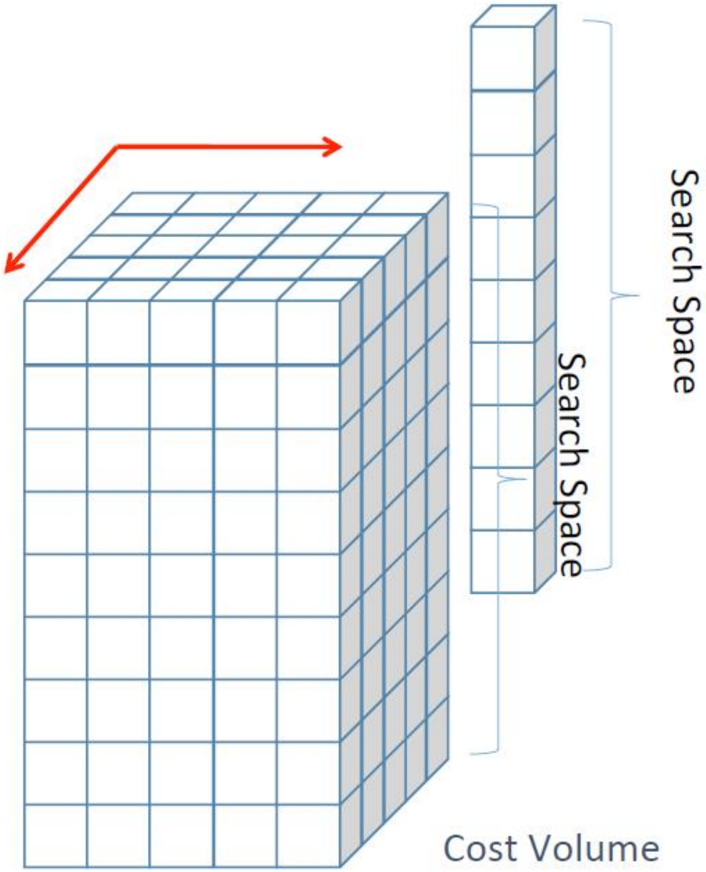
稠密重建(Multiple View Stereo)

聚合(Cost aggregation):

Cost aggregation

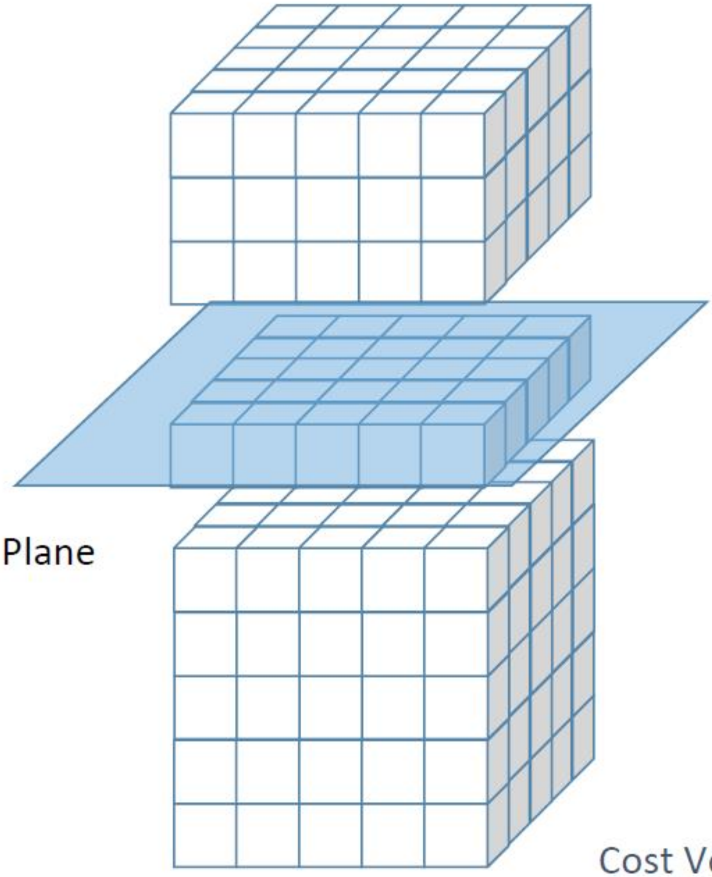
1）单像素计算不稳定

2）通过邻域像素进行一致性聚合



稠密重建(Multiple View Stereo)

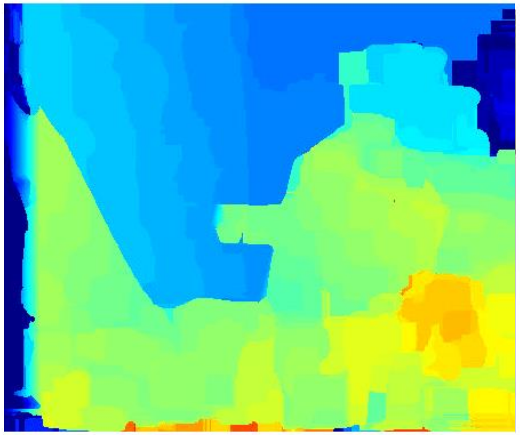
聚合(Cost aggregation)：



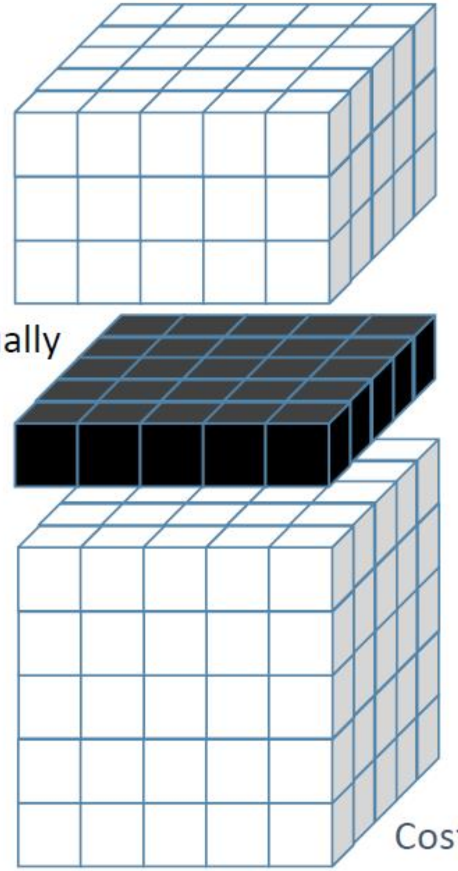
稠密重建(Multiple View Stereo)

聚合(Cost aggregation)：SAD

Treat neighbors equally



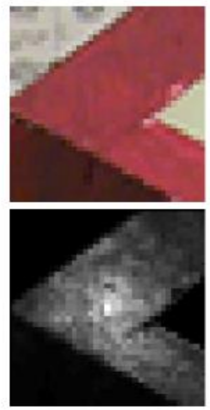
Sum of Absolute Differences (SAD)

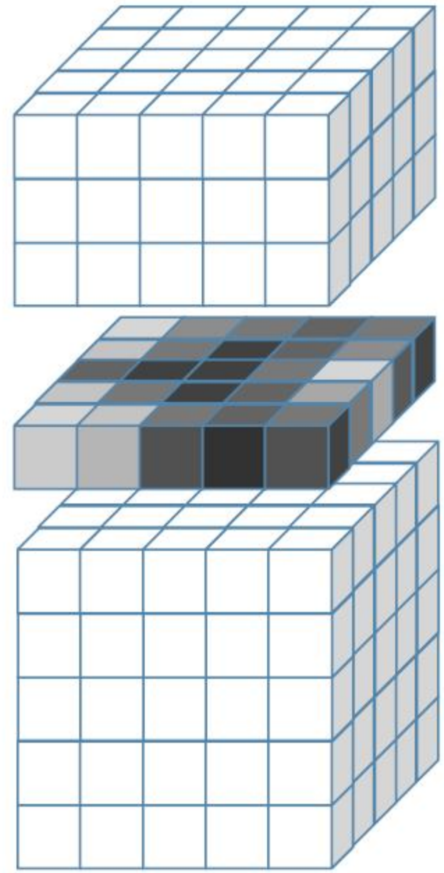


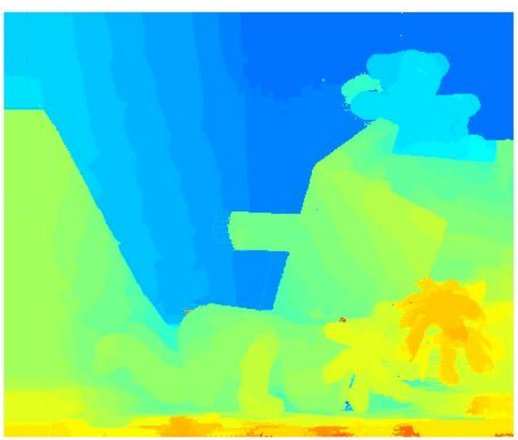
Cost Volume

稠密重建(Multiple View Stereo)

聚合(Cost aggregation)：Adaptive weight





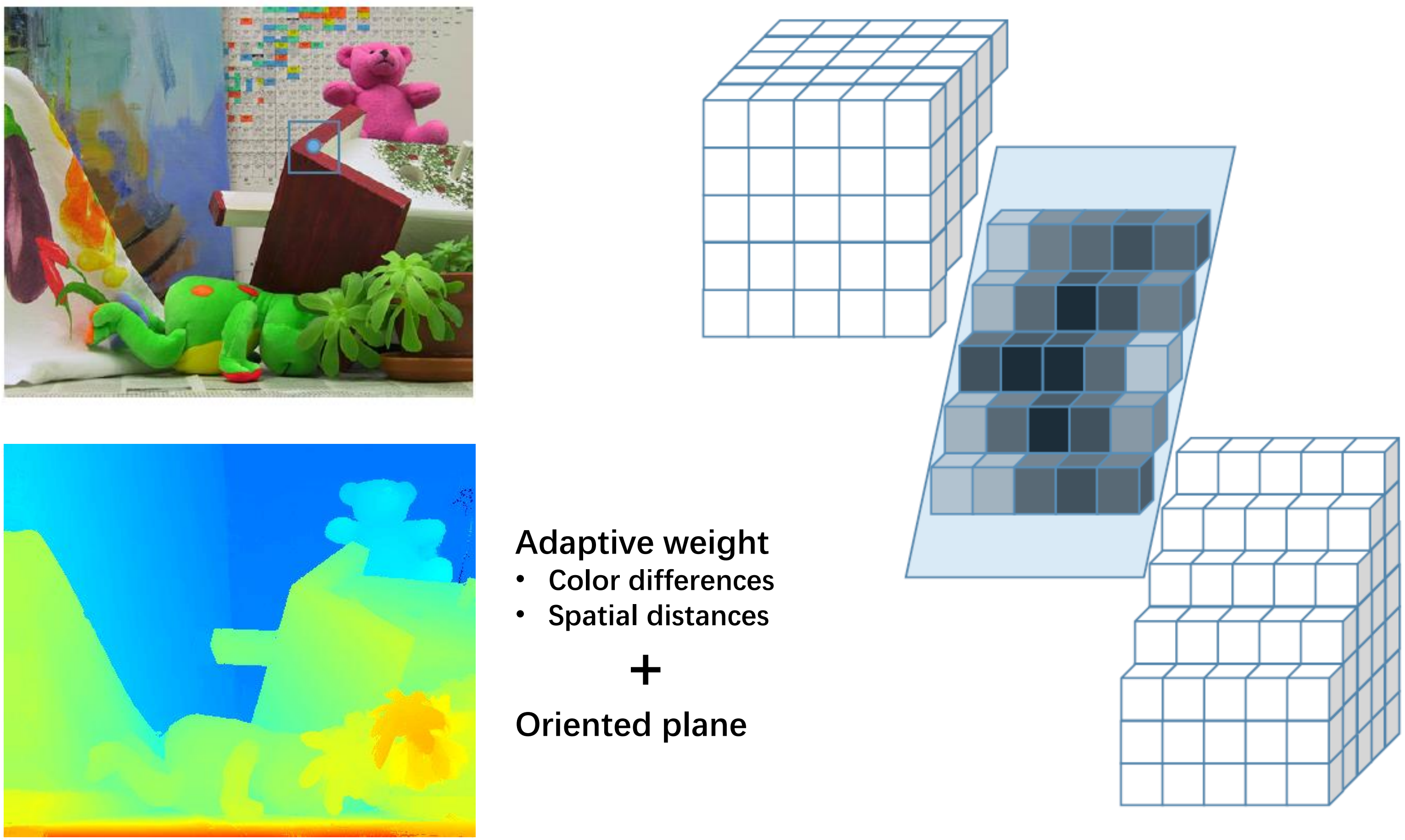


Adaptive weight
- **Color differences**
- **Spatial distances**

K. Yoon, et. al. IEEE T-PAMI 2006.

稠密重建(Multiple View Stereo)

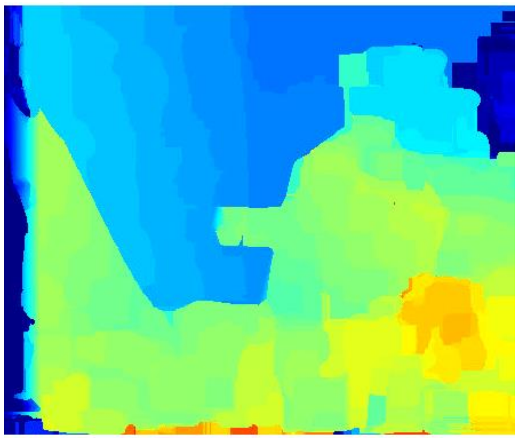
聚合(Cost aggregation)：Adaptive weight + Oriented plane



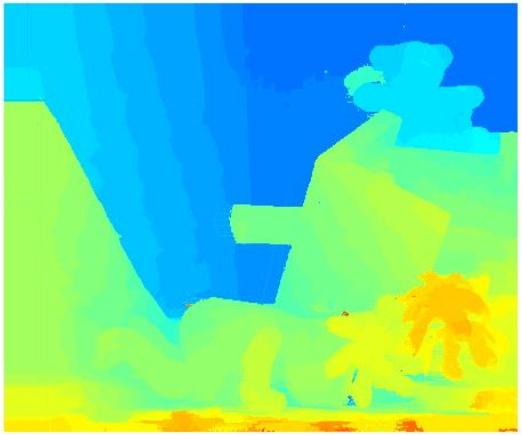
J. Lu, et. al. CVPR 2013.

稠密重建(Multiple View Stereo)

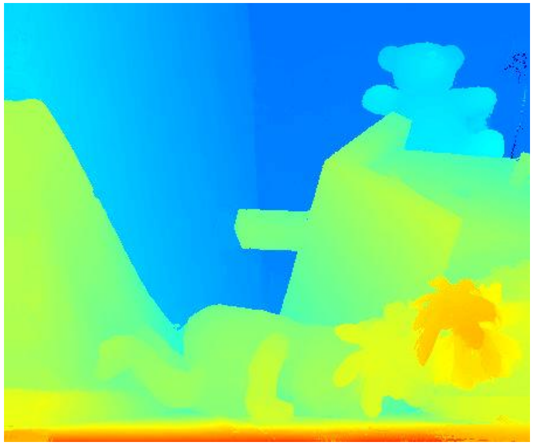
不同聚合(Cost aggregation)方法的比较:



SAD



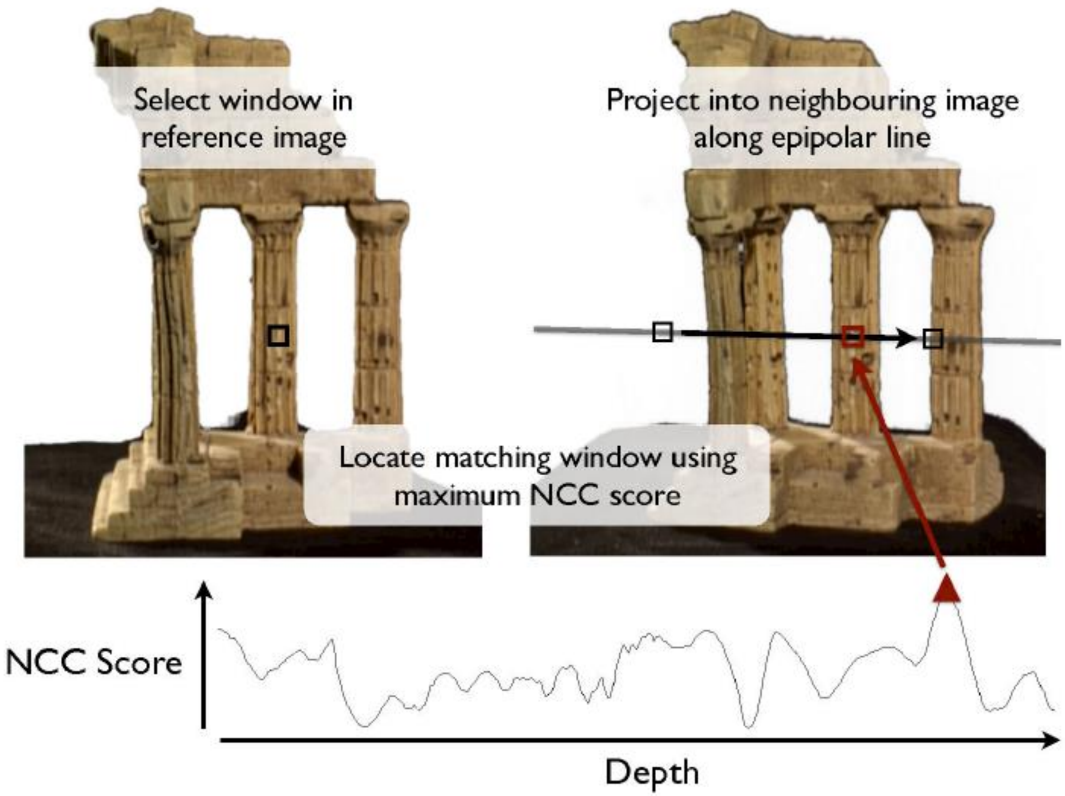
Adaptive weight



Adaptive weight
+
Oriented plane

稠密重建(Multiple View Stereo)

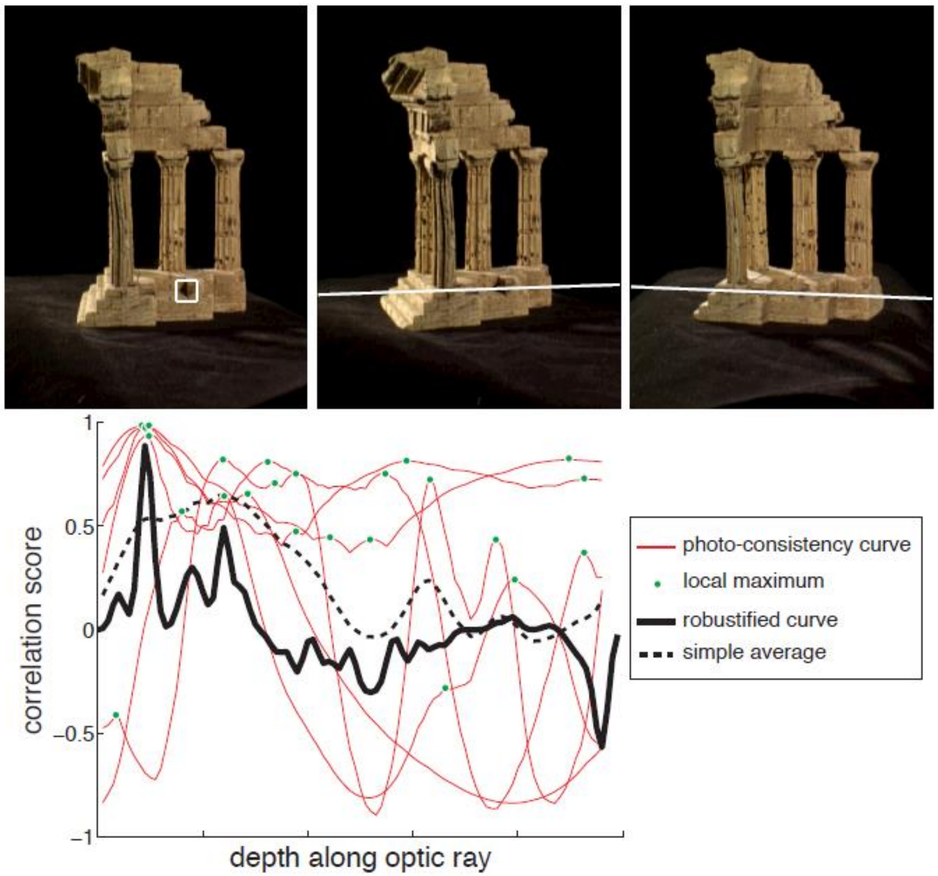
最优Depth的选择：Winner-Takes-All



Y. Furukawa. Foundations and Trends in Computer Graphics and Vision 2015.

稠密重建(Multiple View Stereo)

最优Depth的选择：Robust Photo-Consistency



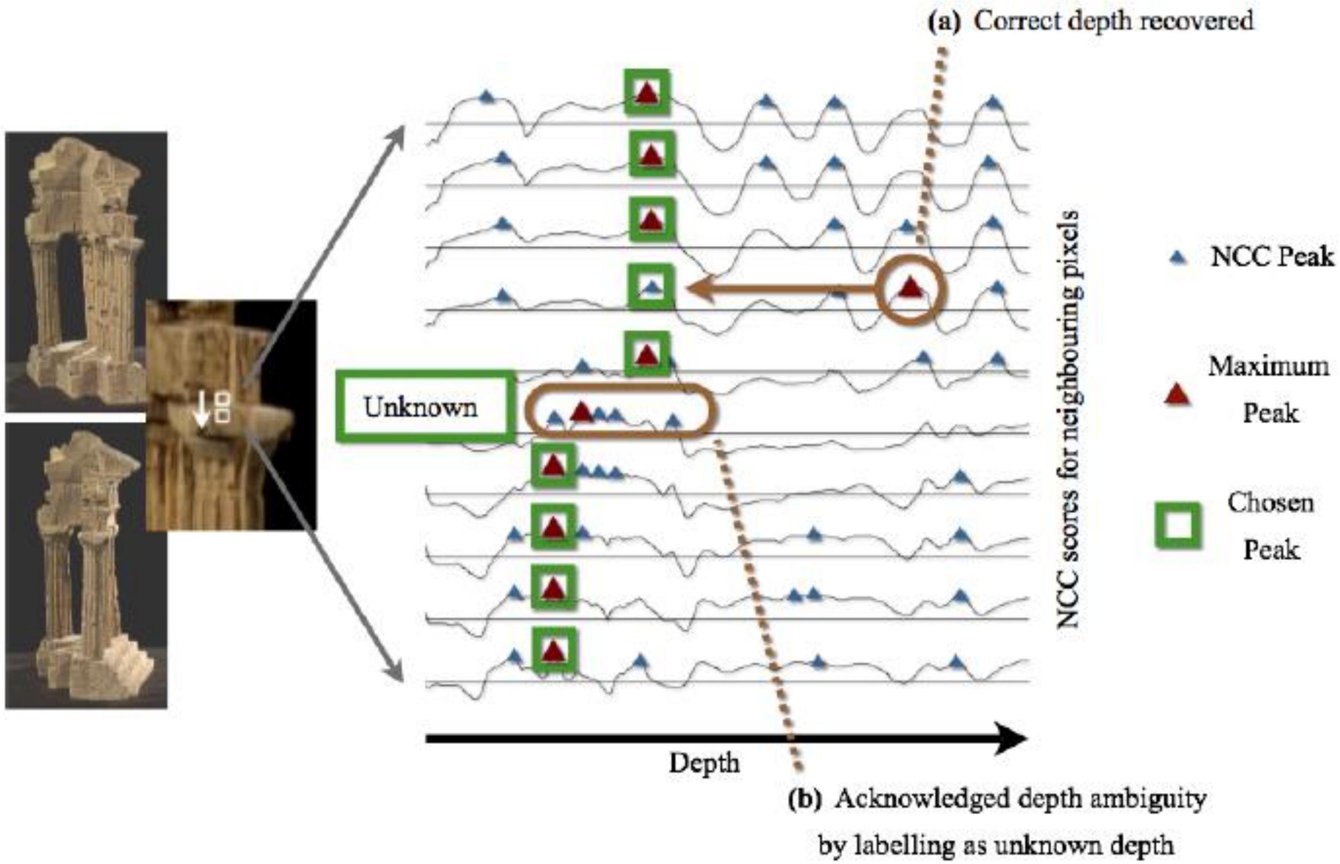
$$C^R(d) = \sum_k C_k W(d - d_k)$$

Parzen窗估计

Y. Furukawa. Foundations and Trends in Computer Graphics and Vision 2015.

稠密重建(Multiple View Stereo)

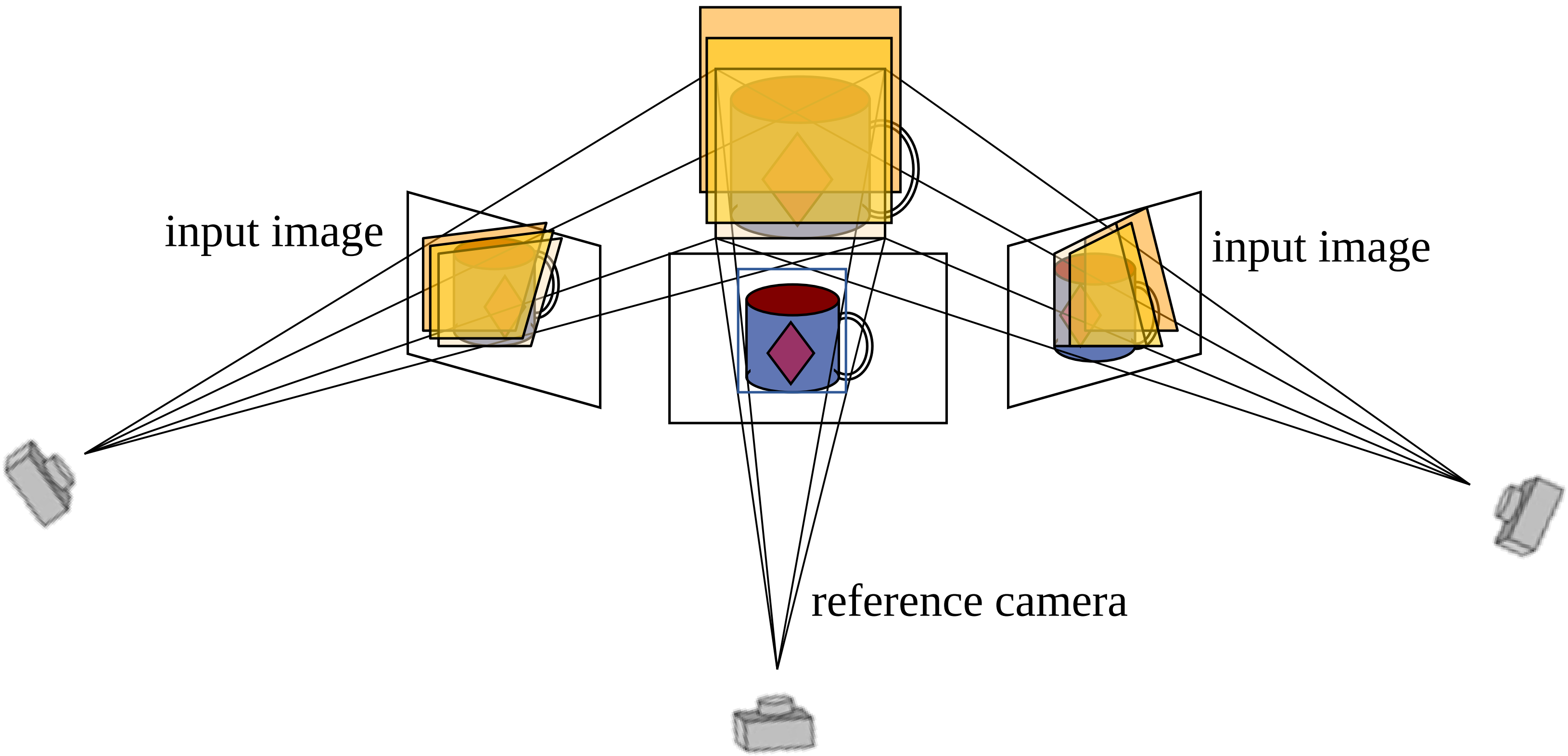
最优Depth的选择：Markov Random Field



Y. Furukawa. Foundations and Trends in Computer Graphics and Vision 2015.

稠密重建(Multiple View Stereo)

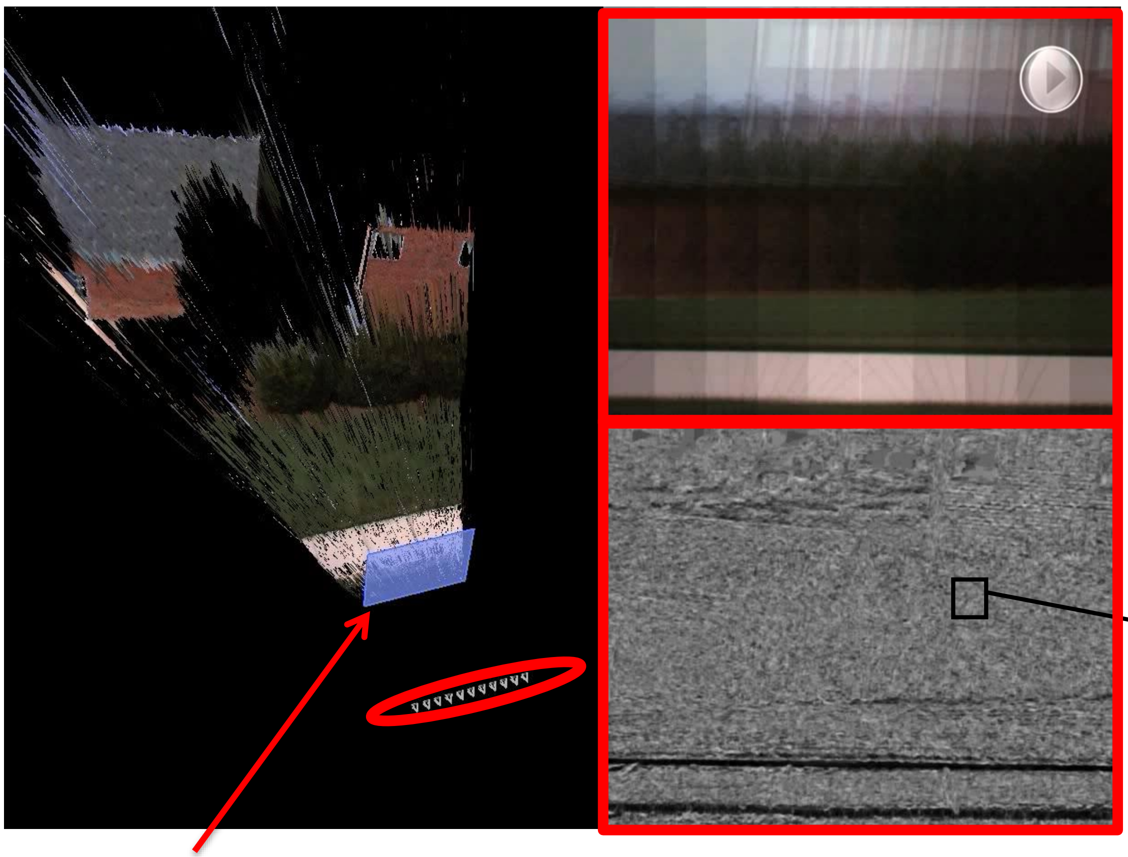
深度图计算方法一：Plane Sweeping



空间平面诱导的单应变换： $H_{ij} = K_j(R_j R_i^{-1} + \frac{R_j(C_i - C_j)n_i^T}{n_i^T X_i})K_i^{-1}$

稠密重建(Multiple View Stereo)

深度图计算方法一：Plane Sweeping



warped images

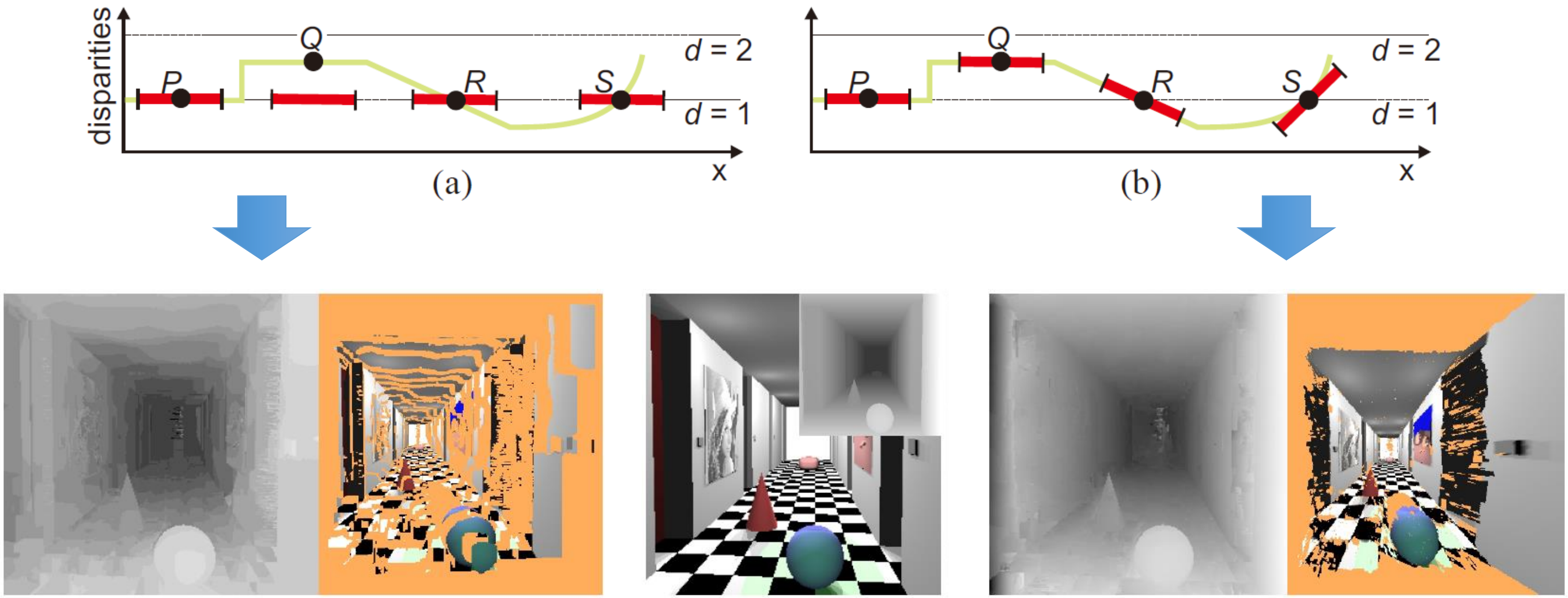
SAD as similarity
(darker is higher
similarity)

3D plane

D. Gallup, et. al. CVPR 2007.

稠密重建(Multiple View Stereo)

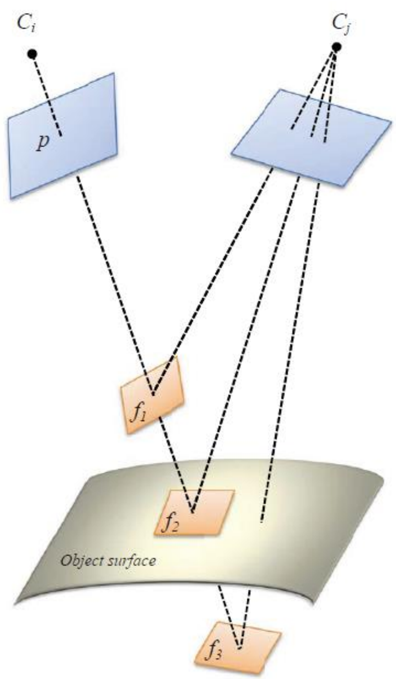
深度图计算方法二：PathMatch

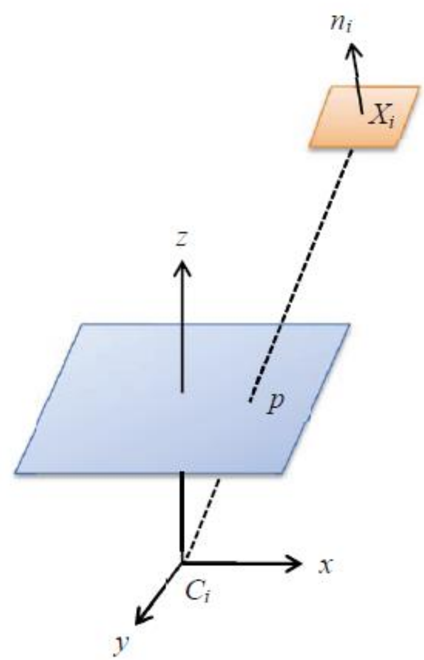


C. Barnes, et. al. SIGGRAPH 2009.
M. Bleyer, et. al. BMVC 2011.

稠密重建(Multiple View Stereo)

深度图计算方法二：PathMatch

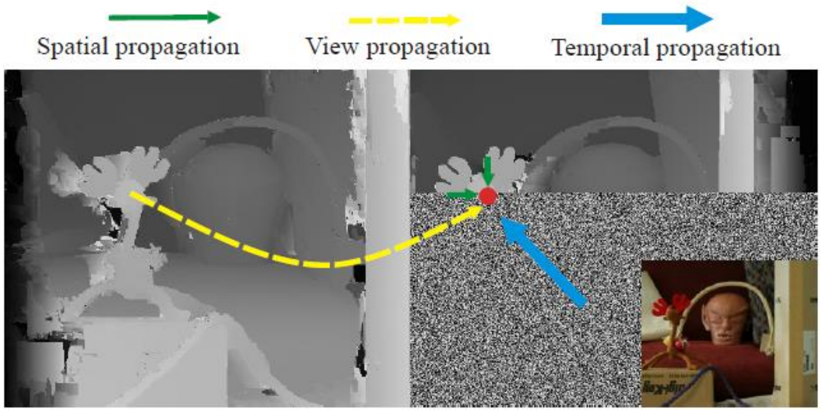




相机C_i坐标系下空间面片(Patch)的表达：d (1自由度), n (2个自由度)

稠密重建(Multiple View Stereo)

两视图PathMatch Stereo：

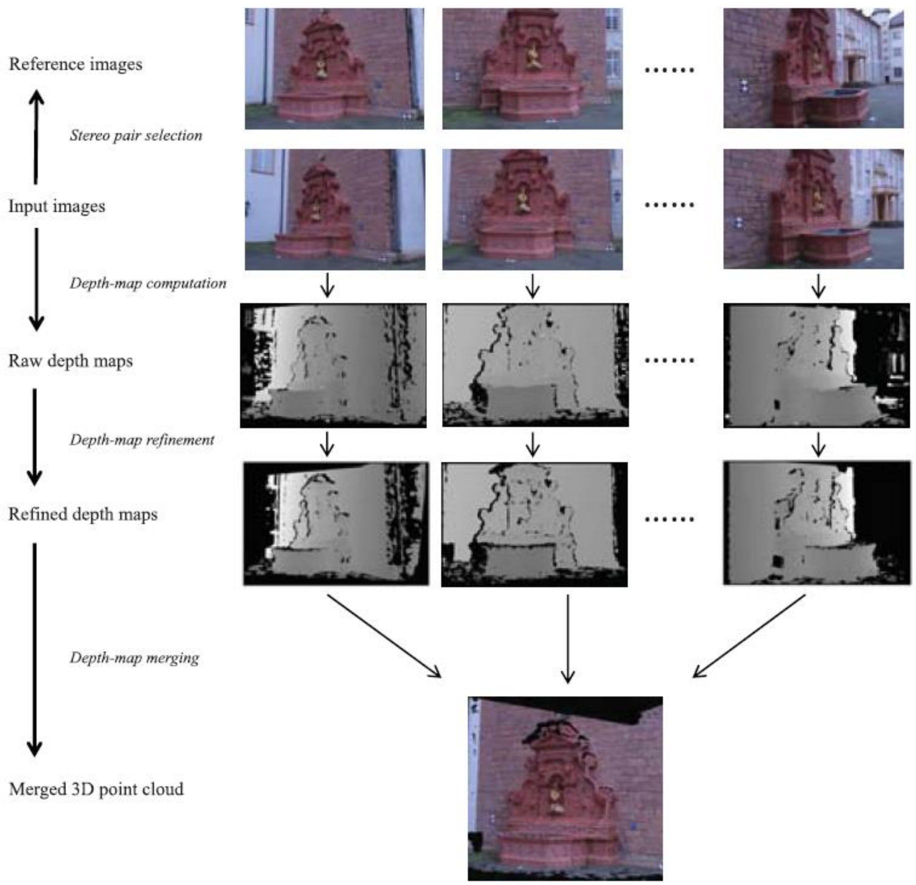


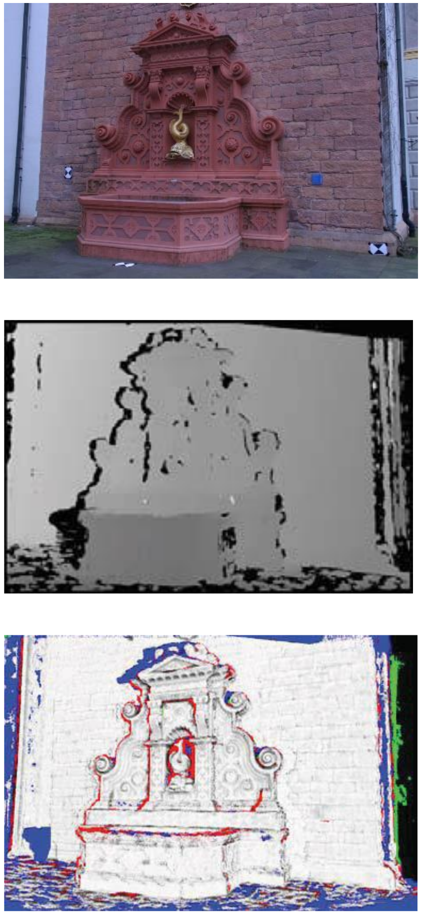
1. 随机生成每个像素点的深度和法向
2. 从左上向右下传播：
   1）检测邻域点的深度和法向是否更好
   2）检测自身随机扰动后的点是否更好
   3）检测立体图像对对应点是否更好
   4）检测前后帧同一位置点是否够更好
3. 从右下向左上再传播一次

Based on the law of large numbers！

稠密重建(Multiple View Stereo)

多视图PathMatch MVS：

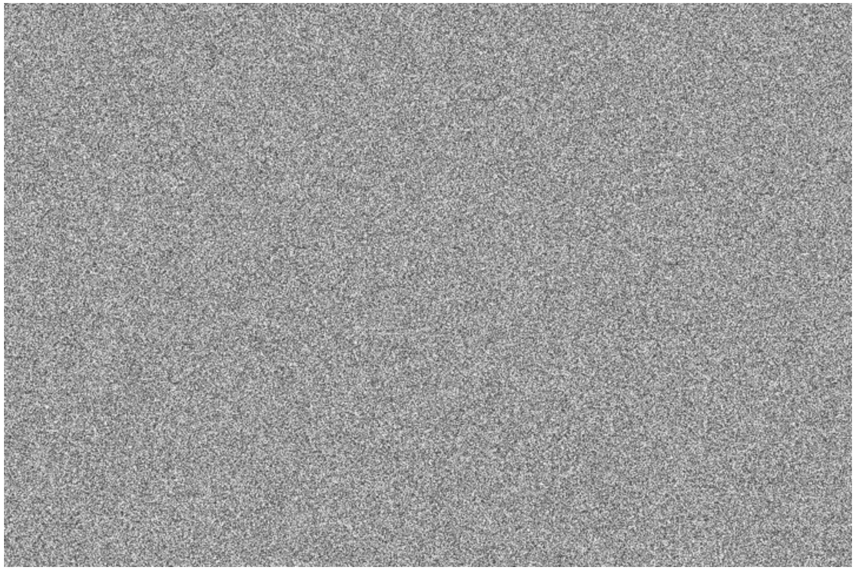




S. Shen. IEEE T-IP, 2013.

稠密重建(Multiple View Stereo)

多视图PathMatch MVS：



稠密重建(Multiple View Stereo)

多视图PathMatch MVS：

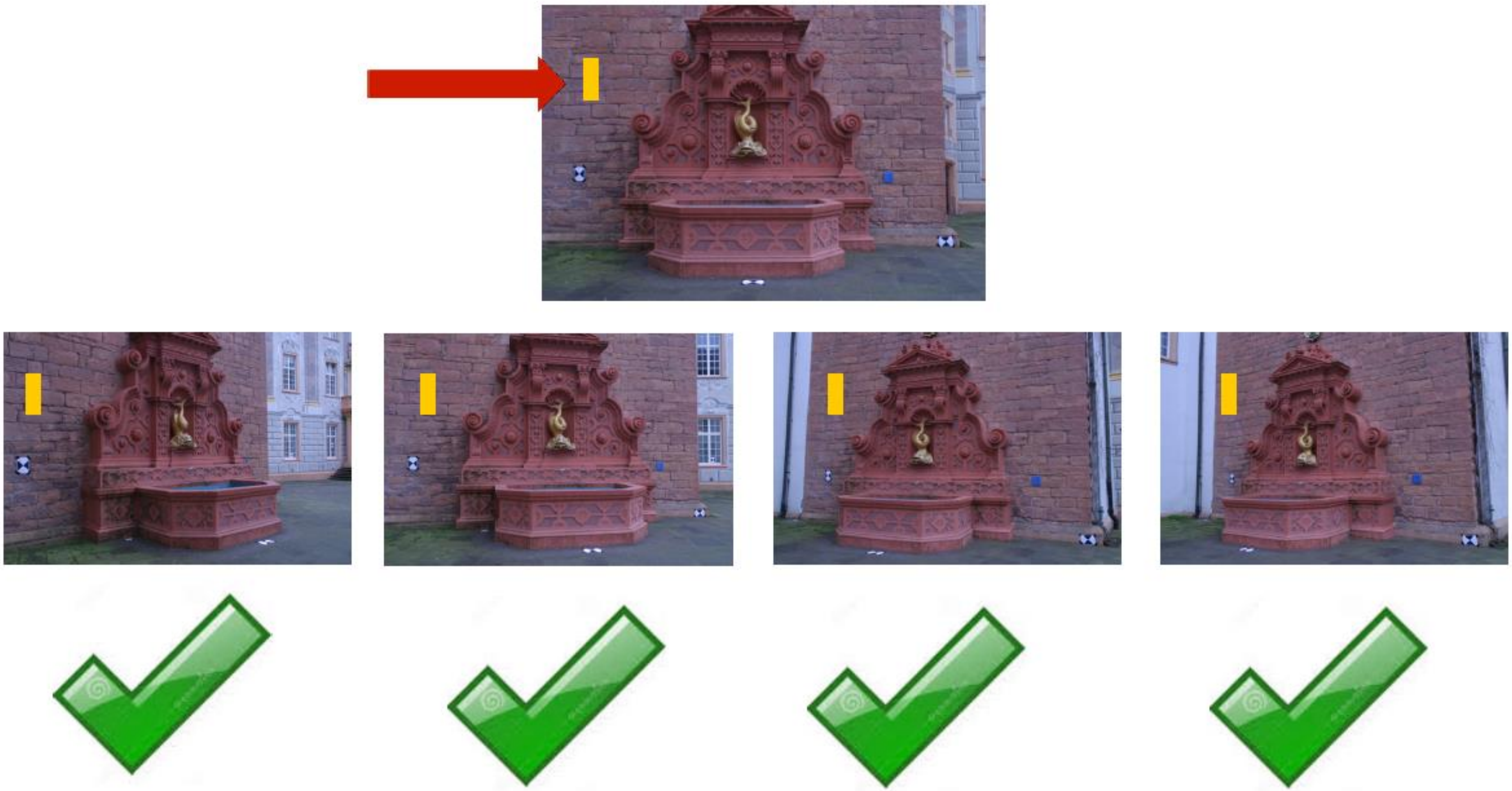
稠密重建(Multiple View Stereo)

逐像素点选择邻域图像组:

E. Zheng, et. al. CVPR 2014.
J. L. Schönberger, et. al. ECCV 2016.

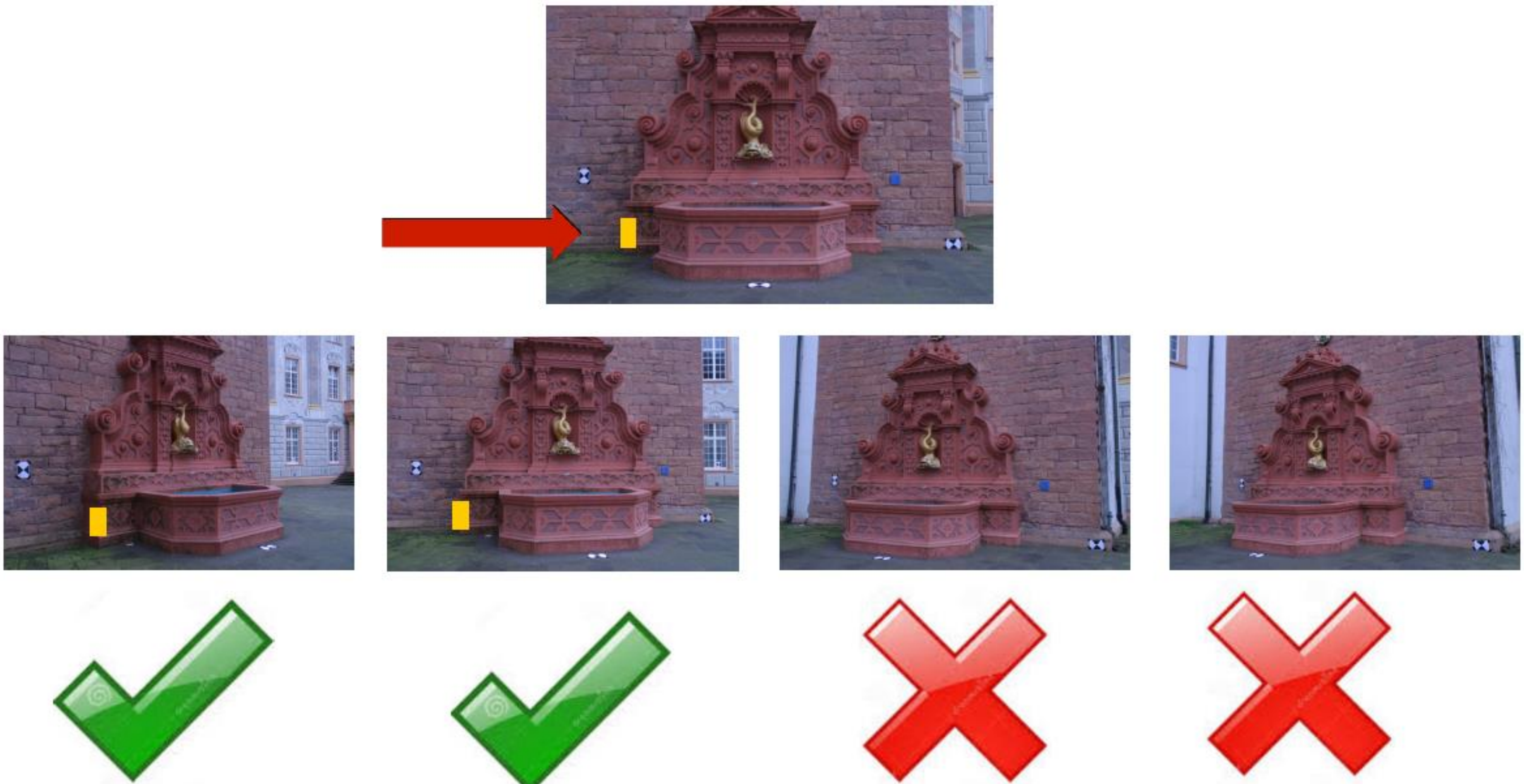
稠密重建(Multiple View Stereo)

逐像素点选择邻域图像组:



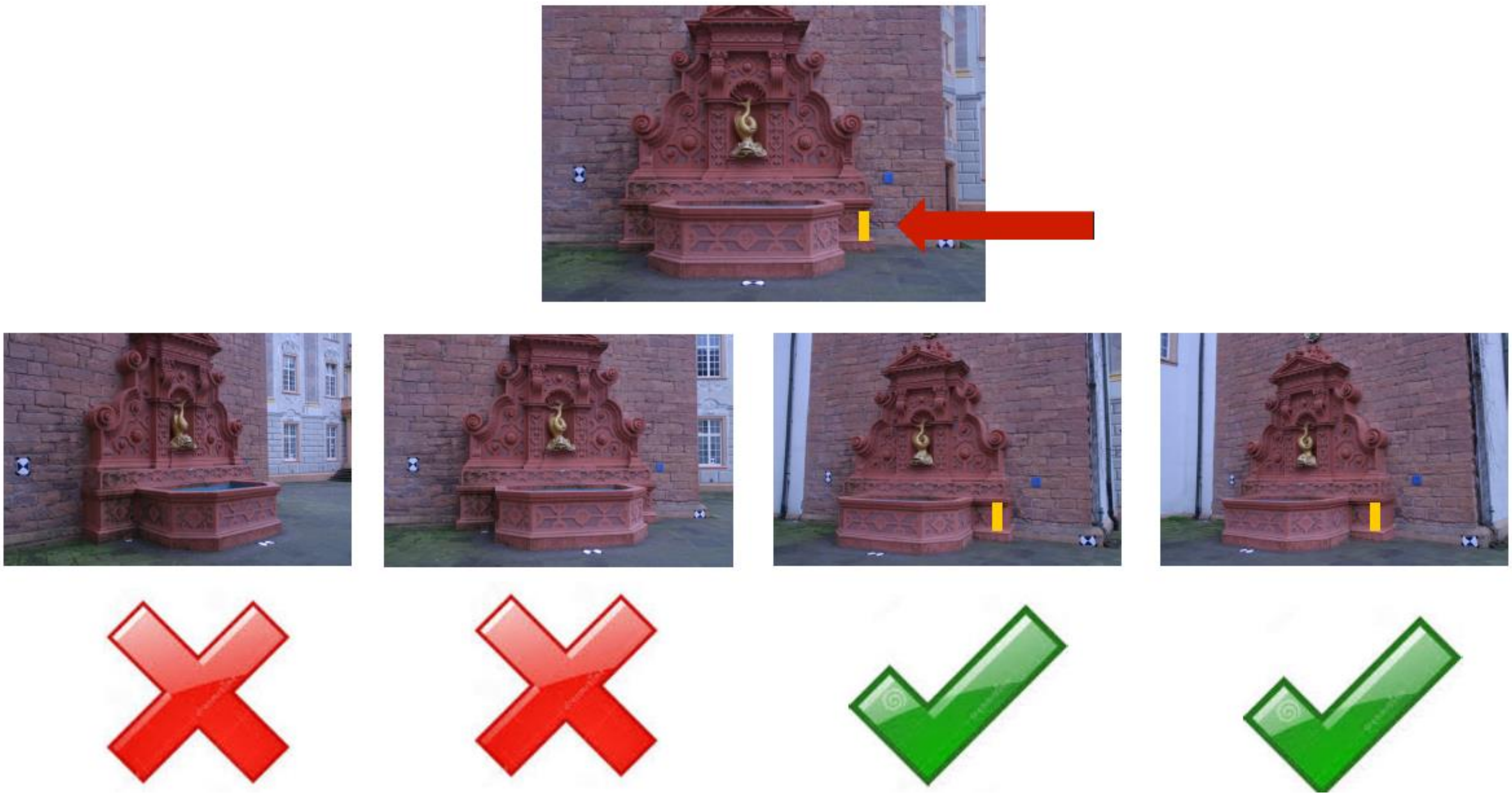
稠密重建(Multiple View Stereo)

逐像素点选择邻域图像组:



稠密重建(Multiple View Stereo)

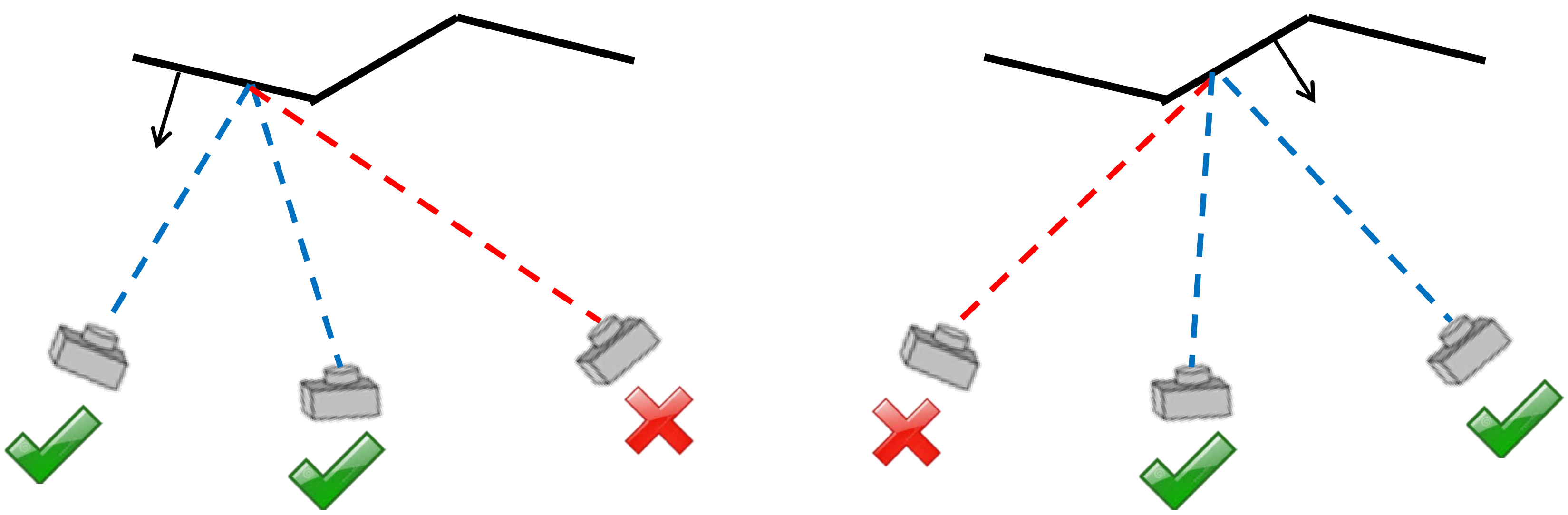
逐像素点选择邻域图像组:



稠密重建(Multiple View Stereo)

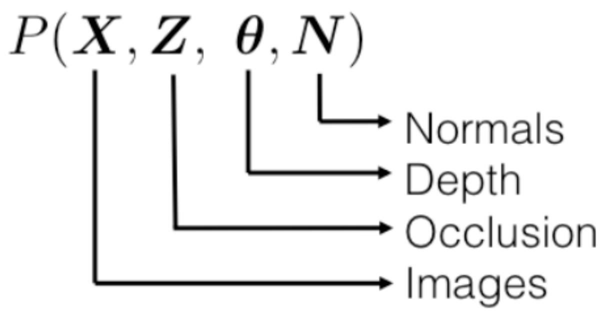
逐像素点选择邻域图像组:

- 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



稠密重建(Multiple View Stereo)

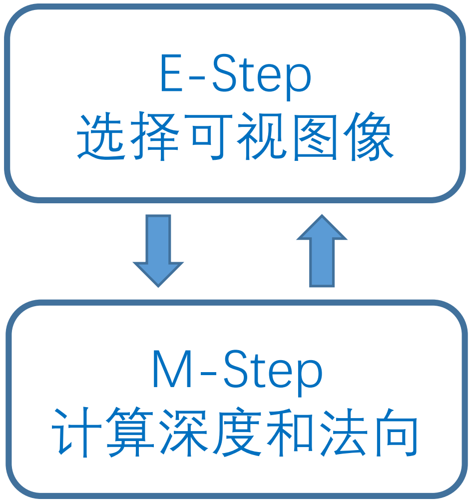
逐像素点选择邻域图像组:



Generalized Expectation Maximization:

- E-Step
  使用变分推断 $\boldsymbol{Z}$

- M-Step
  使用PatchMatch推断 $\boldsymbol{\theta}, \boldsymbol{N}$



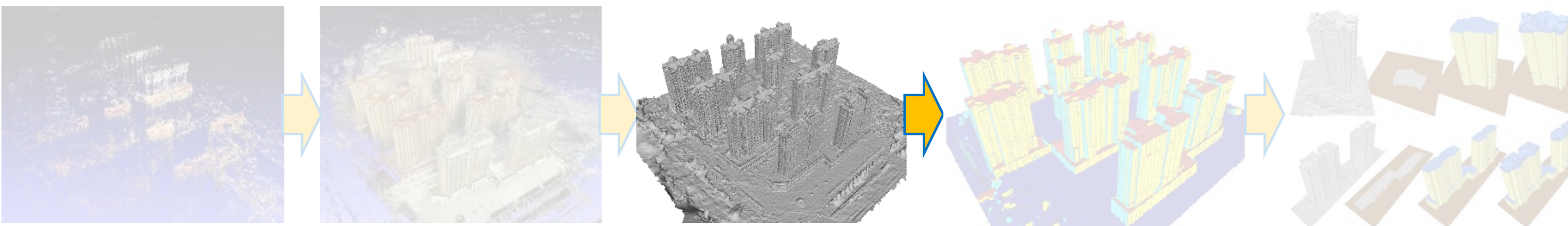
J. L. Schönberger, et. al. ECCV 2016.

相关文献

1. G. Vogiatzis, et. al. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. IEEE T-PAMI, 2007.
2. S.N. Sinha, et. al. Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh. ICCV 2007.
3. D. Gallup, et. al. Real-Time Plane-sweeping Stereo with Multiple Sweeping Directions. CVPR 2007.
4. Y. Furukawa and Jean Ponce. Accurate, dense, and robust Multiview stereopsis.. IEEE T-PAMI, 2010.
5. H. Vu, et. al. High accuracy and visibilityconsistent dense multiview stereo. IEEE T-PAMI , 2012.
6. M. Bleyer, et. al. Patchmatch stereostereo matching with slanted support windows. BMVC, 2011.
7. S. Shen. Accurate multiple view 3d reconstruction using patch-based stereo for large-scale scenes. IEEE T-IP, 2013.
8. S. Galliani, et. al. Massively parallel multiview stereopsis by surface normal diffusion. ICCV, 2015.
9. J. L. Schonberger, et. al. Pixelwise view selection for unstructured multi-view stereo. ECCV, 2016.
10. Q. Xu and W. Tao. Multi-Scale Geometric Consistency Guided Multi-View Stereo. CVPR, 2019
11. P-H. Huang, et. al. DeepMVS: Learning Multi-view Stereopsis, CVPR 2018. (CNN Based)
12. Y. Yao, et al. MVSNet: Depth Inference for Unstructured Multi-view Stereo, ECCV 2018. (CNN Based)
13. Y. Yao, et al. Recurrent MVSNet for High-resolution Multi-view Stereo Depth Inference, CVPR 2019. (CNN Based)

网格建模(Mesh Modeling)

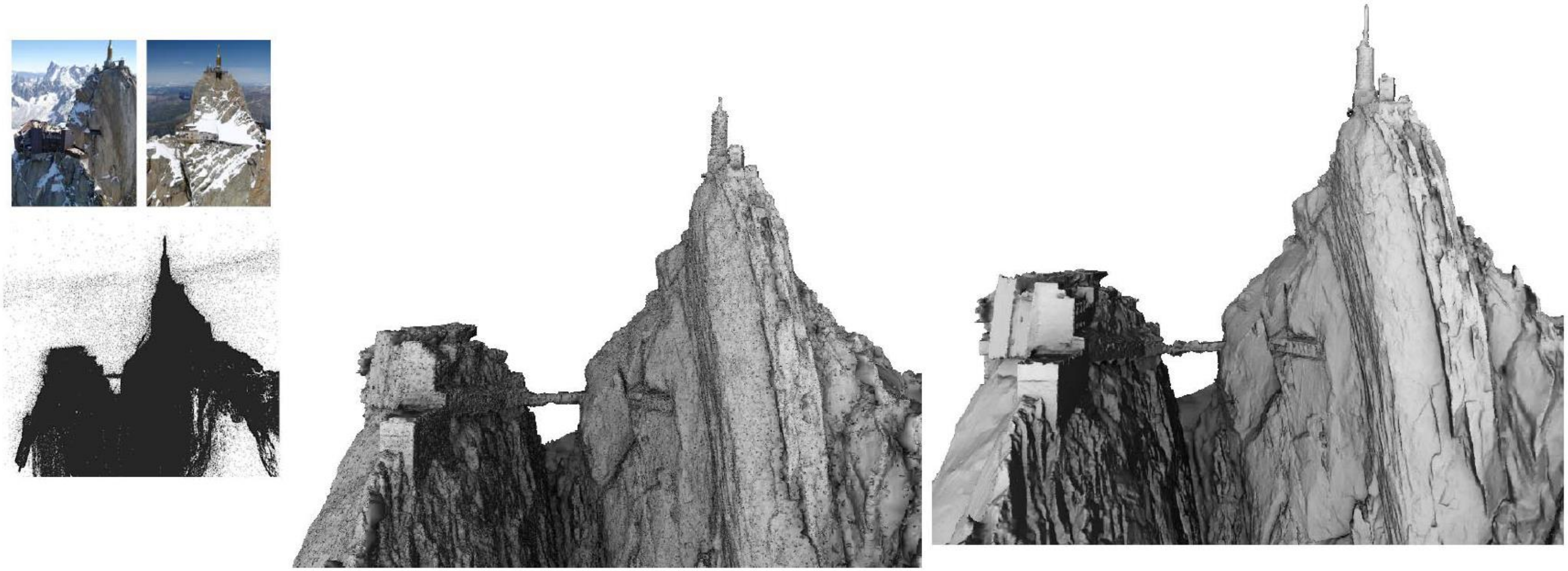
稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模

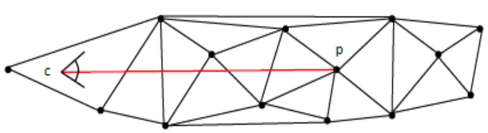


输入：稠密点云、相机位姿

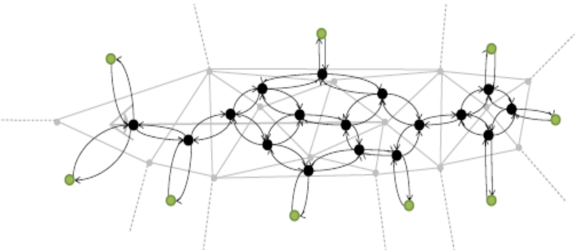
输出：三角网格

网格建模(Mesh Modeling)

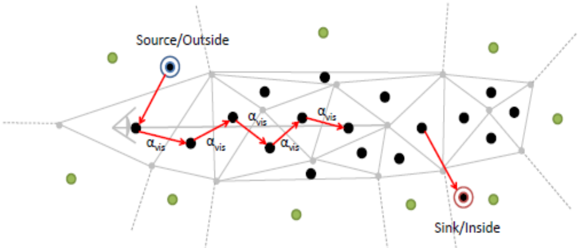




(a) Delaunay Triangulation



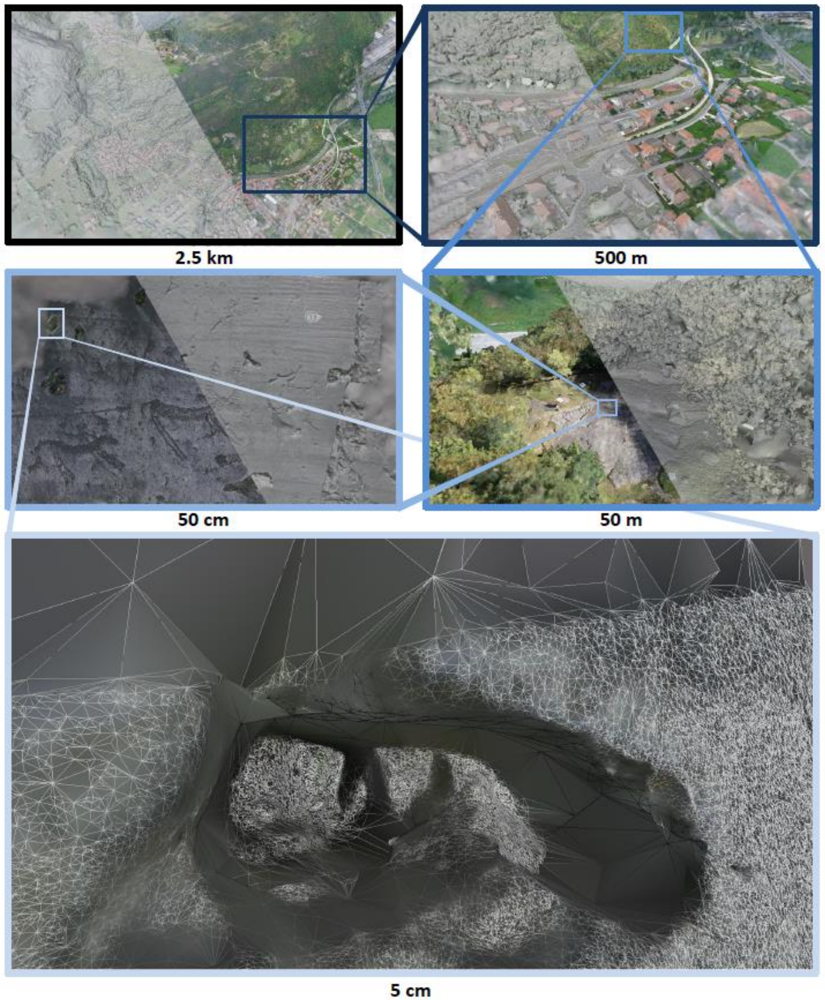
(b) Dual Graph

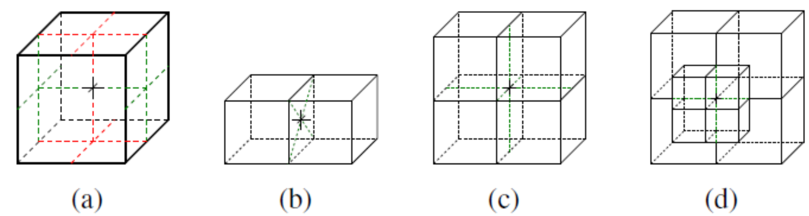


(c) Visibility Terms

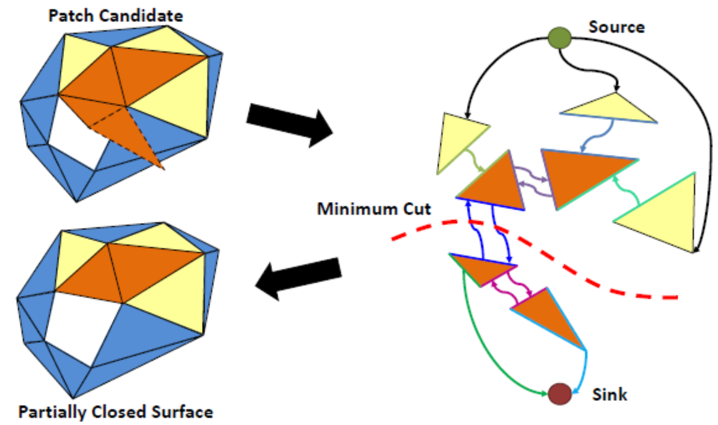
H.H. Vu, et al. IEEE T-PAMI 2012.

网格建模(Mesh Modeling)





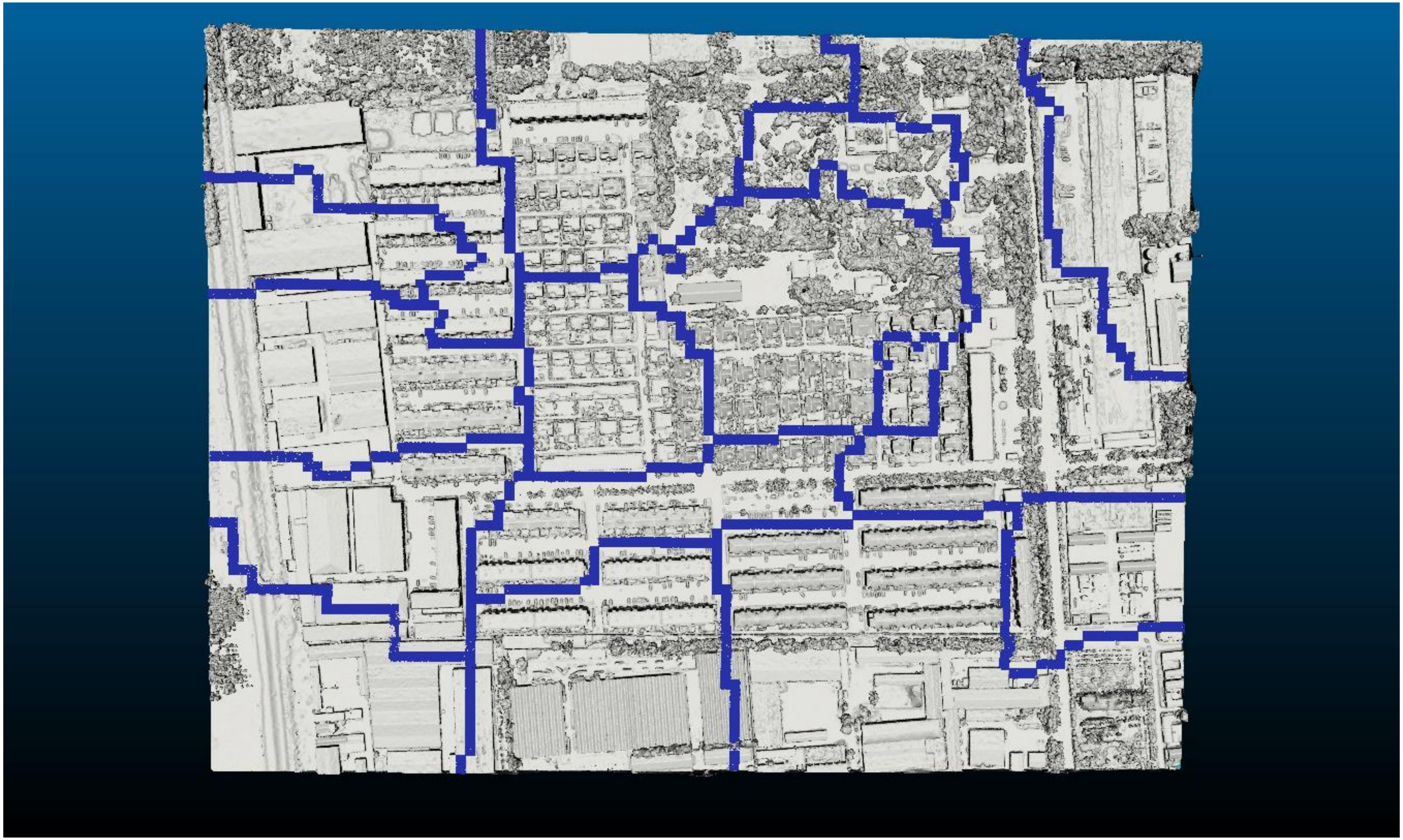
Four types of inner points



Hole filling

C. Mostegel, et al. CVPR 2017.

网格建模(Mesh Modeling)



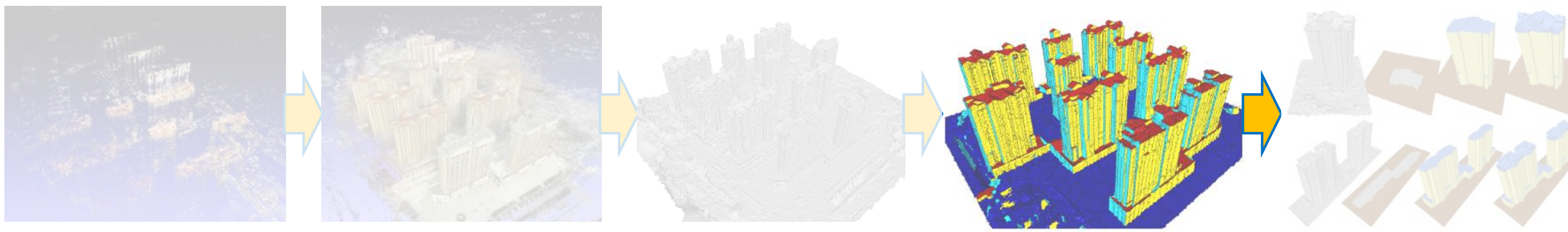
J. Han and S. Shen. Submitted 2019

相关文献

1. H.H. Vu, et al. High Accuracy and Visibility-Consistent Dense Multiview Stereo. IEEE T-PAMI 2012.
2. M. M. Kazhdan, et. al. Screened Poisson Surface Reconstruction. ACM TOG, 2013.
3. S. Fuhrmann and M. Goesele. Floating Scale Surface Reconstruction. ACM TOG, 2014.
4. B. Ummenhofer and T. Brox. Global, Dense Multiscale Reconstruction for a Billion Points. IJCV, 2017.
5. C. Mostegel, et al. Scalable Surface Reconstruction from Point Clouds with Extreme Scale and Density Diversity. CVPR 2017.

语义建模(Semantic Modeling)

稀疏点云重建 → 稠密点云重建 → 点云网格建模 → **三维语义建模** → 三维矢量建模

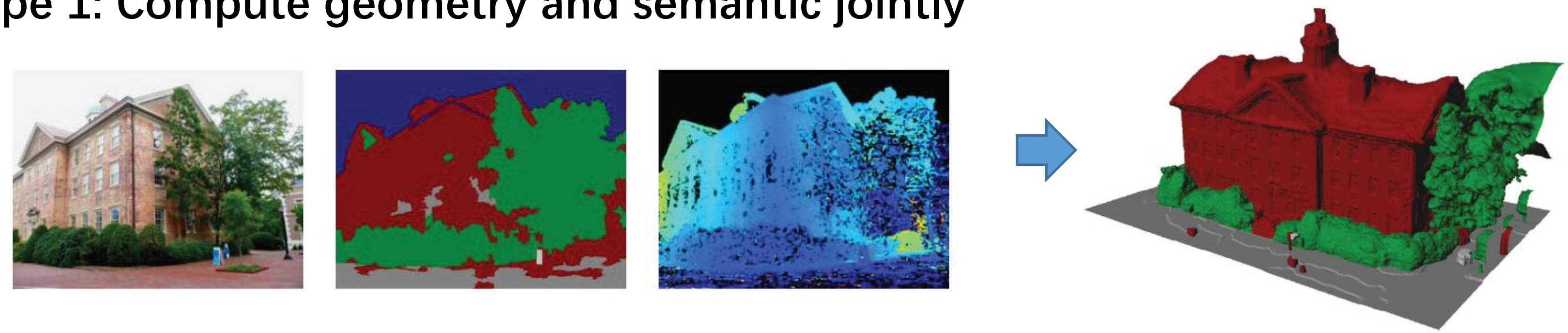


输入： 网格/点云、多视角图像

输出： 网格/点云语义分割

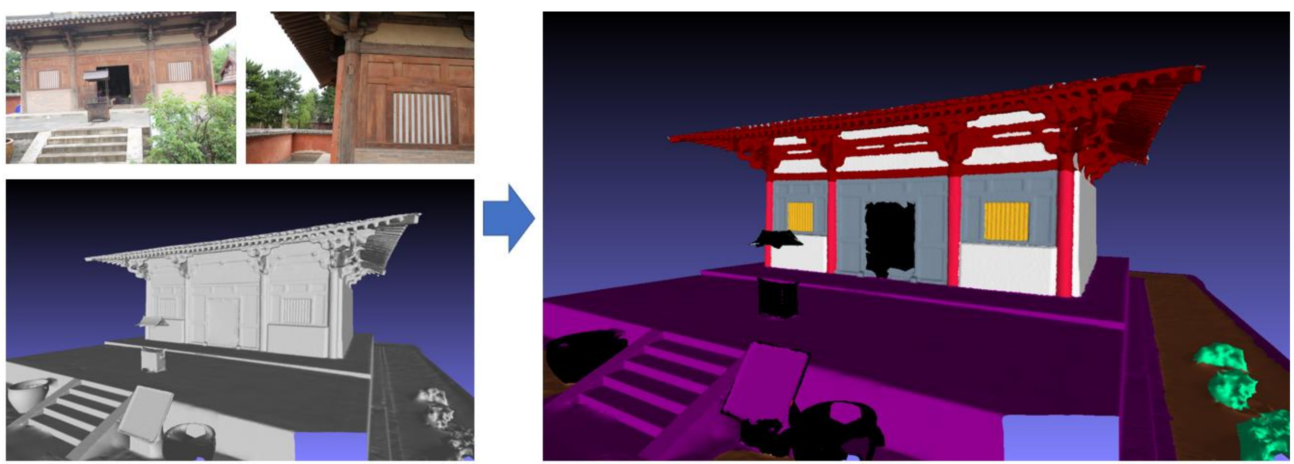
语义建模——类别语义

Type 1: Compute geometry and semantic jointly



C. Hane, et al. CVPR 2013.

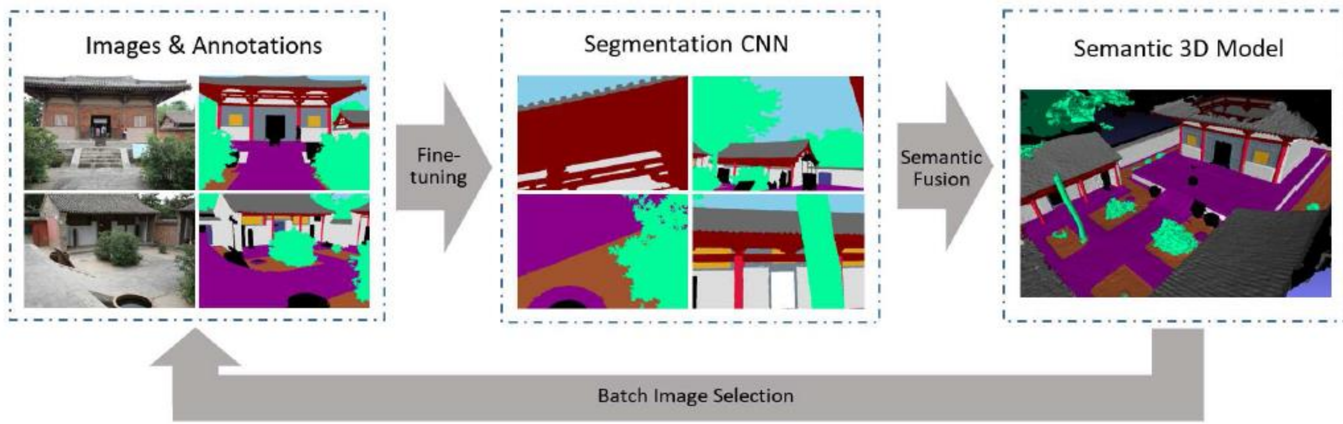
Type 2: Compute geometry first and semantic second



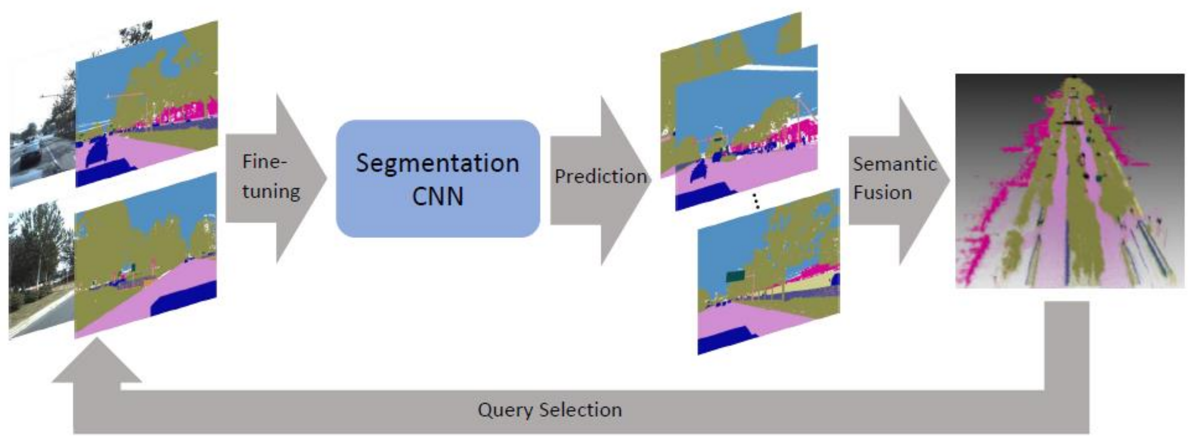
Y. Zhou, et al. 3DV 2018.

语义建模——类别语义

Mesh based:



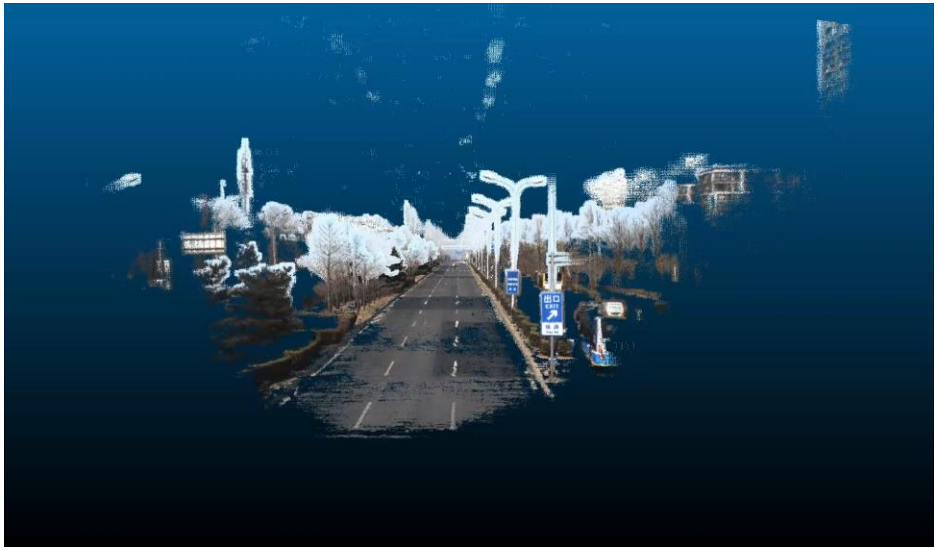
Voxel based:

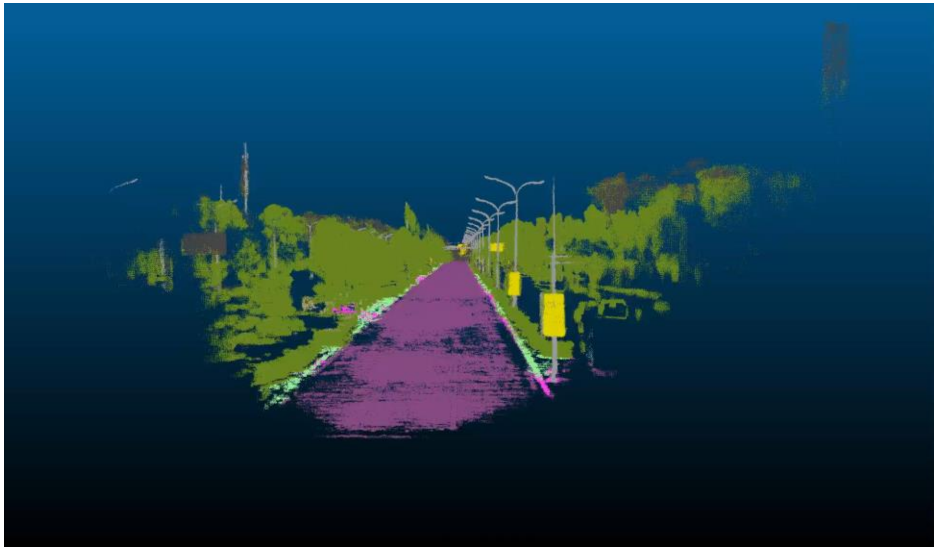


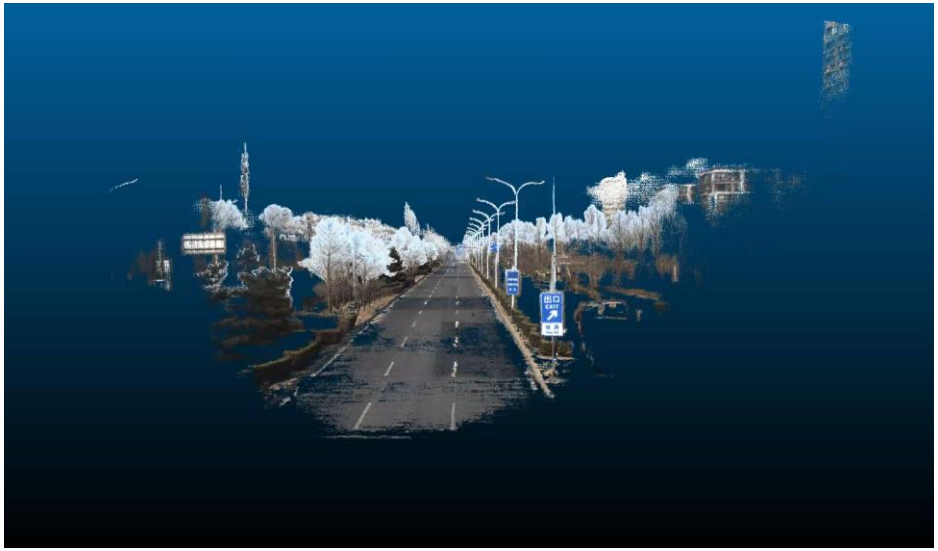
Active Learning Based 3D Semantic Modeling

Y. Zhou, et al. 3DV 2018.

语义建模——类别语义







相关文献

1. C. Hane, et al. Joint 3D scene reconstruction and class segmentation. CVPR 2013.
2. J. P. Valentin, et al. Mesh based semantic modelling for indoor and outdoor scenes. CVPR, 2013.
3. M. Bláha, et al. Large-Scale Semantic 3D Reconstruction: An Adaptive Multi-Resolution Model for Multi-Class Volumetric Labeling. CVPR 2016.
4. N. Savinov, et al. Semantic 3D Reconstruction with Continuous Regularization and Ray Potentials Using a Visibility Consistency Constraint. CVPR 2016.
5. M. Bláha, et al. Semantically informed multiview surface refinement. ICCV, 2017.
6. J. McCormac, et al. Dense 3d semantic mapping with convolutional neural networks. ICRA, 2017.
7. M. Rouhani, et al. Semantic segmentation of 3d textured meshes for urban scene analysis. ISPRS JPRS, 2017.
8. Y. Zhou, et al. Fine-Level Semantic Labeling of Large-Scale 3D Model by Active Learning. 3DV 2018.

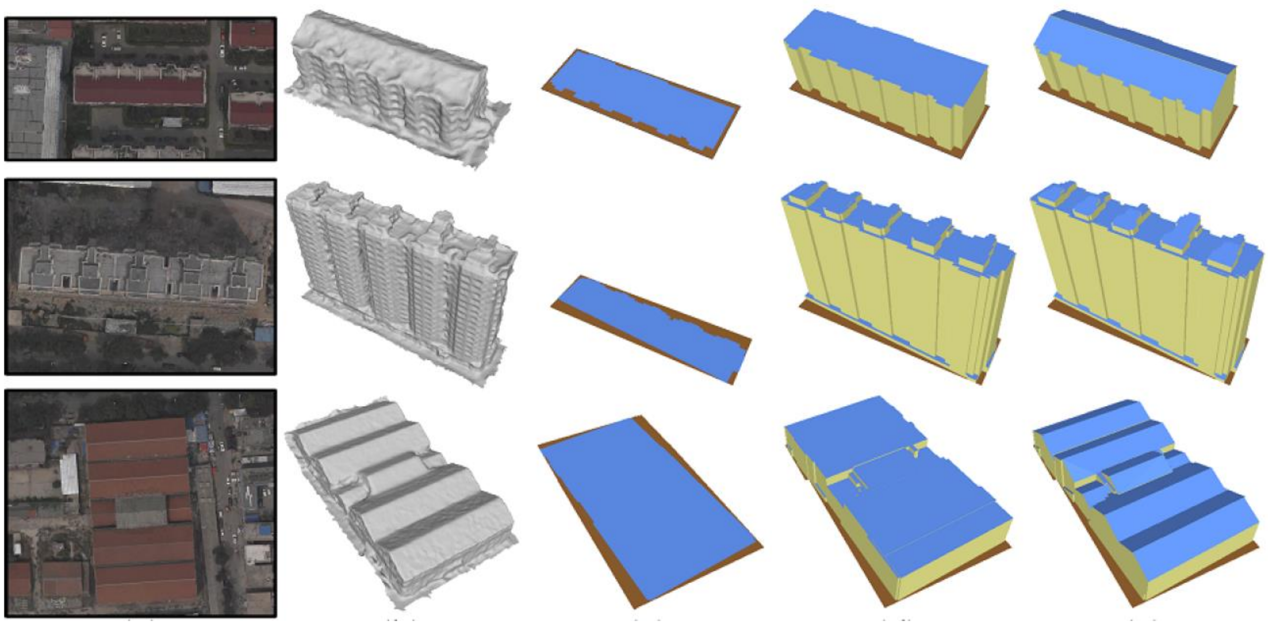
矢量建模(Vector Modeling)

稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模

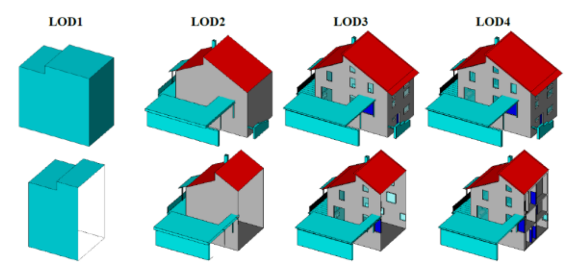
输入：单体部件网格、多视角图像

输出：矢量三维模型

矢量建模(Vector Modeling)

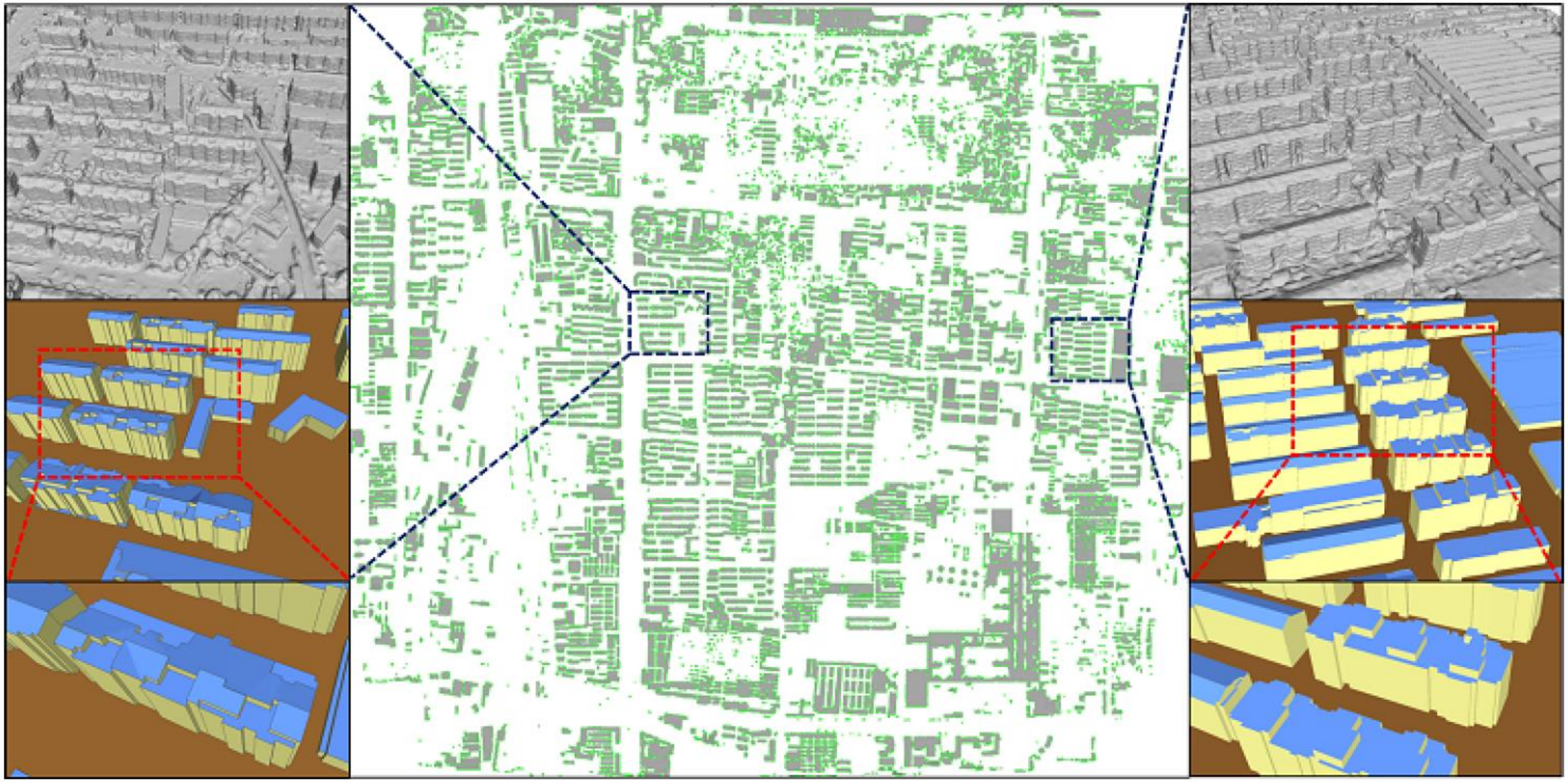


CityGML Level LOD Modeling of City Scene



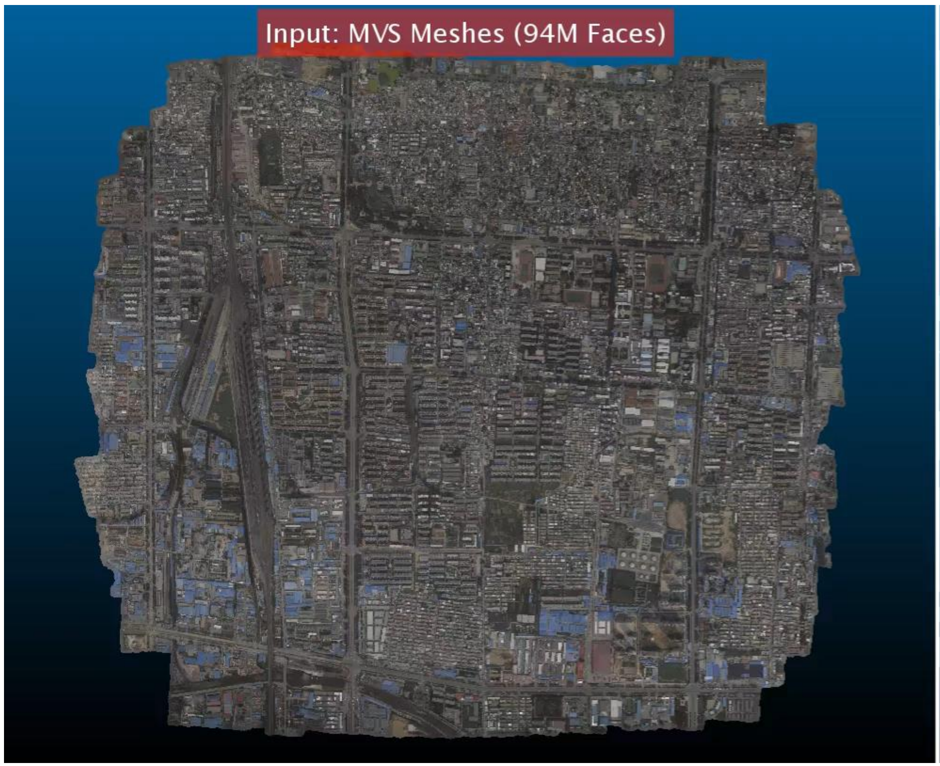
L. Zhu, et al. 3DV 2017.
L. Zhu, et al. ECCV 2018.

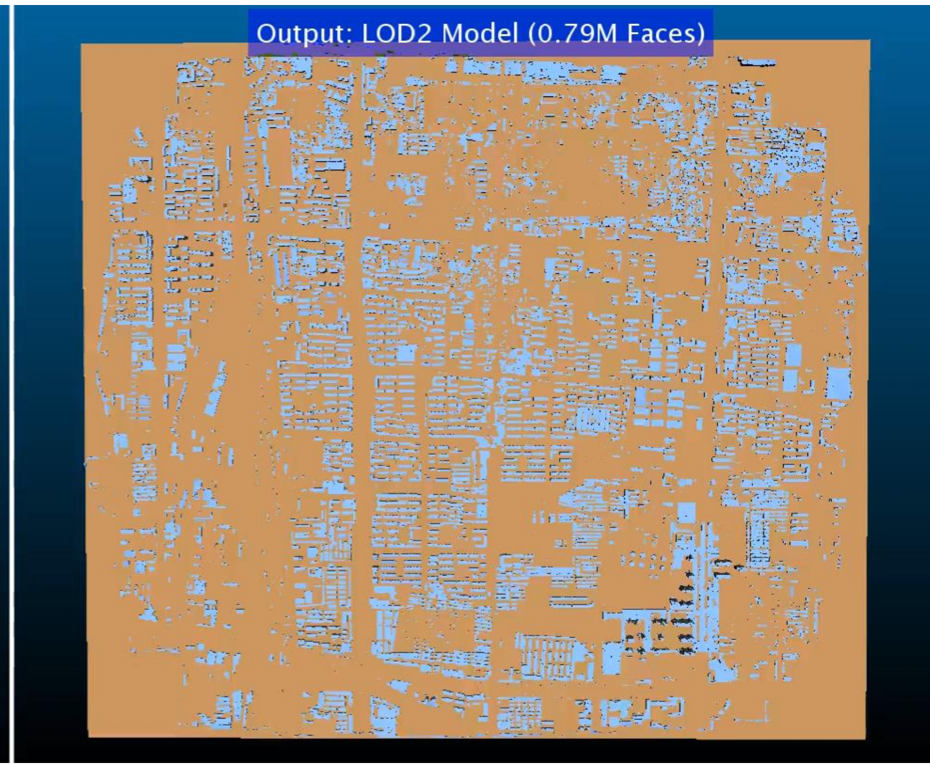
矢量建模(Vector Modeling)



Input: 94M triangle faces covering an area of 12.2km2
Output: 4343 LOD2 buildings with 0.79M triangle faces

矢量建模(Vector Modeling)





Input: 94M triangle faces covering an area of 12.2km2
Output: 4343 LOD2 buildings with 0.79M triangle faces

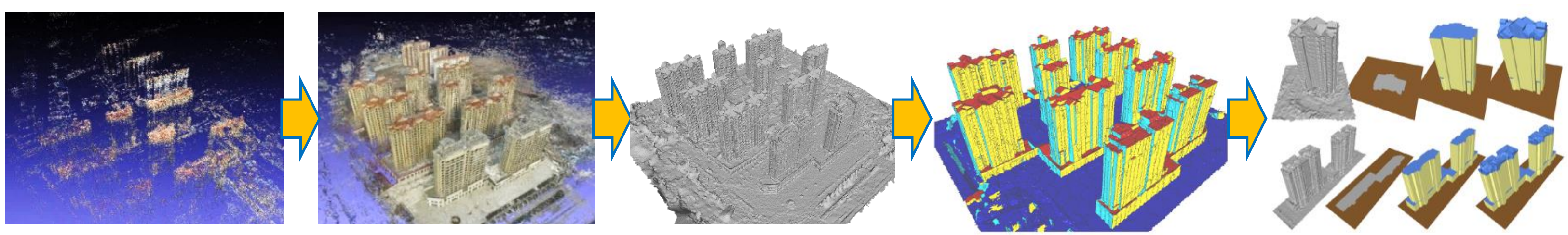
相关文献

1. D. Salinas, et al. Structure-Aware Mesh Decimation. CGF 2015.
2. Y. Verdie, et al. Lod generation for urban scenes. ACM TOG, 34(3): 1–30, 2015.
3. M. Li, et al. Manhattan-World Urban Reconstruction from Point Clouds. ECCV 2016.
4. M. Li, et al. Reconstructing Building Mass Models from UAV images. CAG 2016.
5. H. Thomas, et al. Regularized 3D Modeling from Noisy Building Reconstructions. 3DV 2016.
6. L. Nan, et al. Polyfit: Polygonal surface reconstruction from point clouds. ICCV 2017.
7. T. Kelly, et al. BigSUR: Large-scale Structured Urban Reconstruction. SIGGRAPH Asia 2017.
8. L. Zhu, et al. Variational Building Modeling from Urban MVS Meshes. 3DV 2017.
9. L. Zhu, et al. Large Scale Urban Scene Modeling from MVS Meshes. ECCV 2018.
10. H. Zeng, et al. Neural Procedural Reconstruction for Residential Buildings. ECCV 2018.
11. C. Liu, et al. PlaneNet: Piece-wise Planar Reconstruction from a Single RGB Image. CVPR 2018. (CNN Based)
12. C. Liu, et al. PlaneRCNN: 3D Plane Detection and Reconstruction from a Single Image. CVPR 2019. (CNN Based)

总　结

- 基于图像的三维建模涉及到**图像特征提取、描述、匹配、SfM、MVS、Meshing、Semantic & Vector Modeling**等系列环节，每一个环节都尚有很多待解决的问题；
- 从规范拍摄图像（如航拍倾斜摄影）到**几何模型**的构建相对已经较成熟（商业软件：Smart3D、Pixel4D、Altizure等，开源软件：Colmap、OpenMVG、OpenMVS等）；
- 从几何模型到**语义矢量模型**的自动构建还有很长的距离。

稀疏点云重建 → 稠密点云重建 → 点云网格建模 → 三维语义建模 → 三维矢量建模



Thanks !