



深蓝学院
shenlanxueyuan.com

基于图像的三维模型重建

——稠密点云重建



主讲人 隋博



课程内容

✓ 稠密点云的获取方式

- ✓ Lidar 扫描
- ✓ Kinect
- ✓ 结构光
- ✓ 基于图像的方法-Multi-view Stereo

✓ 基础知识

- ✓ 极线条搜索
- ✓ 光度一致性约束
- ✓ 可视性约束

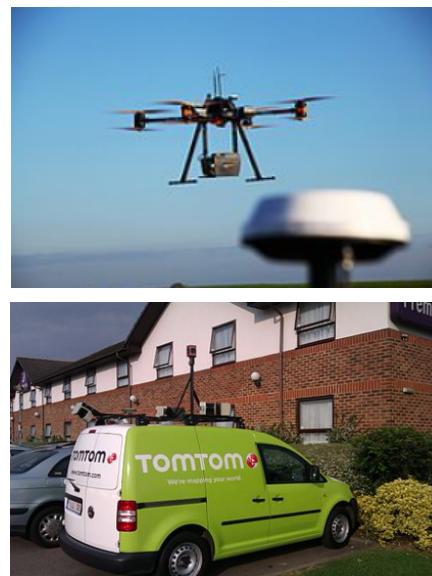
✓ 多视角立体技术

- ✓ 基于体素的方法
- ✓ 基于空间patch 的方法
- ✓ 深度图融合

稠密点云的获取方式

LiDAR扫描

精度高（毫米级别），效率高，有效范围几米到几百米，价格昂贵
高反光，玻璃表面，吸收表面

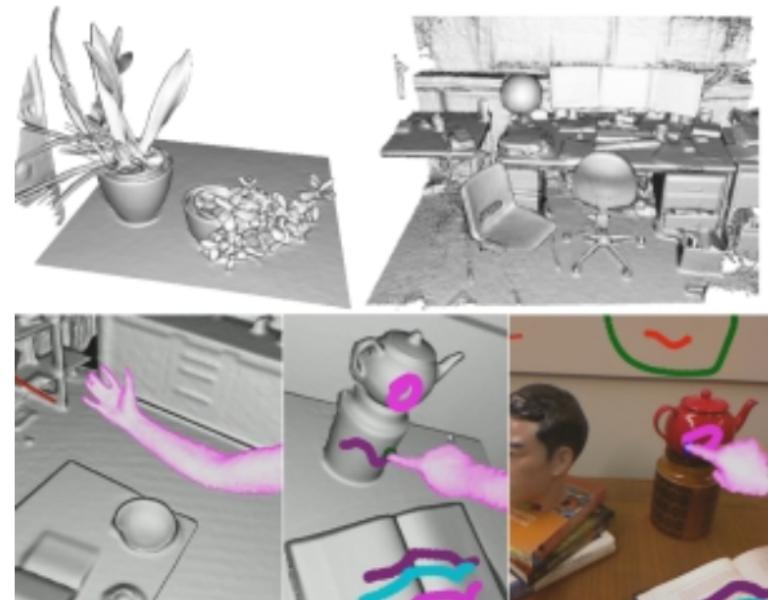


稠密点云的获取方式

Kinect

使用方便，价格适中，速度较快

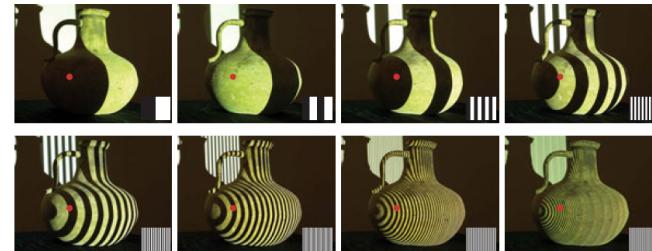
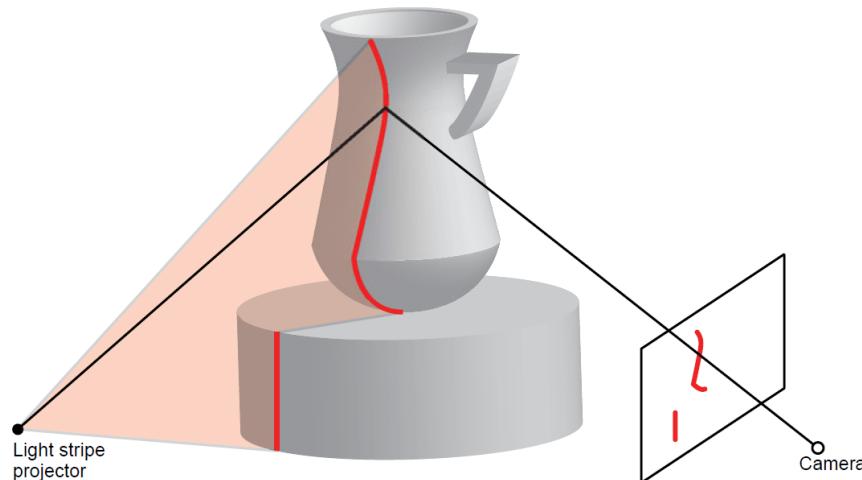
精度较低，有效距离短



稠密点云的获取方式

结构光(Structured Light)

高精度，高效率，近距离数据获取



基于灰度编码的二值样式



Hall-Holt and Rusinkiewicz's



Zhang et al彩色结构光

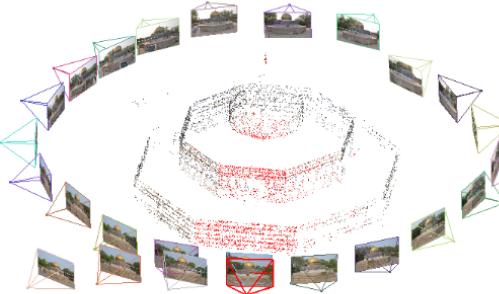
稠密点云的获取方式

基于图像的方法-Multi-view Stereo

无源被动式，成本低，图像来源广，计算速度慢，精度较高



输入无序图像



运动恢复结构(SFM)



多视角立体重建(MVS)

课程内容

✓ 稠密点云的获取方式

- ✓ Lidar 扫描
- ✓ Kinect
- ✓ 结构光
- ✓ 基于图像的方法-Multi-view Stereo

✓ 基础知识

- ✓ 极线条搜索
- ✓ 光度一致性约束
- ✓ 可视性约束

✓ 多视角立体技术

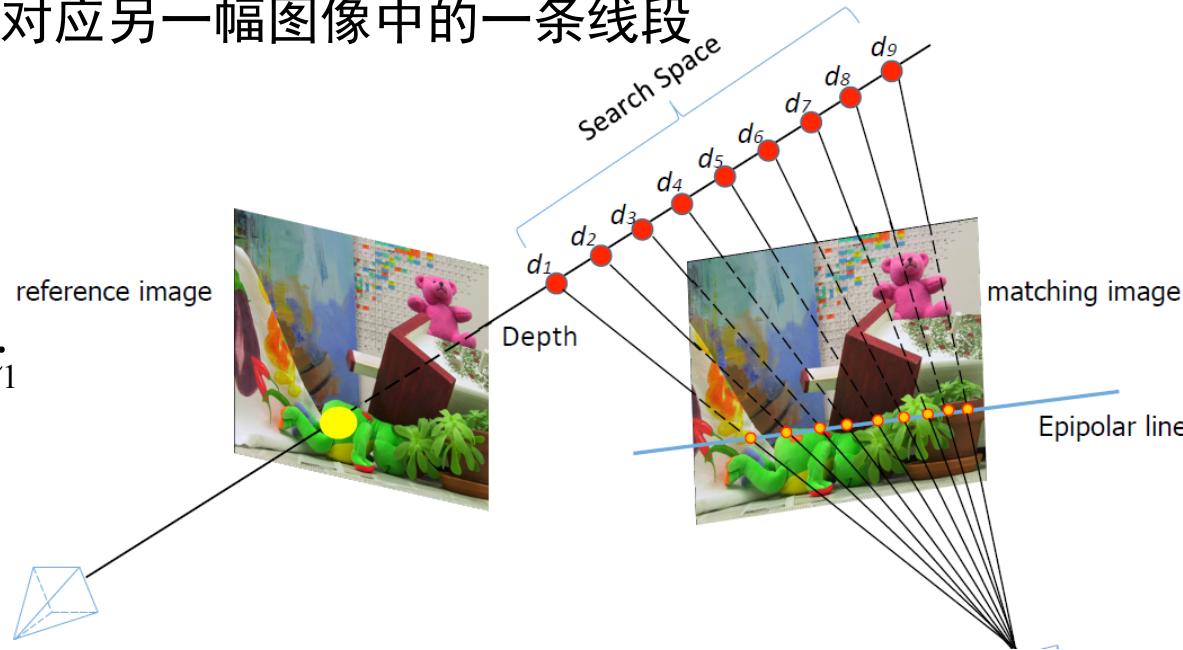
- ✓ 基于体素的方法
- ✓ 基于空间patch 的方法
- ✓ 深度图融合

基础知识

极线搜索

参考图像中的一点对应另一幅图像中的一条线段

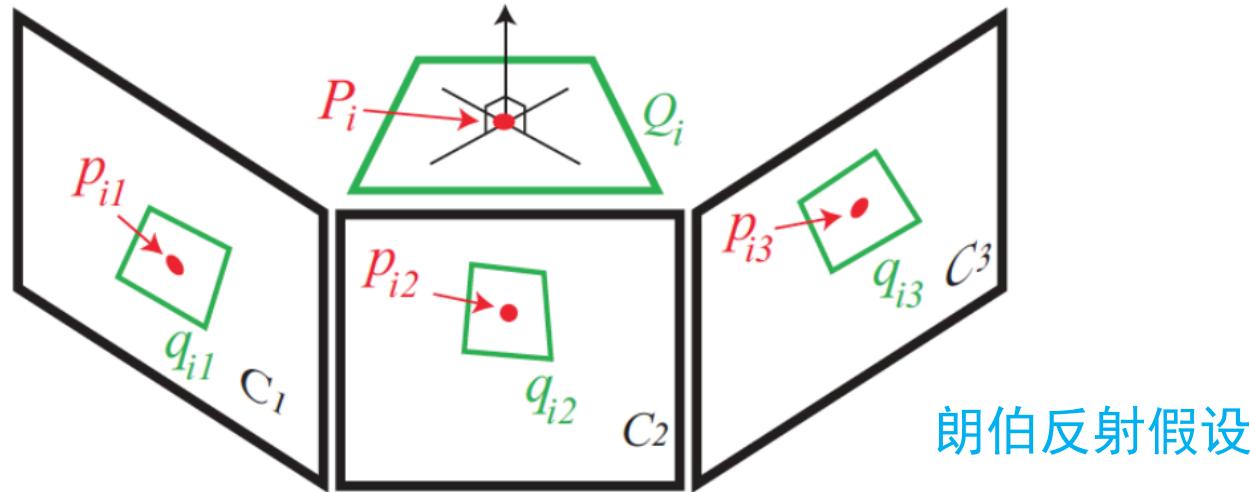
$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 = 0$$
$$l_1 = \mathbf{F}^T \mathbf{x}_2, \quad l_2 = \mathbf{F} \mathbf{x}_1$$



基础知识

光度一致性假设(Photo-consistency)

同一空间的点在不同视角的投影应当具有相同的光度，重建的核心在于恢复空间中具有光度一致性的点



基础知识

光度一致性假设的度量方式

常用的计算区域光度一致性的方法有

- Sum of Squared Differences (SSD)

$$\rho_{SSD}(f, g) = \|f - g\|^2$$

- Sum of Absolute Differences (SAD)

$$\rho_{SAD}(f, g) = \|f - g\|_1$$

- Normalized Cross Correlation(NCC)

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

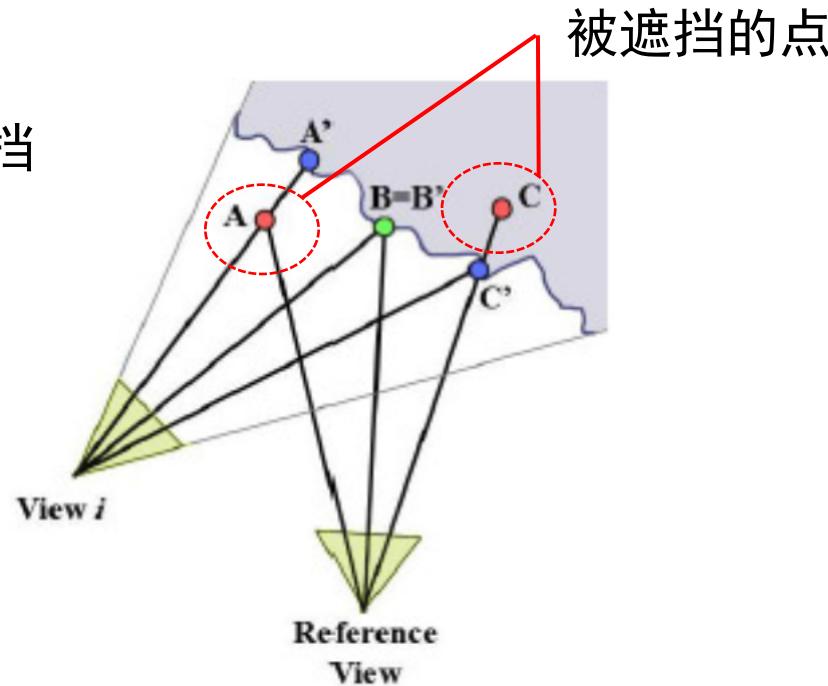
基础知识

可视性约束

图像中出现的点不能被遮挡

重建的点前面不能出现点

不能出现在物体内部



课程内容

✓ 稠密点云的获取方式

- ✓ Lidar 扫描
- ✓ Kinect
- ✓ 结构光
- ✓ 基于图像的方法-Multi-view Stereo

✓ 基础知识

- ✓ 极线条搜索
- ✓ 光度一致性约束
- ✓ 可视性约束

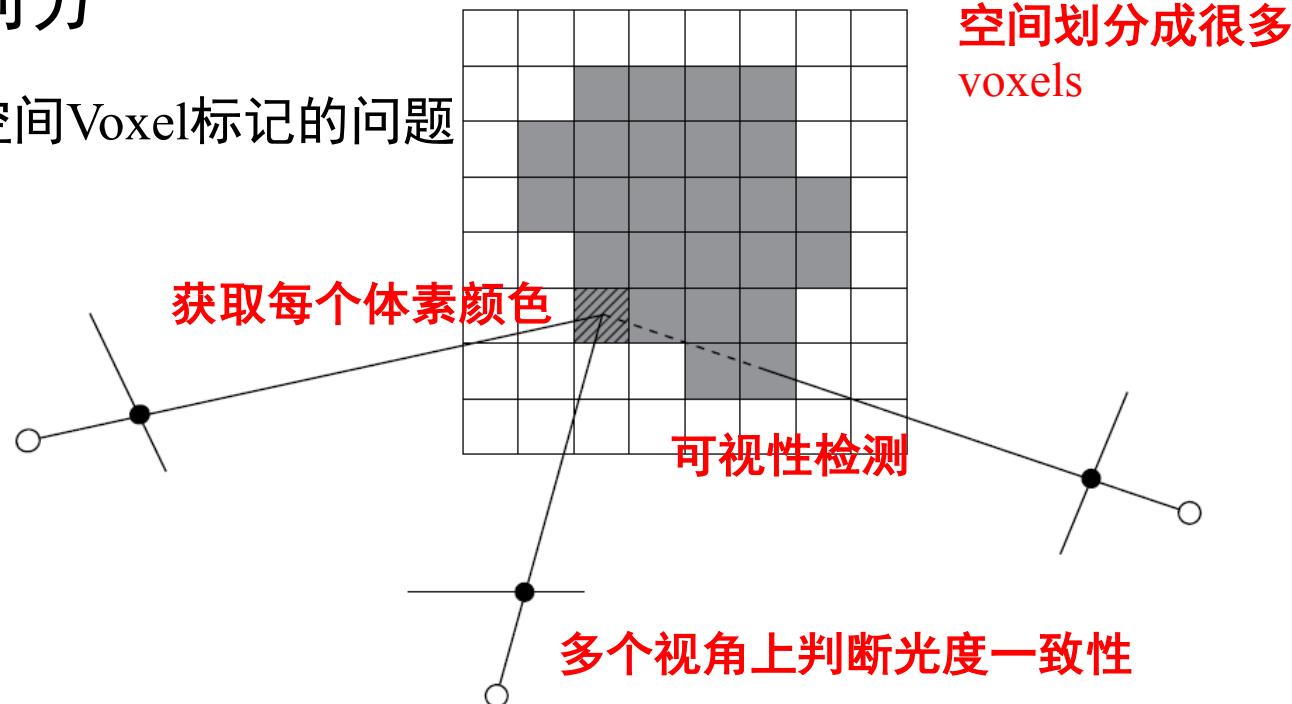
✓ 多视角立体技术

- ✓ 基于体素的方法
- ✓ 基于空间patch 的方法
- ✓ 深度图融合

基于体素的方法

规则的划分

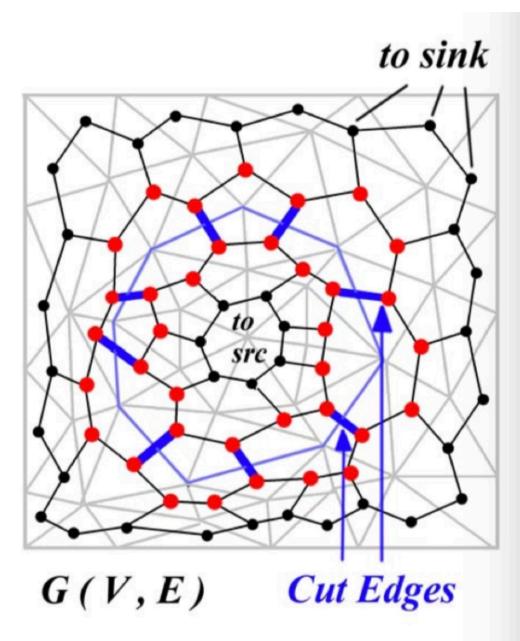
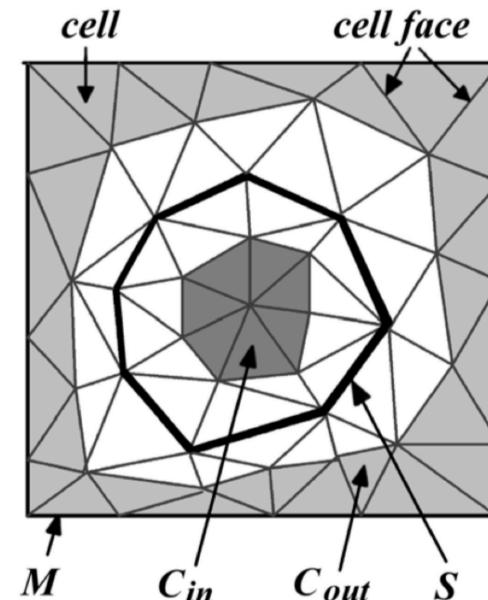
等价于3D空间Voxel标记的问题



基于体素的方法

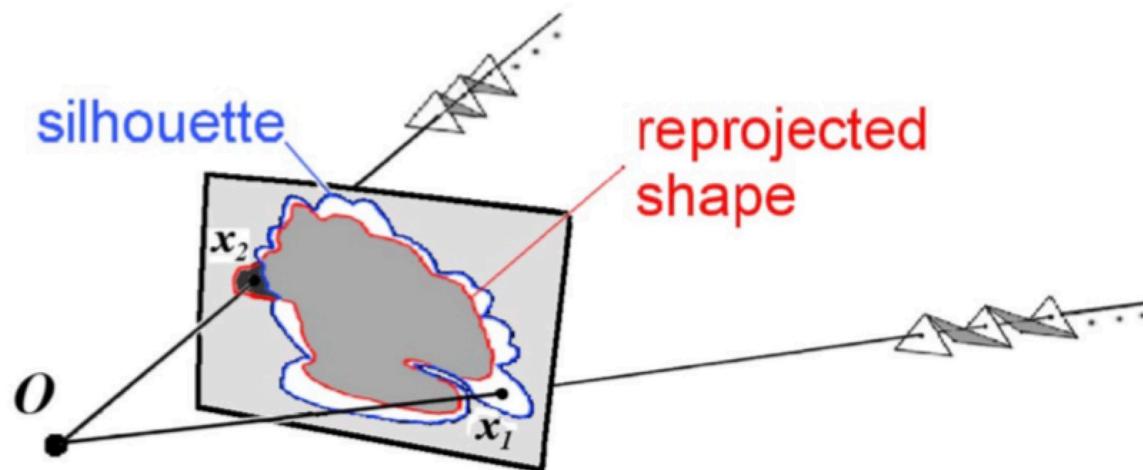
不规则的划分

等价于3D空间四面体标记的问题



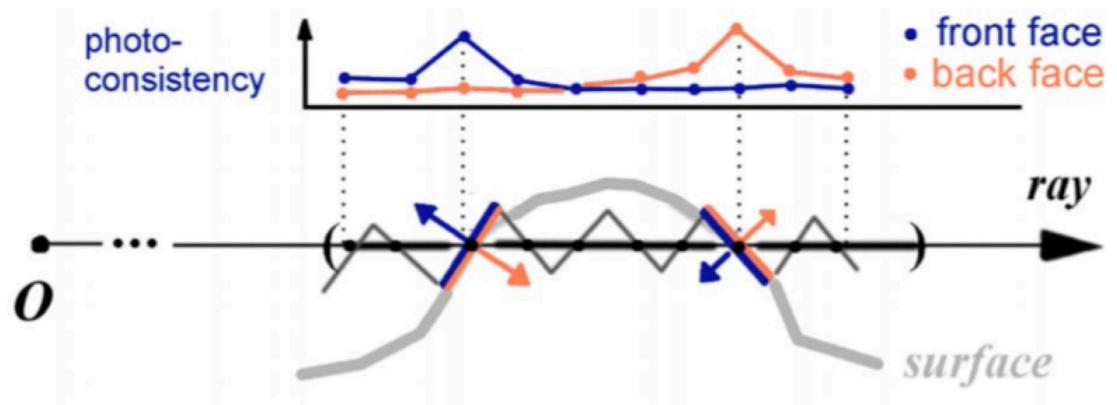
基于体素的方法

图像上的约束-剪影约束



基于体素的方法

图像上的约束-光度一致性约束



基于体素的方法

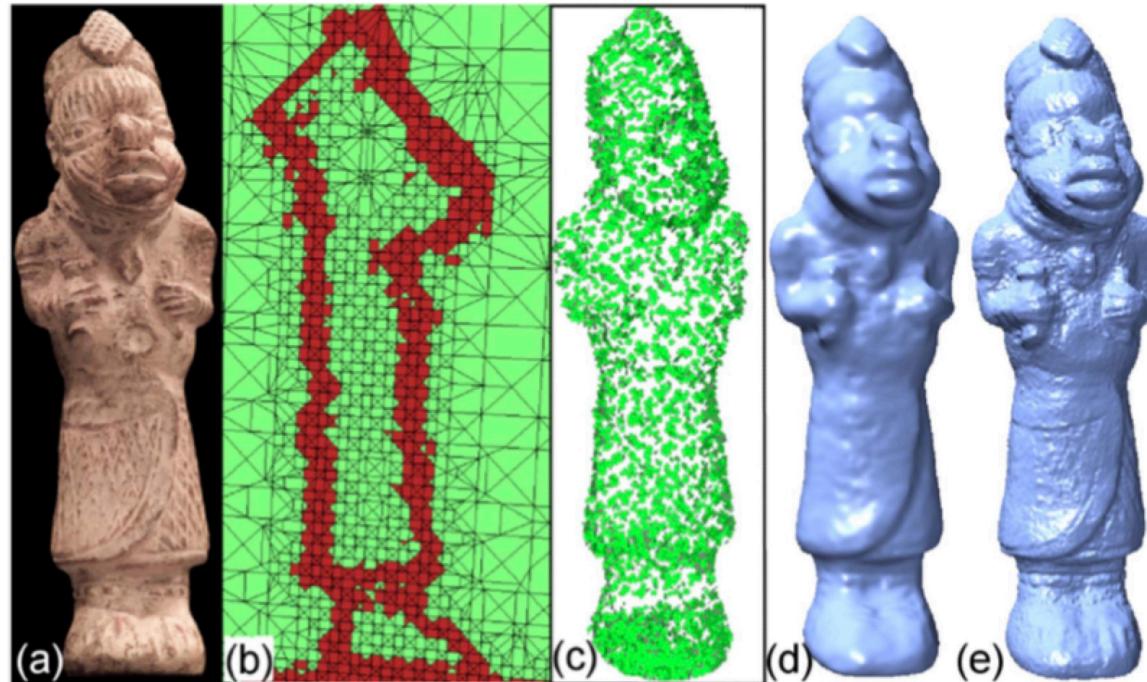
常用的优化模型-典型的MRF离散优化问题

参考图像上每个的像素分配一个标签（内部或者部）

$$E(f) = E_{data}(f) + E_{smoothness}(f) + E_{visibility}(f)$$

- 数据项-光度一致性假设
- 平滑项-邻域假设
- 可视项-可视性约束

基于体素的方法



基于体素的方法

优点：

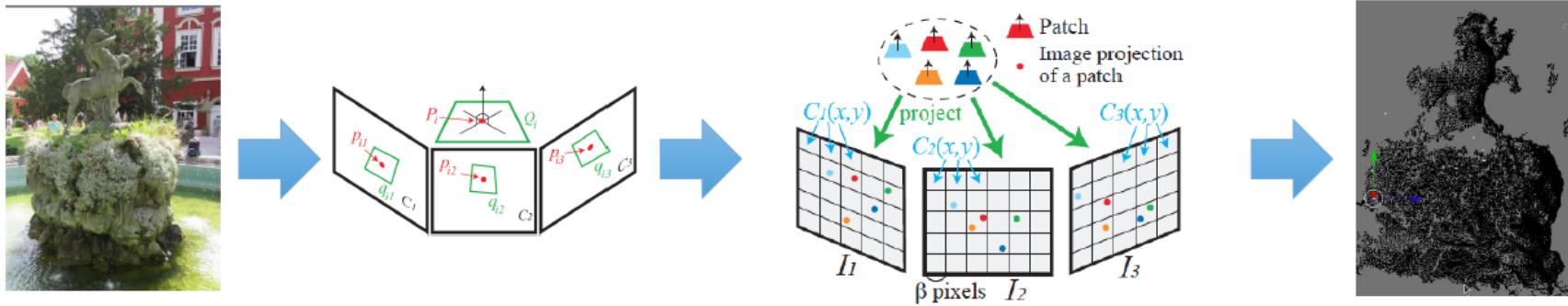
- 生成规则的点云
- 便于提取物体的平面

缺点：

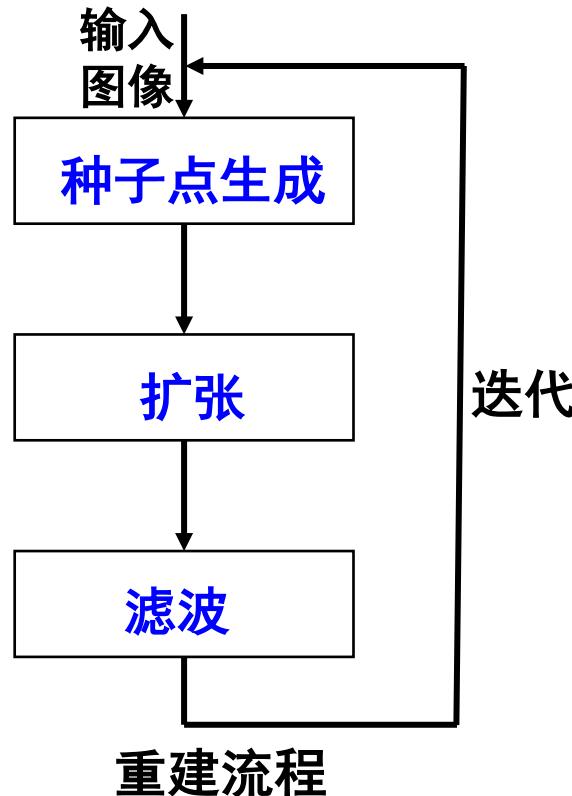
- 精度受到空间划分分辨率的影响
- 难以处理精度高、规模大的场景

基于空间patch扩散的方法

- 假设空间中的3D 矩形patch
- 通过一定规则的扩张方法，使得patch覆盖物体表面



基于空间patch扩散的方法

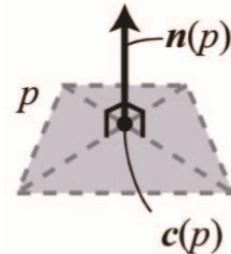


- 初始种子点生成采用SIFT,HOG等特征点
- 扩张过程对已重建三维点的邻域进行匹配
- 滤波过程采用两种约束去除噪声点
- 光度一致性约束
 - 可视性约束

PMVS

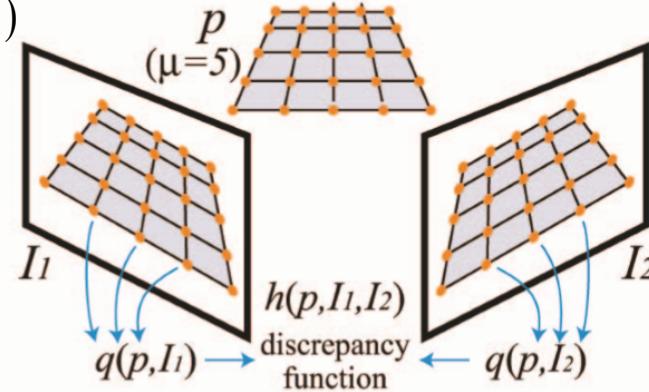
- 3D Patch 定义

位置, 法向量
大小 5×5



- 光度一致性约束 $h(p, I_1, I_2)$

Patch投影到不同图
像上, 计算NCC



PMVS

● 初始3D patch 的生成

1. 在图像上均匀计算HOG/Harris特征
2. 沿极线进行搜索找到匹配特征点
3. 对匹配对，通过三角化建立patch

中心：三角化确立

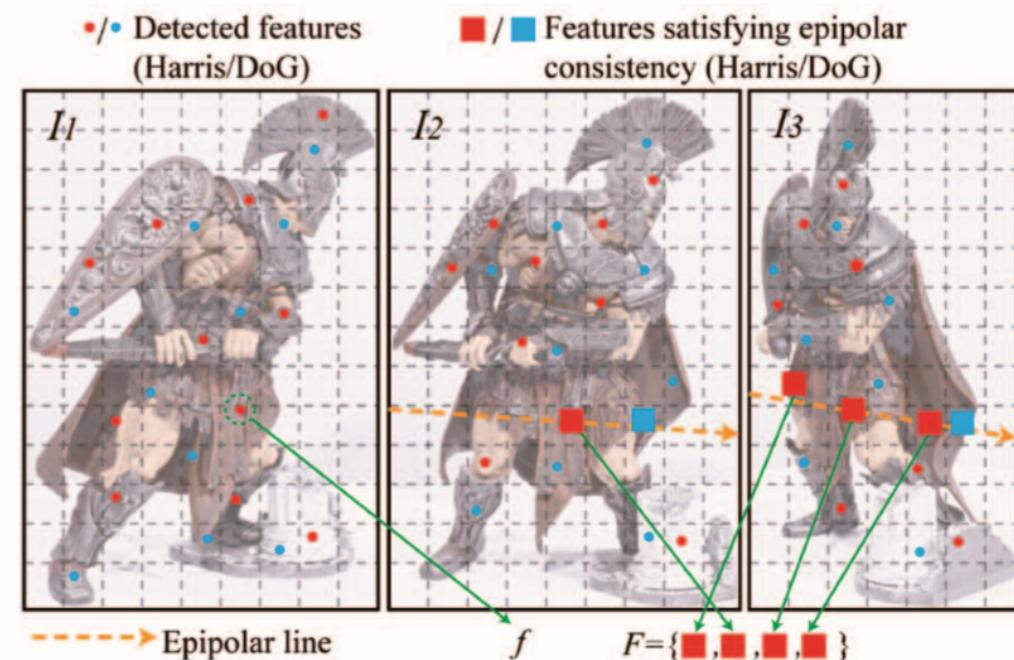
法向量：指向参考图像

可视图像 $V(p)$ ：法向量和视线夹角足够小
通过光度一致性约束对可视图像进行筛选

$$V^*(p) = \{I | I \in V(p), h(p, I, R(p)) \leq \alpha\}$$

4. 对patch 位置和法向量进行优化

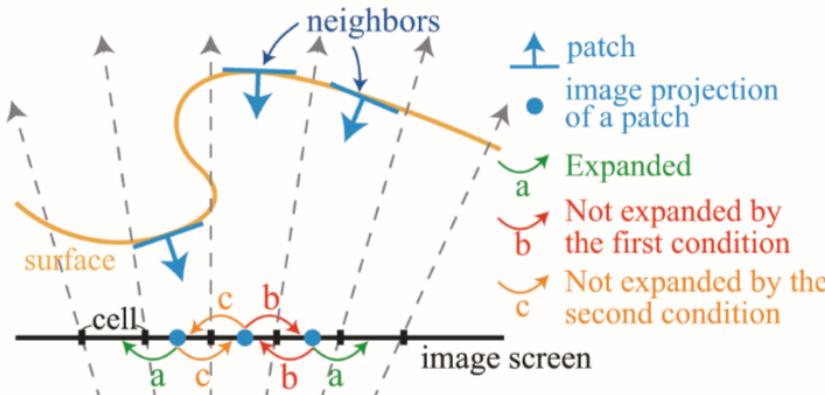
$$\min_{c(p), n(p)} g^*(p) = \frac{1}{|V^*(p) \setminus R(p)|} \sum_{I \in V^*(p) \setminus R(p)} h(p, I, R(p)).$$



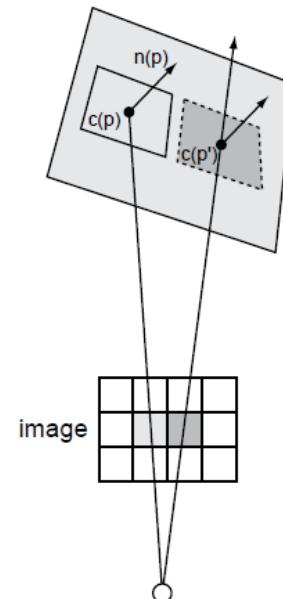
PMVS

● Patch扩展

1. 将三维patch投影到图像上
2. 如果相邻cell没有patch 且深度连续
则建立初始patch



3. 计算初始patch的可视图像，并进行优化



新建patch的可视图像和法向量的
初始值等同于邻域patch

$$n(p') = n(p)$$

$$V(p') = V(p)$$

新建patch的位置设置为当前cell的视
线和邻域patch所在的平面的交点

PMVS

● Patch滤波

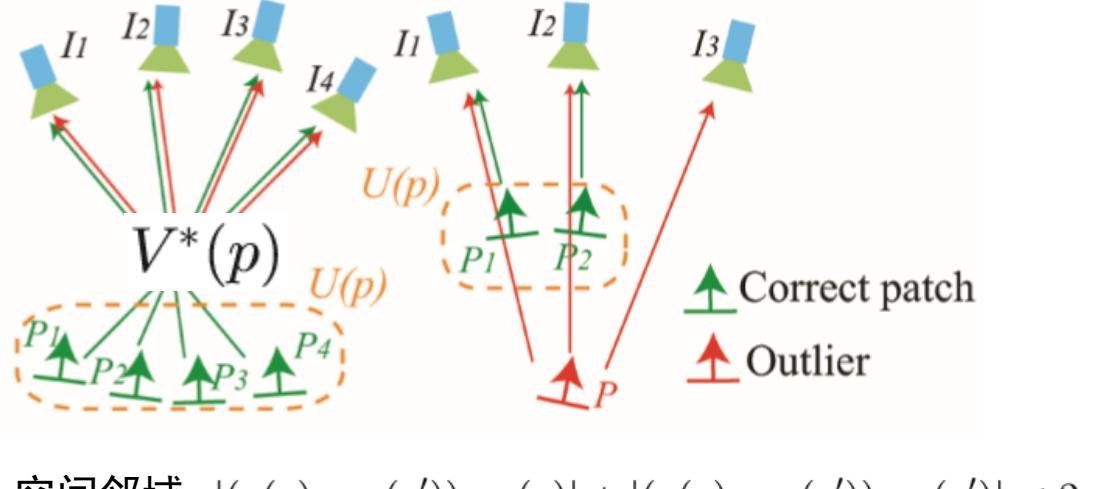
1. 可视性约束

$$|V^*(p)|(1 - g^*(p)) < \sum_{p_i \in U(p)} 1 - g^*(p_i)$$

2. 可视图像个数

$$|V^*(p)| < \gamma$$

3. 图像邻域中的cell同时也是空间邻域的比例小于0.25

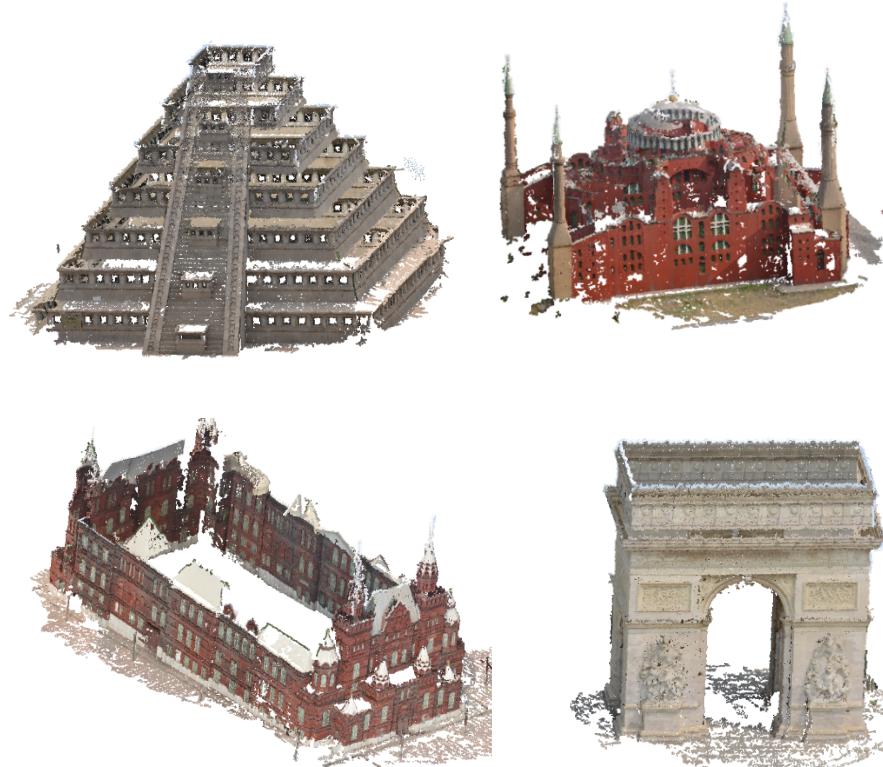


空间邻域: $|(\mathbf{c}(p) - \mathbf{c}(p')) \cdot \mathbf{n}(p)| + |(\mathbf{c}(p) - \mathbf{c}(p')) \cdot \mathbf{n}(p')| < 2\rho_1$

$U(p)$: 三维上不是邻域, 但投影到图像上同一个cell

PMVS

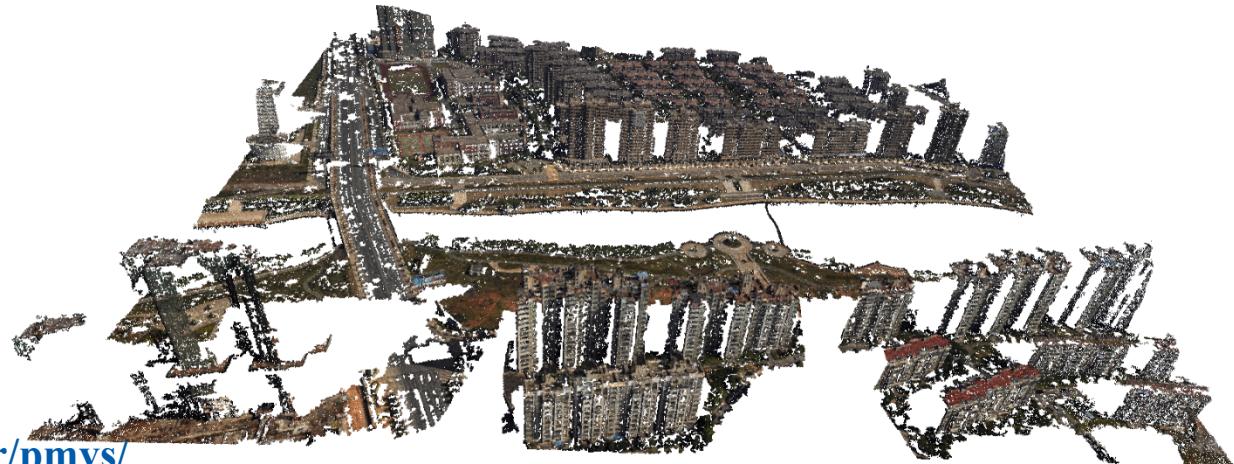
- 算法适用性强
- 使用于各种形状的物体



PMVS



- 朗伯面假设
- 容易产生空洞



PMVS

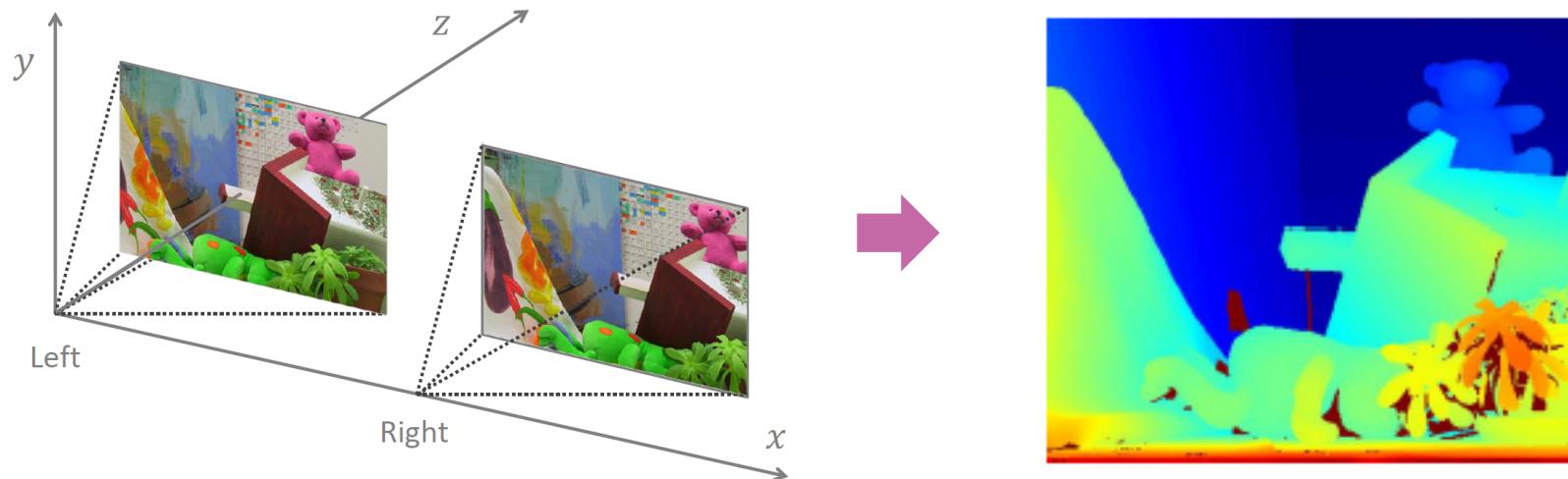


● 计算量非常大

影像集	影像数目	影像尺寸	运动恢复结构(Hour)	稠密匹配(Hour)
鲁甸	407	5616*3744	0.253	4.752
北川	273	5910*3854	0.351	3.545
连云港	782	5616*3744	0.643	7.414

基于深度图融合的方法

人的左右眼立体视觉和深度图



M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz. Multi-view stereo for community photo collections. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.

基于深度图融合的方法

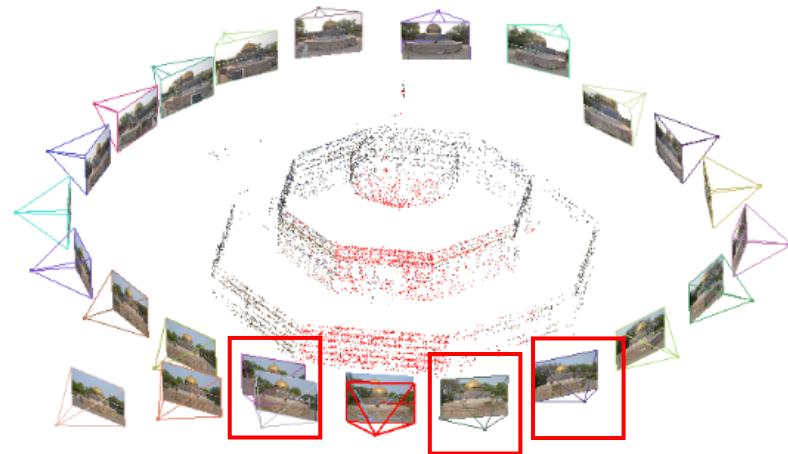


基本流程

1. 为每一幅图像选择邻域图像构成立体图像对
2. 计算每一幅图像的深度图
3. 进行深度图融合

基于深度图融合的方法

视角选择



M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz. Multi-view stereo for community photo collections. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.

基于深度图融合的方法

- 邻域的选择-全局视角选择

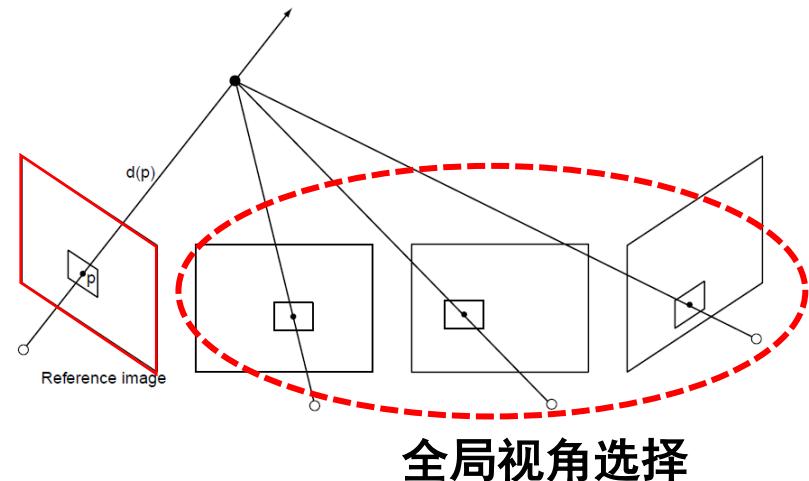
1. 图像具有相同的内容，外观和尺度
2. 图像具有足够大的视差（宽基线）

三角量测夹角适中

$$g_R(V) = \sum_{f \in \underline{F_V} \cap \underline{F_R}} \underline{w_N(f)} \cdot \underline{w_s(f)},$$

图像尺度相似

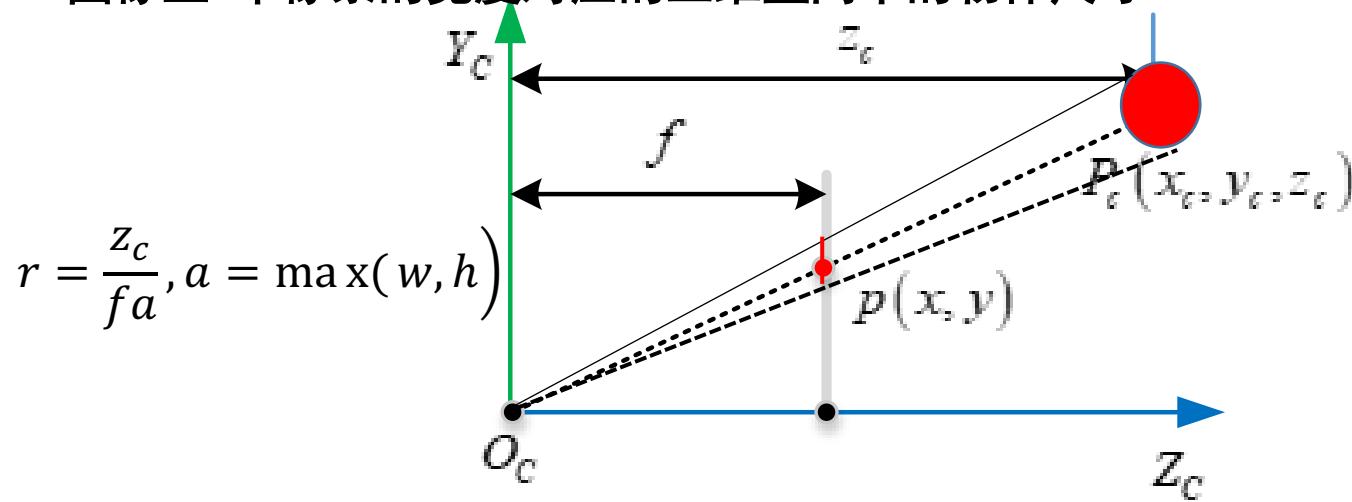
共有的特征点越多越好



基于深度图融合的方法

图像尺度的估计-用于衡量图像的分辨率

图像上1个像素的宽度对应的三维空间中的物体尺寸

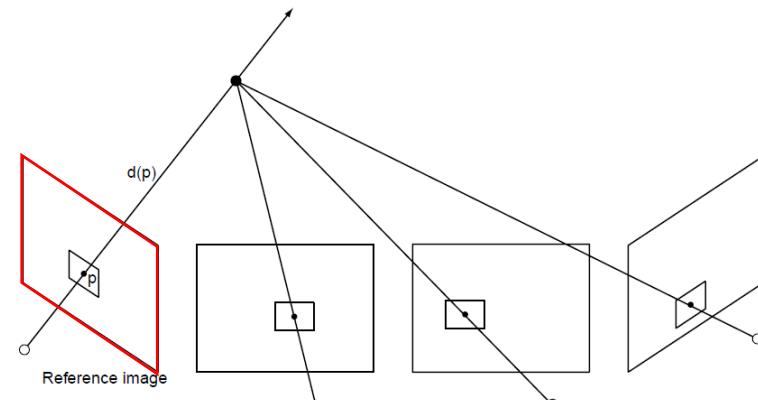


基于深度图融合的方法

- 相关视角的选择-局部视角选择

1. NCC值确定候选视角
2. 视线要足够分散(不共面)

$$l_R(V) = g_R(V) \cdot \prod_{V' \in A} \frac{w_e(V, V')}{\text{视线夹角}}$$

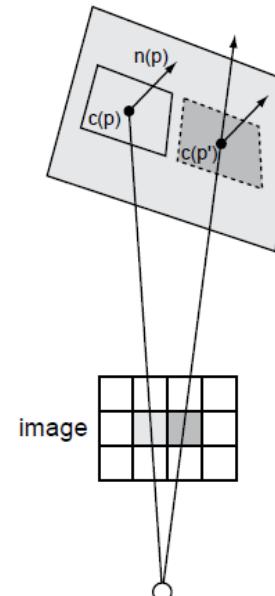


从候选视角中选择
active 视角

基于深度图融合的方法

● 区域生长法扩展

1. 重建的置信度建立优先级队列
2. 从初始的稀疏特征点开始深度估计
3. 对每个种子点进行非线性深度优化
4. 每次优化完后判断以下两种情况，将邻域像素添加到队列中： 1) 邻域没有深度值 2) 当前像素的置信度值高于邻域像素一定范围



新建patch的可视图像和法向量的
初始值等同于邻域patch

$$n(p') = n(p)$$

$$V(p') = V(p)$$

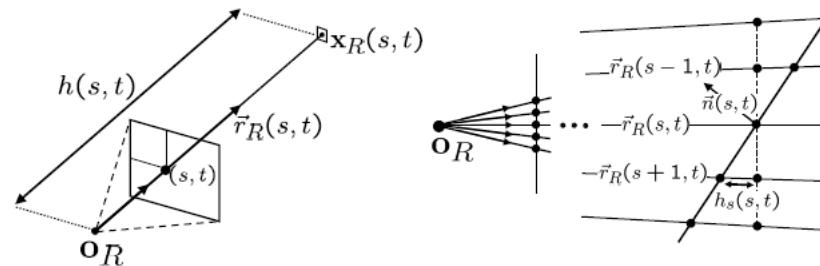
新建patch的位置设置为当前cell的视
线和邻域patch所在的平面的交点

基于深度图融合的方法

- 深度值非线性优化

中心点的三维坐标

$$\mathbf{X}_R(s, t) = \mathbf{o}_R + h(s, t)\vec{r}_R(s, t)$$



Patch 中每个点的三维坐标

$$\mathbf{X}_R(s + i, t + j) = \mathbf{o}_R + \vec{r}_R(s, t)(h(s, t) + i * h_s + j * h_t)$$

h_s 和 h_t 反应了patch的法向量

基于深度图融合的方法

● 深度值非线性优化

颜色模型

为邻域中每个视角 K 分配一个颜色尺度 $\mathbf{c}_k = [c_k^r, c_k^g, c_k^b]^T \in R^{3 \times 1}$, 如果 $h(s, t)$ 估计准确, 那么应该有

$$\mathbf{I}_R(s + i, t + j) = \mathbf{c}_k \cdot \mathbf{I}_k(\mathbf{P}_k(\mathbf{X}_R(s + i, t + j)))$$

为方便, 简写为

$$\mathbf{I}_R(i, j) = \mathbf{c}_k \cdot \mathbf{I}_k(\mathbf{P}_k(\mathbf{X}_R(i, j))) = \mathbf{c}_k \cdot \mathbf{I}_k(i, j)$$

基于深度图融合的方法

● 非线性优化的数学模型

$$E = \sum_{ijk} (I_R(i,j) - c_k \cdot I_k(i,j))^2$$

$i, j \in \left[-\frac{n-1}{2}, \frac{n-1}{2}\right]$ 是 patch 中的采样点， k 表示视角个数

需要优化的变量 $h(s, t), h_s, h_t$ 以及 $\{c_k\}$

基于深度图融合的方法

● 颜色尺度的优化

对 E 求关于 \mathbf{c}_k 的导数

$$\frac{\partial E}{\partial c_k^r} = - \sum_{ij} I_k^r(i,j) I_R^r(i,j) + c_k^r \sum_{ij} (I_k^r(i,j))^2$$

$$\frac{\partial E}{\partial c_k^g} = - \sum_{ij} I_k^g(i,j) I_R^g(i,j) + c_k^g \sum_{ij} (I_k^g(i,j))^2$$

$$\frac{\partial E}{\partial c_k^b} = - \sum_{ij} I_k^b(i,j) I_R^b(i,j) + c_k^b \sum_{ij} (I_k^b(i,j))^2$$

$$\mathbf{c}_k = \left[\frac{\sum_{ij} I_k^r(i,j) I_R^r(i,j)}{\sum_{ij} (I_k^r(i,j))^2}, \frac{\sum_{ij} I_k^g(i,j) I_R^g(i,j)}{\sum_{ij} (I_k^g(i,j))^2}, \frac{\sum_{ij} I_k^b(i,j) I_R^b(i,j)}{\sum_{ij} (I_k^b(i,j))^2} \right]^T$$

基于深度图融合的方法

● 颜色尺度的优化-最小二乘法

分别令上述三个偏导数为零，可以得到以下表达式

$$c_k^r = \frac{\sum_{ij} I_k^r(i,j) I_R^r(i,j)}{\sum_{ij} (I_k^r(i,j))^2}$$

$$c_k^g = \frac{\sum_{ij} I_k^g(i,j) I_R^g(i,j)}{\sum_{ij} (\mathbf{g}(i,j))^2}$$

$$c_k^b = \frac{\sum_{ij} I_k^b(i,j) I_R^b(i,j)}{\sum_{ij} (I_k^b(i,j))^2}$$

基于深度图融合的方法

- $h(s, t), h_s$ 和 h_t 的优化-梯度下降法

$$E = \sum_{ijk} \left(\mathbf{I}_R(i, j) - \mathbf{c}_k \cdot \mathbf{I}_k(i, j) + \mathbf{c}_k \cdot \frac{\partial \mathbf{I}_k(i, j)}{\partial h(s, t)} (dh(s, t) + i * dh_s + j * dh_t) \right)^2$$

令 $\mathbf{b}_{ijk} = \mathbf{I}_R(i, j) - \mathbf{c}_k \cdot \mathbf{I}_k(i, j)$, $\nabla \mathbf{x} = [dh(s, t), dh_s, dh_t]^T$,

$\mathbf{A}_{ijk} = \left(\mathbf{c}_k \cdot \frac{\partial \mathbf{I}_k(i, j)}{\partial h(s, t)} \right) \begin{bmatrix} 1 \\ i \\ j \end{bmatrix}^T$, 则上式可以表达为

$$E = \sum_{ijk} (\mathbf{A}_{ijk} \nabla \mathbf{x} + \mathbf{b}_{ijk})^2$$

基于深度图融合的方法

- $h(s, t), h_s$ 和 h_t 的优化

求 E 关于 x 的梯度，可以得到

$$\frac{\partial E}{\partial \nabla x} = \sum_{ijk} A_{ijk}^T (A_{ijk} x + b_{ijk})$$

令 $\frac{\partial E}{\partial \nabla x} = 0$ 可以得到

$$(\sum_{ijk} A_{ijk}^T A_{ijk}) \nabla x + \sum_{ijk} A_{ijk}^T b_{ijk} = \mathbf{0}$$

令 $A = \sum_{ijk} A_{ijk}^T A_{ijk}$, $b = \sum_{ijk} A_{ijk}^T b_{ijk}$, 则有
 $\nabla x = A^{-1} b$

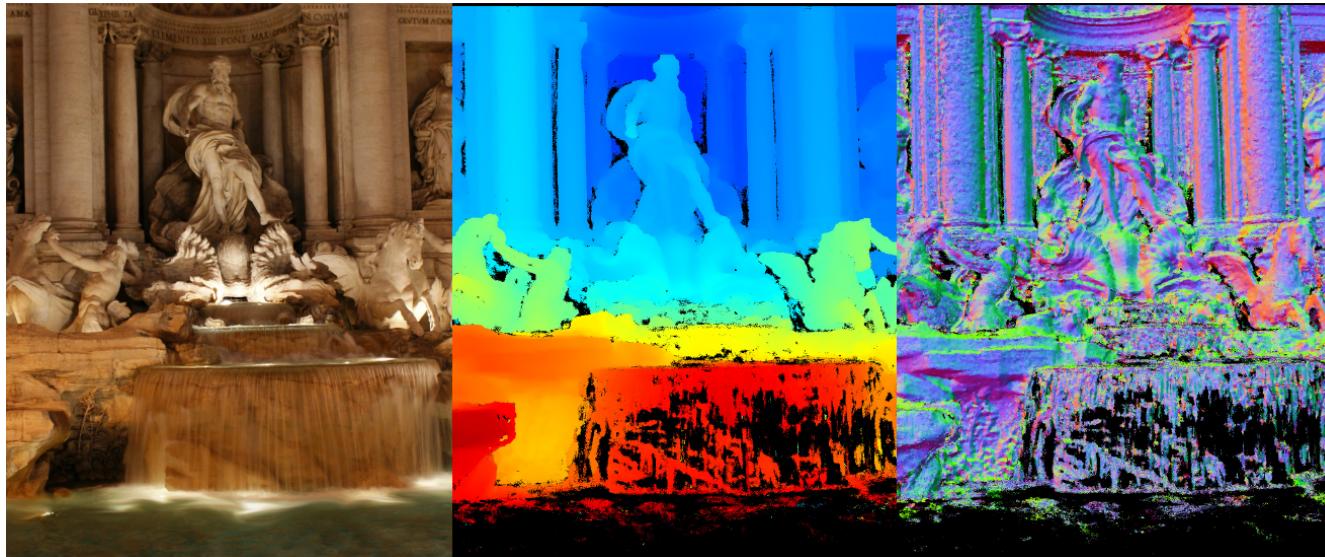
基于深度图融合的方法

● 深度值非线性优化—整体框架

1. 只进行深度 $h(s,t)$ 的优化（迭代4次）
2. While (迭代次数< 20)
3. 每间隔5次迭代，进行 $h(s,t)$, hs , ht 的优化,否则仅优化深度
4. 优化完成后跟踪判断每个视角的置信度
5. 如果视角的置信度太小，或者超过迭代14次尚未收敛，则从局部视角中移除该视角，并重新进行视角选择

基于深度图融合的方法

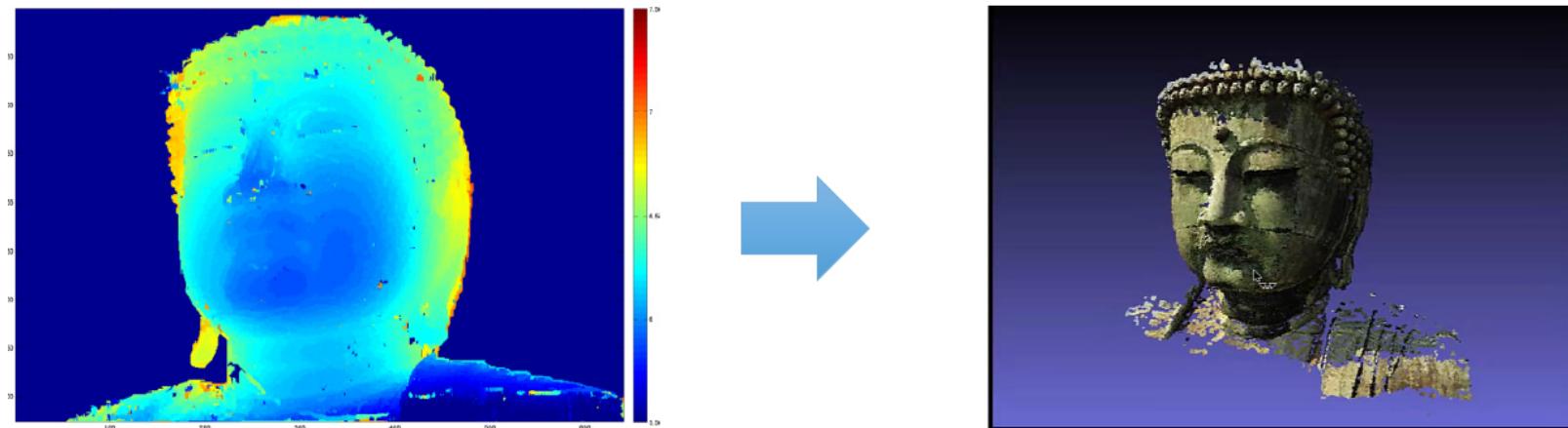
● 深度估计结果



M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz. Multi-view stereo for community photo collections. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.

基于深度图融合的方法

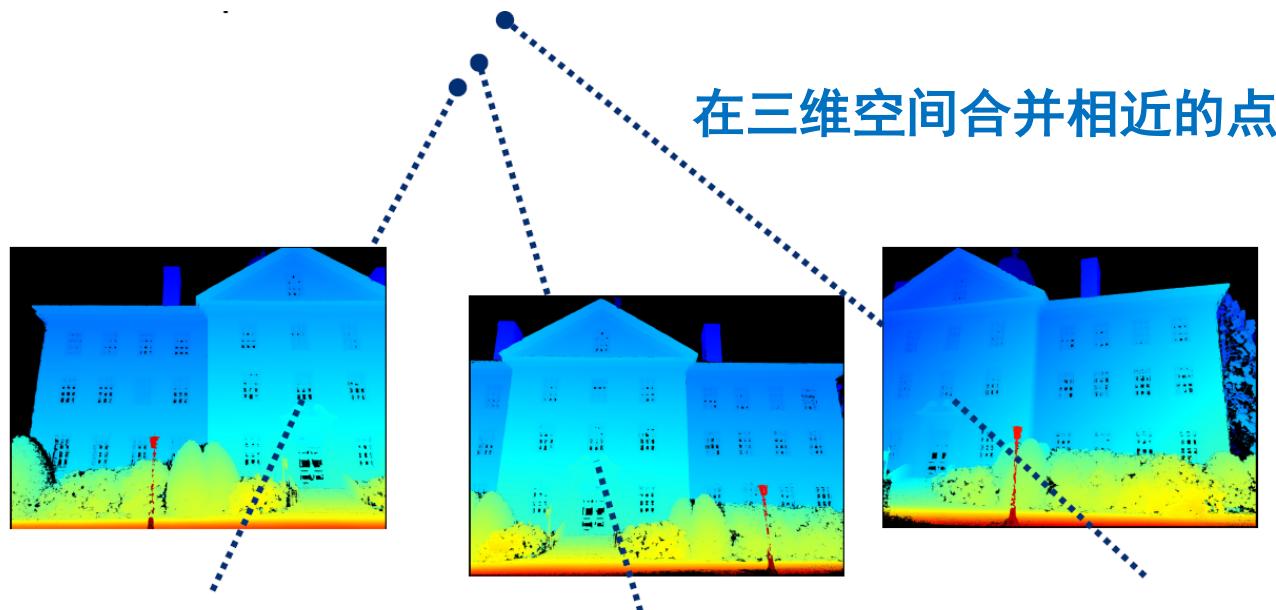
● 深度融合



直接聚合？一致性，可视性

基于深度图融合的方法

- 深度融合-一致性约束

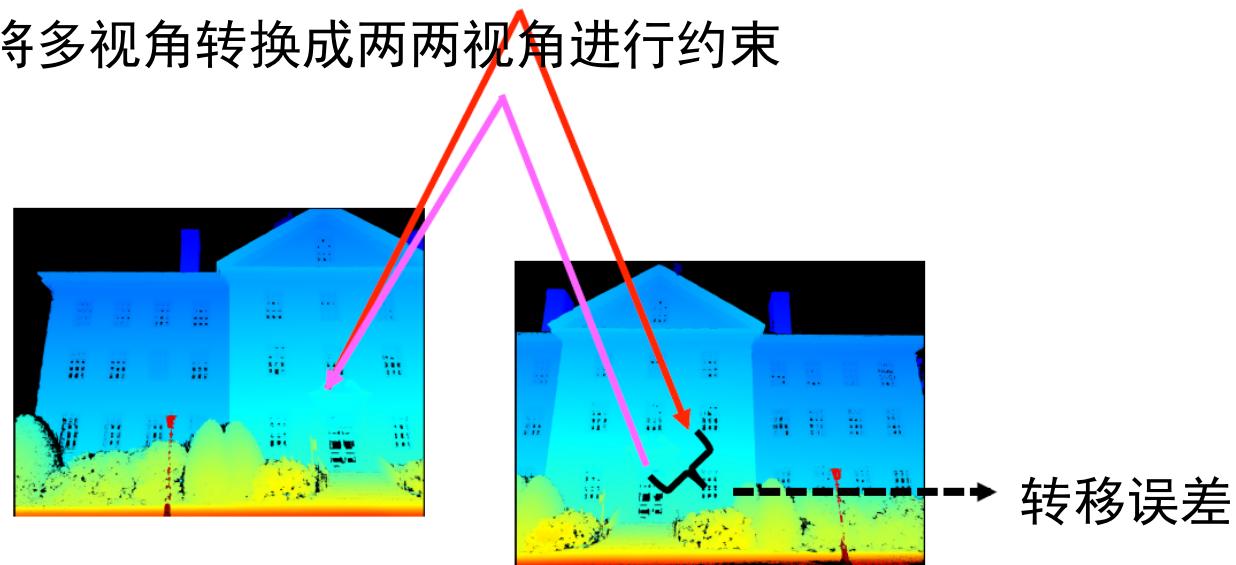


M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz. Multi-view stereo for community photo collections. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.

基于深度图融合的方法

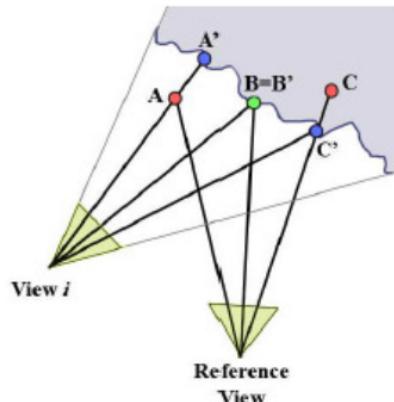
● 深度融合-一致性约束

二维空间的邻域约束-将多视角转换成两两视角进行约束

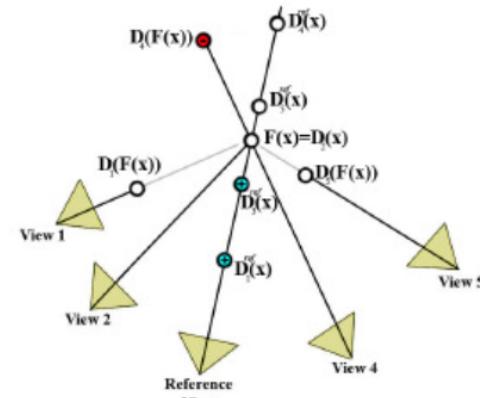


基于深度图融合的方法

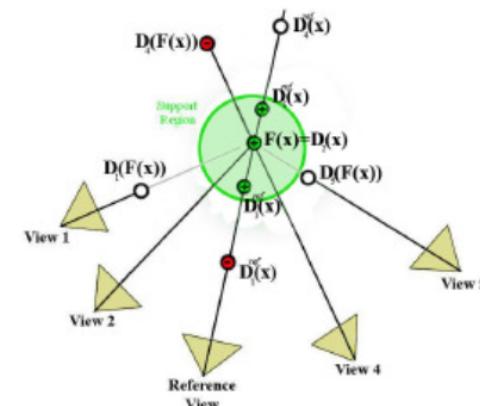
- 深度融合-可视性约束



(a) Visibility relations between points



(b) Stability calculation

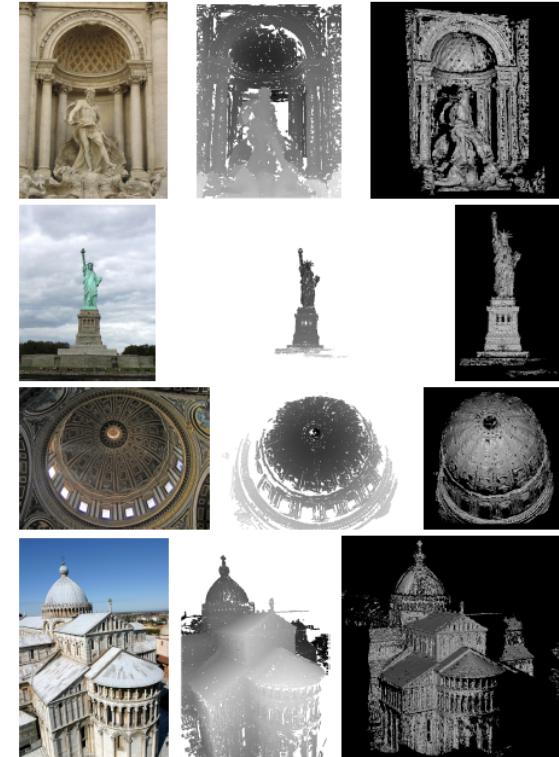


(c) Support estimation

Merrel et al. ICCV 2007

基于深度图融合的方法

- 邻域视角选择使得深度估计准确度提升
- 原理简单，只用到光度一致性约束和可视性约束，适用的场景广泛



M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. Seitz. Multi-view stereo for community photo collections. In *IEEE International Conference on Computer Vision (ICCV)*, 2007.



感谢各位聆听 !
Thanks for Listening !