



城市空间建模与仿真

第二讲 城市空间三维数据获取与模型表达（一）

任课教师：汤圣君

建筑与城市规划学院 城市空间信息工程系

其中部分图片来自互联网和同行专家

目录

CONTENTS

01

从2D迈向3D

02

3D感知特点和面临的难点

03

3D感知相关算法简介

04

挑战项目、软件工具介绍

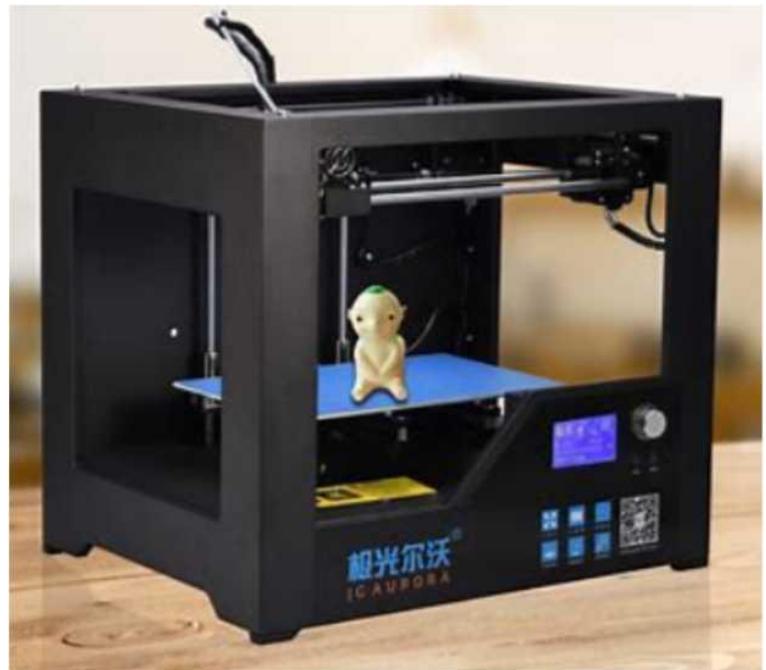
目录

CONTENTS

01

从2D迈向3D

何为3D?

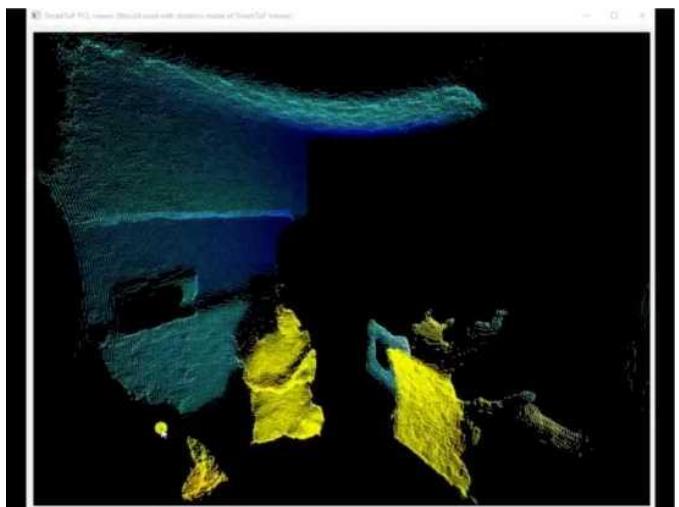


- 很多人的第一印象是3D电影或者3D打印
- 我们讨论的3D传感器数据，它体现在数据上，涵盖了对物体的3D度量

何为3D?—以深度图和点云表示的3D数据



- 深度图显示的房屋一角
- 伪彩色对应了图像内容到镜头的距离



- 3D点云



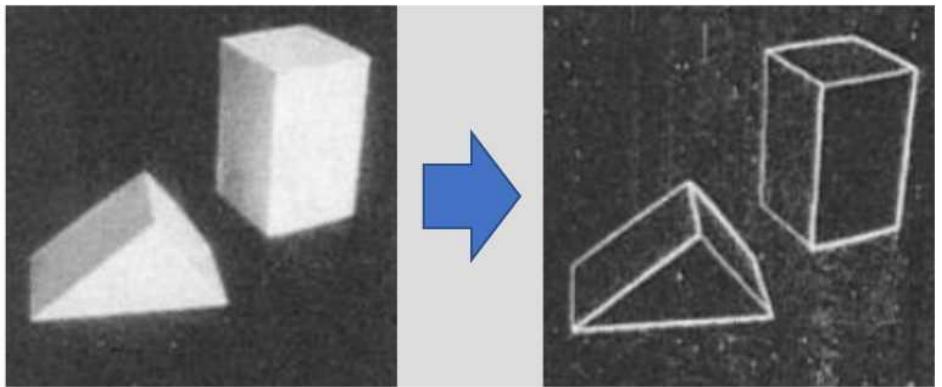
从2D迈向3D

- 2D计算机视觉—从20世纪50年代至今，几十年的发展历史
- 20世纪60年代—美国MIT的Roberts尝试通过计算机提取3D多面体，开始了3D计算机视觉的研究
- 20世纪70年代—马尔计算视觉理论提出之后，各种计算机视觉的理论和方法不断涌现
- 20世纪90年代—多视几何理论的提出，极大地加快了3D视觉的研究进程
- 21世纪—深度学习崛起

从2D迈向3D—通过计算机从2D图像提取3D多面体

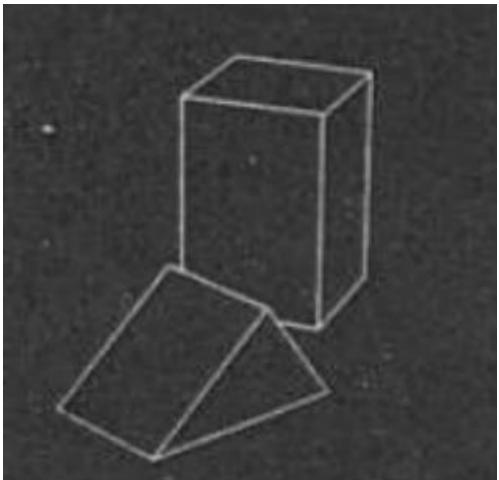


MIT的Larry Roberts



PhD Thesis, 1963, Machine
Perception of Three-Dimensional
Solids

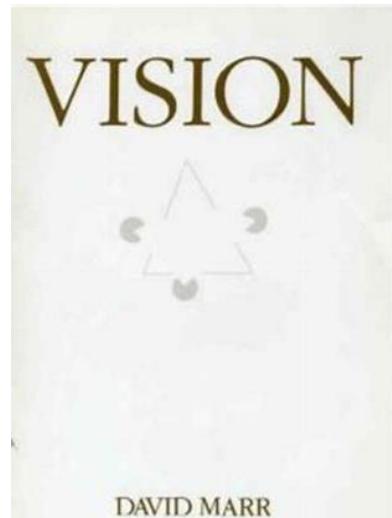
😊-为何说他得到的3D呢?
😊因为他可以旋转得到不同 视
角的图像



从2D迈向3D—计算视觉理论



David Marr



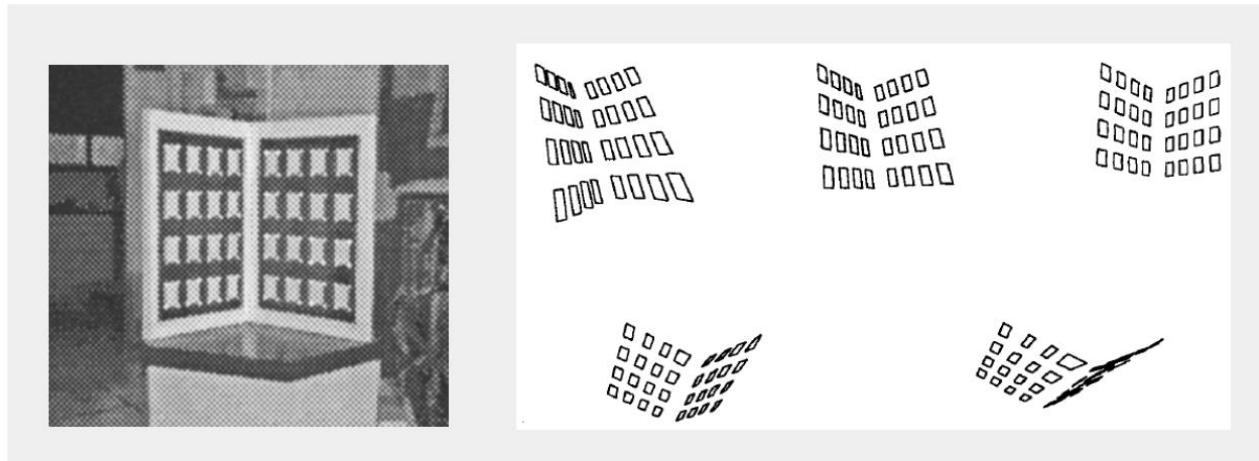
Vision: A computational Investigation
into the human representation and
processing of visual information, W.
H. Freeman and Company, 1982

- 三维重建理论
- 书中提出的视觉计算理论和方法，标志着计算机视觉成为了一门独立的学科

从2D迈向3D—90s



Olivier D. Faugeras ,



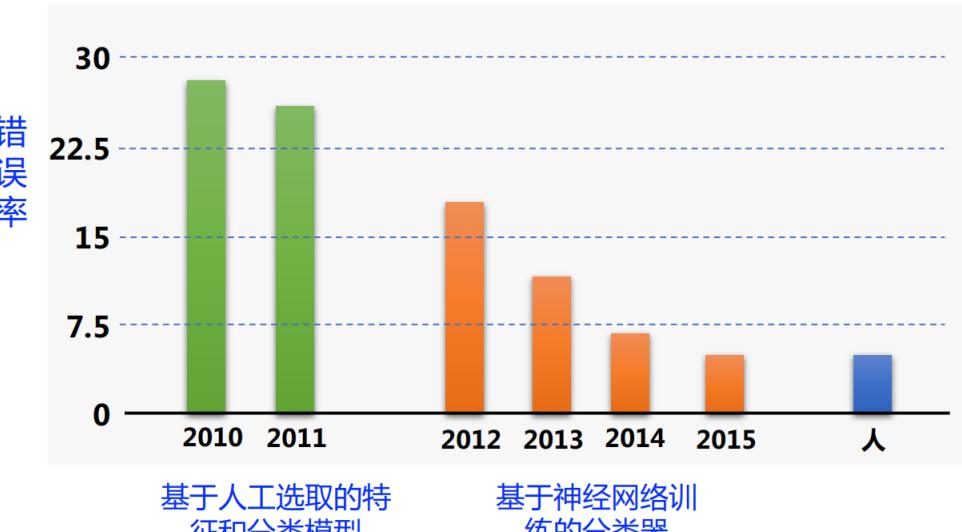
What can be seen in three dimensions with an
uncalibrated stereo rig?
1992 ECCV

多视几何和三维重建

(Multiple View Geometry and Stratified 3D Reconstruction)

从2D迈向3D—深度神经网络

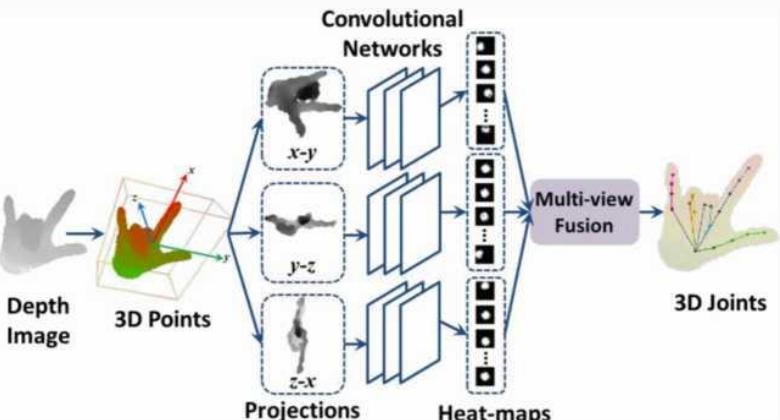
- 2010年起，视觉AI逐步起飞
- 2012年Alexnet达到80.2%分类正确率 (Top-5)，超越之前人工特征选区和分类算法
- 2013年VGG达到91.2%分类正确率 (Top-5)
- 2015年Resnet-50达到93%分类正确率 (Top-5)



从2D迈向3D—3D数据处理神经网络

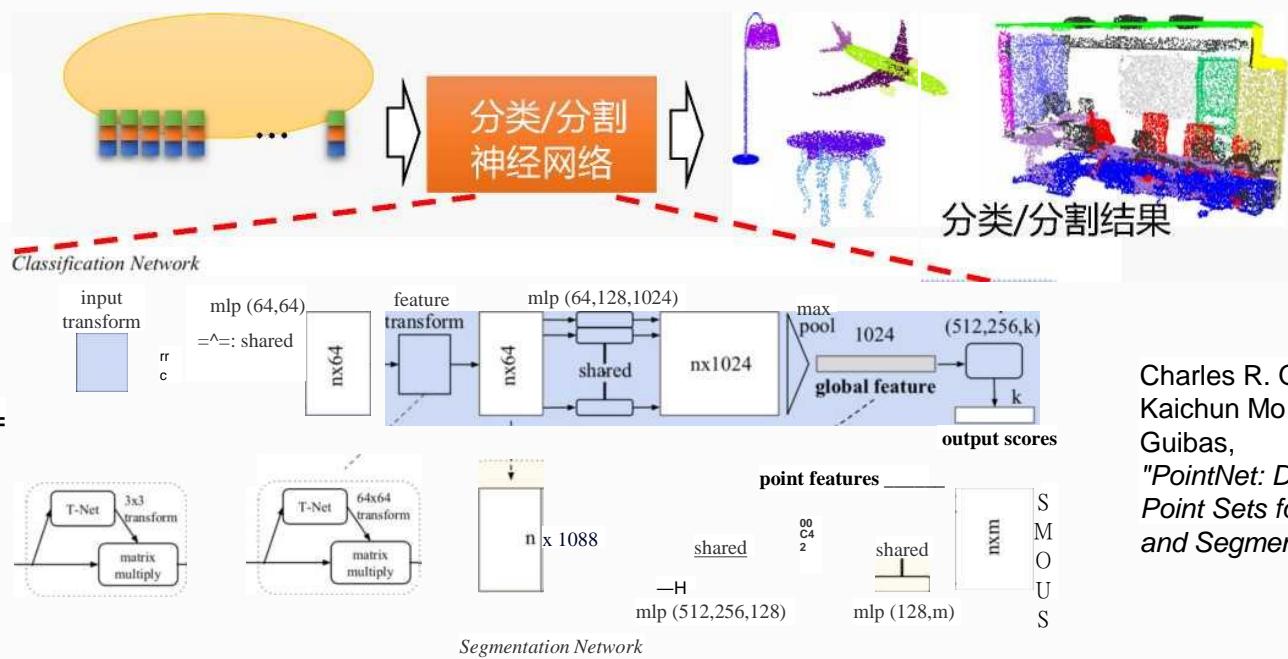
■ 基于3D图像的多视角投影

从3D点云生成特定视角的深度图，
使用和2D图像识别相同的卷积神
经网络结构实现手势识别



Liuhao Ge, Hui Liang, Daniel Thalmann,
*"Robust 3D Hand Pose Estimation in
 Single Depth Images: From Single- View
 CNN to Multi-View CNNs"*

■ 直接利用任意排列次序的点 云



Charles R. Qi* Hao Su*
 Kaichun Mo Leonidas J.
 Guibas,
*"PointNet: Deep Learning on
 Point Sets for 3D Classification
 and Segmentation,"*

目录

CONTENTS

02

3D感知特点和面临的难点

为何使用3D ?



低功耗/低运算量/信息量大/环境、目标感知能力/丰富的应用

为何使用3D？

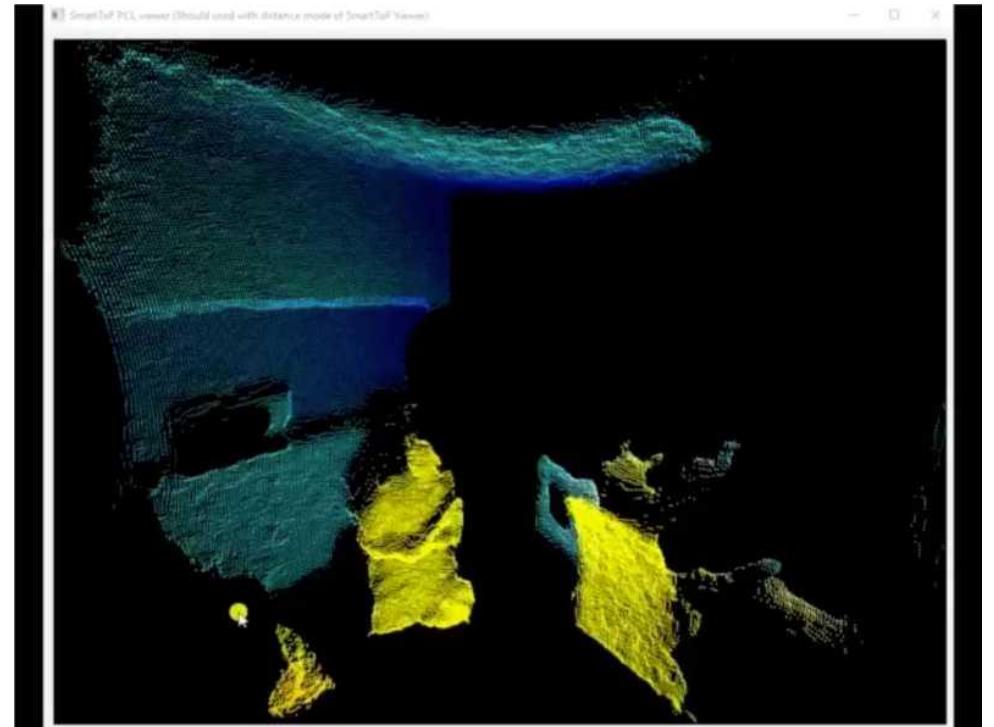
以极低的运算量为机器视觉算法提供更加直接和简洁的解决思路



为何使用3D ?—快速实现视觉对象分割

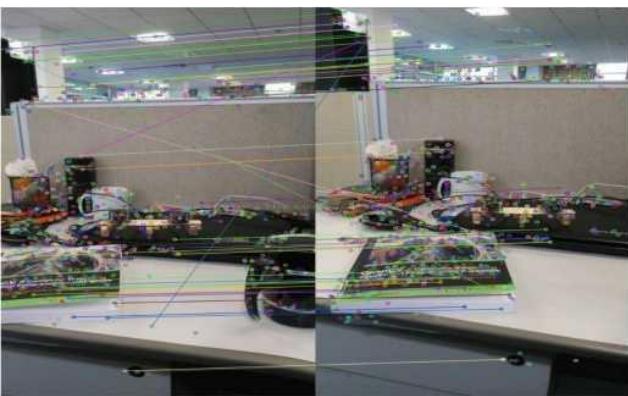


从2D图像目标和环境混合，难以区分

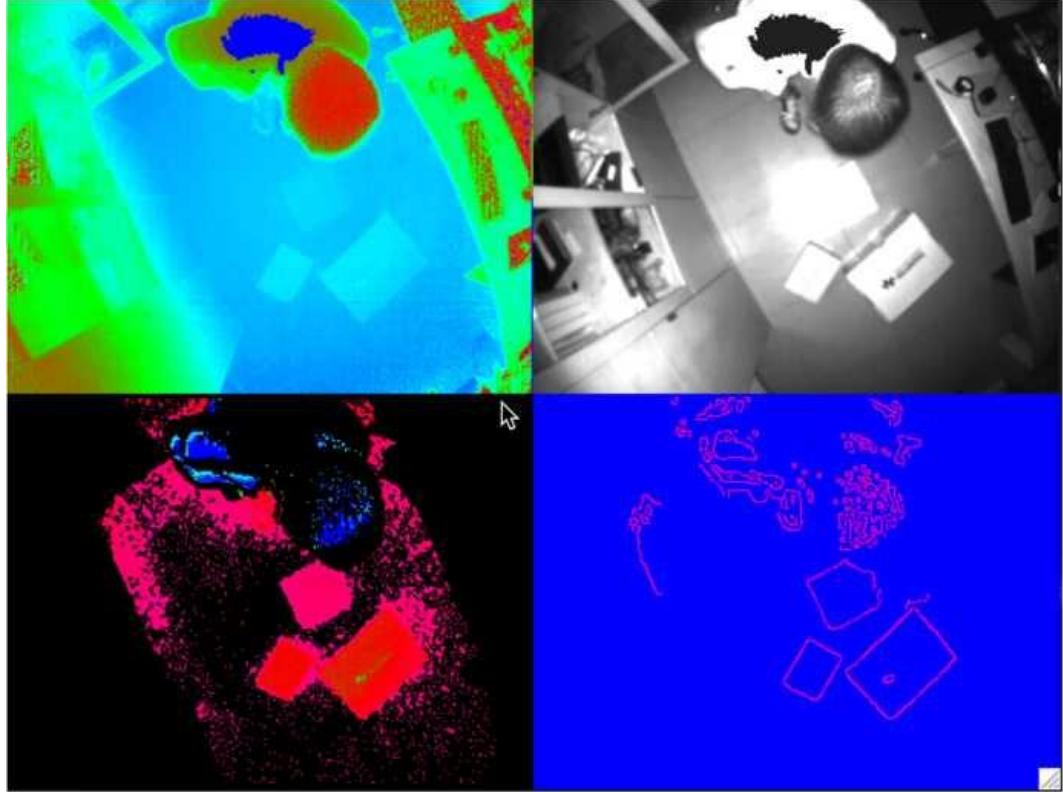


从3D数据根据距离信息直接分割目标

为何使用3D ?—快速识别和测量物体真实尺寸



从2D图像获取尺度信息、困难, 需要大量运算



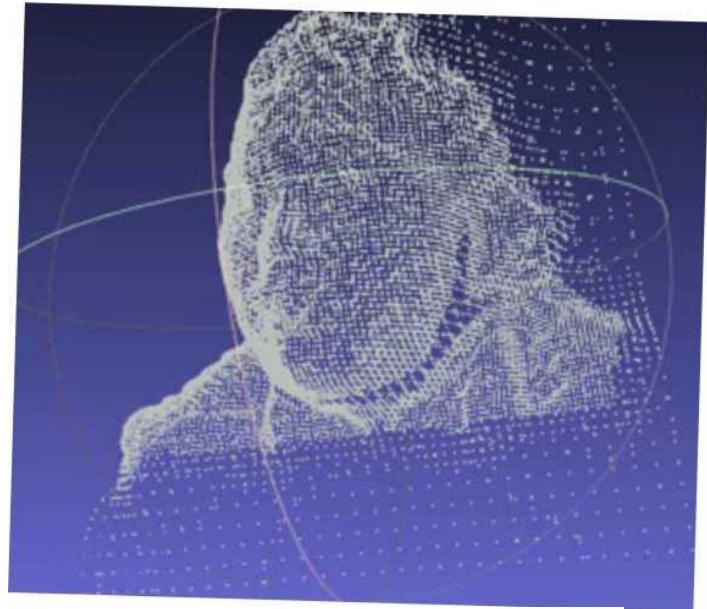
基于3D直接测量获得物理尺寸

不受光照/颜色干扰

为何使用3D ?

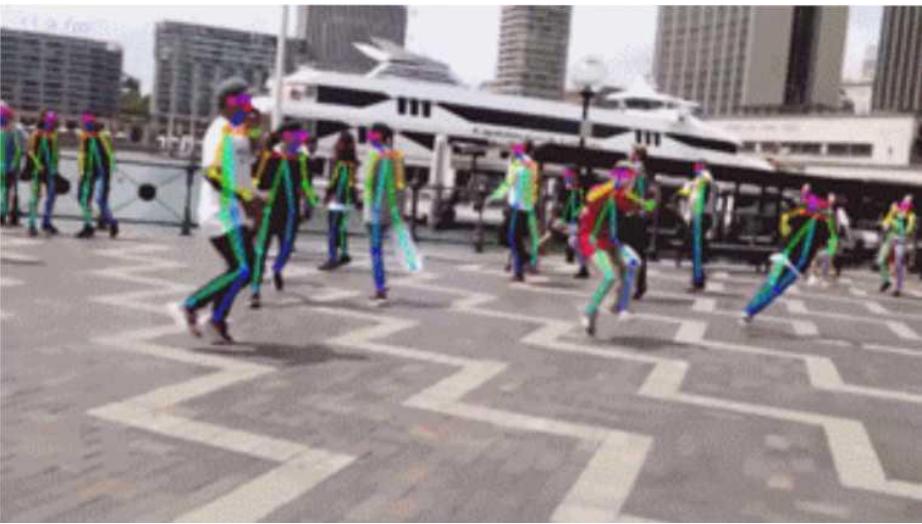


2D图像受光影影响



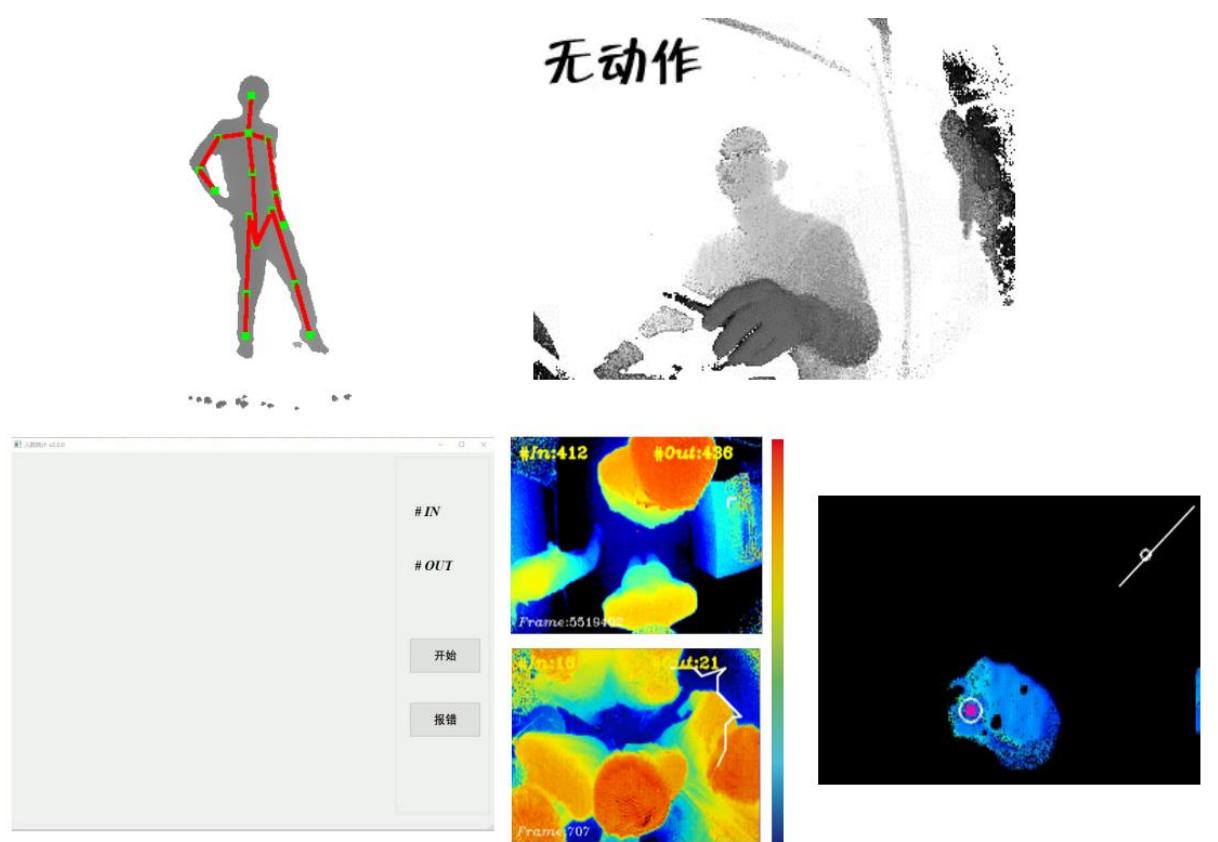
3D数据提供立体人脸数据,
不受光照影响

为何使用3D？



2D人体行为识别需要强大的GPU运算

动作识别



深度图进行3D人体行为识别只需要极低的CPU运算

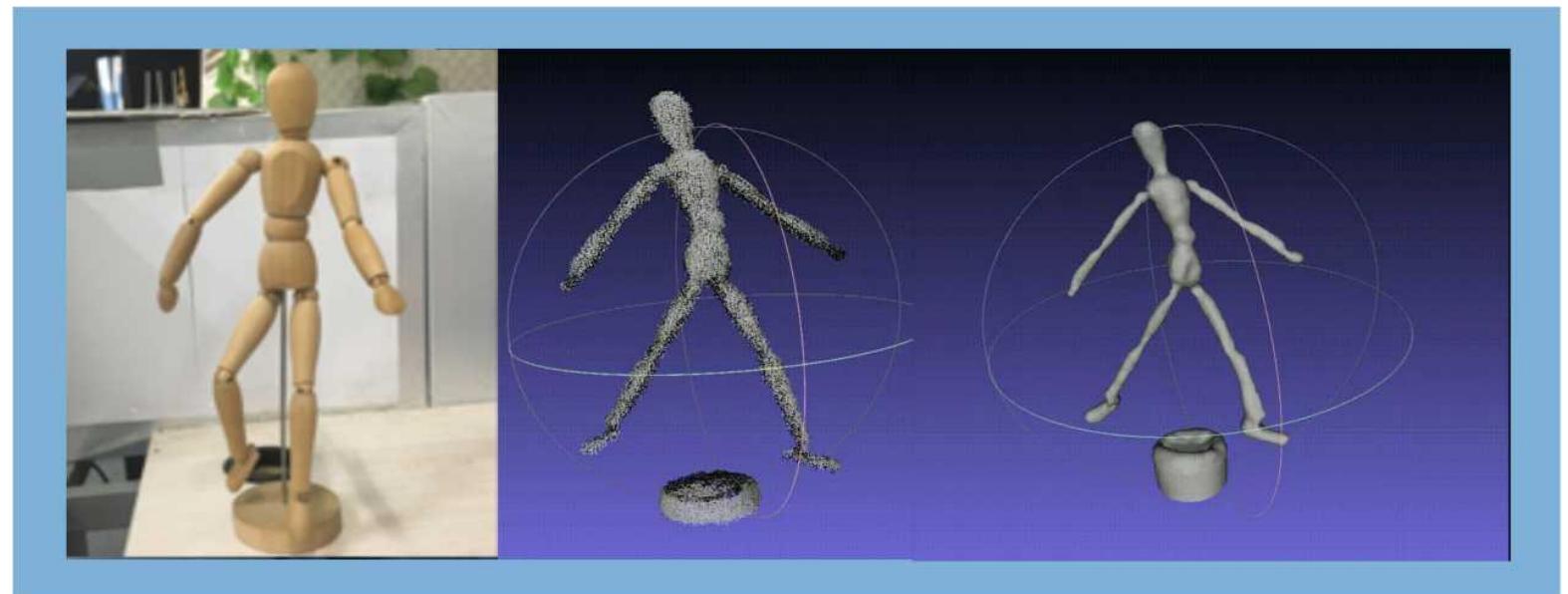
3D数据面临的难点

- 噪声
- 干扰 (比如多径)
- 畸变
- 分辨率
- 结构不完整
- 模糊

理想的3D数据



现实情况



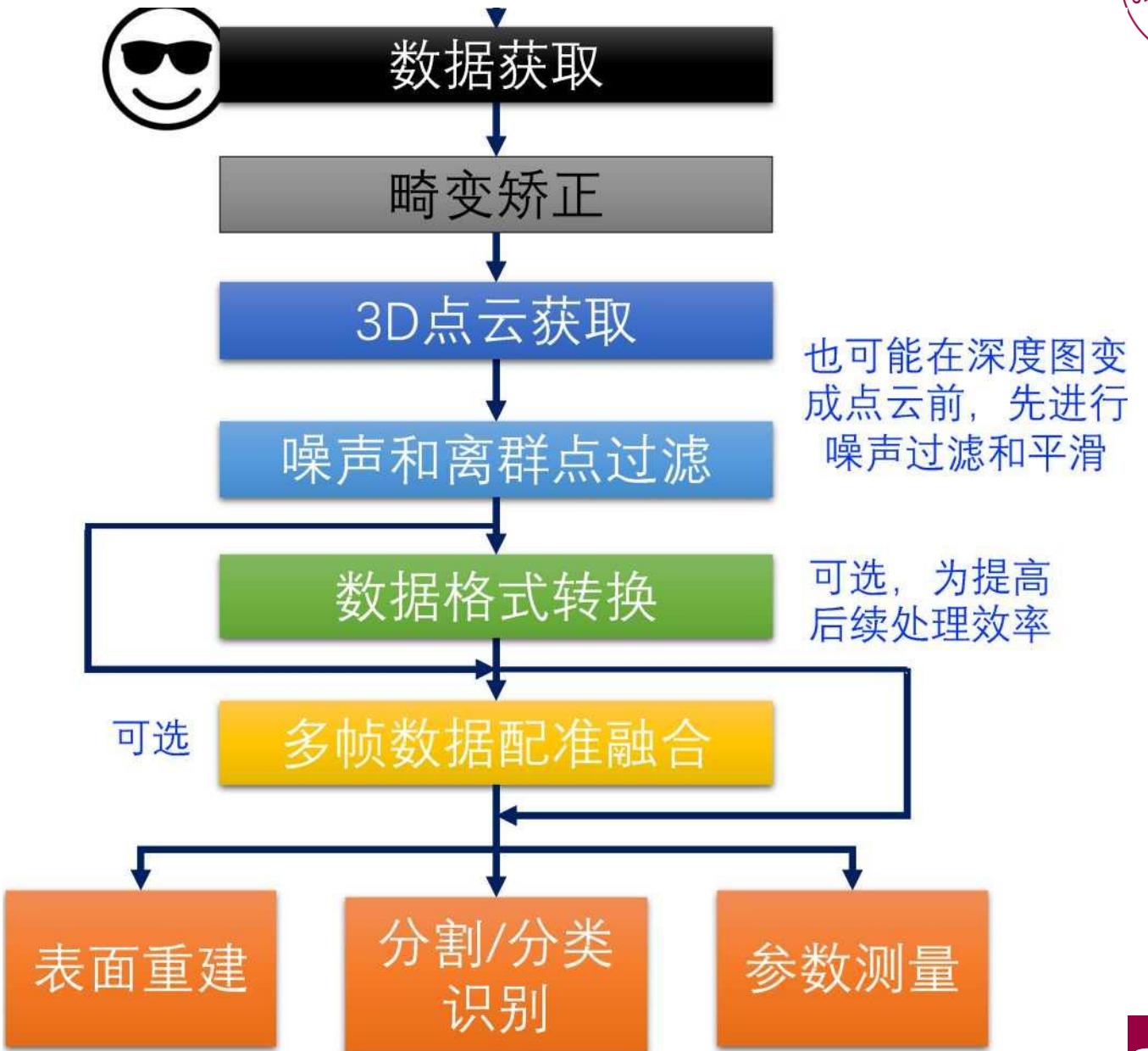
CONTENTS

目录

03 3D感知方法

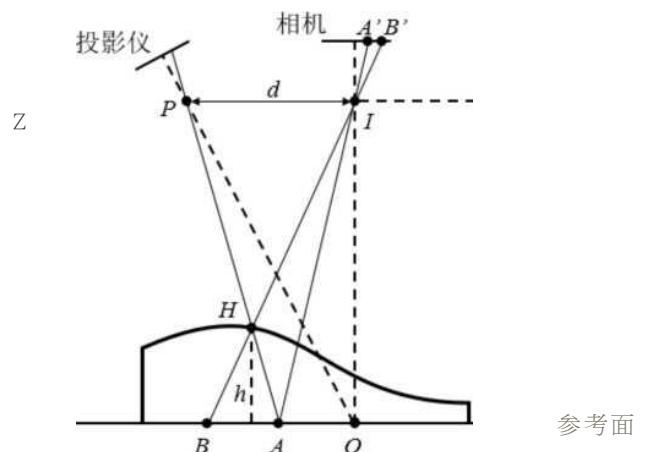
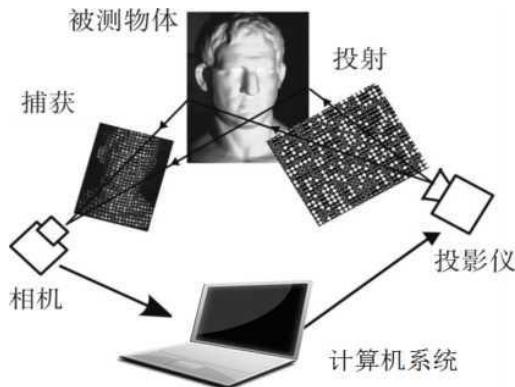
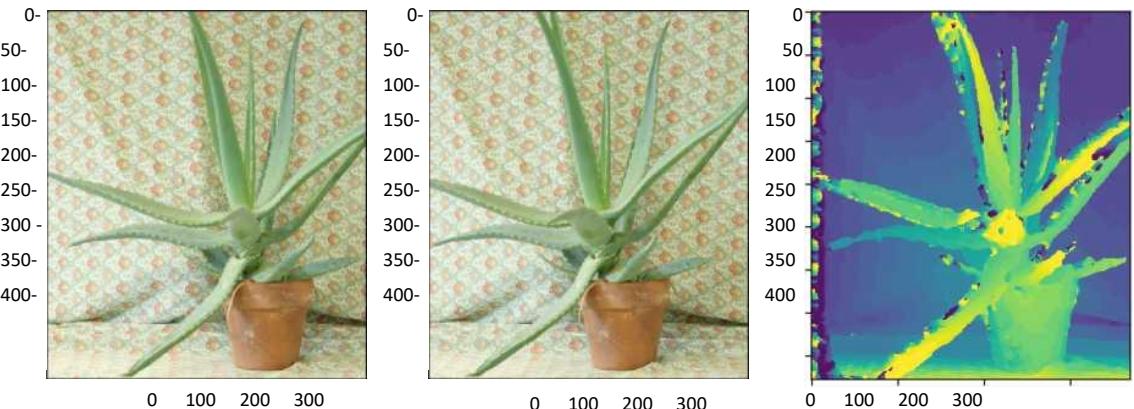
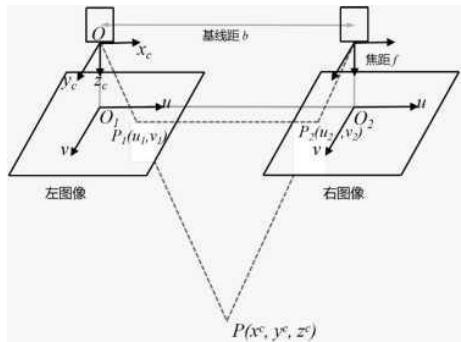
3D数据获取与处理

- 3D建模与应用需要融合多种算法，后面介绍流程中各环节涉及的算法

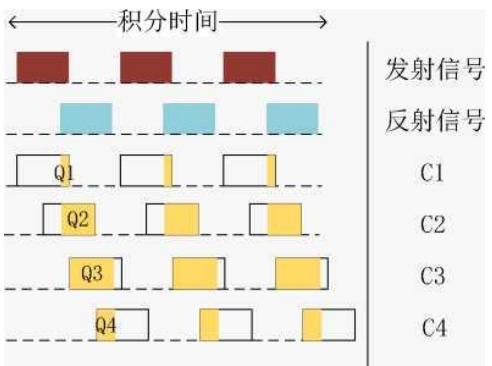
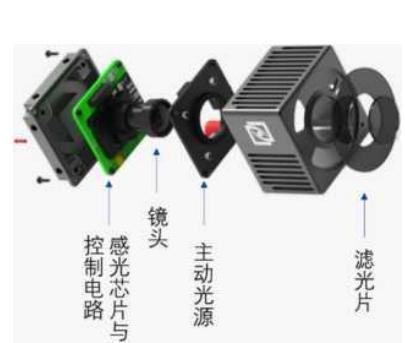


3D数据获取技术

双目匹配3D 成像算法

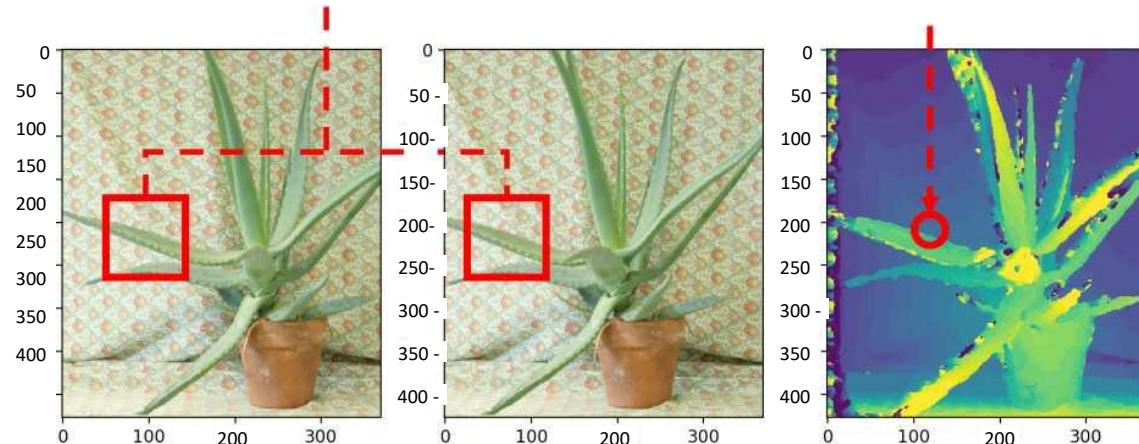


结构光和光编码成像原理



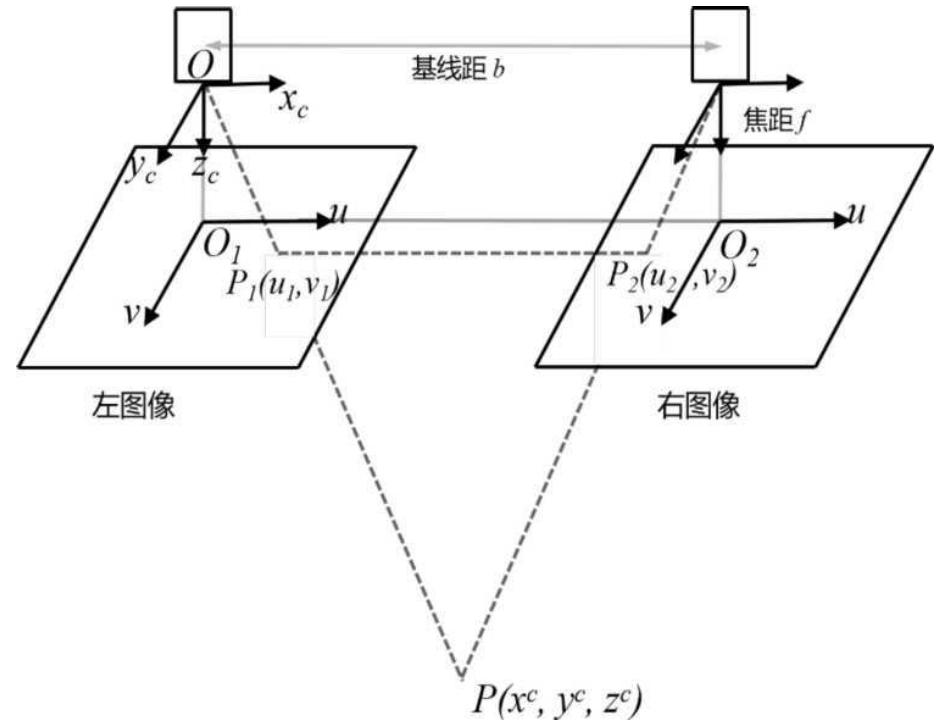
TOF 3D成像原理及信号模型

3D数据获取技术——双目匹配3D成像原理



$$距离 Z = \frac{Bf}{\Delta d}$$

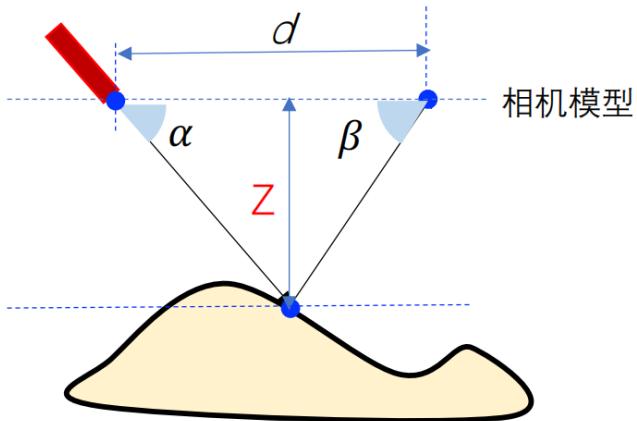
基线距 焦距
视差



双目匹配3D成像原理

3D数据获取技术-结构光3D成像原理

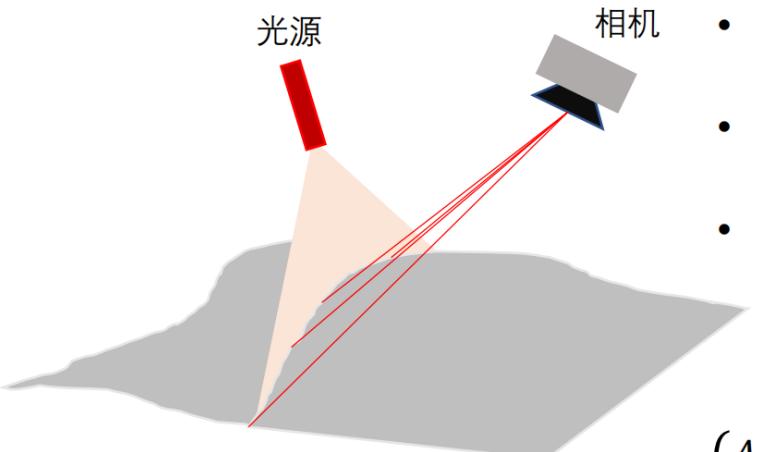
光源



$$\frac{Z}{\tan\alpha} + \frac{Z}{\tan\beta} = d$$



$$Z = d \frac{\sin\alpha \sin\beta}{\sin(\alpha + \beta)}$$



- (x, y) 是相机传感器平面上的点 (已知)
- (X, Y, Z) 是和 (x, y) 对应的空间点的坐标 (未知)
- f 是相机内参

光平面模型

$$\left\{ \begin{array}{l} AX + BY + CZ + D = 0 \\ X = \frac{xZ}{f} \\ Y = \frac{yZ}{f} \end{array} \right. \quad \text{针孔相机模型}$$

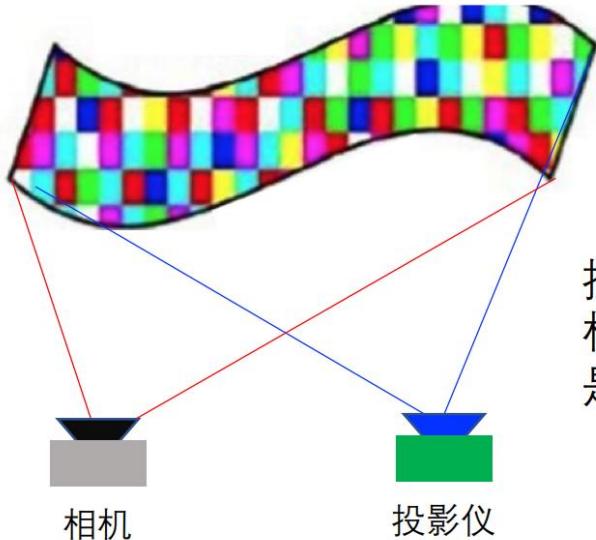


结构光的多种形式

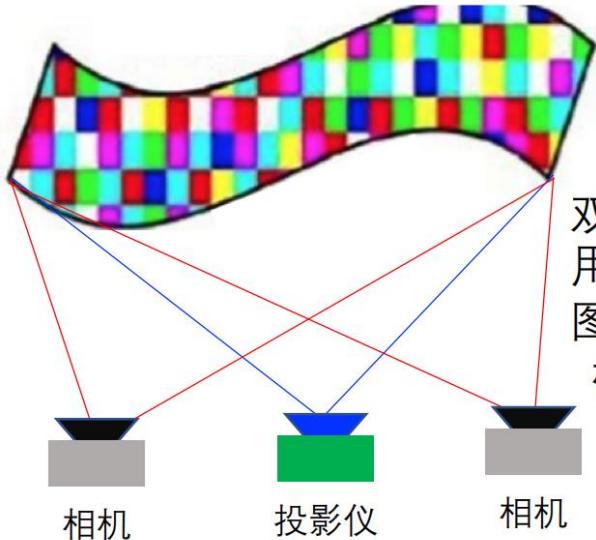
- 点结构光
- 线结构光
- 多线结构光
- 面结构光

$$Z = -\frac{Df}{Ax + By + Cf}$$

3D数据获取技术——光编码成像原理



投影一个面到物体，接收相机通过图案编码识别出是投影仪哪个角度的“射线”，并计算距离



双目立体视觉，利用投影在物体上的图案，帮助两个相机计算双目匹配



Kinect v1

结构光的多种形式

- 点结构光
- 线结构光
- 多线结构光
- 面结构光

3D数据获取技术-激光扫描基本原理



Light Detection and Ranging (LiDAR)

光的探测和测距

LAser Detection and Ranging(LADAR)

激光的探测和测距



具有极高的分辨率

获取的信息量丰富

抗干扰能力强

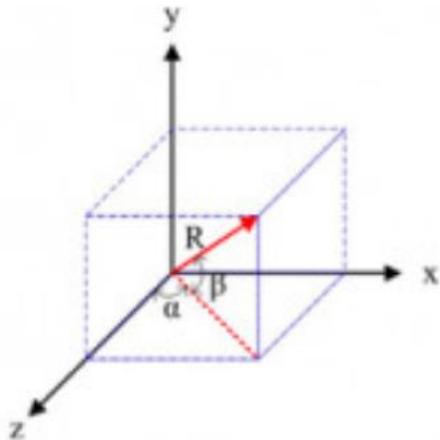
可全天时工作

容易受到大气条件以及
工作环境的烟尘的影响

3D数据获取技术-激光扫描基本原理

按照应用分类：

- 测距激光仪
- 激光三维成像雷达
- 激光交汇对接雷达
- 激光测速雷达
- 激光大气探测雷达
- 激光测风雷达



获取的信息：

- ◆ 距离：R
- ◆ 方向角度： α
- ◆ 俯仰角度： β

$$x = R \cdot \cos \beta \cdot \sin \alpha$$

$$y = R \cdot \sin \beta$$

$$z = R \cdot \cos \beta \cdot \cos \alpha$$

■ 激光三维成像雷达

- 是目前获取大范围三维场景图像效率最高的传感器
- 是目前获取大范围三维场景精度最高的传感器

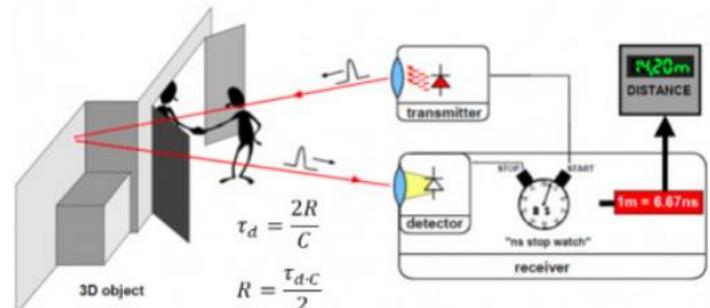
3D数据获取技术-激光扫描基本原理

激光测距方法

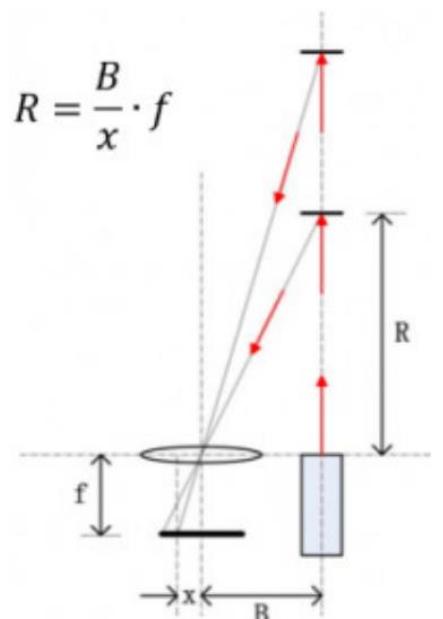
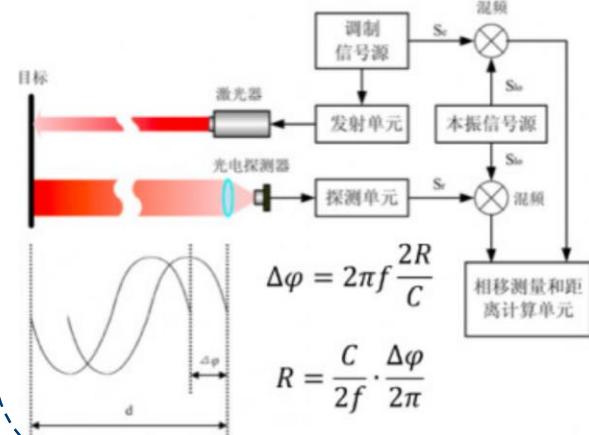
激光飞行时间 (TOF) 法

三角法

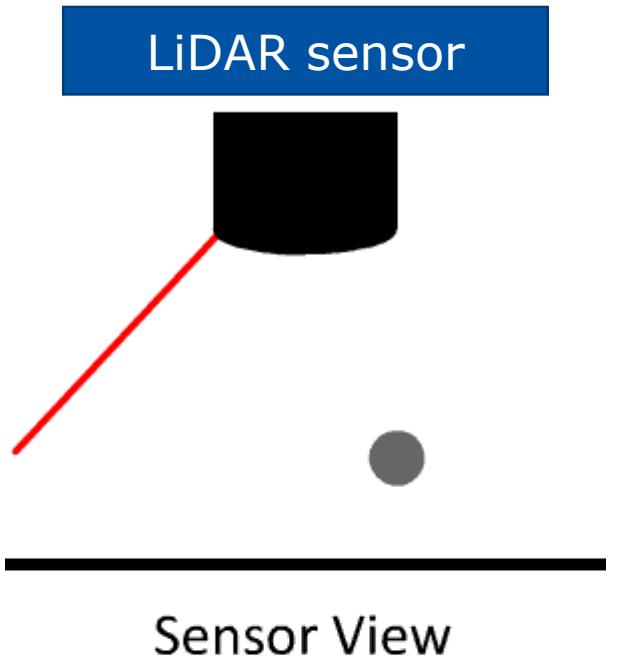
脉冲测距



相位测距



3D数据获取技术-激光扫描基本原理



3D数据获取技术-激光扫描基本原理

Velodyne Lidar®



ibeo
automotive



robosense
速腾聚创

SICK
Sensor Intelligence.



 HOKUYO



3D数据获取技术-常用的激光扫描设备

机载激光扫描



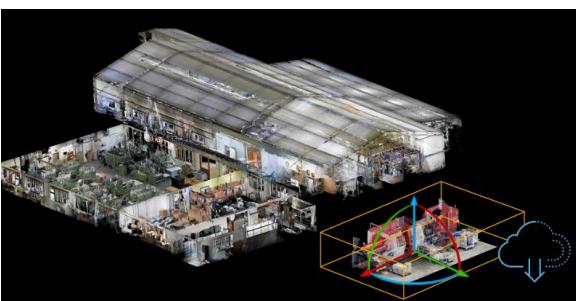
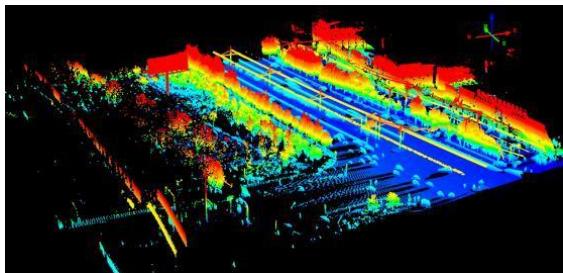
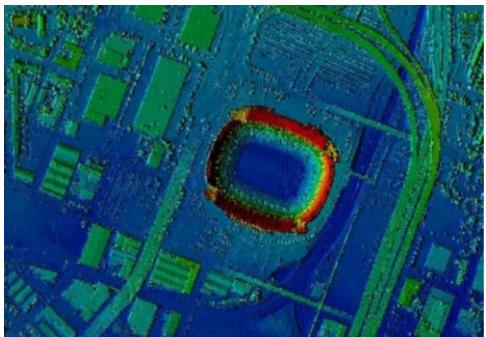
车载激光扫描



手持激光扫描



推车式激光扫描



- 用于获取 DSM、DEM、DTM、建筑物；
- 电力线巡检
- 地表变化监测等

- 用于获取高精地图制作，辅助/自动驾驶

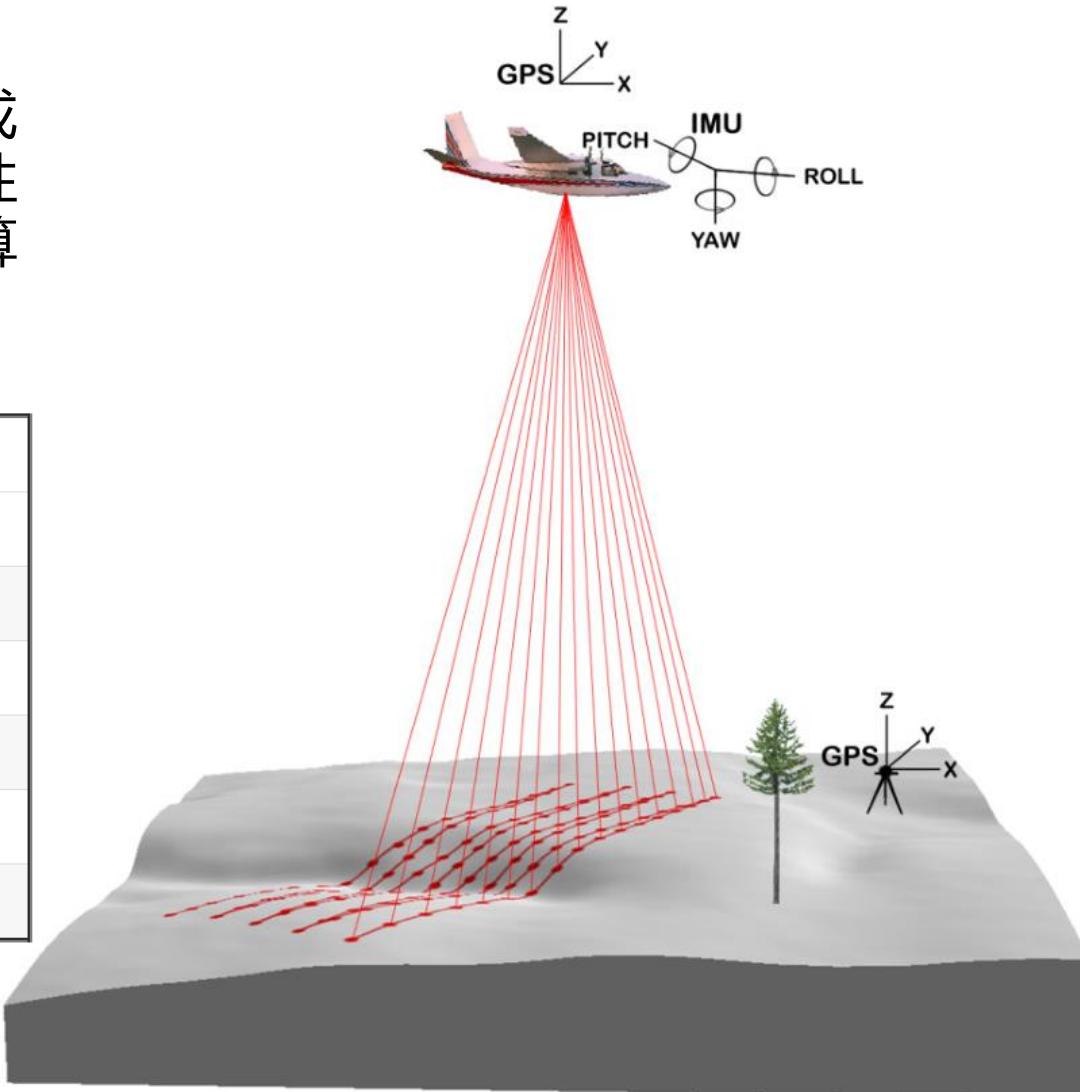
- 用于地下空间、室内空间三维数据扫描

- 用于室内空间三维数据扫描

33D数据获取技术-常用的激光扫描设备

- 机载 LIDAR 系统是一个集多种电子设备于一身的集成系统。其组成包括激光测距仪、动态差分 GPS、惯性测量单元 (IMU) 、激光器机械扫描系统、配套的计算机及其软件

部件	作用
激光测距仪	发射激光脉冲并接收回波
动态差分 GPS	用于飞机导航和遥感器坐标确定
惯性测量单元 IMU	确定激光扫描仪的姿态
计算机及软件	激光扫描数据的在线处理
存储设备	记录作业时获取的各种原始数据
成像设备(可选)	与激光扫描仪同步获取光学航空影像

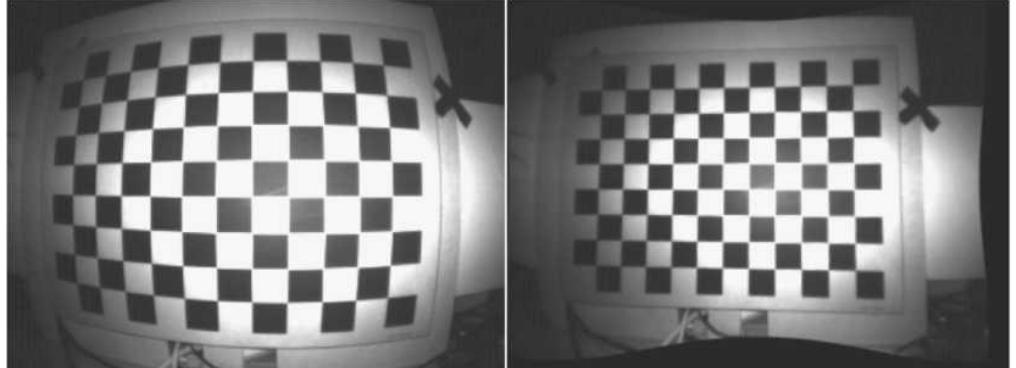


CONTENTS

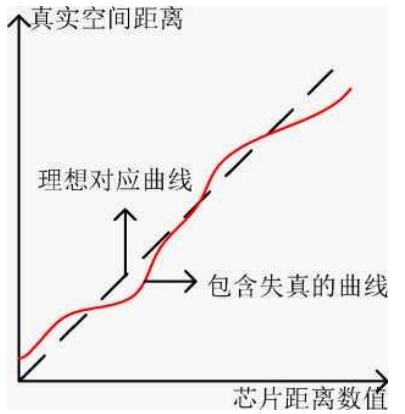
目录

04 3D数据处理方法

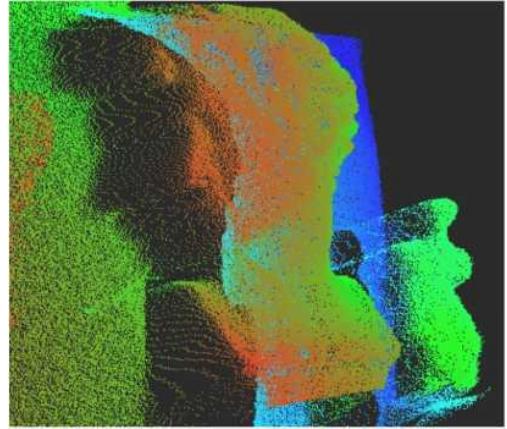
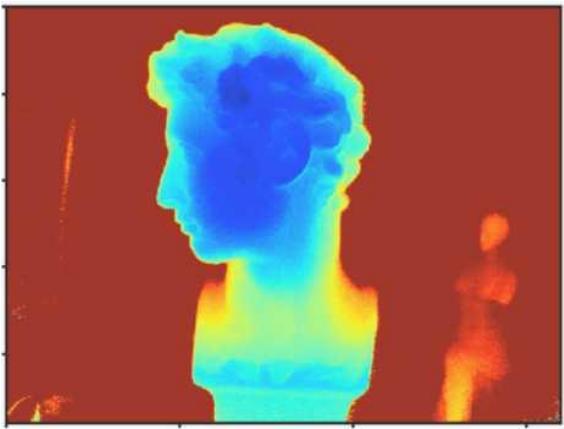
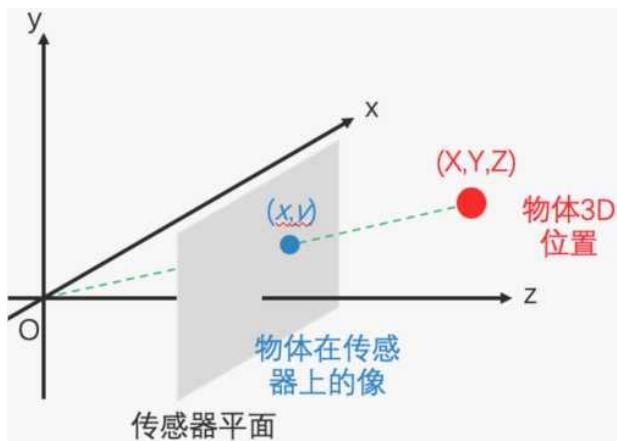
3D信号前处理算法



镜头失真数学模型及畸变校准

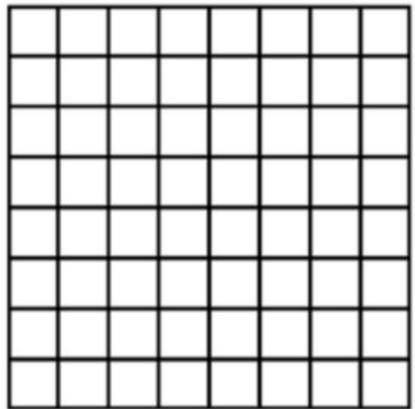


测距失真矫正模型计算

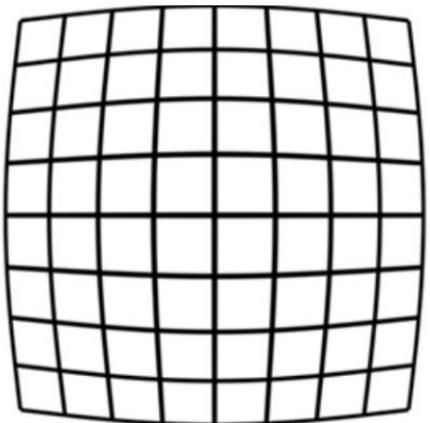


针孔相机模型及深度图到点云的转换计算

3D信号前处理算法-镜头畸变矫正

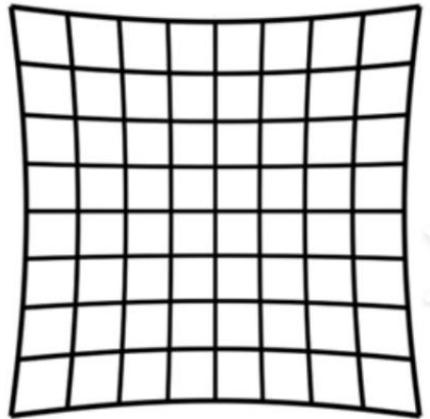


Object



Barrel Distortion

桶形失真



Pincushion Distortion

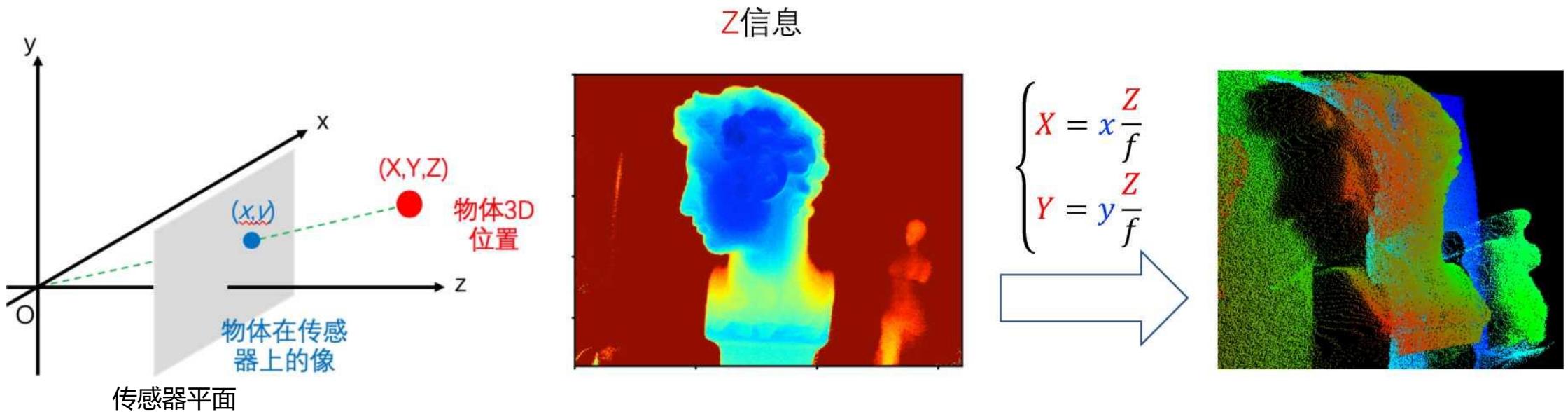
枕形失真

$$\begin{bmatrix} x_{corr} \\ y_{corr} \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2p_1 xy + p_2(r^2 + 2x^2) \\ p_1(r^2 + 2y^2) + 2p_2 xy \end{bmatrix}$$

通过畸变模型
调整像素位置

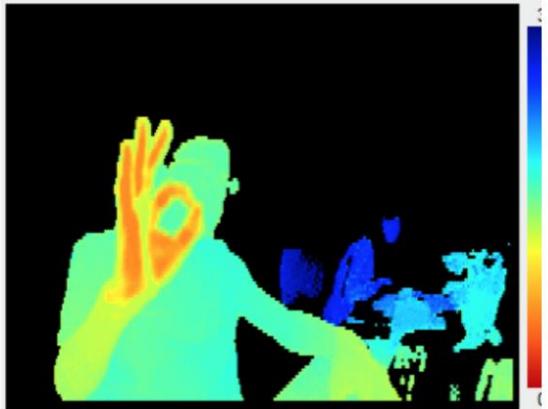
$$r^2 = x^2 + y^2$$

3D信号前处理算法-深度图到点云的转换计算

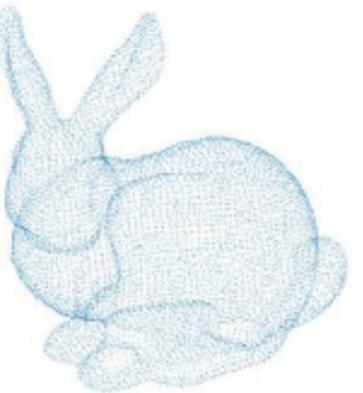


- (x, y) 是相机传感器平面上的点 (已知)
- (X, Y, Z) 是和 (x, y) 对应的空间点的坐标 (未知)
- f 是相机内参

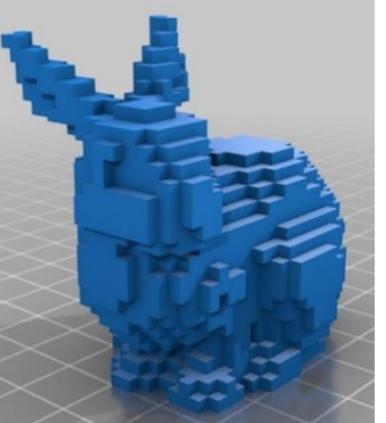
3D数据表示和数据结构



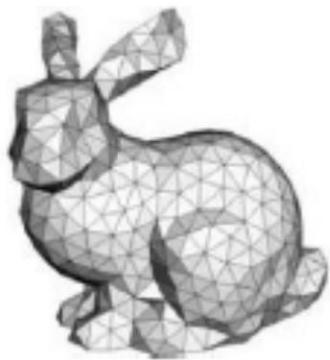
深度图



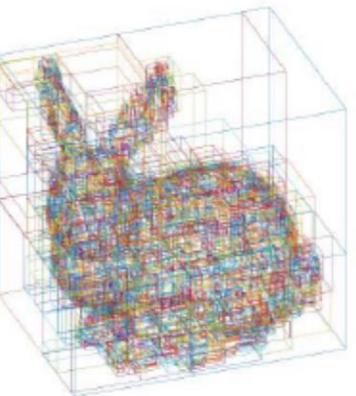
点云



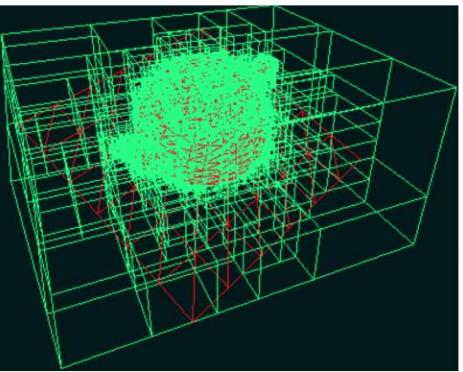
体像素



三角剖分

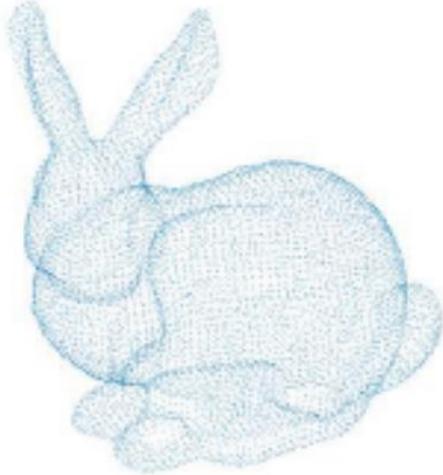


八叉树



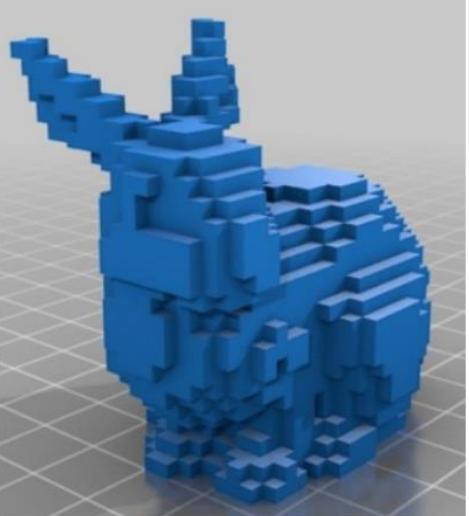
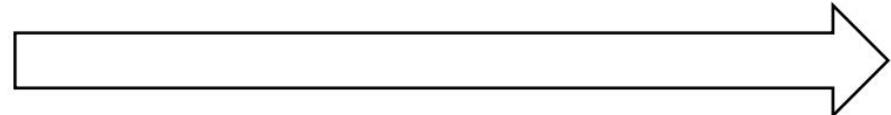
KD树

3D数据表示和数据结构—体像素



点云

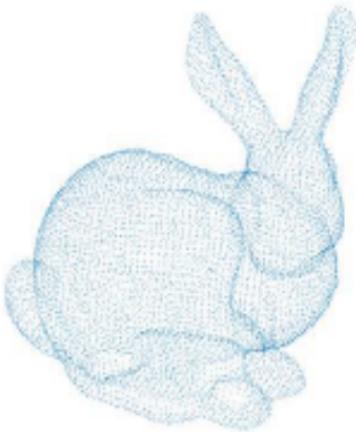
- 将空间划分为立方体,
- 实心立方体是那些内部包含了点云的立方体
- 程序中通过三维数组存储体像素



体像素

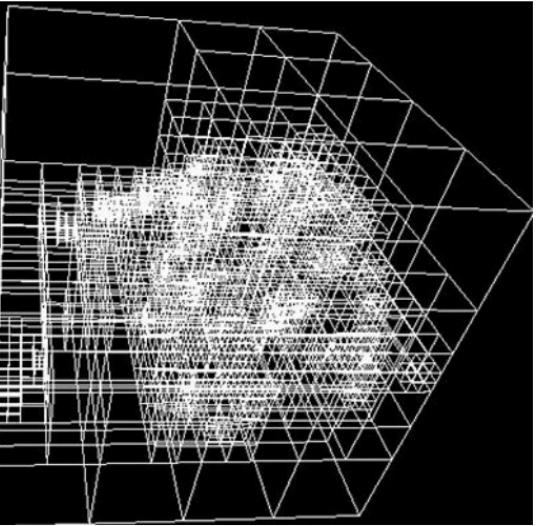
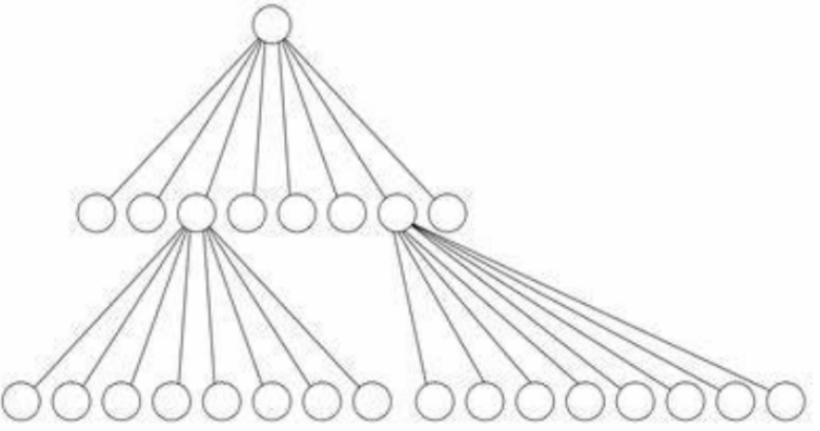
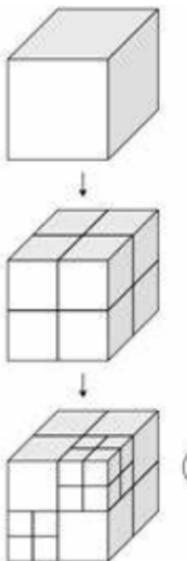
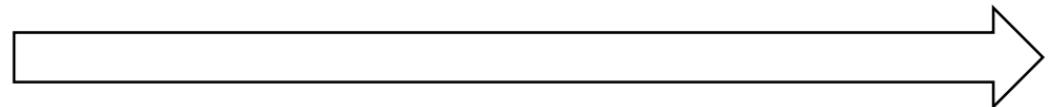
- 存储量很大
- 容易将多个3D集合体进行“逻辑运算”

3D数据表示和数据结构一八叉树



点云

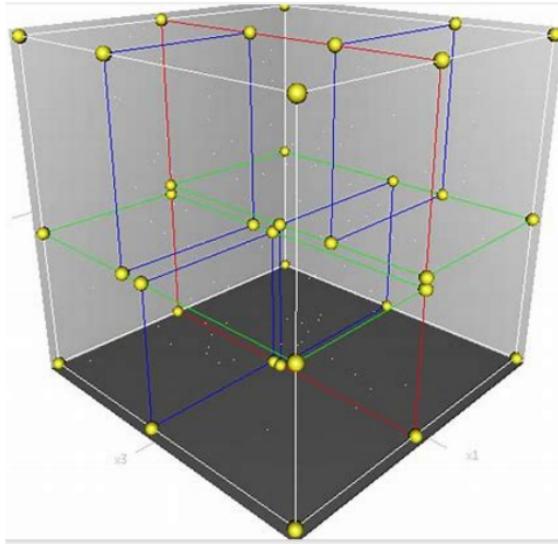
1. 构建一个大立方体，包围所有点云
2. 如果立方体内点云数量超过1时，将其分为8块，并对其中每块重复步骤2，直到每个划分的立方体内点云数量不超过1



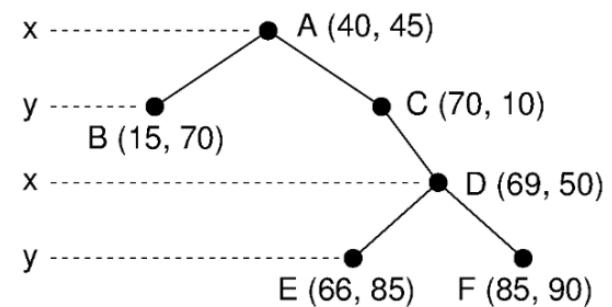
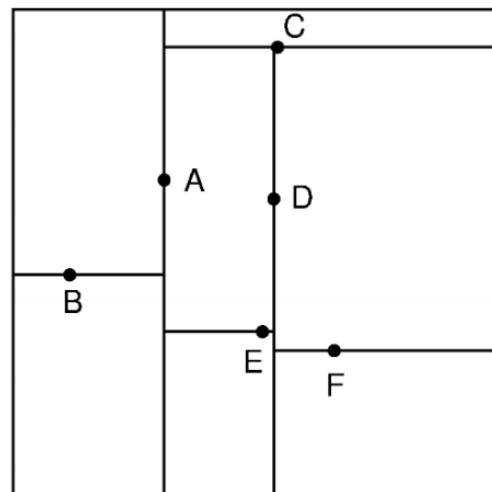
八叉树

- 比体像素更加高效的存储3D物体

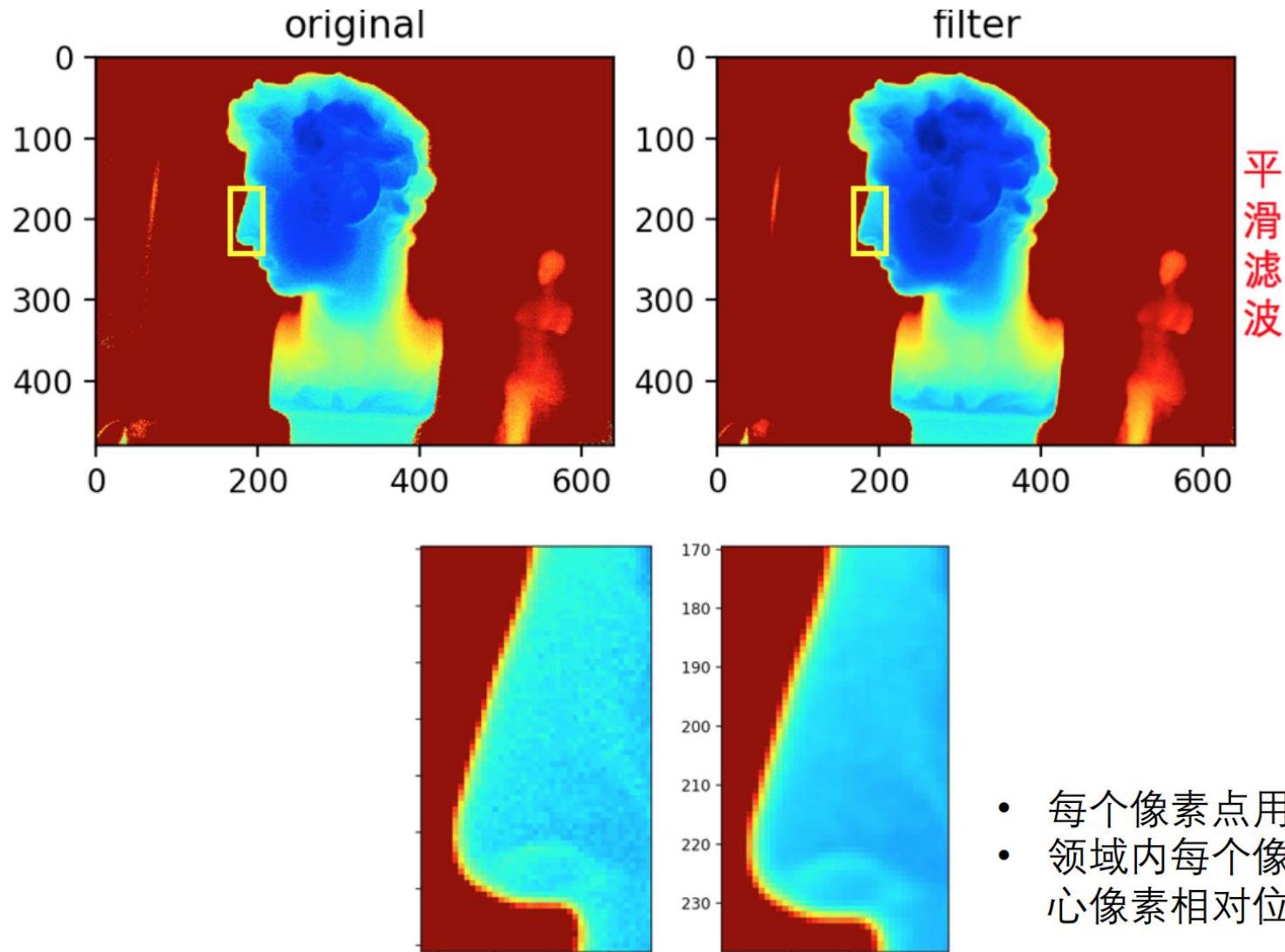
3D数据表示和数据结构—KD-Tree



1. 构建立方体包含所有点云
 2. 对每个包含超多1个点的立方体，构建分割平面，穿过其中一个点，将空间分成两块，分割面平行于XOY或者YOZ或者XOZ平面。
 3. 分割出来的两个子空间和分割平面上的点分别构成树枝的2个分叉及他们的连接点
 4. 对于分割得到的两个子空间，如果内部点的数量超过1，则对其继续分割（即执行2）
- 使用kd-tree是为了实现快速最近邻查找



3D信号处理—深度图平滑滤波



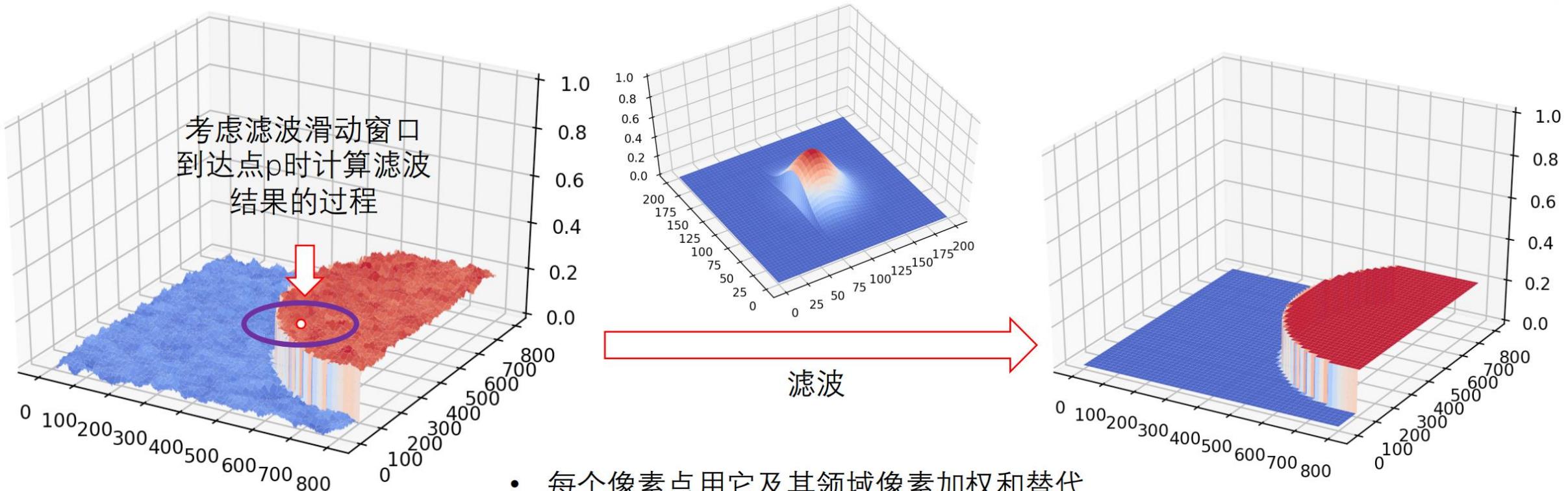
平滑滤波



2D卷积核直接
作用于深度图

- 每个像素点用它及其领域像素加权和替代
- 领域内每个像素对应的加权系数由它和中心像素相对位置决定，是固定的

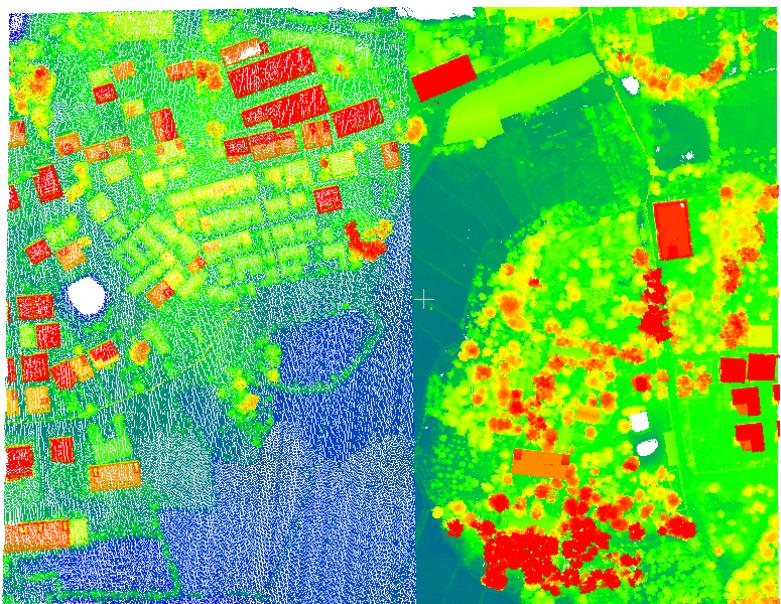
3D信号处理—深度图的双边滤波



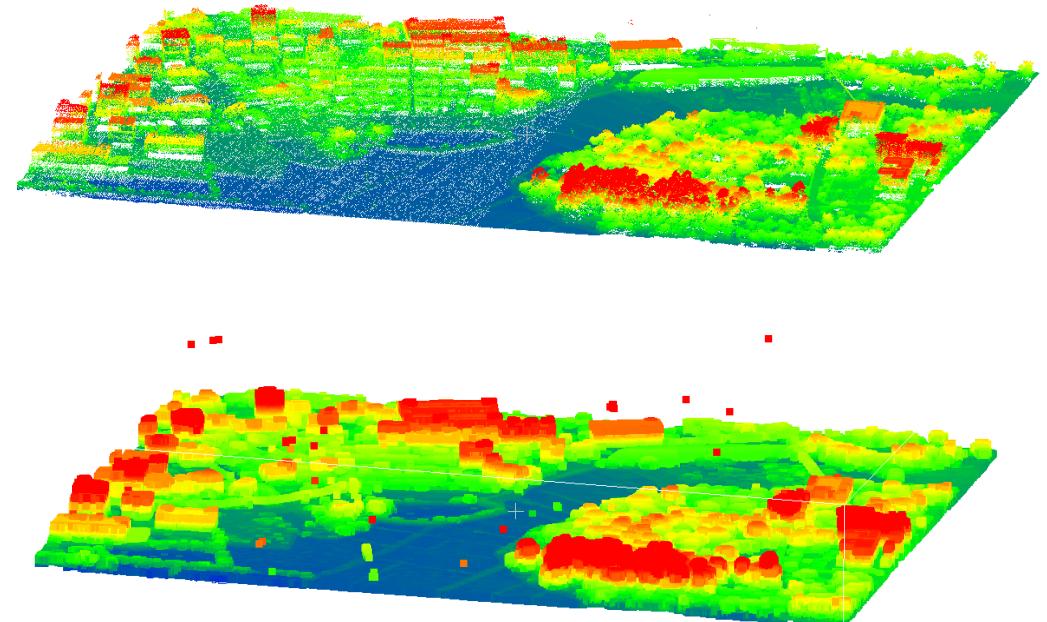
- 每个像素点用它及其领域像素加权和替代
- 加权系数和像素位置有关——相似的像素提高加权系数，不同的降低加权系数

3D信号处理-点云滤波

- 点云数据密度不规则需要平滑处理；
- 因为遮挡等问题造成离群点需要去除；
- 数据冗余或数据量过大需要降采样；
- 噪声数据需要去除；



点云数据密度不一致

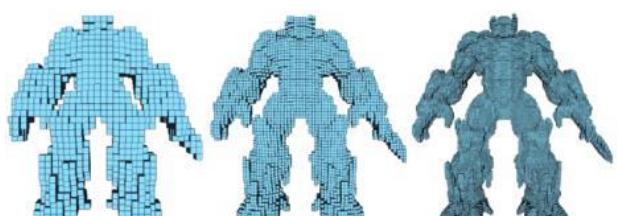


噪点数据

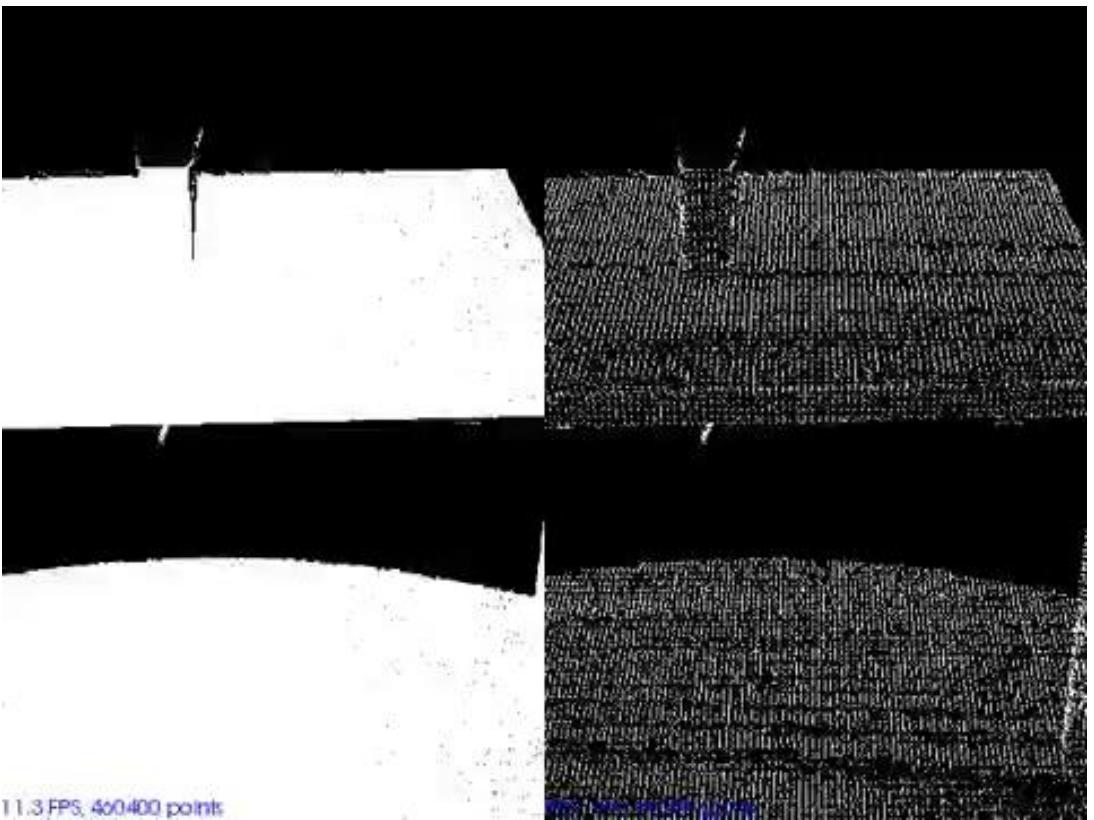
3D信号处理-点云滤波

■ 点云降采样

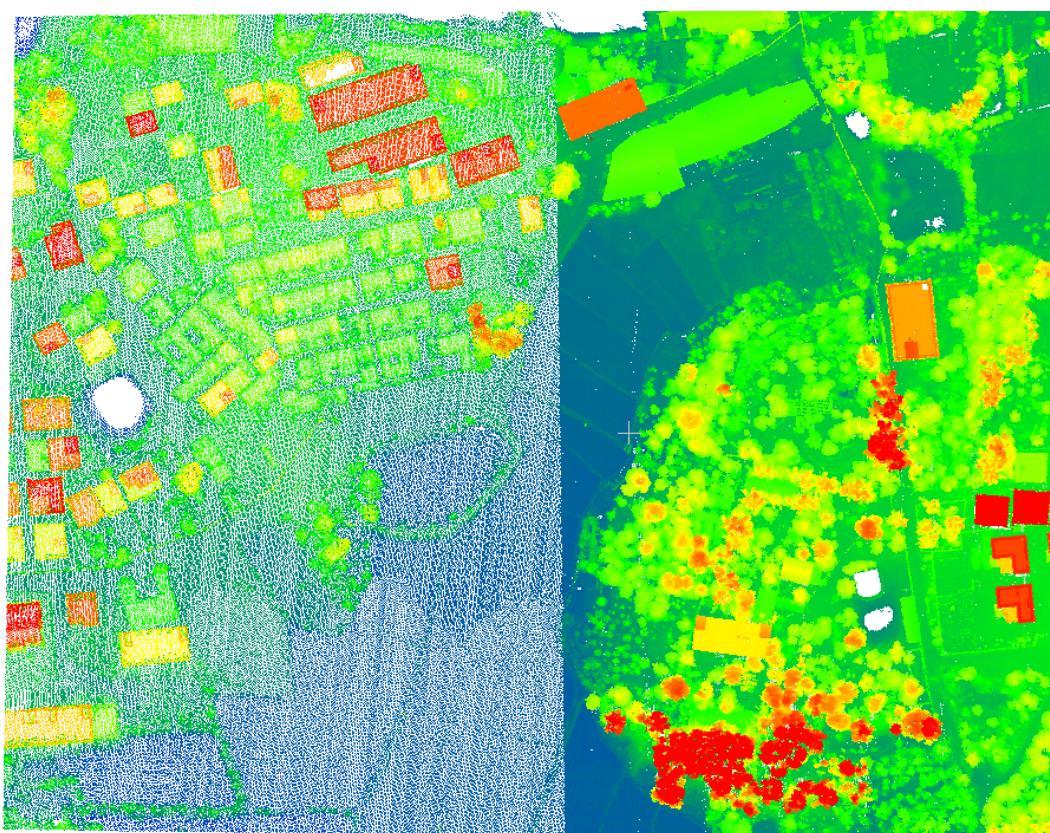
- 点云的数目越大，储存、操作都是个大问题！通过按一定的规则从里面抽取有代表性的样本，可以代替原来的样本，节省计算开销。



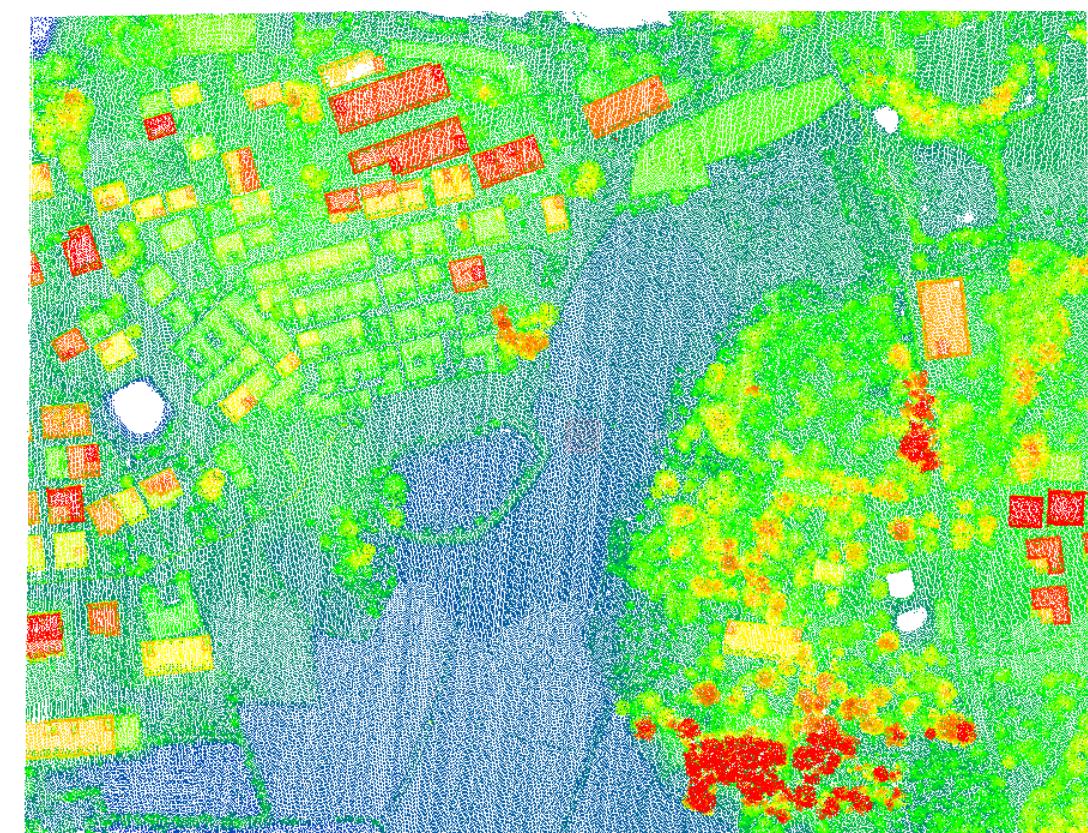
不同大小的体素网格



重采样前



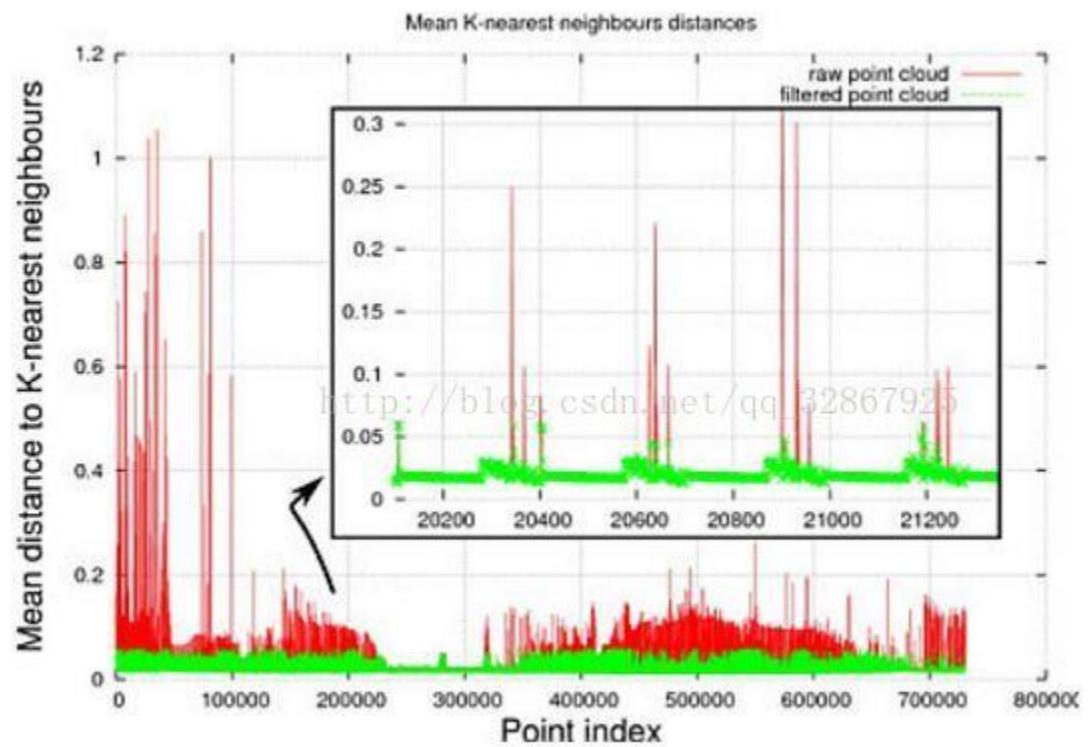
重采样后



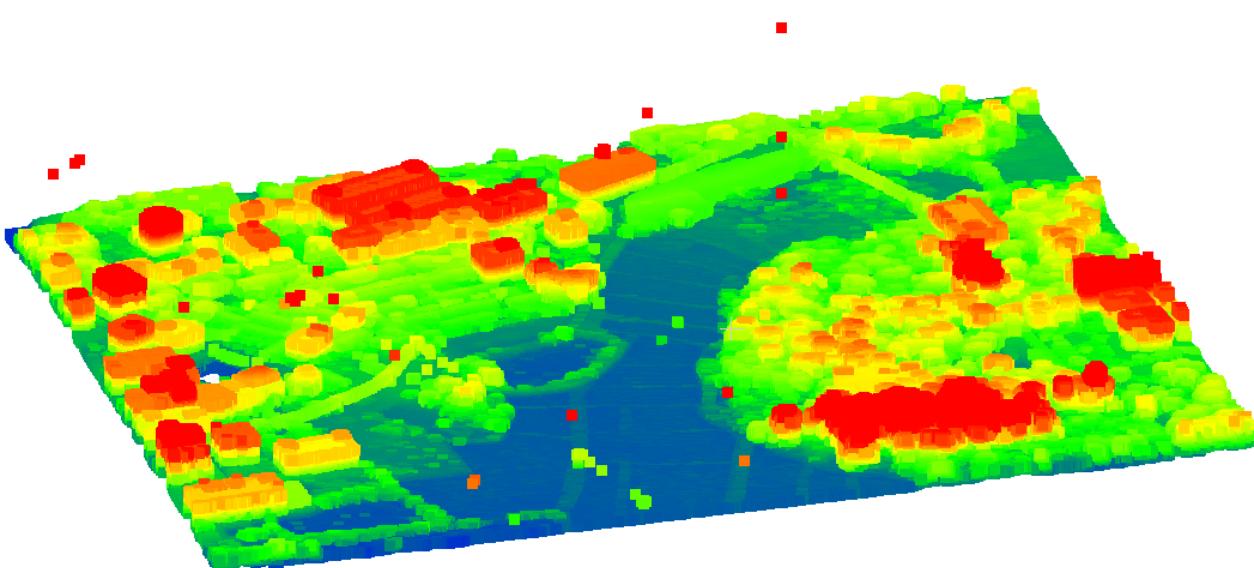
3D信号处理-点云滤波

■ 点云去噪

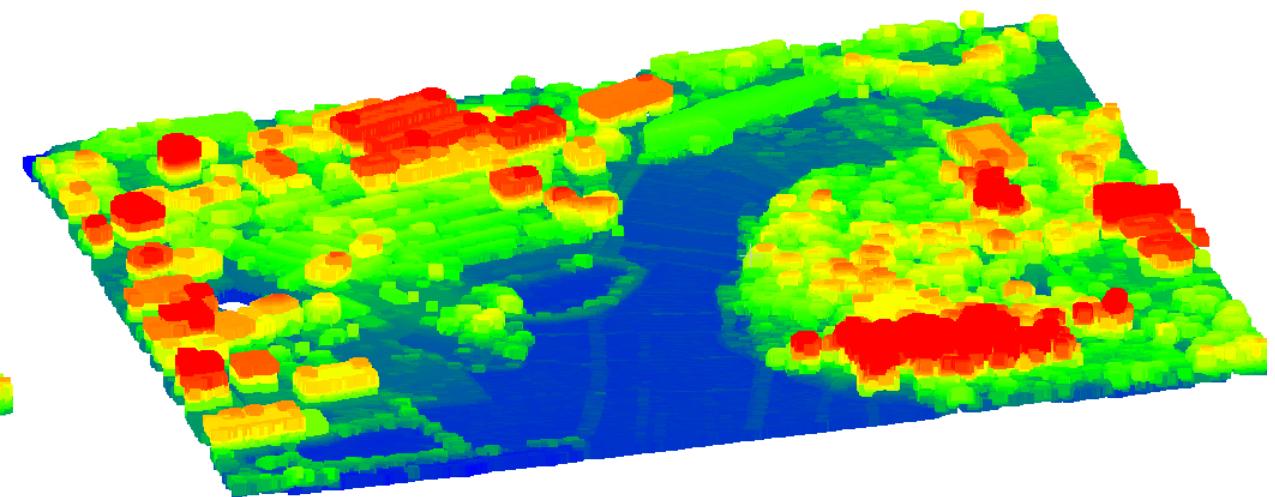
- 点云去噪主要用于剔除离群点，噪声数据，包括高斯滤波（标准差去噪）、条件滤波、直通滤波、随机采样一致性滤波等滤波方法。



去噪前



去噪后

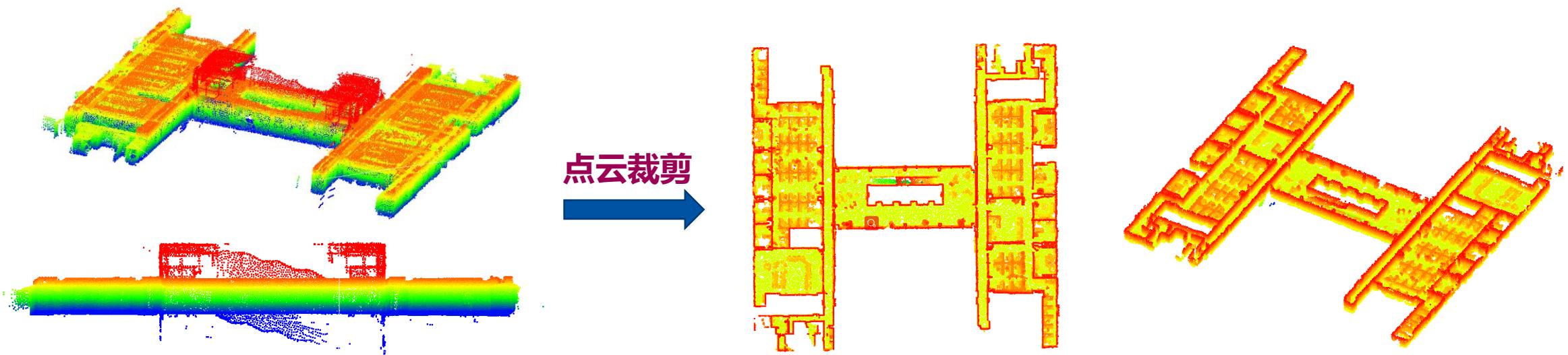


3D信号处理-点云裁剪与拼接



■ 点云裁剪与拼接

- 三维点云数据中存在大量目标无法通过自动化滤波算法过滤的噪点，需要通过手动裁剪将噪点剔除；
- 采用手动裁剪获取特定区域下的三维点云。



3D信号处理-三角网重构

■ 点云三角网重构

■ 实现三维点云三角网重构主要包含泊松重构、贪婪三角投影算法、移动立方体算法。

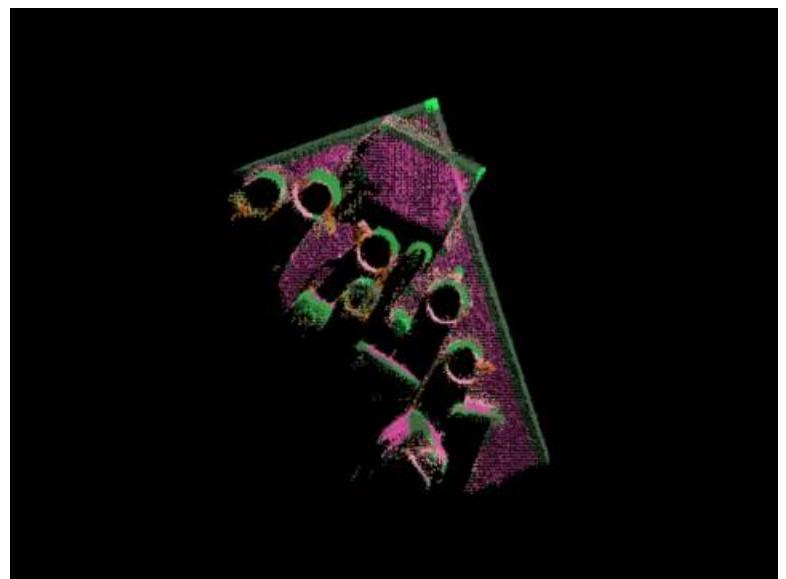
贪婪投影三角化算法

将有向点云投影到某一局部二维坐标平面内

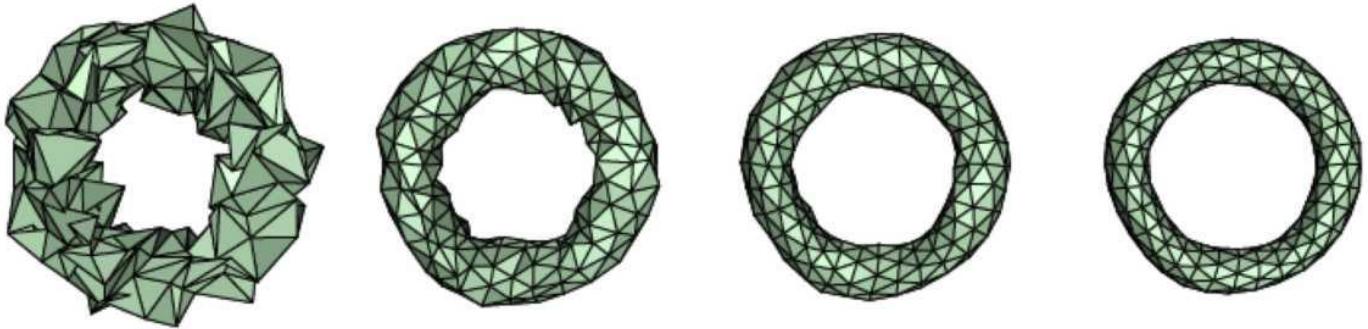
- 优点：可以处理来自一个或者多个扫描仪扫描得到并且有多个连接处的散乱点云；
- 局限：适用于点云来自于表面连续光滑的曲面并且点云密度变化比较均匀的情况；

在坐标平面内进行平面内的三角化

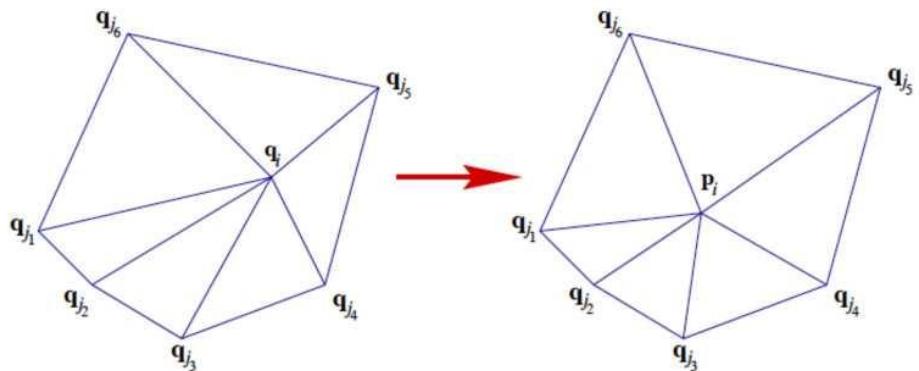
根据平面内三位点的拓扑连接关系
获得一个三角网格曲面模型



3D信号处理—表面平滑（三角网格表述的曲面）*5

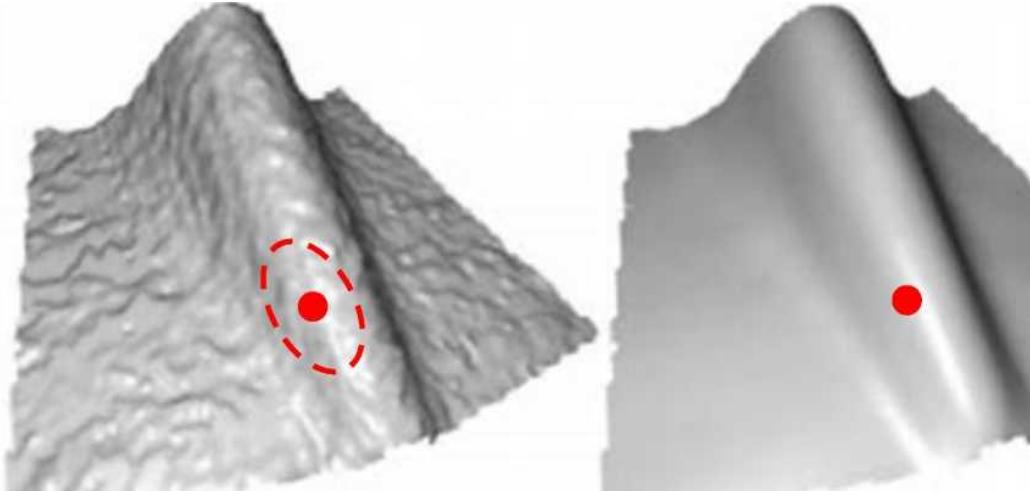
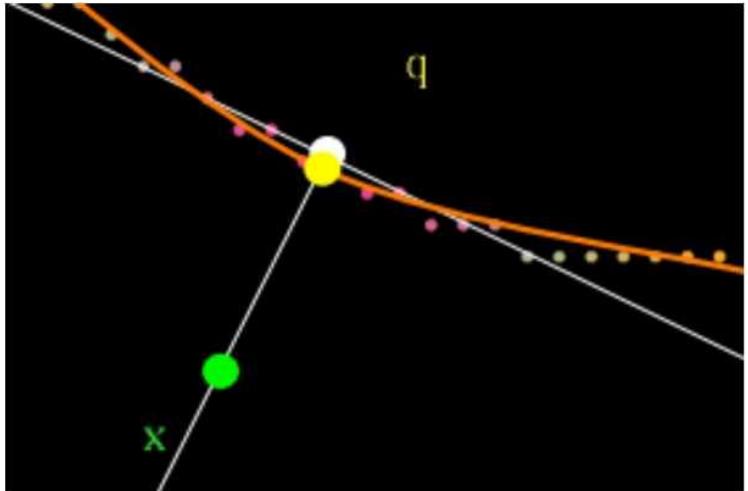


$$\mathbf{p}_i := \begin{cases} \frac{1}{|adj(i)|} \sum_{j \in adj(i)} \mathbf{q}_j, & i \in V_{var}, \\ \mathbf{q}_i, & i \in V_{fix}. \end{cases}$$



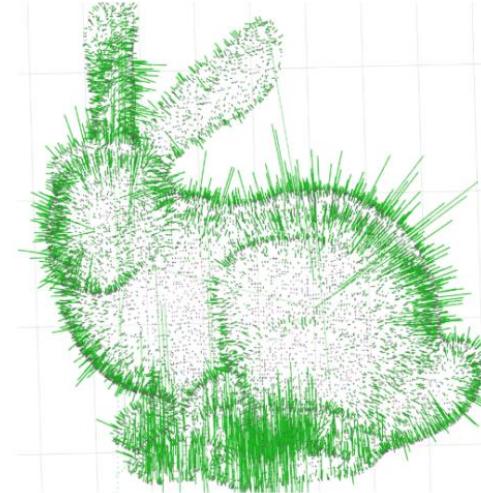
-Laplacian平滑算法——3D网格内的每个顶点，用它相连的顶点的坐标均值取代

3D信号处理—表面平滑

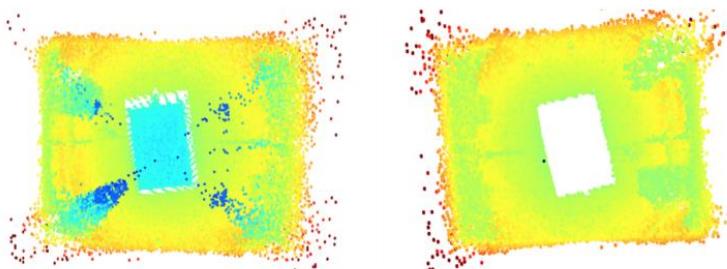
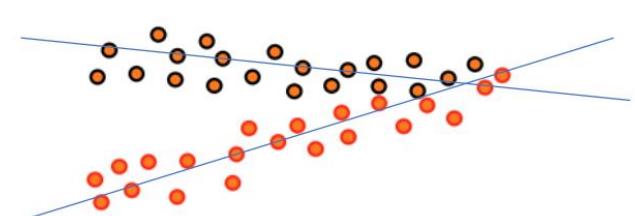


- Moving Least Squares 曲面平滑
- 以曲面上待平滑的点为中心，找到其邻域
- 根据邻域点的坐标，拟合曲面函数，并用曲面中心的函数值（坐标）取代原先点的值（坐标）

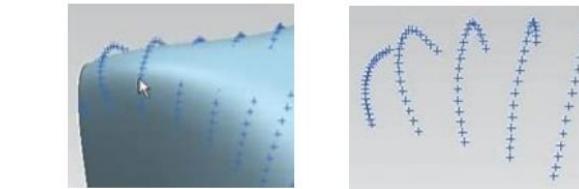
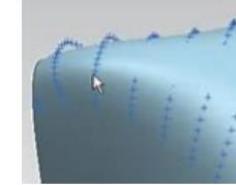
3D几何建模



矩阵分解表面法向量估计

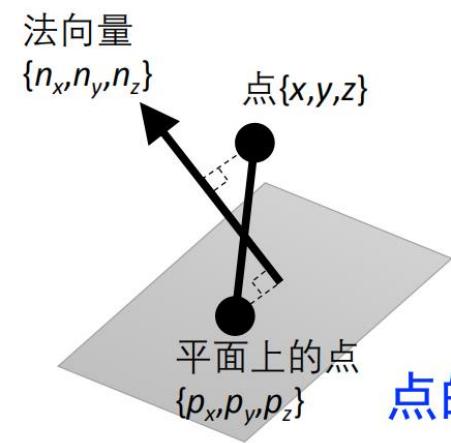


RANSAC点云平面检测

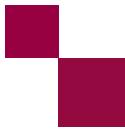


曲面参数提取算法

$$\begin{bmatrix} 1 & x_1 & y_1 & z_1 & x_1y_1 & x_1z_1 & y_1z_1 & x_1^2 & y_1^2 & z_1^2 \\ 1 & x_2 & y_2 & z_2 & x_2y_2 & x_2z_2 & y_2z_2 & x_2^2 & y_2^2 & z_2^2 \\ \vdots & \vdots \\ 1 & x_N & y_N & z_N & x_Ny_N & x_Nz_N & y_Nz_N & x_N^2 & y_N^2 & z_N^2 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \\ w_9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$



点的曲面投影

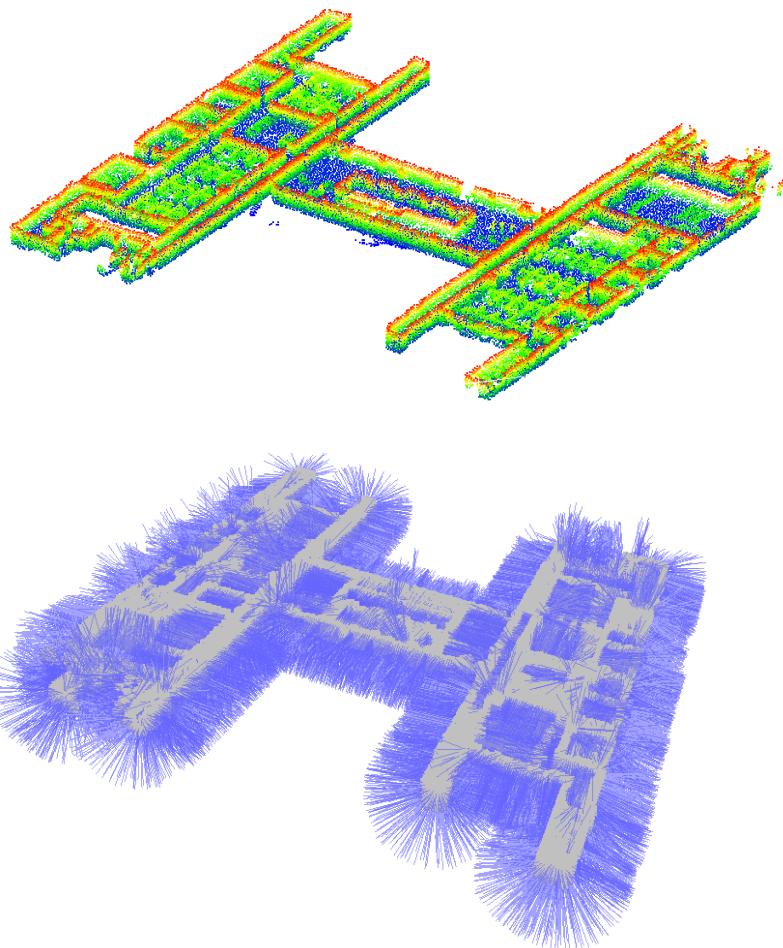
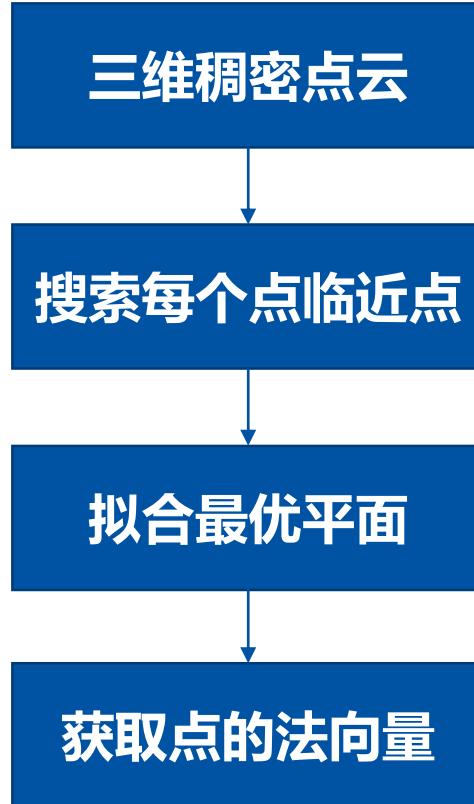


3D几何建模-点云法向量计算



■ 点云法向量计算

- 点云法向量的估计在很多场景都会用到，比如三维点云配准，以及曲面重建等，点云的法向量主要是通过点所在区域的局部拟合的表面进行计算



3D几何建模-平面提取

平面提取的应用例子：

- 在物体测量时，需要测量地平面位置和倾斜角度
- 平直的物体表面测量时，利用平面检测识别并切割出表面，并获得表面边界

$$Ax + By + Cz + D = 0$$

$$\left[\begin{array}{cccc} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_N & y_N & z_N & 1 \end{array} \right] \left[\begin{array}{c} A \\ B \\ C \\ D \end{array} \right] = \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array} \right]$$

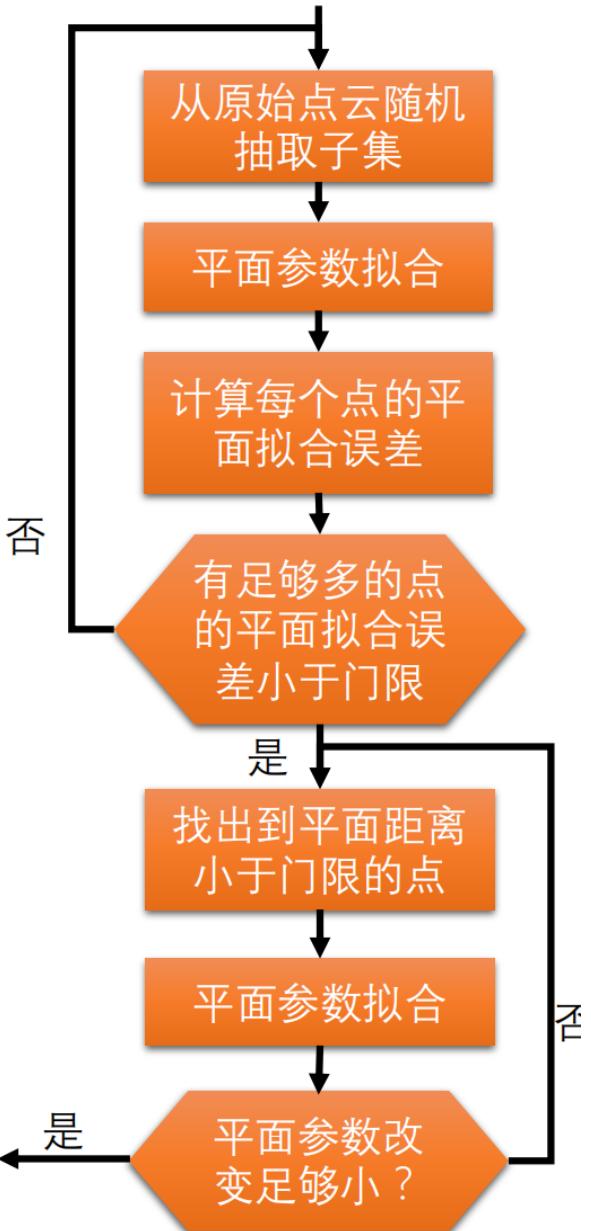
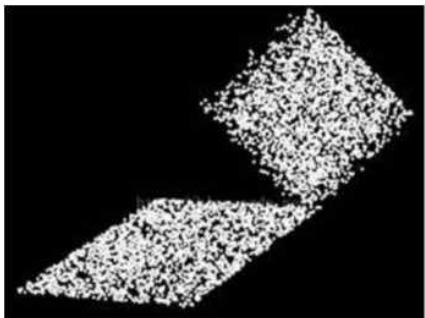
通过SVD计算矩阵行向量的0
空间得到(A, B, C, D)



注意：也可以利用法向量计算算法直接得到
(A, B, C)，再带入点云坐标方程计算D

3D几何建模-RANSAC平面提取

解决实际遇到的困难—非单一点云
-包含非平面上点的数据干扰
-可能有多个平面



3D重建-点云配准

■ 点云配准

- 三维点云数据可以来自于多平台，例如机载激光、车载激光以及手持和推车式三维传感系统，不同平台采集的数据坐标系统难以保持一致，需要进行点云数据配准；

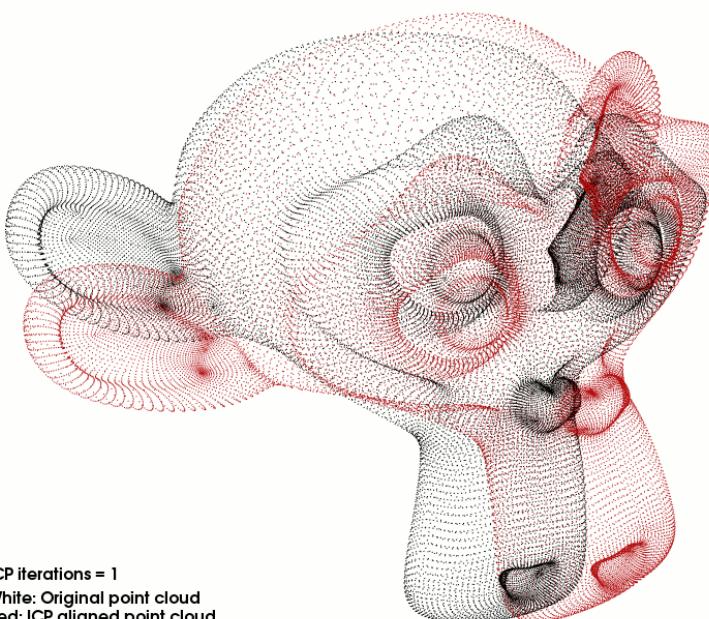
由粗到精细的配准方式

手动或者自动获取不同组点云数据间匹配点

计算粗匹配参数

ICP算法迭代

获取精匹配参数



计算点集Qi中的若干点在点集P中的最近点

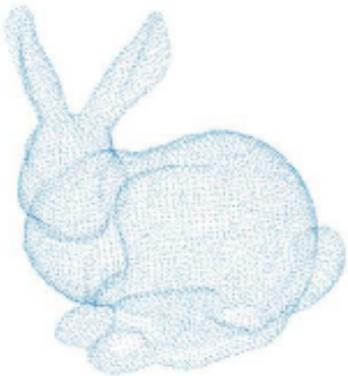
最近点对平均距离最小的旋转变换矩阵RTi

点集Qi使用上述求的RTi进行变换得到点集Qi + 1

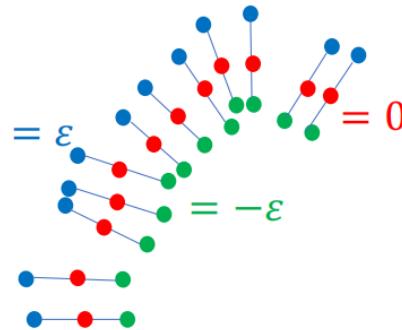
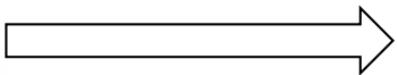
若点集P与点集Qi + 1的某些目标函数小于一定阈值，
 停止迭代

3D重建—SDF表面重建

应用例子：从稀疏的点云中构造出可以3D打印的模型



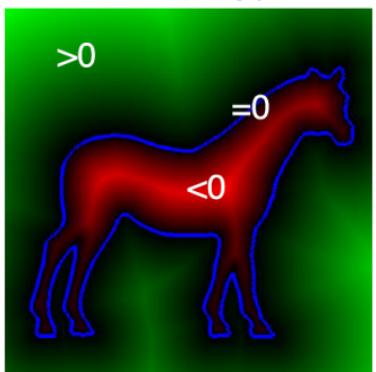
从零散稀疏点云构建成连续光滑表面



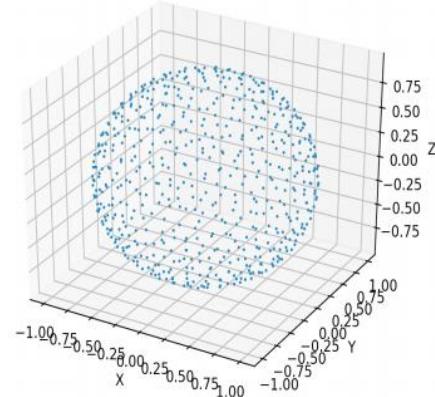
1. 对于点云中每个点，逐点计算法向量
2. 沿每个点的法向量向内和向外扩充点云
3. 构建SDF函数,
 - 该函数描述连续光滑表面
 - 使用若干给定的光滑函数加权和构建
 - SDF的取值满足：
 1. 原始表面=0
 2. 表面外侧沿法向量延伸e距离 := e
 3. 表面内侧沿法向量延伸e距离 := -e

$$F(\mathbf{p}) = \sum_i \mathbf{a}_i \phi(\|\mathbf{p} - \mathbf{p}_i\|) + \sum_i \mathbf{b}_i \phi(\|\mathbf{p} - (\mathbf{p}_i + \varepsilon \mathbf{n}_i)\|) + \sum_i \mathbf{c}_i \phi(\|\mathbf{p} - (\mathbf{p}_i - \varepsilon \mathbf{n}_i)\|)$$

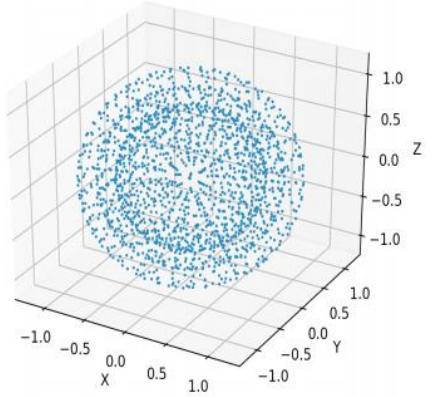
SDF函数特性



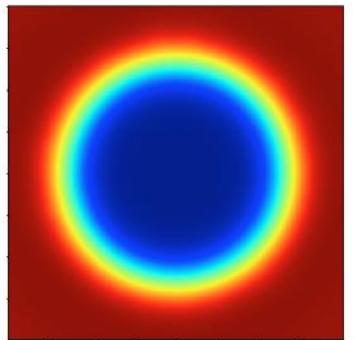
3D重建—Marching Cubes (三角剖分)



原始离散点云



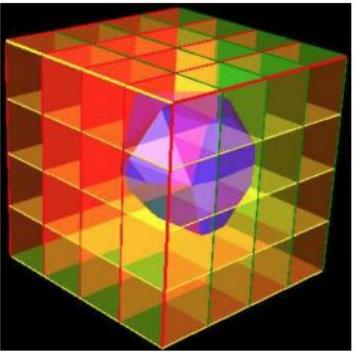
通过法向量拓
展的点云



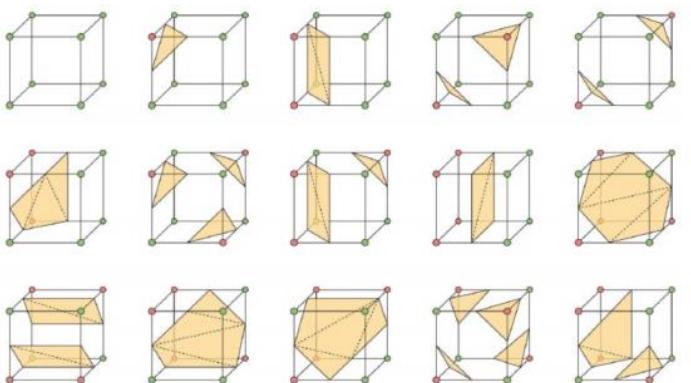
$z=0$ 的横截面
SDF函数值



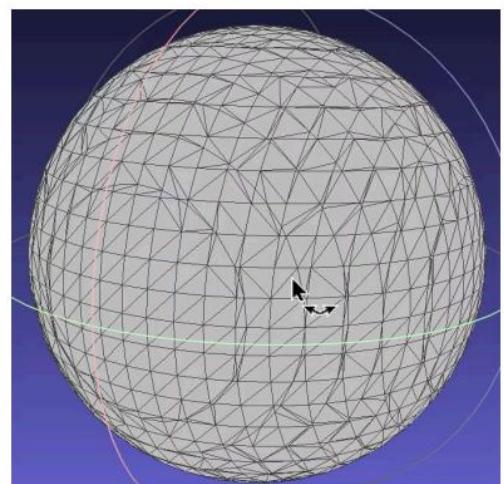
将空间分成大量小立方
格，计算每个立方格8顶
点的SDF函数符号



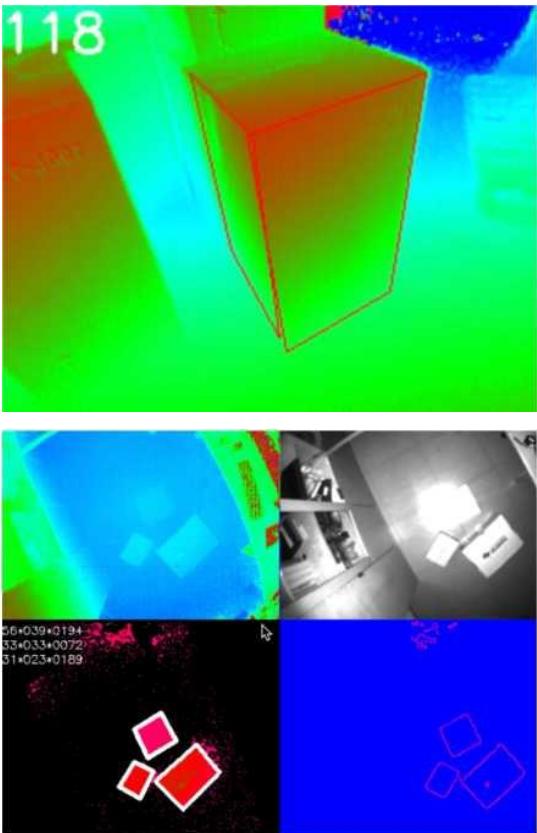
根据下面的规则，
确定是否在该立方
格内生成三角形片



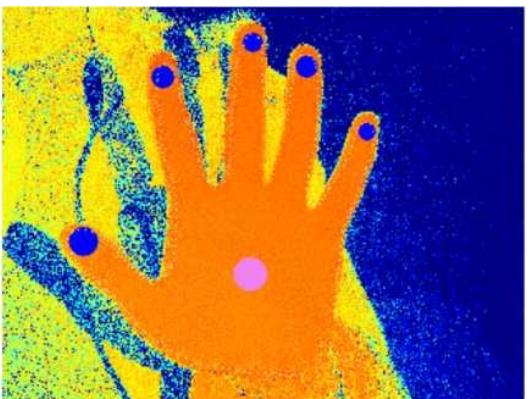
运算结果



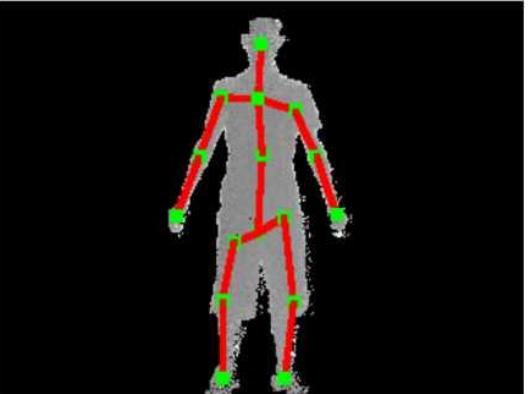
3D物体分割、识别与测量



空间几何测量算法示例



静态手势识别



3D人体骨架识别

基于神经网络的3D识别——两个路线方向



针对3D运算特点构建专用算法

3D Deep Learning:

3D Deep Learning

VoxNet, Octree CNN, Kdnetwork,
PointNet, PointNet++, ...



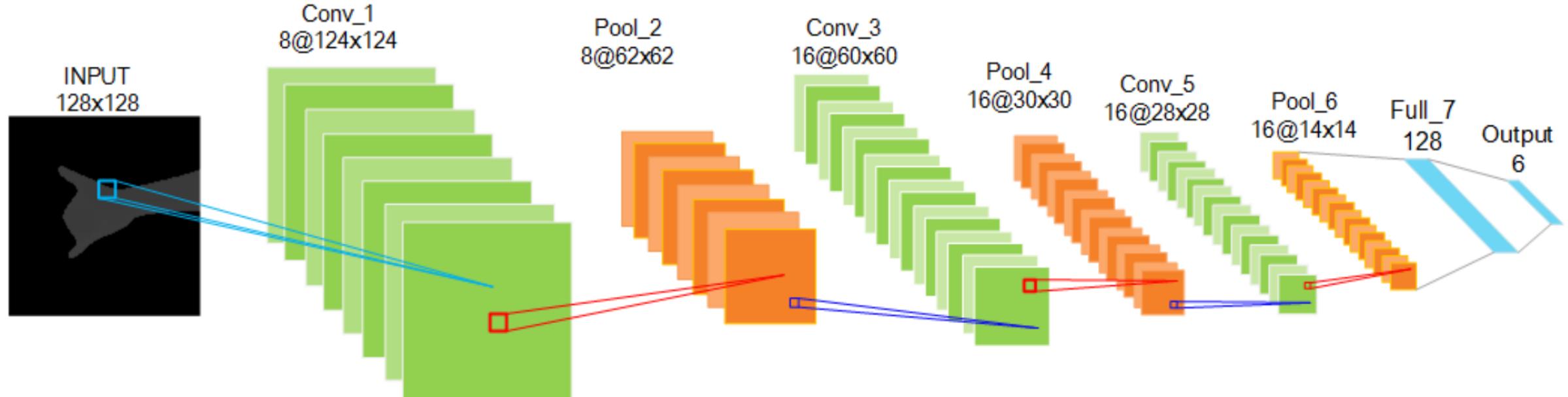
沿用2D的运算架构

2D Deep Learning:

AlexNet, VGG, GoogleNet,
STN, ResNet, DenseNet, ...



基于神经网络的3D识别—深度图作为2D灰度图使用

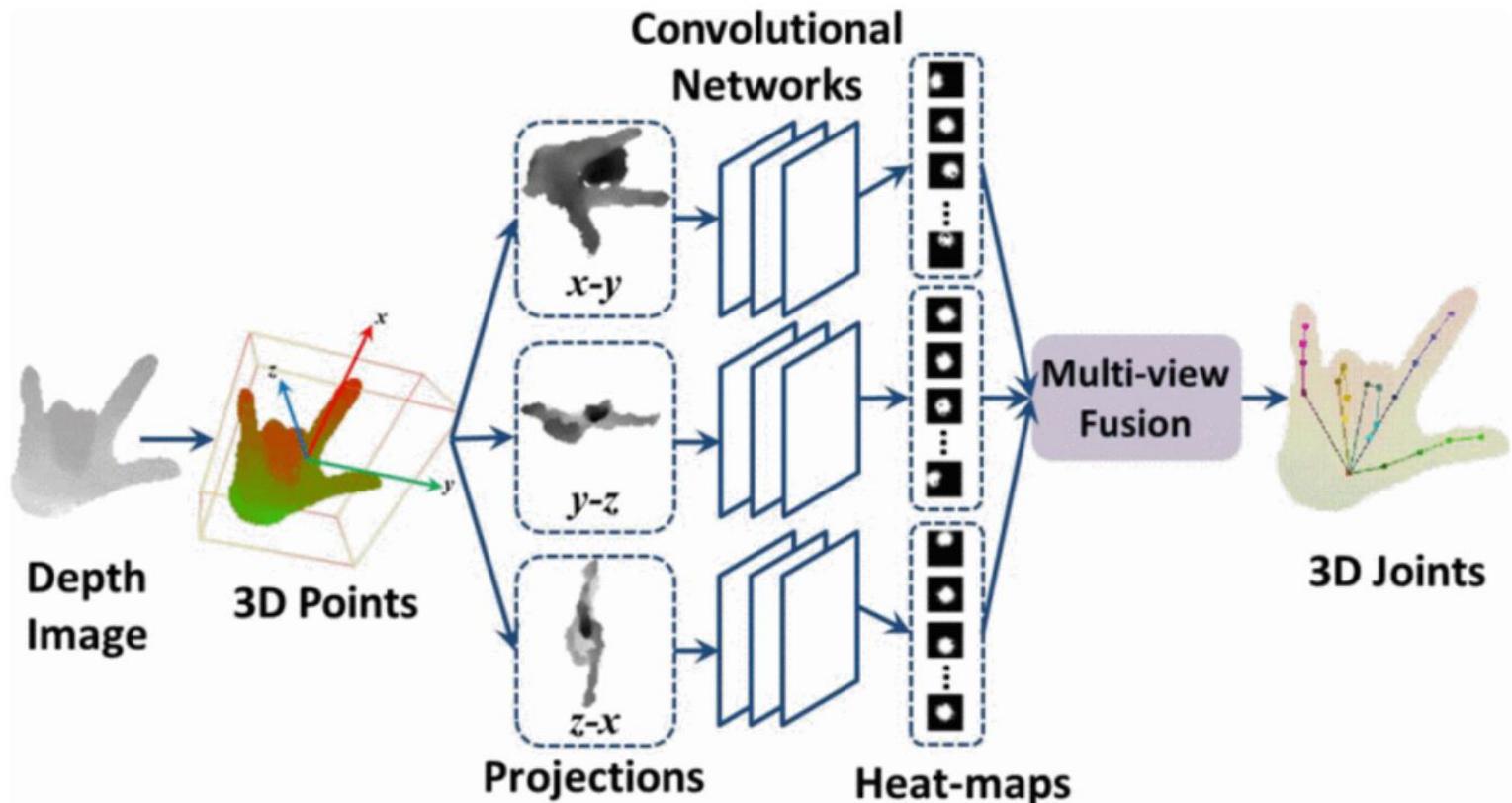


- 把深度图当成2D灰度图处理
- 通过3D对图像进行切割，并对尺寸归一化
- 降低了网络输入的数据的背景“干扰”，并降低了输入数据的“多变性”



手势的卷积神经网络识别示例

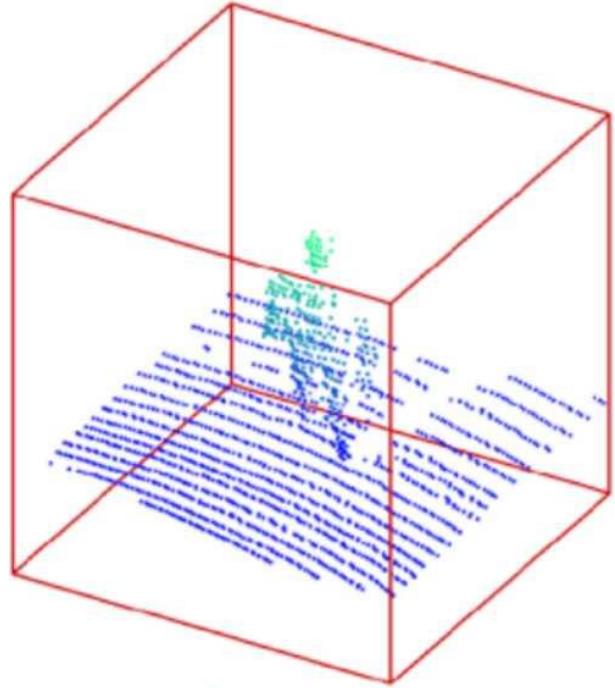
基于神经网络的3D识别-多视角深度图神经网络



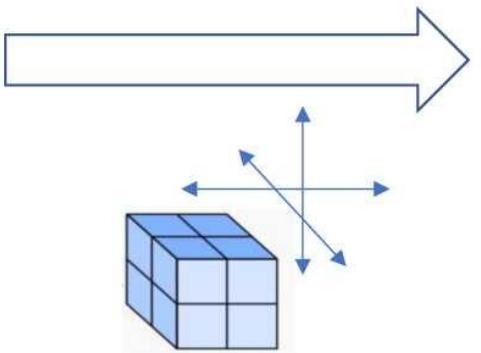
Liuhan Ge, Hui Liang, Daniel Thalmann, "Robust 3D Hand Pose Estimation in Single Depth Images: From Single-View CNN to Multi-View CNNs"

- 直接使用深度图的一个简单扩展
- 对3D点云做不同角度的旋转变换后，重新“投影”到2D平面得到“多视角”深度图

基于神经网络的3D识别—体像素3D卷积



3D的体像素表示



真正的3D卷积核

Daniel Maturana and Sebastian Scherer, "VoxNet: A 3D Convolutional Neural Network for Real- Time Object Recognition"

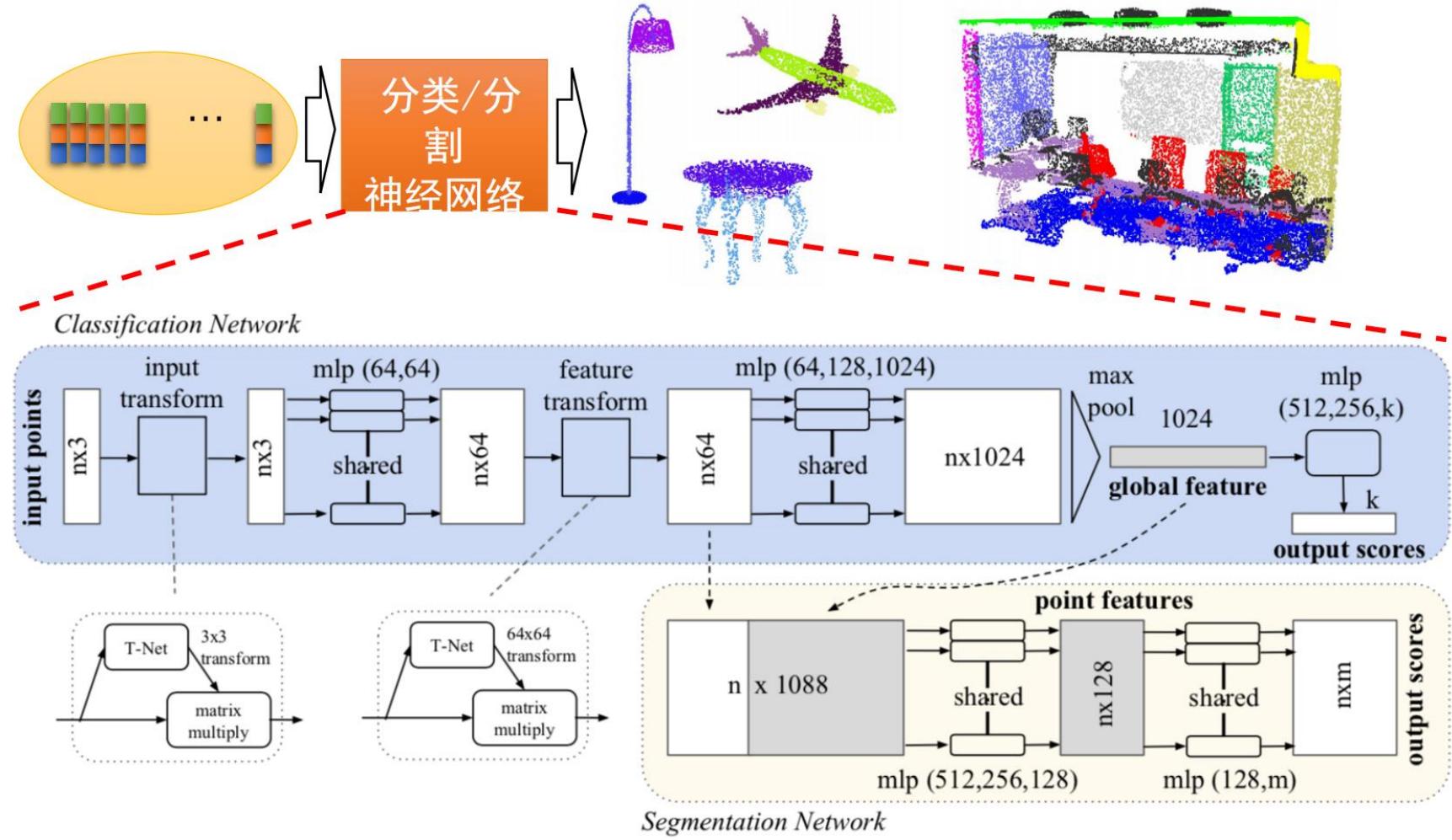
3D卷积核在3个维度滑动，存储和运算量大。（有用八叉树改进的研究，O-CNN）

基于神经网络的3D识别-点云神经网络

架构特点

- 识别结果不受点云排列次序影响
- 识别结果不受点云旋转平移影响

配合训练数据的预处理才能实现网络对排序/旋转平移不敏感的特点



LiDAR 公开数据集与处理软件

Paris-Lille-3D Dataset

<http://npm3d.fr/paris-lille-3d>



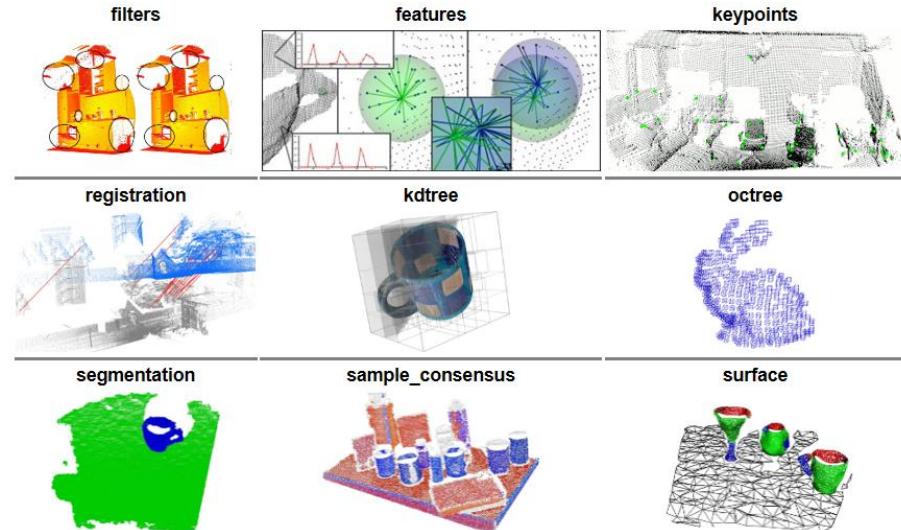
RGB-D Object Dataset

<https://rgbd-dataset.cs.washington.edu/>



PCL官方测试数据集

<https://github.com/PointCloudLibrary/data>



The Stanford 3D Scanning Repository

<http://graphics.stanford.edu/data/3Dscanrep/>





LiDAR 公开数据集与处理软件

处理软件

CloudCompare (免费开源)

<https://github.com/CloudCompare/CloudCompare>

TerraSolid

需要商业License (可以寻找破解版下载使用)

ENVI 的 LiDAR 模块

需要商业License, 可以自动提取 DEM / DSM / 建筑物 / 植被 等的三维模型 (可以寻找破解版下载使用)

FugroViewer

免费的点云快速浏览软件

MCC LiDAR

开源滤波软件

开源库

Point Cloud Library (PCL)

<http://www.pointclouds.org/>

CGAL

<https://www.cgal.org/>

MeshLab

<http://www.meshlab.net/>



项目：3D数据采集与建模

- 随机选取一个物品，通过多角度拍照（倾斜摄影）的方式进行数据采集，并通过VisualSFM或者其他软件进行倾斜建模，通过meshlab进行数据处理，获取三角网模型。



汇报完毕，敬请指正！