



城市空间建模与仿真

第十二讲 城市空间三维数据特征学习与识别-随机森林与人体骨架识别

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



目录

CONTENTS

- 01 随机森林原理**
- 02 深度图随机森林**
- 03 点云随机森林**



目录

CONTENTS

01

随机森林原理

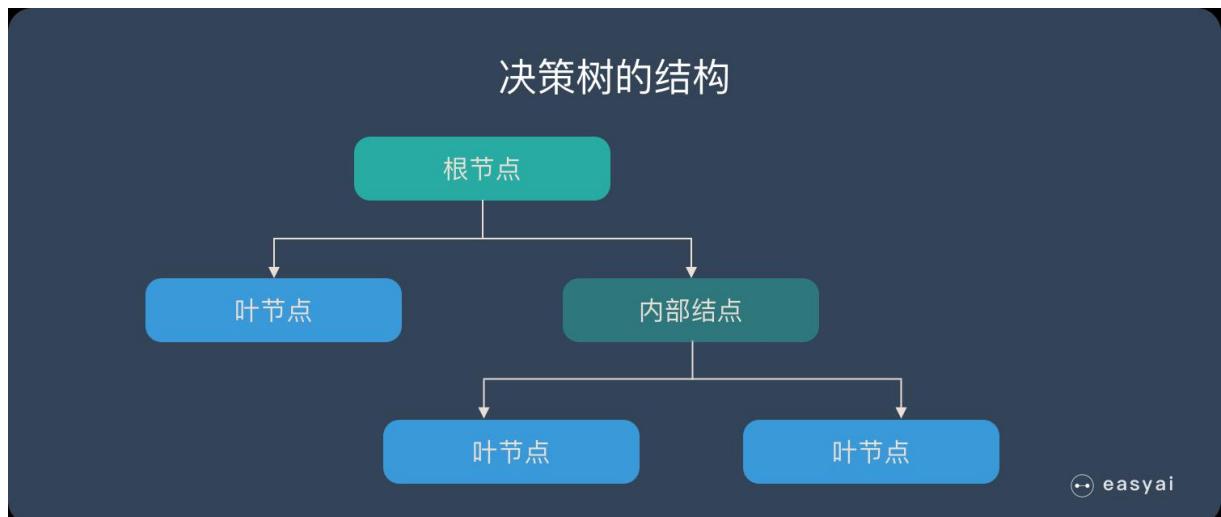
什么是随机森林？

- 在解释随机森林前，需要先提一下决策树。决策树是一种很简单的算法，他的解释性强，也符合人类的直观思维。这是一种基于if-then-else规则的有监督学习算法，下面的图片可以直观的表达决策树的逻辑。



■ 什么是随机森林? - 决策树

- 决策树算法采用树形结构，使用层层推理来实现最终的分类。决策树由下面几种元素构成：
 - 根节点：包含样本的全集
 - 内部节点：对应特征属性测试
 - 叶节点：代表决策的结果





■ 什么是随机森林？-决策树

- 特征选择
 - **特征选择决定了使用哪些特征来做判断。**在训练数据集中，每个样本的属性可能有很多个，不同属性的作用有大有小。因而特征选择的作用就是筛选出跟分类结果相关性较高的特征，也就是**分类能力较强的特征**。
 - 在特征选择中通常使用的准则是：**信息增益**。
- 决策树生成
 - 选择好特征后，就**从根节点触发，对节点计算所有特征的信息增益，选择信息增益最大的特征作为节点特征**，根据该特征的不同取值建立子节点；对每个子节点使用相同的方式生成新的子节点，**直到信息增益很小或者没有特征可以选择为止**。
- 决策树剪枝
 - **剪枝的主要目的是对抗「过拟合」**，通过主动去掉部分分支来降低过拟合的风险。

什么是随机森林？

- 随机森林是由很多决策树构成的，不同决策树之间没有关联。
- 当我们进行分类任务时，新的输入样本进入，就让森林中的每一棵决策树分别进行判断和分类，每个决策树会得到一个自己的分类结果，**决策树的分类结果中哪一个分类最多，那么随机森林就会把这个结果当做最终的结果。**



什么是随机森林？

- 假如有N个样本，则有放回的**随机选择n个样本**(每次随机选择一个样本，然后返回继续选择)。这选择好了的n个样本用来训练一个决策树，作为决策树根节点处的样本。
- **当每个样本有M个属性时，在决策树的每个节点需要分裂时，随机从这M个属性中选取出m个属性**，满足条件 $m << M$ 。然后从这m个属性中采用某种策略(比如说信息增益)来选择1个属性作为该节点的分裂属性。
- 决策树形成过程中**每个节点都要按照步骤2来分裂** (很容易理解，如果下一次该节点选出来的那个属性是刚刚其父节点分裂时用过的属性，则该节点已经达到了叶子节点，无须继续分裂了)。一直到不能够再分裂为止。注意整个决策树形成过程中没有进行剪枝。
- 按照**步骤1~3建立大量的决策树**，这样就构成了随机森林了。





什么是随机森林？

- 优点
 - 它可以处理高维度（特征很多）的数据，并且不用降维，无需做特征选择
 - 它可以判断特征的重要程度
 - 可以判断出不同特征之间的相互影响
 - 不容易过拟合
 - 训练速度比较快，容易做成并行方法
 - 实现起来比较简单
 - 对于不平衡的数据集来说，它可以平衡误差。
 - 如果有很大一部分的特征遗失，仍可以维持准确度。
- 缺点
 - 随机森林已经被证明在某些噪音较大的分类或回归问题上会过拟合。
 - 对于有不同取值的属性的数据，取值划分较多的属性会对随机森林产生更大的影响，所以随机森林在这种数据上产出的属性权值是不可信的

什么是随机森林？

- 随机森林可以在很多地方使用：

- 对离散值的分类
- 对连续值的回归
- 无监督学习聚类
- 异常点检测

随机森林的应用方向



分类



回归



聚类



异常检测



目录

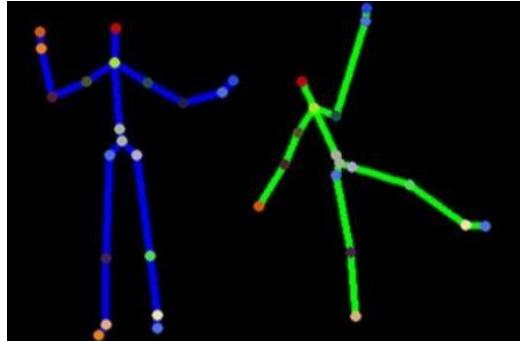
CONTENTS

02 基于随机森林的人体骨架识别

基于随机森林的人体骨架识别

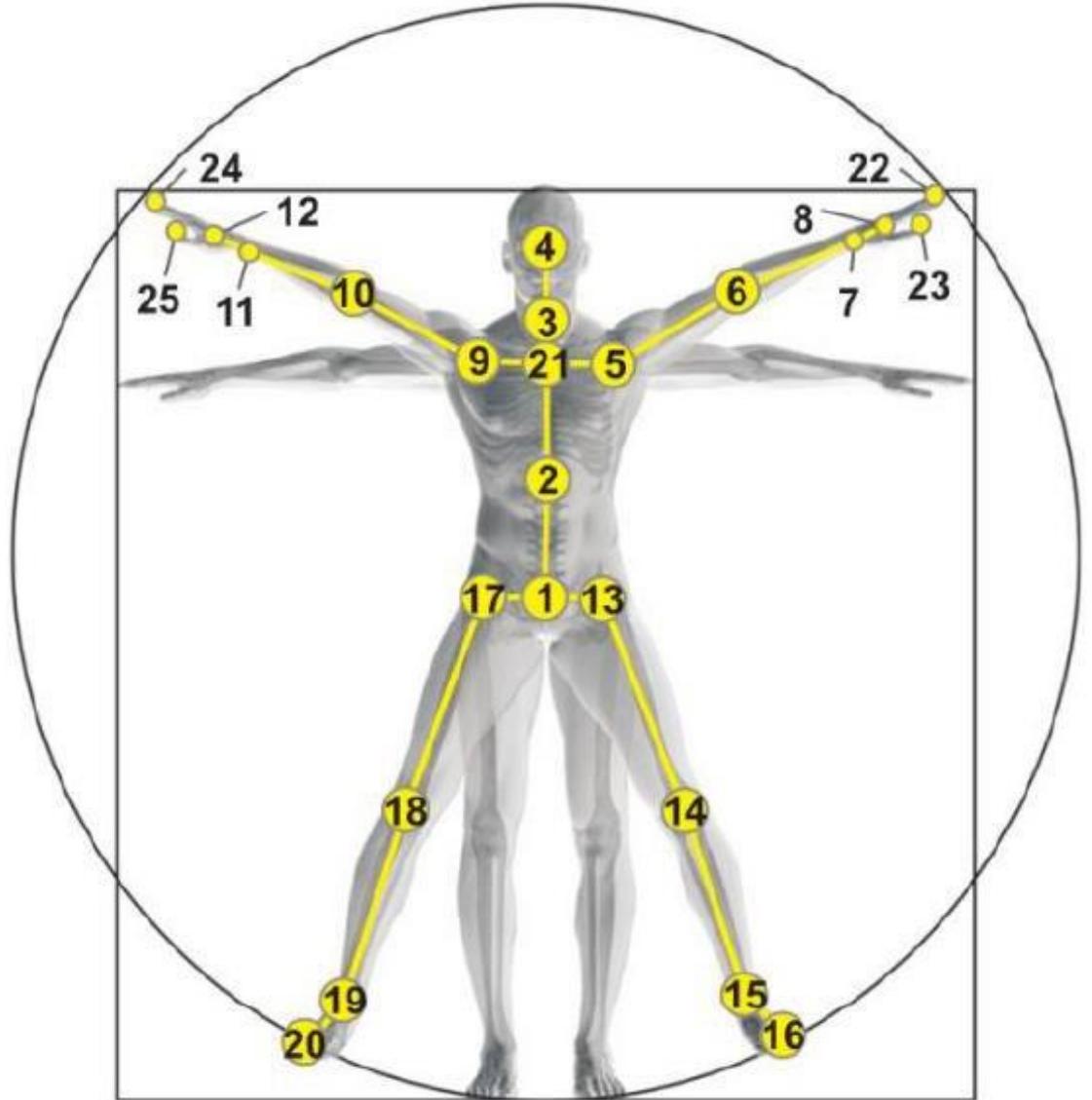
- 计算机视觉邻域有大量和人体动作识别有关的应用

- 行为识别
- 运动员训练
- 人机交互
- 体感游戏/实时VR合成
- 护理/医学诊疗



基于随机森林的人体骨架识别

- 两种动作识别的途径
 - 直接法——由点云获得动作
 - 间接法——先从点云得到骨架再通过骨架得到动作
- 后面的内容围绕骨架提取方法进行介绍
- 骨架的模型
 - 不是严格医学意义上的骨架
 - 25关节的简易“骨架模型”



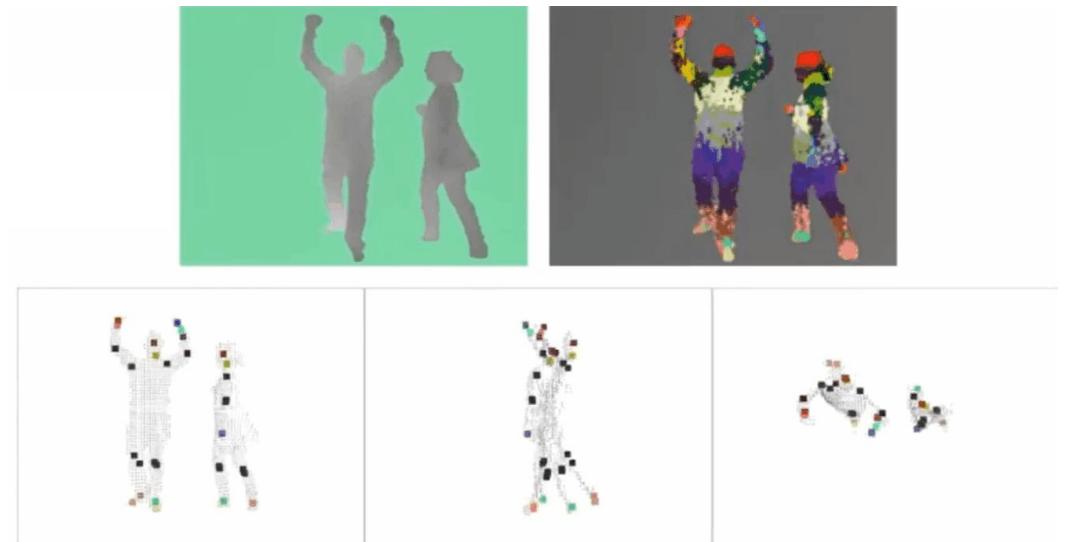
■ 人体骨架提取-基于RGB和神经网络的方法

- CMU的OpenPose是比较成功的基于RGB的骨架提取算法，能够同时识别多人动作，它使用深度神经网络实现
- 运算量很大，需要GPU实现，运算效率低、但RGB数据获取方便



■ 人体骨架提取-基于深度图

- 相比RGB人体骨架识别方法，从深度图能够更加高效的获取人体骨架数据，可以在CPU甚至嵌入式系统上实现运算
- 为什么能够高效？
 - 简单快速前背景分离
 - 得到真实物理尺寸，不受距离影响
 - 利用人体表曲面信息
- 右图识别人体时，去除了背景，使得问题简化
- 通过3D空间点云分割，分离两个人的点云，分别进行识别

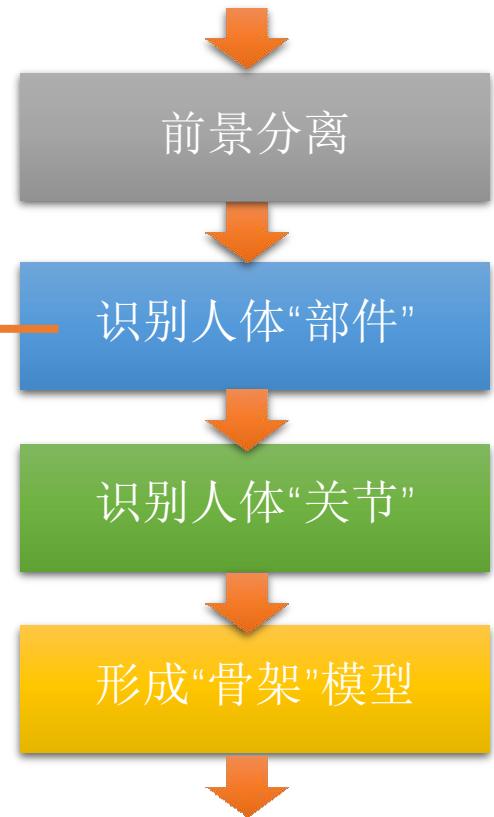


■ 人体骨架提取-基于深度图

Real-Time Human Pose Recognition in Parts from Single Depth Images

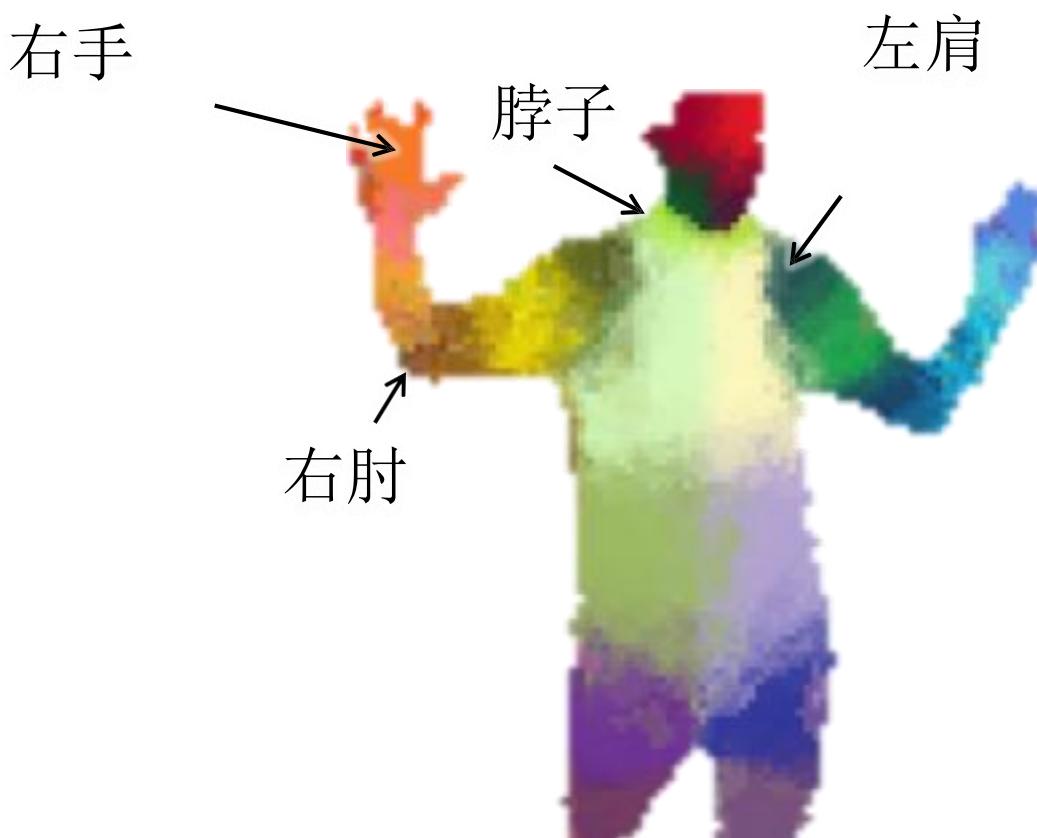
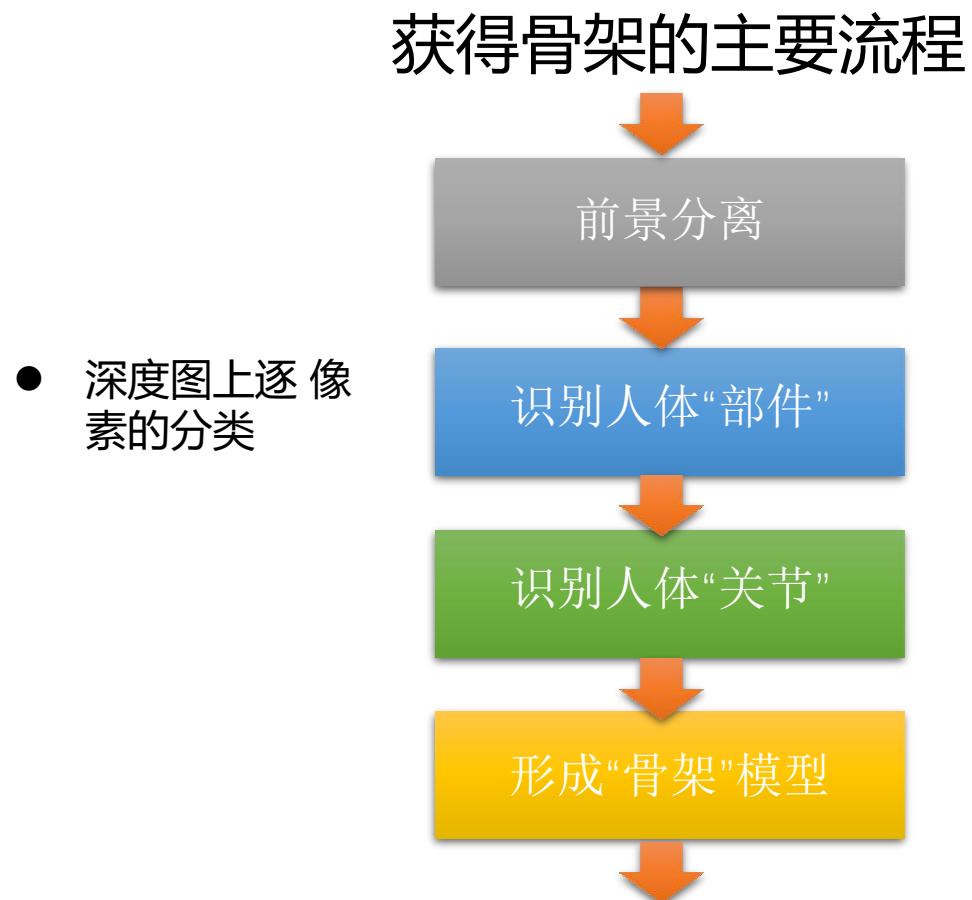
- 获得骨架的主要流程
- 右边看到每个环节对应的数
据可视化效果

- 这一部分看
上去最困难
- 如何实现的?



■ 人体骨架提取-基于深度图

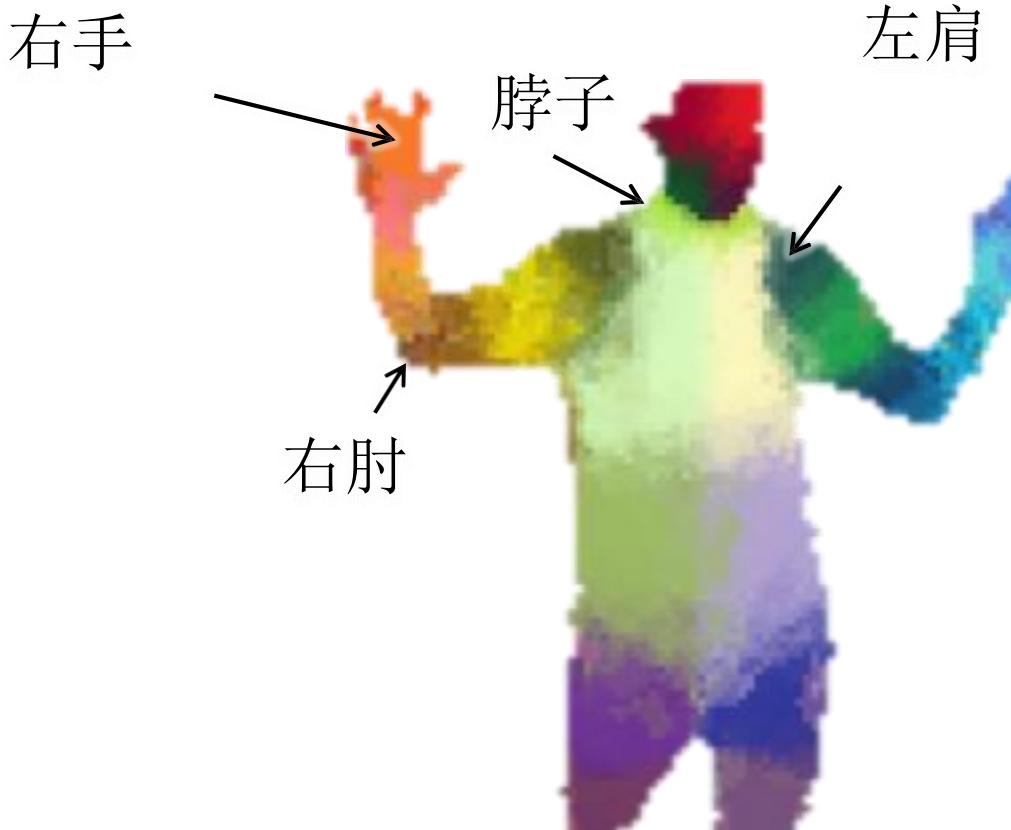
Real-Time Human Pose Recognition in Parts from Single Depth Images



■ 人体骨架提取-基于深度图

- 深度图上逐像素的分类，基本思路：

- 对每个前景像素，估算其属于人体各个“部件”的概率（见右图）
- 通过训练分类器的方式实现上述功能
- 对于每个像素，哪些信息能够用于对他分类、求出它属于不同身体部件的概率？



■ 人体骨架提取-基于深度图

Real-Time Human Pose Recognition in Parts from Single Depth Images

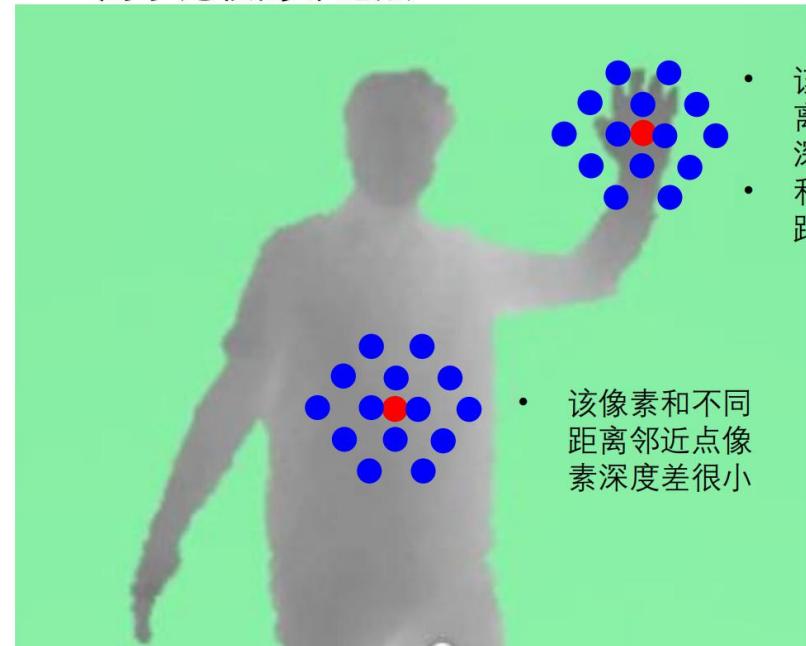
对于每个像素，哪些信息能够用于对他分类？

- 利用当前像素和周围特定位置像素的深度差分类用线索
- 计算“点对”的深度差特征： $f(I, \mathbf{x}; \mathbf{v})$

$$f(\hat{I}, \mathbf{x}; \mathbf{v}) = d_I(\mathbf{x}) - d_I(\mathbf{x} + \Delta)$$

待分类像素深度 特定邻近点像素深度
 $\Delta = \frac{\mathbf{v}}{d_I(\mathbf{x})}$
 深度图 像素坐标
 深度图中像素 \mathbf{x} 和它附近 Δ 位 置像素深度差
 邻近像素相对方向

以下面红色像素点分类为例，蓝色点是用于比较的邻近点



位移距离反比于深度，用于补偿近大远小现象

注意：1) \mathbf{v} 是2元素向量，分别代表移动方向的x/y分量

2) 这里的 \mathbf{x} 和 \mathbf{v} 是指深度图上的位置和移动（不是3D空间）

Jamie Shotton, Real-Time Human Pose Recognition in Parts from Single Depth Images, CVPR 2011

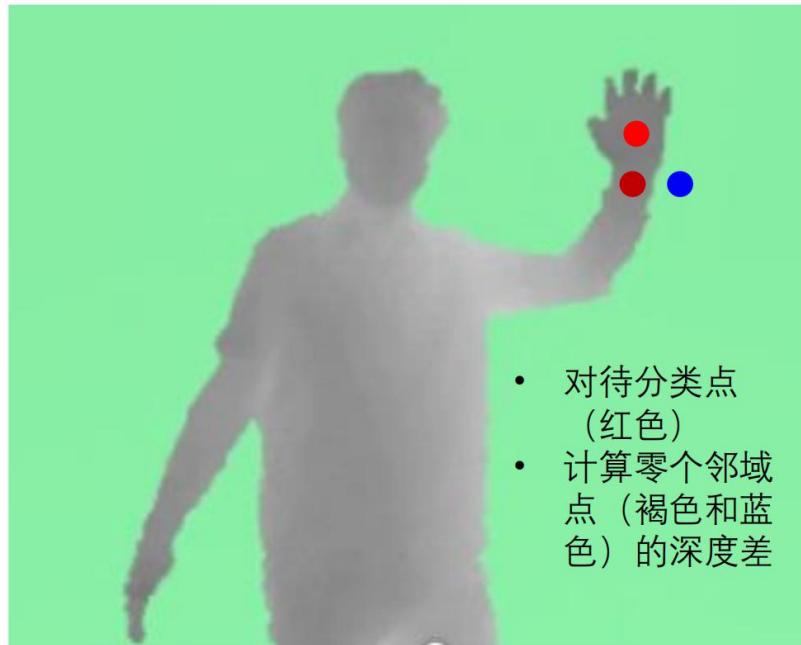
■ 人体骨架提取-基于深度图

Real-Time Human Pose Recognition in Parts from Single Depth Images

- 计算“点对”深度差方法的拓展
- 计算邻域两点之间的深度差（而不是计算待分类像素和邻域的深度差）

$$d_I \left(\underbrace{\mathbf{x} + \frac{\mathbf{u}}{d_I(\mathbf{x})}}_{\text{点x分别沿着u方向移动 (指深度图上移动)}} \right) - d_I \left(\underbrace{\mathbf{x} + \frac{\mathbf{v}}{d_I(\mathbf{x})}}_{\text{点x分别沿着v方向移动 (指深度图上移动)}} \right)$$

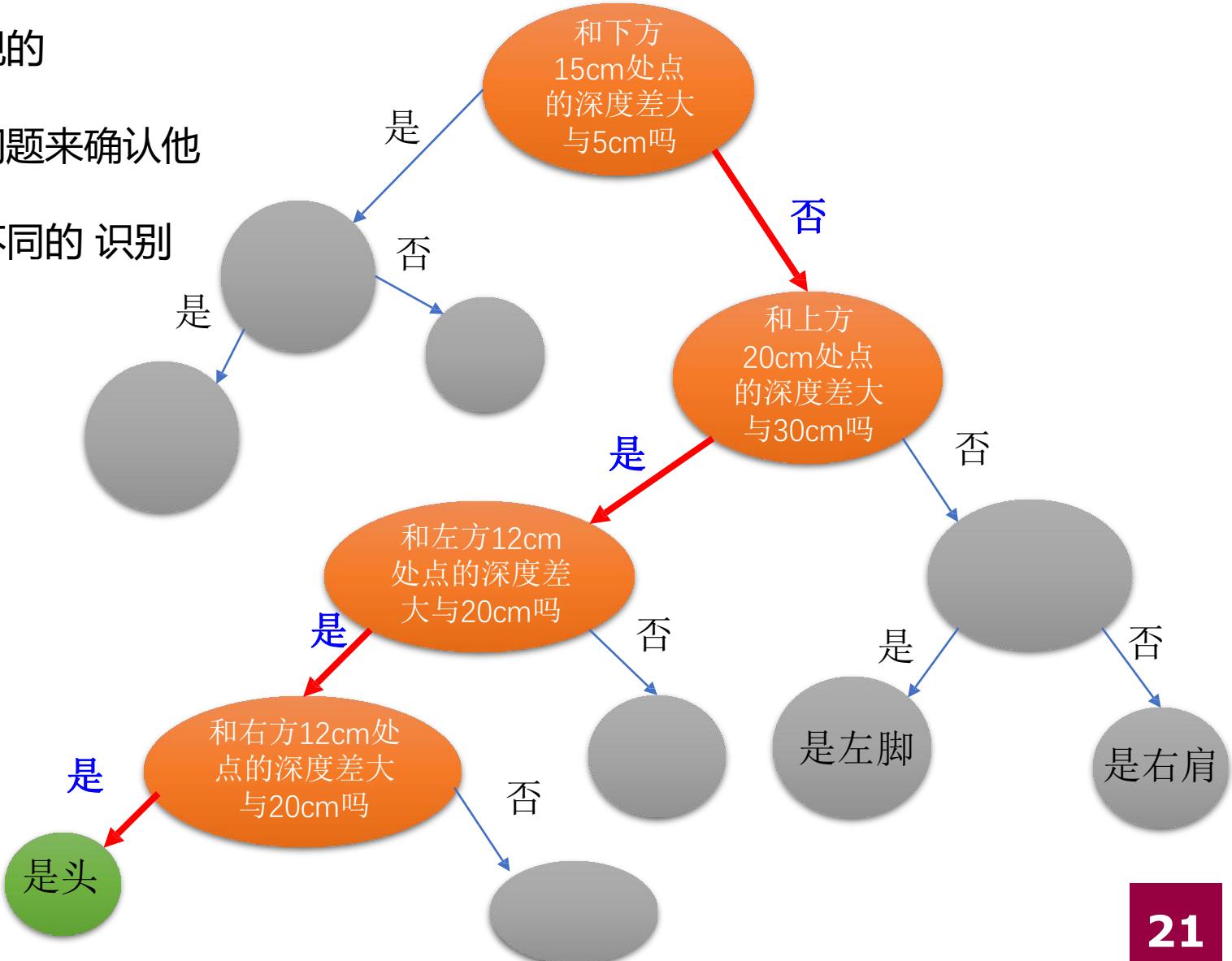
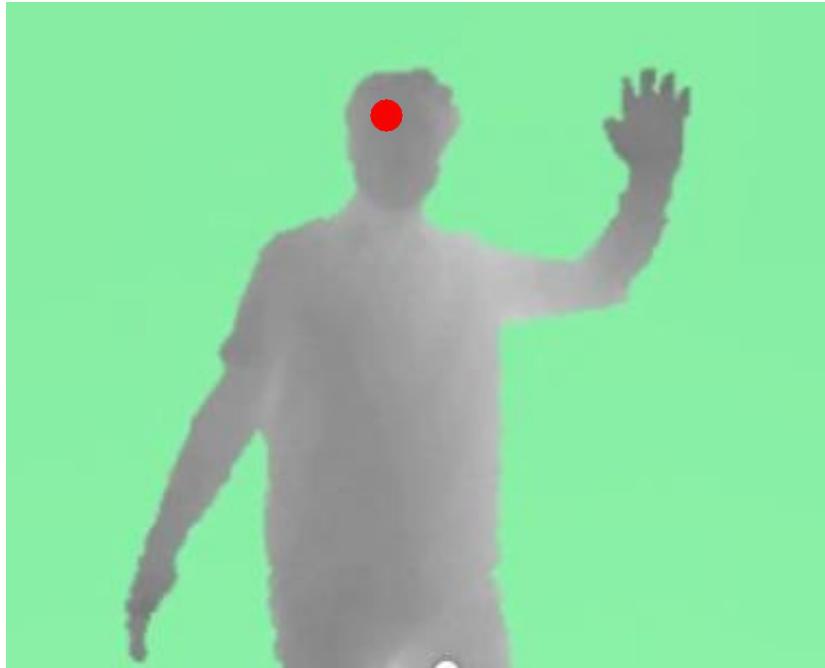
点x分别沿着u和v方向移动 (指深度图上移动)



Jamie Shotton, *Real-Time Human Pose Recognition in Parts from Single Depth Images*, CVPR 2011

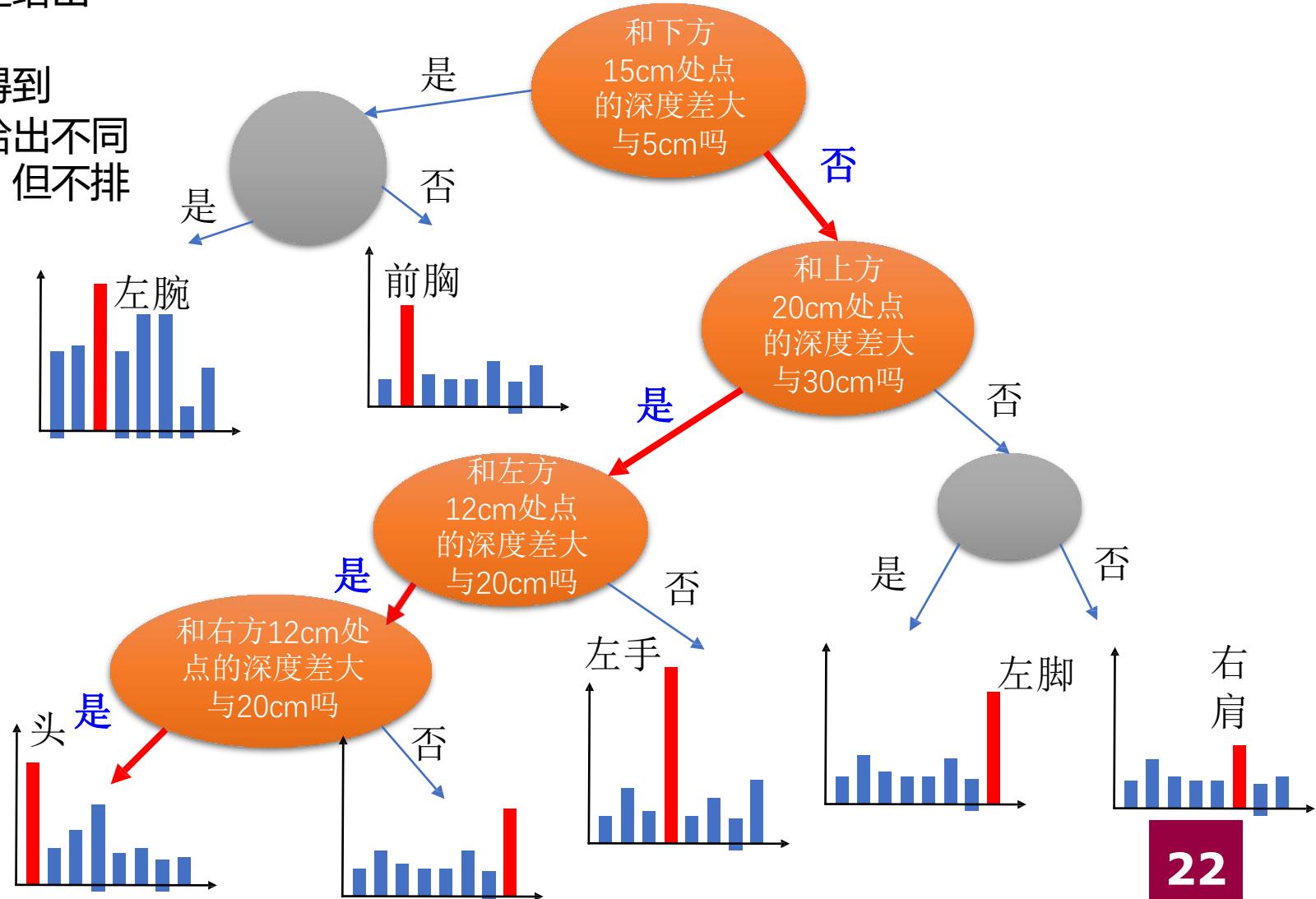
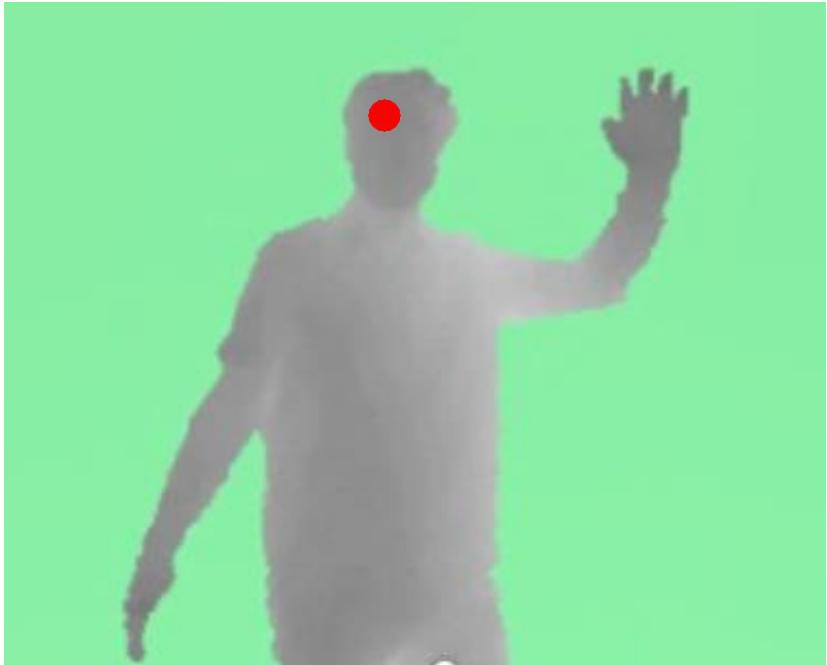
■ 人体骨架提取-决策树

- 分类过程是通过构建一颗“提问树”实现的
- 用过有限的问题知道待分类像素的类别
- 比如下图红色像素，通过右图所示几个问题来确认他是头部
- 不同的回答导向不同的“叶子”，对应不同的识别结果



■ 人体骨架提取-决策树

- 改进：叶节点不对像素进行硬分类，而是给出“类型概率”
- 叶节点的“类型概率”可以从训练数据得到
- 看右边例子，对于一系列问题的回答，给出不同身体部位的可能性，其中头部可能最大，但不排除其他可能

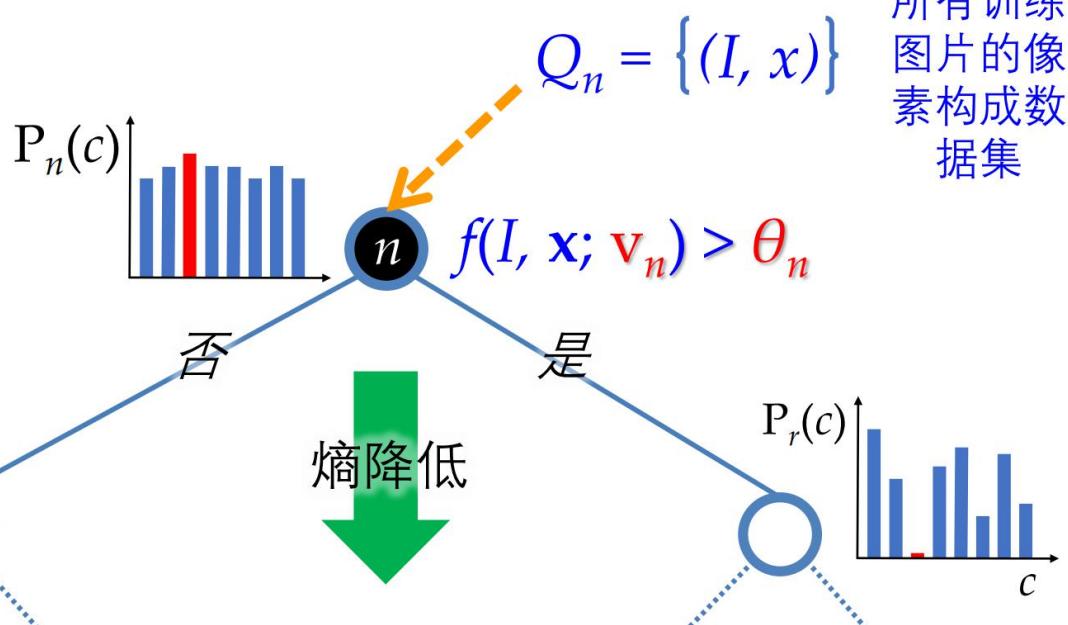


■ 人体骨架提取-决策树

- 前一页的问答树——决策树
- 如何设计决策树?
 - 确定每个节点应该问什么
 - 选择问题的内容使得“信息增益”最大



- 所选的问题 (v, θ) 将上一层节点的(训练)数据集 Q_n 集分为两部分 Q_l 和 Q_r
- 好的问题使得 Q_l 和 Q_r 中不同身体部件的样本差异尽可能大——我们可以用熵衡量
- $E(Q)$ 是数据集 Q 的离散熵的估计算法 $= - \sum_{i=1}^n P_i * \log(P_i)$



选择 (v, θ) 最大化信息增益

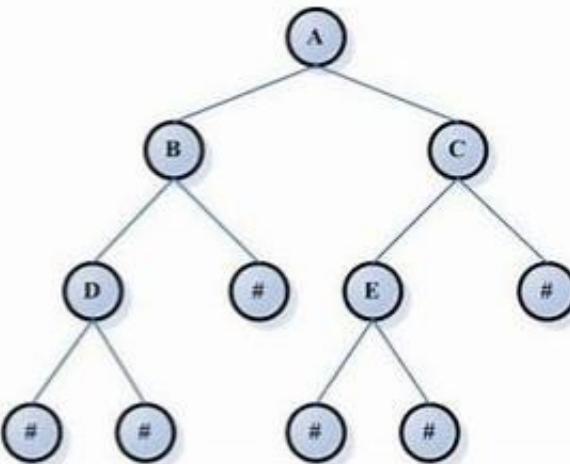
$$\Delta E = -\frac{|Q_l|}{|Q_n|} E(Q_l) - \frac{|Q_r|}{|Q_n|} E(Q_r)$$

i 是训练集 Q 中, 属于身体部件 i 的数据相对 Q 的总数据量比例

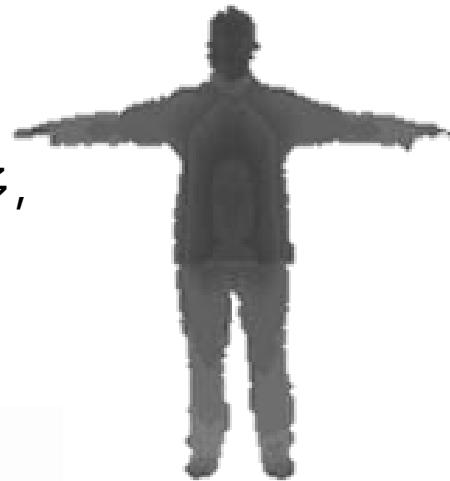
■ 人体骨架提取-决策树

- 下面考虑决策树的设计的另一个问题—深度

- 问的问题层数越多，分类越准确
- 树越深越好？



输入深度图



真实答案



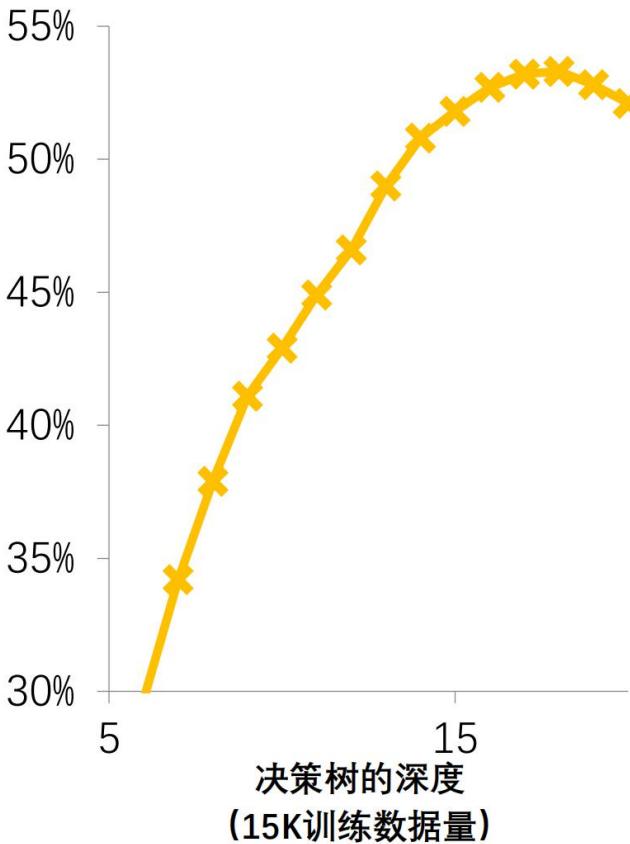
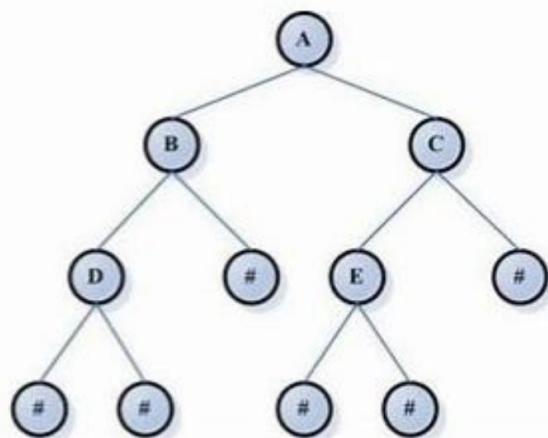
决策树深度 2



随着深度增加，分类（用色彩标注）约精细，越正确

■ 人体骨架提取-决策树

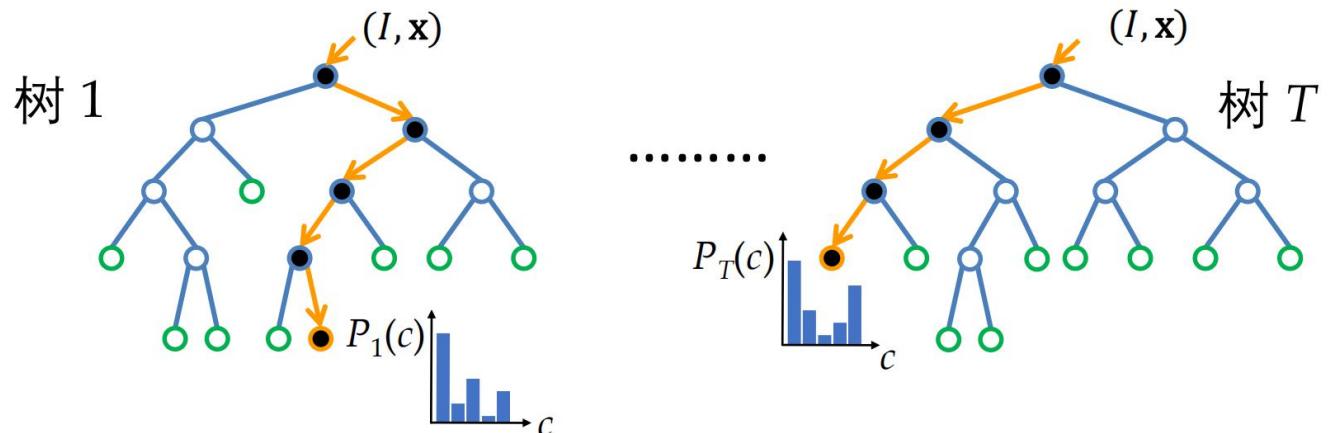
决策树的深度不能无约束地增加



- 训练数据分类正确率随着树的深度上升而提高
- 验证数据当树的深度高于特定值之后，反而下降

■ 人体骨架提取-随机森林

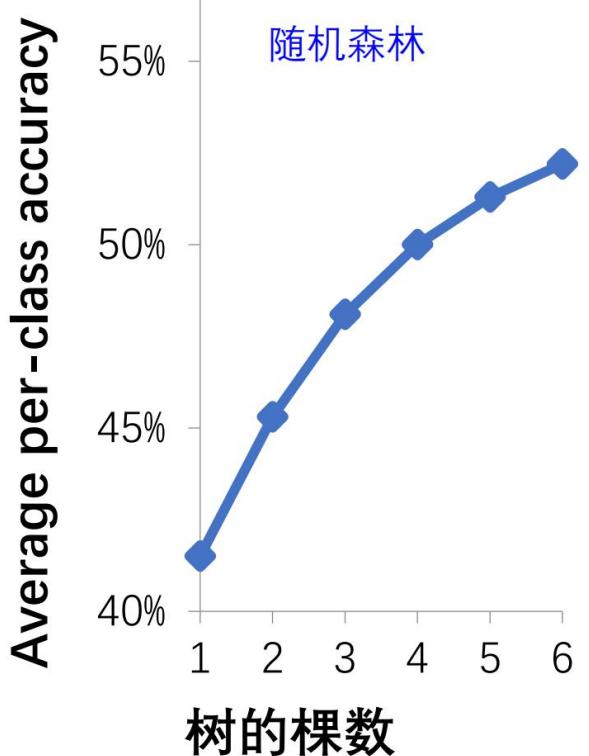
- 如何提升正确率，避免单棵决策树的“过拟合”？
- 建立多棵决策树——**随机森林**
- 每棵决策树需要在构建时确保他们的“独立性”（如何实现？如下）
 - 每棵树训练使用的样本从一个总的训练样本集合中随机抽取
 - 建树过程中，每个节点允许使用的特征集合（这里是问题集）是随机选取的一个特征子集，从这个子集中找出节点分裂的最优问题（使得信息增益最大的问题）



- 最后输出的分类结果(像素归属不同身体部件的概率)是所有决策树分类结果的平均

$$P(c|I, \mathbf{x}) = \frac{1}{T} \sum_{t=1}^T P_t(c|I, \mathbf{x})$$

■ 人体骨架提取-随机森林



真值



不同数量随机森林的分类效果

1 棵树



3 棵树



6 棵树



(不同颜色表示叶节点最高平均概率对应的
身体部件类型)

■ 人体骨架提取-随机森林

- 再次检查一下分类效果
 - 逐个像素的分类正确率能最高到60%左右
 - 看上去不很高
 - 反映在分类结果上呈现“类别噪声”
 - 如何去除“类别噪声”？
 - 这个问题会同关节定位同解决



■ 人体骨架提取-从像素分类到关节定位

分一下几步实现关节定位

每个像素的类型“扩散到空间”

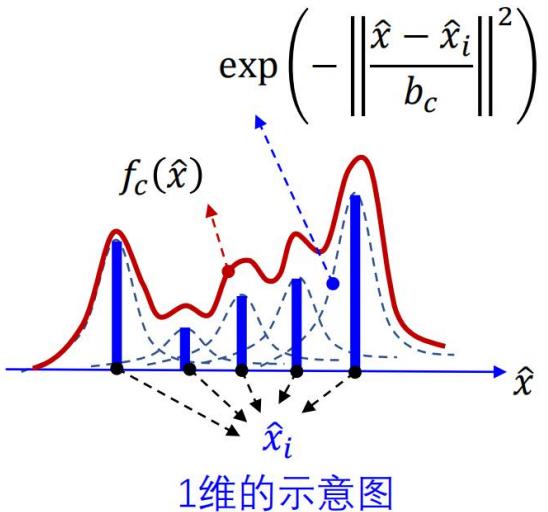
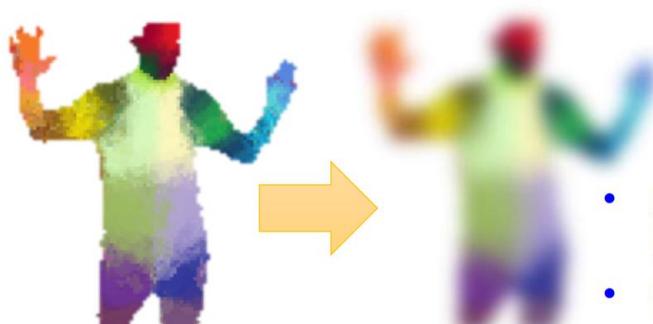
用Meanshift, 找到从空间“漫布”的类型(身体部件)的中心点

根据身体部件的中心点位置, 加上骨架比例约束, 找到关节点

$$f_c(\hat{\mathbf{x}}) \propto \sum_{i=1}^N w_{ic} \exp\left(-\frac{\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_i\|^2}{b_c}\right)$$

对外扩散到的空间位置 深度图像素 x_i 对应空间坐标
 训练获得的常数, 对应扩散程度

- 像素 x_i 对类别“扩散”的贡献
- 和像素 x_i 归属类别 c 的概率以及他的尺寸 (面积) 有关的权重



$$w_{ic} = P(c|I, \mathbf{x}_i) \cdot d_I(\mathbf{x}_i)^2$$

像素 x_i 归属类别 c 的概率 深度平方和像素对应空间曲面面积成正比

- 对于每个类别 c , 空间任何一点 \hat{x} 对应的类型“权重”为 $f_c(\hat{x})$
- 每个身体部件, 对应空间“一团彩色烟雾”
- 身体部件通过计算“烟雾”浓度最高的区域的到 (下一页meanshift算法)

■ 人体骨架提取-从像素分类到关节定位

分一下几步实现
关节定位



每个像素的类型“扩散
到空间”

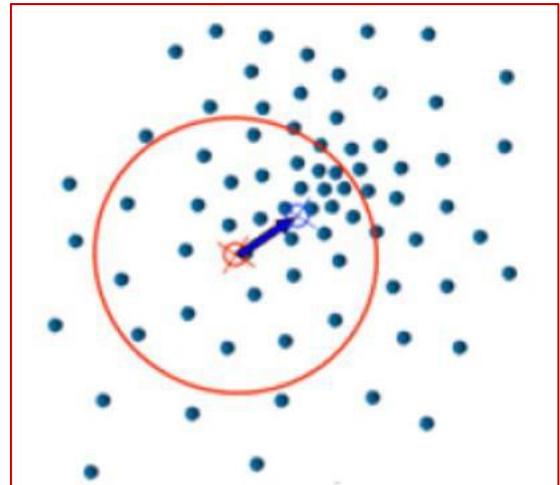


用Meanshift，找到从
空间“漫布”的类型(身
体部件)的中心点

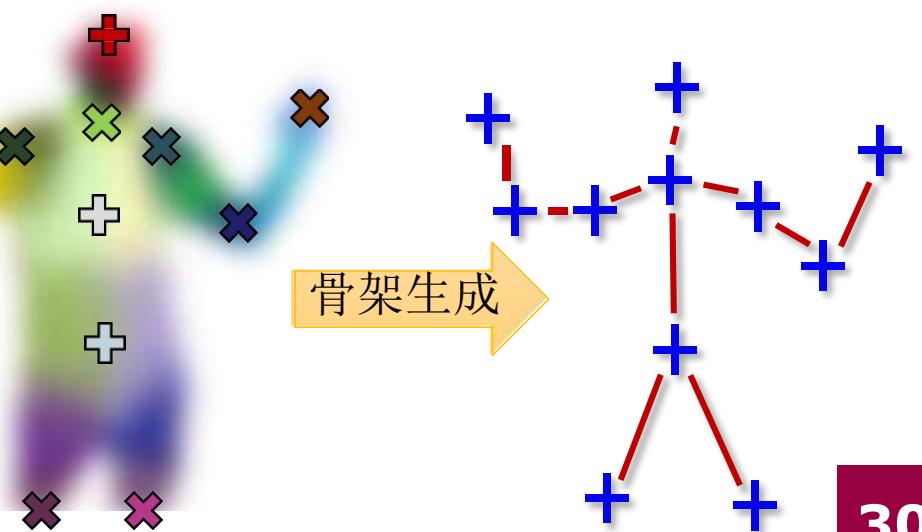


根据身体部件的中心
点位置，加上骨架比
例约束，找到关节点

Meansift算法示意图



- 设图中的点对应类型c的身体部件
- 我们希望估算出身体部件c的中心
- 1. 初选一个中心，以它为球心，画球
- 2. 计算球内点的重心
- 3. 移动球心到上一步计算得到的重心
- 4. 重复步骤2/3直到球心不再移动
- 我们使用的算法和上述meanshift还
有少量差别——重心是从“漫布”在球
内的类型c的权重 $f_c(\hat{x})$ 计算得到的



■ 人体骨架提取





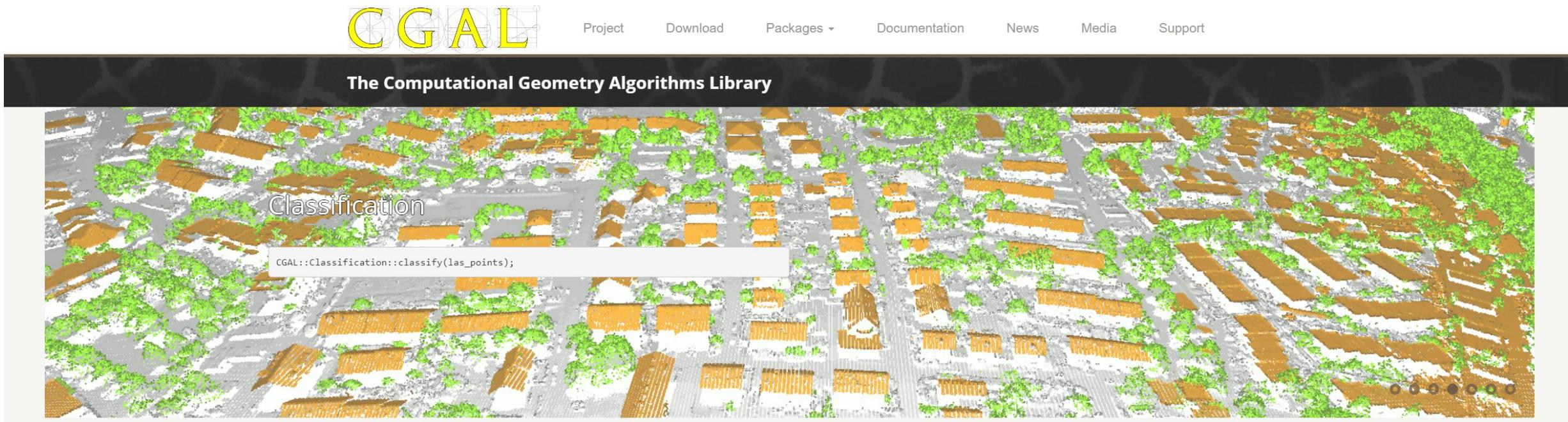
目录

CONTENTS

03

点云随机森林

点云随机森林框架



点云随机森林框架

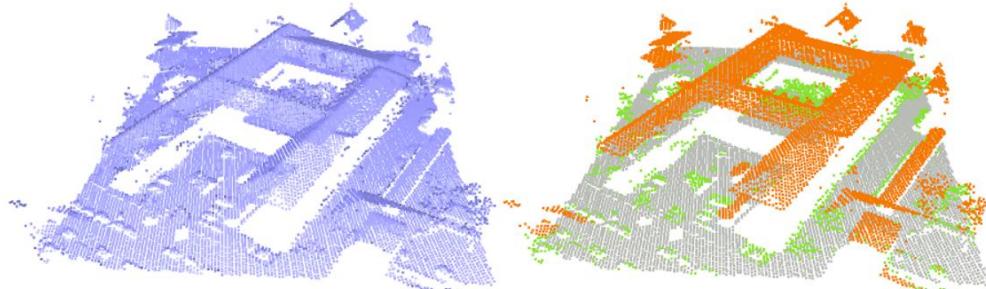
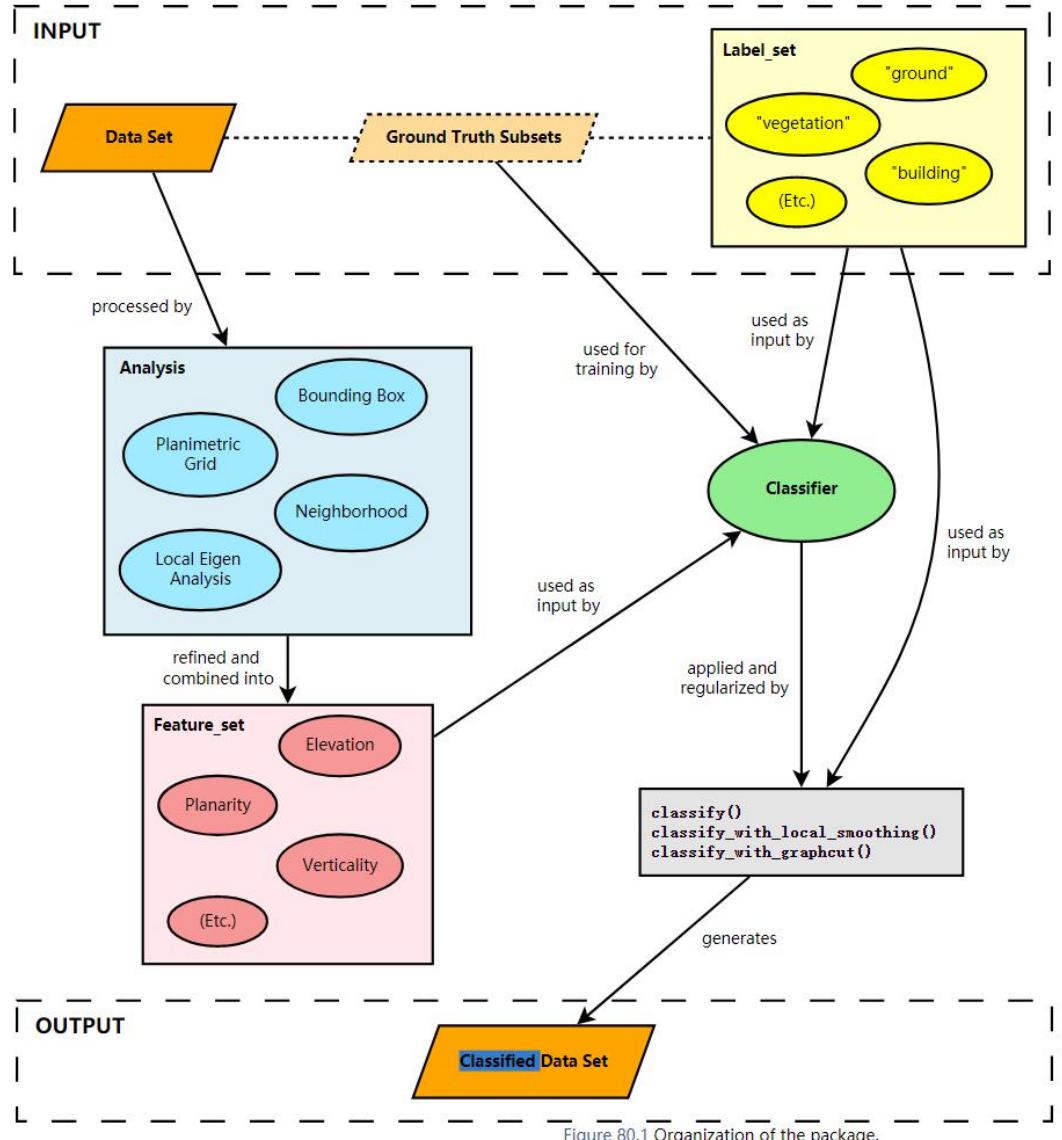


Figure 80.2 Example of point set classification (left: input, right: output). Ground is grey, roofs are orange, vegetation is green.



Figure 80.3 Example of mesh classification (left: input, right: output). Ground is grey, roofs are orange, vegetation is green.

Distance_to_plane measures how far away a point is from a locally estimated plane;
Eigenvalue measures one of the three local eigenvalues;
Elevation computes the local distance to an estimation of the ground;
Height_above computes the distance between the local highest point and the point;
Height_below computes the distance between the point and the local lowest point;
Vertical_dispersion computes how noisy the point set is on a local Z-cylinder;
Vertical_range computes the distance between the local highest and lowest points;
Verticality compares the local normal vector to the vertical vector.

https://doc.cgal.org/latest/Classification/index.html#Chapter_Classification



谢谢