

指导教师： 杨涛

提交时间： 2016/3/17

# CVPR2015 Paper Translation

No: 01

姓名： 杨佳佳

学号： 2013302607

班号： 10011306



## 图像分割中的 20 个问题

**概要：**考虑下面的场景在一个人类用户和计算机之间。给定一个图像，用户认为在这张图片中分割的对象，但是只被允许提供二进制输入给计算机，在这些情况下，计算机可以通过询问精心挑选的问题来给用户猜出这个隐藏分割。我们为电脑采用了一个战略去增加其猜测最少数量问题的准确性。在每一轮，目前有关应答的信念是编码在所有可能的分割中经由一个概率分布的贝叶斯方式。为了有效的解决这个巨大的空间，分布通过抽样使用 Metropolis-Hastings 算法的改编版代表分割近似。其建议移动建立在一个测地距离变换分割方法。经过二分搜索，减半加权组样品的问题是最后选中，和所提供的答案是用于更新信念，为即将到来的回合。这种策略的性能评价与多样化的视觉特性 3 可公开获得的数据集。我们的方法展示出了是一个易于处理和非常自然地解决了这个问题。

**1, 介绍：**20 个问题是一个经典的双人游戏包括一个发问者和预测。在这游戏开始之前，这个预测要考虑一些事情（例如一个对象，动物或字符）那种我们可以叫答案，和调查问卷被允许提出一系列二进制问题猜出答案是什么。虽然本场游戏原本涉及人类在预测中的作用和另一个人的球员，人工智能技术的发展导致软件所在的计算机试图猜测人类用户心中的想法。打算在同一个方向，我们在这项工作中考虑计算机和人类用户玩这二十个问题游戏的情况下，那里的预期的答案是一个二元分割给定图像的我们引入一个战略，为计算机以提供一个猜测尽可能准确在有限数

量的的问题。另外，该方法能也可以看作是一个交互式分割的任务。关键我们的方案的特征是，用户交互用机被限制为二进制输入（是/否），而不是通常的涂鸦或边界框的[10]。这个设置为通用提供交互式反馈有关图像检索[19, 24, 26]，但从未考虑在分割的情况下。此外，如果结合例如语音识别系统，我们的方法可以最终提供免提分割技术，对于其中替代的解决方案，提出了通过 Sadeghi 等。[21] 谁使用的眼动仪种子位置。这有潜力应用中，比如消毒医疗数据的分割手术室环境中，物理相互作用与对象有被避免。我们的方法可以概括如下。在每个回合，要提出的问题在于，询问是否一个很好的选择像素是隐藏的分段内或没有。该在地点的选择询问被制成最大化由答案所带来的信息。更精确地说，我们限定为任何可能的分割的概率有被甲骨文采摘。每个分段的分数是根据在两个方面：在图像分割方面其内在质量（即，分割的对象均匀性），及其与答案兼容性已收集。因此，该定义在该组的概率分布的可能的分割  $S$ 。在理论上，理想的问题选择将遵循一个分而治之的办法，并在大约相等的概率两部分减半第然而， $S$  的体积庞大排除了详尽的计算上述的分布，这是我们近似的而不是通过采样一系列分割的按比例来他们已选定的概率由 Oracle。这种采样是通过马尔科夫进行链蒙特卡罗（MCMC）框架。他建立在这个算法和包括探索同时花更多的时间在高概率的区域。

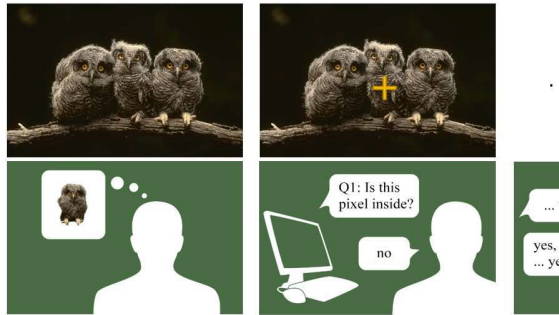


图 1: 二十个问题分割场景概述。给定一个图像时, 对象段是秘密由人的用户选择。在每一个步骤, 所述计算机询问目标像素是否位于期望的分割内部。在预测的很多问题后, 它返回其有关应答的推测。

最翔实的问题可以计算从易处理的方式抽取的样本, 和知识由应答带来被包括在概率模型对未来的回合。我们评估我们的方法上 3 可公开获得的数据集的分割与不同视觉属性和几个基准进行比较。该实验结果表明, 我们的做法是一种很有前途的溶液到交互式对象分割问题

只用二进制输入。现在的问题选择策略足够快 (约 1 秒), 以在一个软件被结合在实践中, 不依赖于任何离线训练一步, 这也使得整体架构非常适应到不同类型的图像。相关工作在本段中, 我们简要回顾我们与现有方法的工作连接。二十题设置已经在计算机中提到通过布兰森等工作愿景社区人。[5] 和由华等其扩展名。[28] 用于交互式细粒度图像分类。这些作品考虑情况下既人类用户和机器忽略图像的标签, 但 (i) 该机器知道哪些问题是重要的问找出答案和 (ii) 人是能回答是基于这些问题场景的视觉方面。因此, 结合的专业知识计算机与人类的视觉能力的允许协作找到隐藏的图像标签。以外在任务 (图像分类方面 VS 差异分割), 这个设置是根本不同于我们的, 其中待分割对象是完全已知由人类用户, 同时由计算机预测。

交互式分割技术通常依赖于双方

[4, 10, 20] 或包围盒 [11, 16] 是手动在或围绕感兴趣对象放置由人类用户。更贴近我们的工作, 有几个方法 [3, 8, 13, 14, 23] 保持在循环中的人类用户和提示信息最多的地区下一个标签, 在活跃的学习方式。我们的方法的一个重要方面是固有歧义图像解析任务, 因为一个无法预料语义层由 Oracle 选择了分割。在这个方向上, 涂朱 [27] 提出驱动的数据基于能够主动轮廓 MCMC 框架以产生同一图像的几个 parsings。最近的替代品确定在一组候选相关对象的现场 [6, 7, 12, 15] 通过学习似是而非的对象露面或形状。在我们的情况下, 不执行离线训练因此该方法可以是任何种类的图像或地面实况。

## 2, 方法

### 2.1 问题陈述和记号

图像  $I$  被过晶格  $\{1, \dots, H\} \times \{1, \dots, W\}$  其中  $H$  和  $W$  分别表示高度和一宽我们定义的二元分割图像  $I$  作为一个函数  $s: \rightarrow \{0, 1\}$  表示是否图像的像素  $(x, y)$  是内部 ( $S(x, y) = 1$ ) 或外 ( $S(x, y) = 0$ ) 的分段区域。取决于上下文,  $S$  也可以看作是一个矢量  $s \in S = \{0, 1\}^{|I|}$ 。

我们的问题可以如下所示。之前, 给定一个固定的形象我, 甲骨文 (即人类用户) 决定在具有由发现一个分割小号  $\in$  小号电脑。要做到这一点, 计算机要问到 oracle 一系列的形式 (P) 的二进制的问题: “是解散对象内播放的位置  $p$  要细分? ”。选择最佳的问题要问达找到最翔实的位置页。作为回报, 直接回答提供了真正的标签  $L(\text{对}) \in \{0, 1\}$  在此位置。k 问题已经提出之后, 所收集的答案提供了两个可靠的集  $\cdot K^- ?$  和  $k^+$  的背景和前景分别为种子, 与  $K^- + K^+ = k$  的。我们也表示  $? K = \cdot K^- ? \cup k^+$  集收

集可靠的种子。这种知识是通过贝叶斯后验编码概率  $P(S = S | ? K)$  超过设定的分割小号说明它是如何可能的分割  $s$  已被最初给出透露已知的种子用户选取在答案已收集。我们将表示这个概率  $P(S | ? K)$  在本文的其余部分明确指出，这个概率被看作是  $s$  的函数和作为一个概率在分配  $S$ 。如果这后可以计算对于  $S$  中每一个可能的分割，最佳的鸿沟，征服战略将减半每次导一组可能的细分到每个概率 0.5 的两个子集。然而，集  $S$  具有非常大的尺寸，以理论上  $2^{||}$  可能性，排除了详尽磷计算 (送  $| ? K$ ) 在整个集合  $S$  为了克服此，我们提出在每次迭代  $k$  以近似后  $P(S | ? K)$  由一系列样本的  $SK1$ 。 。 。 ，  $SKN \in$  小号根据马尔可夫链蒙特卡罗画 (MCMC) 方案，该方案中详细描述。 2.2。经过这些  $N$  个样本被抽中，我们选择了最翔实的由以下的问题基于这些样品的问题选择方法以秒曝光。 2.3。这两个取样和问题选择步骤，然后迭代，直到预定达到允许量的问题。

输入：图像 许多被允许的问题  $K$

```

 $\Sigma_-^0 \leftarrow \emptyset; \Sigma_+^0 \leftarrow \emptyset;$ 
for  $k \leftarrow 0$  to  $K - 1$  do
  Sample  $s_1^k, \dots, s_N^k$  from  $p(.|\Sigma^k)$ ;
  Find most informative location  $p_k$  w.
  Ask question  $Q(p_k)$  and receive true
  if  $l(p_k) = 0$  then
     $\Sigma_-^{k+1} \leftarrow \Sigma_-^k \cup \{p_k\};$ 
     $\Sigma_+^{k+1} \leftarrow \Sigma_+^k;$ 
  else
     $\Sigma_-^{k+1} \leftarrow \Sigma_-^k;$ 
     $\Sigma_+^{k+1} \leftarrow \Sigma_+^k \cup \{p_k\};$ 
  end
end

```

G GDT-分割种子  $K$ ;

输出：猜测甲骨文分割

算法：I 我们的方法概述。在每一个步骤，问题选择策略包括采样分割根据他们的已概由 Oracle 采摘，并询问

相关问题最翔实位置相对于这些样品。与所提供的标签一起形成此位置

这是添加到两种种子  $\Sigma_-^0$  or  $\Sigma_+^0$  取决于其标签。

2.2. 与抽样 MCMC 可能的分割. 在本节中，我们介绍我们的抽样程序从后验概率代表分割分布  $p(. | K)$  条件的现有知识。继经典的都市 - 黑斯廷斯算法 [18] 而在引进涂朱 [27] 原来的想法图像解析成多个区域的情况下，我们定义们的采样过程作为马尔可夫链在空间用转移概率分割的定义如下。鉴于当前的分割  $S$ ，一个新的细分候选人  $S'$  是根据提案建议分布化

$q(s' | \Sigma^k, s)$  是马尔可夫链的新状态。此举被接受的概率有

$$\alpha = \frac{p(s' | \Sigma^k) q(s | \Sigma^k, s')}{p(s | \Sigma^k) q(s' | \Sigma^k, s)}, \quad \text{其中,}$$

$p(? | k)$  表示后验概率分布根据我们想要绘制样本，即概率对于细分都被接走用户鉴于目前已知的种子  $\cdot K$ 。从初始开始

分割  $SK0$  中，生成分割的一连串和滑雪， $1 \leq I \leq N$  被以固定的选择作为样品在这个探索过程率。此外，一个老化步骤开始采集样品，使之前进行在初始状态  $SK0$  的依赖强烈降低。

要完成分段取样的说明方法，人们必须小心定义后验分布并建议分布。我们对这些分布设计

将二段描述。 2.2.2 和 2.2.3 分别。它们都建立在组的参数化的分割 ■ 哪些我们现在要介绍。

### 2.2.1 状态空间参数

为了方便有效提议分布的设计中式。

1，我们需要一种方法来变形分割成另一种。

在他们的数据驱动的图像解析 MCMC 框架，涂朱提出了变形活动轮廓 [27]。

在我们的情况下，MCMC 范式有两个问题之间发生在人/机交互的上下文中。因此，它保持两个问题之间的时间是至关重要的。越小越好。出于这个原因，我们引入一个替代鉴于我们分割空间的基础上测地线距离变换 (GDT)，最近被证明是非常高效在对象提议[15]的上下文。我们定义  $X = \{1, \dots, C\} \times [0, B] \times P(\cdot)$  作为我们的状态空间和考虑功能 GDT:  $X \rightarrow S$  是同伙向的向量  $x = (C, +, -) \in X$  中的分割? GDT ( $X$ )  $\in \mathbf{■}$  通过计算测地距离获得改造上: 在 CTH 图像通道, 用高斯模糊标准偏差, 与套正的和负种子+和-? 分别。C 表示的数量图像中的色彩通道, 和 B 的最大的从伯克利 300 集[17]的图像。B) + D) 外 (橙色的逼近测地距离) 和内部 (绿色) 分别种子。C) 最终分割

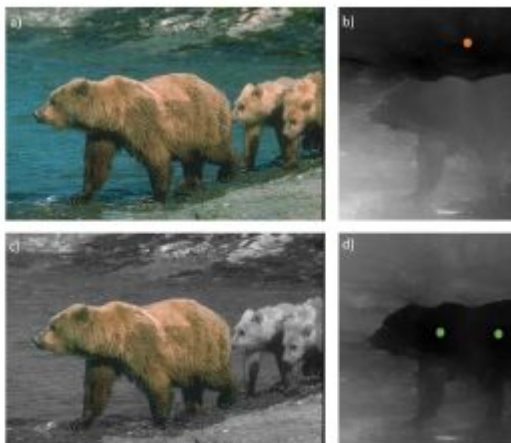


图 2: 采用测地距离的分割例子变换。

允许标准偏差。在 GDT 通过计算得到的每个像素的最短距离到设定种子像素。通常, 两个相邻像素之间的距离即, 图像图形的边缘的权重) 被定义为一个欧几里德距离的混合物之间的梯度这两点。以释放对种子的依赖感兴趣的对象内的位置, 我们使用的平方只有强度差这是因为, 特别重要的但不保证该种子将在中心对象, 因为用户不再手动放置种子。以产生最终

分割, 每个像素接收其最接近的种子标签。系数  $\in [0, B]$  被引入以控制到边缘的灵敏度的图像中, 并且隐含产生 MCMC 过程中图像金字塔流程 (见节 2.2.3)。请注意, 任何其他种子基于交互式分割算法可以包括在这代替框架的阶段。我们的主要动机背后测地距离变换的选择是事实它们可以在线性时间[25]来近似, 并且因此, 非常快速地计算。基于 GDT 分割过程在图进一步说明。2. 用于我们分段采样的马尔可夫链打算在状态空间的水平, 即, 在采取行动 GDT 基于分段参数和种子。式。1 变成

$$\alpha = \frac{p(\phi^{GDT}(x')|\Sigma^k)q(x|\Sigma^k, x')}{p(\phi^{GDT}(x)|\Sigma^k)q(x'|\Sigma^k, x)}$$

这两个下一小节将要暴露我们的设计

$$p(\phi^{GDT}(x)|\Sigma^k) \text{ and } q(x'|\Sigma^k, x)$$

### 2.2.2 后验概率

概率  $P(S|? k)$  的规定分割的概率已经给出的用户已经初步回升该设置 k 种子  $\mathbf{x}$  已被回答揭示第 k 第一个问题。都市斯接受概率 (公式 2) 的一个重要特征是事实上只有概率  $P(GDT(x')|? | k)$  的比和  $P(? GDT(xk))$  出现。因此, 归一化因子这个概率分布没有起到任何作用, 我们可以设计这种分布不考虑到这一点。要定义这个概率来看, 我们建议取决于是否存在至少区分 2 案件一个后台种子和  $\bullet K$  一个前景种子。

案例 1:  $K+=\emptyset$  或  $\bullet K-=\emptyset$  这种情况下, 通常发生中的第一个问题, 对应于不存在的

种子的两个标签中的至少一个。需要注意的是它发生一个至少在两个第一回合, 我们的知识包括的分别为 0 和 1 的种子。我们定义了分割概率

$$p(s|\Sigma^k) \propto \frac{1}{1 + \text{Var}(\{I(p), p \in s^{-1}(1)\})}$$

$\text{Var}(\{I(p), p \in s^{-1}(1)\})$  表示的方差以上定义的集合前景位置的图像值由分割第这方差求和所有颜色通道。直观地说，我们鼓励分割其中划定同质区域的前景。

情形 2:  $\Sigma_+^k \neq \emptyset$  and  $\Sigma_-^k \neq \emptyset$  一旦至少有一个种子内外对象是已知的，我们可以使用在 GDT 分割算法打造一个更准确的估计的背景和前景区域的视觉属性。对于这一点，我们执行基于 GDT 分割对每个颜色通道和计算的已知的种子背景和前景强度直方图  $H_{k-}$  和  $H_{k+}$  聚集在颜色通道。虽然建设直方图，各强度值由逆加权测地距离进行编码的事实，即信心降低与对种子的增加（测地线）的距离。至分数分配给新的分割  $S$ ，我们计算同样  $S$  的背景和前景的直方图表示  $H_-(S)$  和  $H_+(S)$  和测量它们的失配目前的估计  $H_{k-}$  和  $H_{k+}$  通过卡方距离：

$$p(s|\Sigma^k) \propto \frac{1}{1 + \frac{1}{2} \sum_{\delta \in \{-, +\}} \chi(H_{\delta}^k)}$$

### 2.2.3 建议分布



我们的分割是通过一组参数生成  $X = (C, +, -) \in \text{点 } x$  作为输入发送到

### GDT 分割

算法。这种表示的主要优点是空间  $X$  允许移动更自然的方式从状态到另一个状态，并促进了提案的设计分布  $Q(\cdot | K, X)$ 。在实践中，我们维持两个集种子+和 - 同时包含已众所周知，因此包含在可靠的种子  $\cdot K+$  和  $\cdot K-$ （固定种子）以及由 MCMC 创建了一些其它的种子过程只（移动种子）。从给定的状态  $x$ ，一状态  $x'$  被均匀绘画和执行下面提到的五个动作之一：

- 1, 更改图像通道：图像通道  $c$  已重绘均匀。
- 2, 更改：平滑参数重绘均匀。
- 3, 添加种子：移动种子在随机加入（非种子）的位置和随机加到+或-。
- 4, 删除种子：移动的种子是从移除了+或-（如果不离开这一套空）。
- 5, 移动种子：移动种子在空间移动到（非种子）根据正态分布像素它完全定义建议分布  $Q(\cdot | K, X)$ 。以避免种子发散的数量，我们平衡采摘移动 3 和 4，这样前进的概率落后的举动同样有可能。

### 2.3. 选择问题

之后  $k$  问题都得到了提问和回答 ( $K \geq 0$ ) 在二段中记载的方法。2.2 绘制  $\tilde{n}$  分割滑雪， $1 \leq i \leq n$  表示近似的概率分布  $P(S | K)$  的分段  $s$  是期待已久的答案

由鉴于目前的知识甲骨文（由编码  $K$ ）。从这些样品，我们必须在最佳的决定

要问的问题给用户。要做到这一点，我们首先执行由相交的所有的图像的过度分割滑雪（图 3）。这提供了图像的一个分区成多个区域  $(R_j) 1 \leq j \leq ?$  根据样品滑雪。由于滑雪是唯一可用的信息，所述信息携带的每个像素是通过一个单独的段常数。该概率为一

个区域的  $R_j$  属于感兴趣的对象  $P$  (的  $R_j \subset S$ ) 可以从我们的样本为近似  $P$  (的  $R_j \subset S$ )  $= |P \cap R_j| / |R_j| = |P \cap R_j| / |R_j|$ 。 (5) 最不确定的段是一个针对此概率是最接近 0.5。然而，每个段具有由它的大小进行加权以帐户的事实，每个像素将同样由携带的信息中获益该段属于。因此，我们选择了段  $R_j$  最大化  $|R_j| (1 - |1 - 2P(R_j \subset S)|)$  作为最翔实。在这个阶段，我们可以淡然任意选择位置  $P_k \in R_j$  可确定选择的问题  $Q(P_k)$ 。

在实践中，我们选择了  $P_k$  的最内点  $*j$  (图 3)。实际上，这导致在那些在一般的问题更容易为大多数室内点一般在视觉上回答远离图像内的边缘所在。



### 3. 实验

#### 3.1. 实验装置

我们描述了该方法的参数设置，即是三个数据集是相同的，在更多的细节。我们用 CIELAB 色空间和最大模糊水平，即高斯平滑的标准偏差施加在 MCMC 过程(秒 2.2.1) 被设置为  $B=6$  个像素。该 MCMC 老化一步，其目标是争取独立从初始状态  $SK_0$ ，包括 1000 迭代。

250 马尔可夫链的动作是两者之间进行样品附图以确保之间有很强的相关性两个连续的样本。采样分割的总数在每圈设置为  $N=32$ 。这是根据经验发现是足以产生之间有足够的各种

的分割，使得过分割不变得过于细粒度不超过之间 1 秒两个问题。的主要瓶颈是计算近似 GDT 的，需要在每个 MCMC 步骤，这是在像素的数量线性。因此，总复杂度为  $O(NK)$ 。Markov 链上运行图像的下采样版本，以加快计算的样品，但每拉伸分割之间 250 步保持完整图像上计算的，以确保一个细粒度分割。该方法是为最多评价以  $K=30$  个问题。在我们的实验中，人类用户由地面实况分割，这使得模拟一个广泛的评估在一个大组图像。所有实验是在英特尔 i7-4820K3.7GHz 的执行中央处理器。该算法运行四个独立的马尔可夫链在平行下。

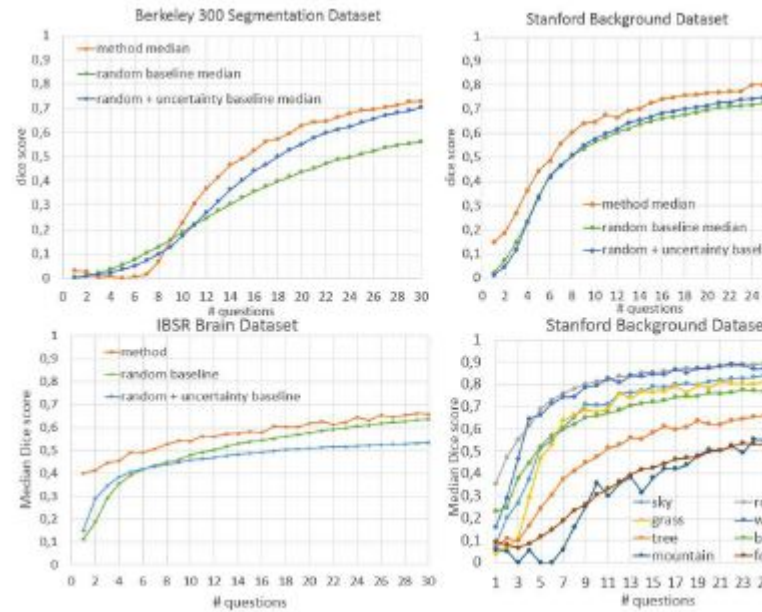


图 5: 平均得分散子与演变问题数量。性能显示在三个数据集: 伯克利分割数据集 300[17],

斯坦福大学的背景数据集[9]和 IBSR 脑数据集。我们还表明,在斯坦福大学的背景数据,个人表现为每个类的标签。少量数据集的山标签说明上的噪声相应的曲线。

#### 3.2. 结果

要测量分割质量,我们使用索伦森指数或骰子得分[22],它测量分割与地面之间的真相重叠。我们比较对两名直观的基线我们的方法。该首先通过随机绘制相应的选择每个问题像素位置。我们的第二个基础是一种进步第一个的:一旦种子已发现里面隐藏的分割外, GDTbased 分割对这些种子进行。该算法使用结果像素明智的信心来选择其中,标签不确定性是最大的区域内的问题。换句话说,这种方法可以被看作是一个迭代分割边界细化。实验在三个不同的数据集进行评价。



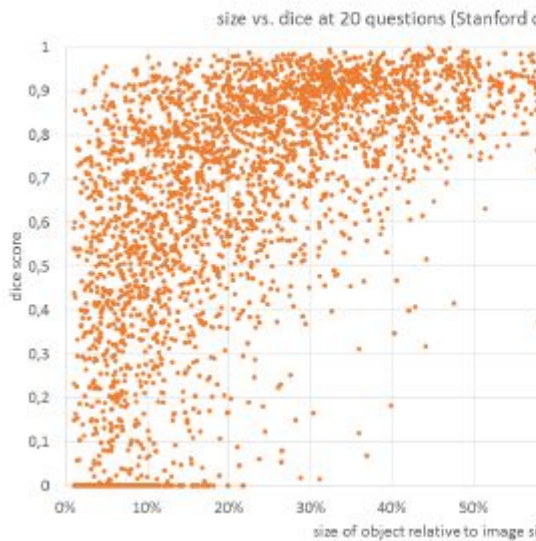


图 6: 隐藏的分割大小的关系和性能。显示的是一个散点图相对对象的大小 (X 轴) 与 20 后骰子得分 (Y 轴) 的问题。大面积的比小的更容易猜到。

伯克利分割数据集 300 伯克利分割数据集 300 [17]是一组 300 自然的图像。对于每个人,几个地面实况人类注解限定该图像的解析成几个区域分别集。这个数据很好地说明了复杂性

发现在 oracle 认为的区域中,由于 parsings 的通常从注释器到另一个不同。的平均数量的区域中的每个单独地事实是 20.3。每到这个数据评估我们的方法,包括在地面实况 parsings 的一个区域被选择甲骨文分割。图 5 显示了平均骰子得分超过所有运行时允许问题的数量增加,为我们的方法和上述基线。在图的曲线。5 开始在低高原性能。这对应于阶段,没有可靠前景种子已经发现的是,并且因此采样细分是非常嘈杂的,因为他们只是

基于从马尔可夫链移动种子。图。五示出所需的题中位数找一个可靠的种子是约 7。这是有趣的,因为该数据集包含每个图像 20 不重叠的区域平均。因此,即使该池 20 可能地

真理分割事先是已知的,随机种子位置仍然需要大约 10 个问题找到一个前景种子。这

表明,我们的问题选择程序更有效地识别区。斯坦福背景 DataSet 使用的第二个数据集评估是斯坦福大学的背景数据集[9]是 715 自然的图像不同的收藏品从不同的取其他数据集。在这种情况下,图像已被标记的根据八种不同的语义范畴:天空,道路,草,山,水,树,建设和前景目标。这有力地从伯克利分割不同

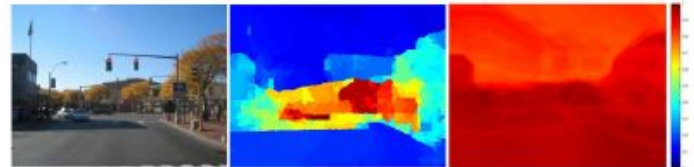


图 7: 我们基于 MCMC-提案之间的比较和测对象的建议。热图对应

到每个像素的发生频该组样品分割,为测对象建议方法[15] (中) 和我们的 MCMC 采样

方案 (右)。而前者更侧重具体在某些领域,我们的探索更多的分割空间。

数据集,其中图像解析成任意区域这是本质上更加暧昧。结果是上图所示。5.为了让更多的见解,我们还显示性能为每个单独的标签 (图 5) 和关系骰子之间得分后,20 个问题和大小目标分割 (图 6)。更大的细分 (分别为一个典型的主导类的分割) 似乎更容易比较小的猜测。这种相关性对象的大小和分割质量之间是相当自然和直接来自问题的类型被问。这一点在在二段更详细讨论。

### 3.3。

IBSR 脑 MR 数据,以评估进一步的一般性我们的方法,我们评估它的医疗 IBSR 脑部 MR18 核磁共振脑部扫描数据集组成用 4 个不同的大脑结构的标签一起。我们分解扫描为二维片具有较大的数都运行于我们的评价。这些数据和他们的手工分割是由中心提供的形态在分析美国马萨诸塞州总医院和

有可在 <http://www.cma.mgh.harvard.edu/IBSR/>。标记的脑结构是一个非常复杂的任务这需要很高的专门知识。我们的方法的结果在 IBSR

数据集可以在图中看到。5.如所预期的，在绝对性能比上的两个其他数据集低其中有一个更丰富的视觉内容，但一个显著改善经过 20 个问题仍实现。

用测对象的建议最后的比较，我们分割的比较我们基于 MCMC 抽样描述以秒。2.2 随着近期短程目标的提案 (GOP) 由克雷恩布尔和 Koltun[15]，其引入目标是在图像中建议感兴趣的对象。通过使用对于 GOP 可用 (预先训练) 的代码，我们样本与 GOP 和 N 分割每个图像都 n 对象的建议与我们的抽样方法，其中 N 是自动由共和党算法推断。我们报告杰卡德

最佳匹配地面实况样本的指数 (ABO 评分 [15])。共和党获得 66.0 的平均分在斯坦福背景数据集和 43.1 的 ISBR，而我们的方法分别得到 59.6 和 51.0。GOP 出现有对斯坦福的背景数据集的优势，大概是因为他们依赖于种子配置技术这是在室外场景精确教训。上相反，我们的方法不依赖于任何离线训练步骤，即使在图像内容没有假设，并且仍然执行对大脑图像更好，它说明我们的框架的灵活性。图 7 示出了另一有趣这两个提案框架之间的区别。没有当前图像上的先验知识，在分割由共和党提出的比我们更多的冗余，这是一个很好的特性，提出最有可能的对象，但适合交互式猜谜场景更少的上下文。

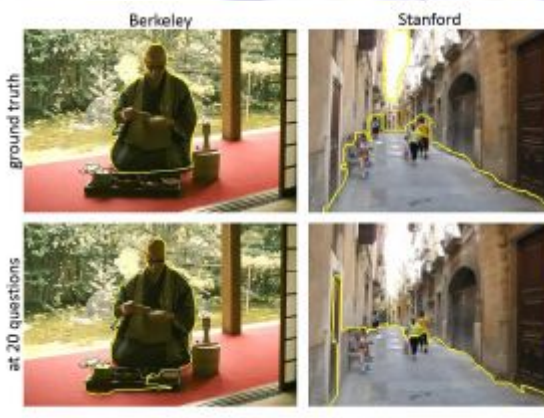


图 8: 定性结果。甲骨文分割的例子和相应的分割猜测之后 20 个问题。从左至右依次为: 伯克利分割数据集, 斯坦福背景数据集, IBSR 脑数据集。

我们的方法分别得到 59.6 和 51.0。GOP 出现有对斯坦福的背景数据集的优势，大概是因为他们依赖于种子配置技术这是在室外场景精确教训。上相反，我们的方法不依赖于任何离线训练步骤，即使在图像内容没有假设，并且仍然执行对大脑图像更好，它说明我们的框架的灵活性。图 7 示出了另一有趣这两个提案框架之间的区别。没有当前图像上的先验知识，在分割由共和党提出的比我们更多的冗余，这是一个很好的特性，提出最有可能的对象，但适合交互式猜谜场景更少的上下文。

Dataset	Method	10Q	20Q	30Q
Berkeley	Random	30.9 / 19.0	41.7 / 43.7	48.9 / 56.2
	Uncertainty	32.5 / 17.5	46.9 / 55.0	55.9 / 70.3
	Ours	34.7 / 23.8	48.8 / 62.0	56.1 / 73.2
Stanford	Random	49.8 / 56.5	60.2 / 69.8	65.6 / 74.7
	Uncertainty	50.6 / 57.6	61.9 / 71.4	67.9 / 77.7
	Ours	52.6 / 63.9	63.9 / 75.8	67.9 / 79.8
IBSR	Random	43.1 / 47.8	52.0 / 57.6	58.1 / 63.6
	Uncertainty	41.7 / 45.8	46.2 / 50.7	49.2 / 53.3
	Ours	51.5 / 53.7	58.4 / 60.1	62.9 / 64.7

表 1: 定量结果。平均数和中位数骰子在三个数据集 10, 20 和 30 分后的问题我们的方法和二段描述的两个基线。

3.2 不准确的答案，探讨的稳健性提出的方法相对于不准确的答案给问题，我们尝试了另一种实验性的设置上对 SBD 数据集，使得当一个问题是在内构成所需的对象的边界的 3 个像素，答案是随机决定 (50% 是 50% 否)。出现这种情况有关的 20% 的时间，主要是对以后的问题。平均，需要 1.2 其他问题来达到同样的骰子得分与完美的答案。因此，所提出的法是相当强大的，以错误的

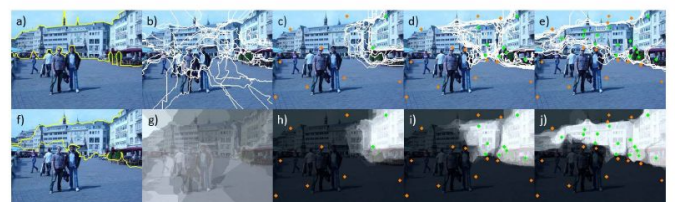


图 9: 与收集答案数目的分割信念演变。a) 示出原始图像和由 Oracle 选择的分割。下

面, F) 显示了在 30 个问题由计算机提出的最终分割。右边的图象显示与采集 (二种子数分割信念的演变: 0, C: 10, D: 20, E: 30)。所有采样分割的两个交点 (鼻内的过分割在 2.3 描述并示于图 3) 和象素可能性的热图中示出, 分别在顶部和底部行。热图是由产生计算, 对每一个像素, 即包含其取样分割的比例 (类似于等式 5)。正 (相应。从收集到的答案负) 种子以绿色 (RESP. 橙色显示)。注意的问题是如何逐步问向左, 其中的不确定性是最高的, 使整个系列的建筑物可以最终分割。

### 3.3. 讨论

我们开始这个讨论用一个例子来说明进一步该方法的行为。图 9 示出一个典型的实例的场景。神谕认为一个对象, 其中是在这种情况下, 在后台建筑物的行。它为分割算法一个具有挑战性的目标对象由于其结构的复杂性, 事实上, 它对应于一个相当高的水平语义。计算机产生第一组可能分割的。由于没有种子产品尚未推出, 统一的区域如天空, 或路面在此阶段的青睐。第一个问题是问在将根据最信息的位置在抽取的样本。由于从 Oracle 的答案是渐进收集, 我们可以看到, 该空间采样分割急剧变化。该算法建议在问题的边界位置, 并扩大一步一步向左的初始猜测。甲骨文的一个很好的估计分割后的 30 个问题, 最终提供的。

结果整体, 我们的方法演示更困难段的小物件, 特别是发现在这些情况下, 第一前景种子。虽然好像比较直观的, 它需要更长的时间才能找到小比大的领域, 我们问问题的性质是该部分负责为好。我们调查这个由改变样的问题问。例如, 一个更有效的解决方案, 以找到小物件会显示一个突出显示的区域给用户, 并问: “是目标分割充分显示区域内?” 如果答案是肯定的, 所有未突出显示像素可以被添加到外部种子和搜索空间将被大大降低。然而, 一个没有能够提供几乎没有信息, 因为它不仅会状态的对象中

的至少一个 (未知) 像素以外所显示的区域。因此, 这种类型的问题, 患有从相反的问题: 它们是有效的找到小对象, 但在发现路数差, 而且有附加缺点了否定的回答实际上是很难以充分利用。这主要是什么促使我们选择基于位置的问题类型, 但我们相信, 这两种类型的问题的组合将是理想

适合寻找图像中的小型和大型物体。

### 4. 结论

我们推出了能猜到一个分割的方法通过问简单的二元问题的任意图像用户。该问题是由一个近似计算后超过设定分割的概率分布与一测地线 MCMC 采样算法建设距离变换分割。我们的方法显示是温顺与两者之间的等待时间实际使用少于一秒的问题。没有关于假设给定图像的类型是必要的, 因为框架确实不依赖于任何离线训练的一步。实验评价在三个非常不同的数据集进行演示该方法提供了一个全面有效的解决方案到了这个问题。我们今后工作的方向包括研究混合几种题型的可行性并自动学习的最相关的隐藏的分割。由于回答问题二是振振有词比人工标注更有吸引力, 我们还在考虑使用这种方法生成大量的训练在一个众包的方式种子展示位置的数据。致谢的支持下 TECHNISCHE 的 Universit "atM " unchen - 高级研究所, 资助德国卓越计划和欧洲根据赠款协议联盟第七框架计划无 291763, 并与 DFG 的部分资金支持合作研究中心 824: 成像选择, 监测和癌症治疗的个性化参考:

### References

- [1] 20q.net: The neural-net on the internet ([www.20q.net](http://www.20q.net)).
- [2] Akinator the web genie (<http://en.akinator.com>).

- [3] D. Batra, A. Kowdle, D. Parikh, J. Luo, Heidelberg, 2010.  
and T. Chen. *icoseg: Interactive co-segmentation with intelligent scribble guidance.* In *CVPR 2010*, pages 3169–3176. IEEE, 2010.
- [4] Y. Boykov and M.-P. Jolly. *Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images.* In *ICCV 2001*, volume 1, pages 105–112 vol.1, 2001.
- [5] S. Branson, C. Wah, F. Schroff, B. Babenko, P. Welinder, P. Perona, and S. Belongie. *Visual recognition with humans in the loop.* In K. Daniilidis, P. Maragos, and N. Paragios, editors, *Computer Vision ECCV 2010*, volume 6314 of *Lecture Notes in Computer Science*, pages 438–451. Springer Berlin
- [6] J. Carreira and C. Sminchisescu. *CPMC: Automatic Object Segmentation Using Constrained Parametric Min-Cuts.* *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.
- [7] I. Endres and D. Hoiem. *Category-independent object proposals with diverse ranking.* *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 36(2):222–234, Feb 2014.
- [8] A. Fathi, M. F. Balcan, X. Ren, and J. M. Rehg. *Combining self training and active learning for video segmentation.* In *Proceedings of the British Machine Vision Conference (BMVC 2011)*, volume 29, pages 78–1,

2011. *Recognition*, 2009.
- [9] S. Gould, R. Fulton, and D. Koller. *CVPR 2009. IEEE Conference on*, pages 1030–1037, June 2009.
- Decomposing a scene into geometric and semantically consistent regions. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1–8. IEEE, 2009.
- [10] L. Grady. Random walks for image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(11):1768–1783, Nov 2006.
- [11] L. Grady, M.-P. Jolly, and A. Seitz. Segmentation from a box. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 367–374. IEEE, 2011.
- [12] C. Gu, J. Lim, P. Arbelaez, and J. Malik. Recognition using regions. In *Computer Vision and Pattern Recognition, 2009. IEEE Conference on*, pages 1030–1037, June 2009.
- [13] A. Kowdle, D. Batra, W.-C. Chen, and T. Chen. imodel: interactive co-segmentation for object of interest 3d modeling. In *Trends and Topics in Computer Vision*, pages 211–224. Springer, 2012.
- [14] A. Kowdle, Y.-J. Chang, A. Gallagher, and T. Chen. Active learning for piecewise planar 3d reconstruction. In *CVPR, 2011*, pages 929–936. IEEE, 2011.
- [15] P. Krahenbühl and V. Koltun. Geodesic object proposals. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, editors, *Computer Vision ECCV 2014*, volume 8693 of *Lecture Notes*

- in *Computer Science*, pages 725–739. Springer International Publishing, 2014.
- [16] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp. Image segmentation with a bounding box prior. In *ICCV*, number MSR-TR-2009-85, 2009.
- [17] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.
- [18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [19] C. Nastar, M. Mitschke, and C. Meilhac. Efficient query refinement for image retrieval. In *CVPR* 1998, pages 547–552, Jun 1998.
- [20] B. Price, B. Morse, and S. Cohen. Geodesic graph cut for interactive image segmentation. In *CVPR* 2010, pages 3161–3168, June 2010.
- [21] M. Sadeghi, G. Tien, G. Hamarneh, and M. S. Atkins. Hands-free interactive image segmentation using eyegaze. In *SPIE Medical Imaging*, pages 72601H–72601H. International Society for Optics and Photonics, 2009.
- [22] T. Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of

- species and its application to analyses of the vegetation on danish commons. *Biol. skr.*, 5:1–34, 1948.
- [23] C. Straehle, U. Koethe, G. Knott, K. Briggman, W. Denk, and F. Hamprecht. Seeded watershed cut uncertainty estimators for guided interactive segmentation. In *CVPR*, 2012, pages 765–772, June 2012.
- [24] K. Tieu and P. Viola. Boosting image retrieval. In *CVPR*, 2000. *Proceedings*, volume 1, pages 228–235 vol.1, 2000.
- [25] P. J. Toivanen. New geodesic distance transforms for grayscale images. *Pattern Recognition Letters*, 17(5):437–450, 1996.
- [26] S. Tong and E. Chang. Support vector machine active learning for image retrieval. In *ACM International Conference on Multimedia*, pages 107–118, New York, NY, USA, 2001. ACM.
- [27] Z. Tu and S.-C. Zhu. Image segmentation by data-driven markov chain monte carlo. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, 24(5):657–673, May 2002.
- [28] C. Wah, S. Branson, P. Perona, and S. Belongie. Multiclass recognition and part localization with humans in the loop. In *Computer Vision (ICCV)*, 2011 *IEEE International Conference on*, pages 2524–2531, Nov 2011.