

# 基于轮廓线统计量的前景分割 Markov 随机场模型

汤慧旋<sup>1</sup> 危辉<sup>1</sup>

**摘要** 提出了一种基于轮廓线统计量的前景分割 Markov 随机场 (Markov random field, MRF) 模型. 和 Grabcut 等以往模型不同, 本文模型通过在分割标签的编码中加入对轮廓线方向的考虑, 将 Gestalt 知觉组织的原则加入分割约束中去, 从而使分割边界更为平滑. 作为前景分割和 Gestalt 知觉组织原则研究的基本框架, 本文模型的系统结构分为前景分割、注意力选择和信息整合三个子模块, 与相关神经生理研究的结论相一致. 最后, 分别给出了基于本文模型的自动和半自动前景分割实现, 结果好于 Grabcut 等相关算法的结果.

**关键词** 神经视觉, 前景分割, 轮廓线, 自然统计量, 格式塔原则  
**中图分类号** TP18

## Figure-ground Separation by Contour Statistics and Markov Random Field Model

TANG Hui-Xuan<sup>1</sup> WEI Hui<sup>1</sup>

**Abstract** In this paper, we propose a Markov random field (MRF) based representation for the Gestalt law, and suggest using a message passing-like scheme to infer the segmentation. Different from other grabcut models, our MRF function is specially encoded to consider orientations along the contour, thus the Gestalt law is embedded into the inference. As a basic framework of the research in figure-ground separation and Gestalt law, our system is designed in reference to neurophysiology, and the architecture is composed of three modules: primal visual cortex (V1), extra-striate cortex (V2), and the interest selected region. To validate our method, we conduct experiments in both auto and interactive segmentation algorithm. The results are better than those of grabcut and other related algorithms.

**Key words** Neural vision, foreground segmentation, contour, ecological statistics of contours, Gestalt law of perceptual grouping

计算机视觉研究的目的是将数字图像解释为它所表示的物理场景. 每幅数字图像都是一个有序的像素集合, 每个像素表示一种颜色. 只有当所有的像素汇聚到一起, 图像才具有物理意义. 视觉的基本任务之一是前景分割 (Figure-ground separation), 即把感兴趣的主体从图像中分离出来. 前景分割是一系列重要视觉感知能力的基础: 它把图像的表达上升到了对象的层次, 大脑对于动态视觉刺激所需的存储空间大大减少<sup>[1]</sup>; 同时, 前景分割具有的整体性表达能力, 运动分析时所面临的多义性问题也在很大程度上被弱化<sup>[2]</sup>; 此外, 由于图像的背景通过分割机制与前景分离开来, 注意力迁移所需要的计算代价也减少<sup>[3]</sup>.

和其他视觉任务一样, 前景分割的主要困难也

在于图像的多义性. 视网膜上的成像所提供的信息是极为有限的, 仅仅依靠这些信息很难完成某些看似简单的智能任务, 比如视觉心理学中的著名图像《狗》的例子. 心理学先驱 Helmholtz 指出无意识推理 (Unconscious inference) 在视觉中的关键作用, 即只有结合图像信息和相关的先验知识才能顺利地完对图像的解释分析工作<sup>[4]</sup>.

在前景分割中最重要先验知识被称为 Gestalt 知觉组织原则 (Gestalt law for perceptual grouping). 它指的是大脑在视觉处理过程中倾向于把感知到的信息解释为简洁、单一的整体. 这一思想是由心理学家 Ehrenfels 首先提出的<sup>[5]</sup>. 视觉心理学家 Wertheimer 首先指出了该原则在视觉中的决定性作用<sup>[6]</sup>. 神经生理实验<sup>[7]</sup> 发现前景分割主要包括三个部分:

1) 图像边缘检测. 这一模块检测出图像灰度或纹理不连续的位置. 这里使用文献 [8] 中基于学习的灰度和纹理边缘检测模型.

2) 全局信息整合. 这一模块修补 (Fill in) 图像边缘检测错检或漏检的边缘. 第 2 节和第 3 节将描述该区域的建模和它与另两个区域的衔接.

3) 预注意前景选择. 这一模块引导信息整合机制选择观察者目前感兴趣的位置. 本文使用基于谱

收稿日期 2008-06-18 收修改稿日期 2008-12-01  
Received June 18, 2008; in revised form December 1, 2008  
国家重点基础研究发展计划 (973 计划) (2010CB327900), 国家自然科学基金 (60303007) 和上海市科委重点科技攻关项目 (08511501703) 资助

Supported by National Basic Research Program of China (973 Program) (2010CB327900), National Natural Science Foundation of China (60303007), and Shanghai Science and Technology Development Fund (08511501703)

1. 复旦大学计算机科学与工程系认知算法模型实验室 上海 200433  
1. Department of Computer Science, Laboratory of Algorithm for Cognitive Model, Fudan University, Shanghai 200433  
DOI: 10.3724/SP.J.1004.2009.01033

残差的预注意模型或用户笔画来完成这部分的实现.

显然, Gestalt 知觉组织原则主要在第 2) 步, 即全局信息整合中发挥作用. 然而, Gestalt 知觉组织原则所强调的“整体”目标是如何在大脑这样一个并行分布式的组织结构上进行? 由于计算模型的缺乏, 这一问题始终没有得到回答. 这也使得对 Gestalt 知觉组织原则的研究止步于定性的分析. Berkeley 的 Malik 等研究了 Gestalt 知觉组织原则下的一系列学习算法以用于图像边缘提取<sup>[9]</sup> 和边界修补<sup>[10]</sup>. 同时, York 大学的 Elder 等采用了相似的方法研究数据集的生态统计特征<sup>[11]</sup>. Elder<sup>[11]</sup> 和 Malik<sup>[8-10]</sup> 的模型都是由 Canny 边缘构造的, 因此只能给出 Canny 边缘上各线段的显著性, 而不能解决 Canny 线段定位不准确或漏检的问题. 虽然 Malik 和 Elder 等在建立全局优化的计算模型方面做出了卓越的贡献, 但这些工作无一具有生物可行性.

在 Gestalt 原则神经基础研究方面, Sokolov 等为 Gestalt 原则的神经基础提出了两种可能的假设<sup>[12]</sup>: 1) Gestalt 模型是视觉系统层次化的组织结构的结果; 2) Gestalt 模型是神经动力系统同步化震荡的结果. 2000 年, Li 设计了前景预分割的同步震荡模型<sup>[13]</sup>, 并提出分割是由回归网络中神经动力学中产生的. 该网络通过神经元间水平连接的侧抑制 (Lateral inhibition) 作用形成相互竞争, 并最终收敛到一系列可能的解. 近年, Ehrenstein 等神经生物学家开始倾向于认为 Gestalt 知觉组织是视觉系统感受野层次组织的结果<sup>[14]</sup>. Grossberg 等提出的 FACADE 模型<sup>[15]</sup> 包括边界轮廓系统 (Boundary contour system, BCS) 和特征轮廓系统 (Feature contour system, FCS) 两个互补的部分. Kokkinos 等简化了该模型并通过分析该模型的神经动力学特性<sup>[16]</sup>, 发现 FACADE 模型的合作竞争机制事实上起到局部正归化的作用.

综上, 本文试图为视觉皮层 V1 和 V2 区中的前景分割机制建立一套数学模型, 从而:

1) 以前景分割为例, 探索中层的智能任务如何由基础的神经计算完成. 计算的基本步骤应与基本的神经计算机制相一致, 且具有一定程度的可扩展性.

2) 为边界检测和纹理分割提供一个一致的、符合生理实验的计算模型. 为了给纹理分割、错觉轮廓等相关的认知研究提供计算基础. 该模型应具有通过局部计算对全局信息进行优化的能力.

3) 通过对解空间的重新表示, 更完善地在神经网络模型上表达 Gestalt 原则. 在文献 [8-11] 已有的模型中, 仅表示了 Gestalt 原则中的相似律, 接近律和连续律.

## 1 Markov 随机场模型的设计和求解

由于图像分割的主观因素较强, 因此近年也提出了通过人为干预辅助图像分割的算法. 这类方法或者要求用户首先在图像上标注少量的样本点<sup>[17]</sup>, 或者要求用户同时分割多张图像<sup>[18]</sup>. 这类方法作为一种辅助设计手段, 为图像分割问题提供了可行的解决途径. 本文使用两种方式引导前景提取 (如图 1): 1) 基于注意力选择机制的自动前景分割使用基于谱残差 (Spectral residual) 的方法<sup>[19]</sup>, 即首先给出图像的显著性图谱, 然后再通过本文中的推理算法导出整幅图像的分割; 2) 基于用户笔画的半自动前景分割首先由用户在图像上输入笔划 (User stroke) 以确定一些像素 (前景和背景像素) 的标签, 再通过推理算法导出分割.



图 1 本文的两种应用

Fig. 1 Two types of applications in this paper

这两种不同的引导前景提取方法对应于 V4 区注意力选择功能的两种不同实现. 自动场景选择算法通过对高频信号定位的方式决定了前景物体的位置. 而半自动算法将前景选择的任务交给了用户. 假定 V4 区的注意力显著性是空间独立的, 并且上述两种前景显著性猜测是准确的. 此时本文的算法可转化为这样一个问题: 给定在网格  $\Omega$  上的图像  $I(x, y)$  和各像素被标注为前景的概率, 决定  $\Omega$  中每一格为前景/背景标签, 求该图像的前景  $R \subset \Omega$ . 一般而言, 对前景/背景标签有以下要求:

1) 每个像素是前景的概率  $P((x, y) \in R)$  与用户笔画给出的约束条件或是前景显著性一致.

2) 前景  $R$  的边界轮廓线  $\partial$  落在图像灰度或纹理变化最显著的位置.

3) 前景  $R$  的边界轮廓线  $\partial$  符合自然界中轮廓线的统计特征, 如闭合、连续、平滑等.

等价的, 对输入图像的每个像素赋予一个定义在网格  $\Omega$  上的标签  $L(x, y)$ , 如图 2(a) 和 (b) 所示.

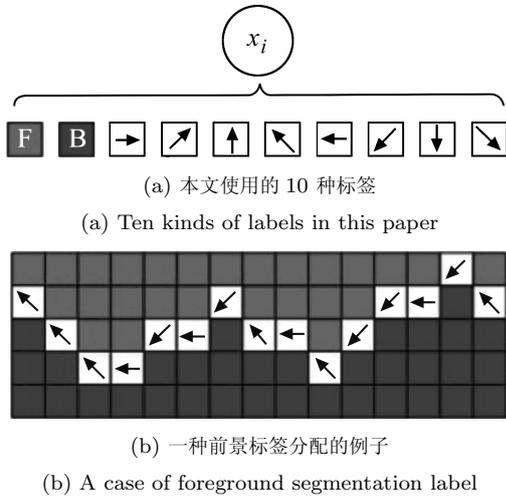


图 2 前景分割标签

Fig. 2 Foreground segmentation label

和以往模型不同, 本文使用 10 种标签来表示图像的前景/背景分配. 每个像素不仅可能被指定为前景或背景, 也可能被指定为轮廓线. 轮廓线的标签分 8 个不同的方向. 给定一条轮廓线, 该轮廓线上某点的标签由它沿顺时针方向的切线方向决定. 作为表示分割的一种方式, 标出轮廓线局部的朝向是冗余的. 然而, 它有利于将 Gestalt 法则编码到 Markov 随机场对应目标函数中去. V2 区同时受到来自 V1 和 V4 区输入的驱动. 此外 Gestalt 知觉组织原则同时也作为一个不可见的约束参与了 V2 区的信息整合. 由于该原则的作用, 在整合过程中各变量的取值不是独立的, 而与其他变量之间协同作用. 基于此, 使用 Markov 随机场模型来表示这种复杂的协同关系. 一个 Markov 随机场是一个具有空间相关性的随机变量集合. Hammersley 和 Clifford 证明了 Markov 随机场的概率密度函数可以写作一些边际概率的乘积<sup>[20]</sup>:

$$P(X) = \prod_C P(X_C) \quad (1)$$

其中,  $X_C$  是图中第  $C$  个团中的随机变量集合.

为了简化 Markov 随机场的复杂性, 仅考虑低

阶的统计相关性:

$$P(X) = \prod_i P(x_i) \prod_{i,j} P(x_i, x_j) \quad (2)$$

在式 (2) 表示的 Markov 随机场中只考虑两个变量之间的相互关系. 上式中  $x_i$  表示随机场中第  $i$  个变量. 对 Markov 随机场概率密度的最大似然估计常被转化为对概率负 log 值最小化问题:

$$\Phi(X) = \sum_i \phi(x_i) + \sum_{i,j} \varphi(x_i, x_j) \quad (3)$$

这样, 目标函数中的各项概率就以线性方式迭加起来. 在式 (3) 中,  $\sum_i \phi(x_i)$  称为数据代价, 表示推理过程中由数据驱动的部分, 在前景分割模型中, 用它来表示由注意力机制控制的部分. 式 (3) 中的  $\sum_{i,j} \varphi(x_i, x_j)$  称为平滑代价, 表示对两个不同变量的赋值 (或标签) 与先验知识的拟合程度. 在前景分割模型中, 用它反映 Gestalt 知觉组织原则所驱动统计规律.

图 3 给出了本文所使用的网络模型. 它的主要推理结构由实线圆所表示的变量和它们之间的连接构成. 每个实线圆都表示一个像素对应的标签, 它们受到 V4 区模型的前景显著性的约束, 即数据代价的部分. 变量之间的实心方块表示 Gestalt 知觉组织原则所调控的平滑代价. 值得注意的是, Gestalt 知觉组织原则并不完全独立于输入图像, 例如其中的相似律就收到图像中灰度边缘显著性等特征的驱动 (虚线方块部分).

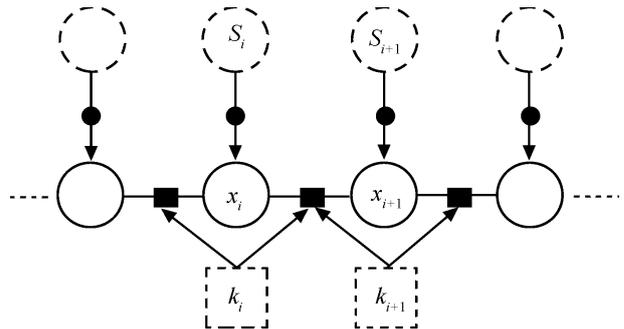


图 3 本文网络模型 (虚线圆表示注意力选择, 虚线方形表示边界显著性特征, 实线圆表示变量, 实心圆表示数据约束, 实心方形表示平滑约束)

Fig. 3 The network model in this paper (The circles with dotted line show the attention choice; the squares with dotted line show the significant features; the circles with solid line show variables; the solid rounds show the data constraint; and the solid squares show smoothing constraint.)

本文中的 Markov 随机场中每个变量的取值是离散的. 求解离散 MRF (Markov random field) 的

算法可以分为三类: 迭代条件模式方法、信息传递方法和二值优化方法. 其中, 信息传递方法和二值优化方法在视觉应用中往往具有较好的优化能力. 对本文模型, 我们选择使用信息传递算法求解上述 MRF 模型. 这首先是由于二值优化问题一般适用于处理模块化的问题<sup>[21]</sup>, 而本文的模型却往往无法具有良好的模块化结构. 其次, 信息传递模型的计算机制与神经计算机制十分吻合. 例如, Rao 指出 Message passing 算法中每个变量上积累的信度由神经元膜电位编码, 变量之间传递的信息则是突触之间相互传递的信号<sup>[22]</sup>. Popple 甚至指出 Gestalt 知觉组织原则的实现正依赖于神经元之间的协同工作, 而这恰好是信息传递算法的最主要计算特性<sup>[23]</sup>. 最后, 当图像存在多个解时, 信息传递算法往往能够找到多个解, 而二值优化算法一般只能找到一个解. 尽管多义图像的分割不在本文的讨论之列, 但选择 Message passing 类的算法求解本文模型有利于这方面后续研究的展开.

信息传递算法包括一个迭代的过程. 在每一次迭代中更新从每个结点  $s$  到它的每个邻居  $t$  的信息:

$$m_{st}(k) \leftarrow \min_{j \in L} \left\{ \phi_s(j) + \sum_{u \in N_s, u \neq t} m_{us}(j) + \varphi_{st}(j, k) \right\} \quad (4)$$

当迭代收敛后, 计算每个结点上收到的信息

$$b_s(k) = \sum_{t \in N_s} m_{st}(k) \quad (5)$$

即可得到该结点的标签:

$$l_s = \min_{k \in L} b_s(k) \quad (6)$$

Kolmogorov 等建议对 MRF 最终的解选择应随着信息的更新完成<sup>[24]</sup>. 因此, 每个变量的赋值仅取决于它当前所接受到的信息. 即如果解的选择是在前向传输过程中完成的, 那么在它之后被更新的变量对该变量没有影响. 由此可以看出, 信息传递的排序十分重要.

为了使其他变量的信息尽可能地影响每个变量的解选择, 本文按照解选择的不确定性对变量进行排序. 解的不确定性越高, 变量更新的时间就越晚, 它所接受的信息也越多. 使用一种类似于信息熵的度量来表示解选择的不确定性:

$$UC_i = \frac{\sum_k b_x(k) \exp\{-b_i(k)\}}{\sum_k \exp\{-b_i(k)\}} + \log_2 \left( \sum_k \exp\{-b_i(k)\} \right) \quad (7)$$

在实验中我们发现, 距离用户笔画更近的像素受到更大的约束, 因此不确定性更低, 所以首先进行信息更新. 边界清晰处像素的解选择更容易, 因此也相对更早地更新信息. 而图像中的角点、不连续位置等, 则具有最大的歧义性, 直到最后才确定标签. 这种最大约束优先的机制使得本文算法的计算资源集中到最关键的变量和标签上, 从而增加了前景分割结果的有效性和效率.

至此, V2 区的信息整合算法就确定下来. 它首先迭代地进行信息更新, 直到神经元上的信息收敛. 每次信息更新包括一轮前向信息传播. 最后, 再进行一轮信息传播并最终决定标签.

## 2 Gestalt 知觉组织原理的定量分析

Gestalt 知觉组织原理的统计表现确定了 Markov 随机场的目标函数, 由贝叶斯概率公式给出

$$P(l_i, l_j | g_i, g_j; N_{ij}) = \frac{P(g_i, g_j | l_i = p, l_j = q; N_{ij}) P(l_i = p, l_j = q; N_{ij})}{\sum_{p, q} P(g_i, g_j | l_i = p, l_j = q; N_{ij}) P(l_i = p, l_j = q; N_{ij})} \quad (8)$$

由此, 可将 Markov 随机场的目标函数分为两部分. 先验概率  $P(l_i = p, l_j = q; N_{ij})$  描述了轮廓线平滑性和封闭性的自然统计量, 即 Gestalt 知觉组织原理中的平滑律和闭合律. 条件概率  $P(g_i, g_j | l_i = p, l_j = q; N_{ij})$  描述了轮廓线上边缘显著性的平滑特点, 即 Gestalt 知觉组织原理中的相似律. 可用非参数的形式描述闭合律和连续律并通过统计分割标签的共生矩阵 (Co-occurrence matrix) 来表示标签的先验分布:

$$P(l_i = p, l_j = q; N_{ij}) = \frac{1}{N} \sum_{i, j \in N_{ij}} \sigma(l_i = p, l_j = q) \quad (9)$$

其中,  $l_i$  和  $l_j$  表示第  $i$  个和第  $j$  个变量赋予的标签,  $N$  是正则化参数,  $N_{ij}$  表示两个变量所属的邻域集合.

观察该先验概率 (如图 4), 发现它的概率集中在连续的方向上. 可见该概率模型很好地反映了连续律的作用. 特别地, 由于轮廓线表现为直线链码的形式, 直接反映了闭合的硬约束, 该约束也表现在统计数据中. 在以往的前景分割中往往都不涉及对闭合律的建模, 这导致了在前景/背景分配、边缘提取等相关工作往往不能取得很好的结果. Kovacs 等指出该约束能大大增加前景分割的鲁棒性<sup>[25]</sup>. 相似律要求被组合到一起的基元具有相似的特性. Elder 的研究发现同一条轮廓线上相近的线段具有相似的对比度和平均亮度<sup>[11]</sup>, 在这里, 通过类似的统计手段

分析边缘显著性在不同标签下的条件分布, 图 4(a) 给出了灰度不连续特性的条件概率直方图, 对每一对标签  $N_{ij}$ , 分别统计在该标签分配下对应的条件概率, 用不同亮度区分各方向下的概率分布. 图 4(b) 给出了灰度不连续特性边界显著性直方图, 对每一对标签  $N_{ij}$ , 分别统计在该标签区域或者边界上不同的概率分布. 纹理显著性直方图的形状与之类似. 本文的数据与文献 [7] 中的研究相似, 发现轮廓线上边界显著性的差异与它的大小不是独立的, 因此更适合以边界显著性的大小为控制变量, 使用高斯混合尺度模型<sup>[26]</sup> 进行建模. 在实验中, 由于训练数据较多, 非参数化的概率模型形式与参数化概率模型形式得到的结果没有明显差异.

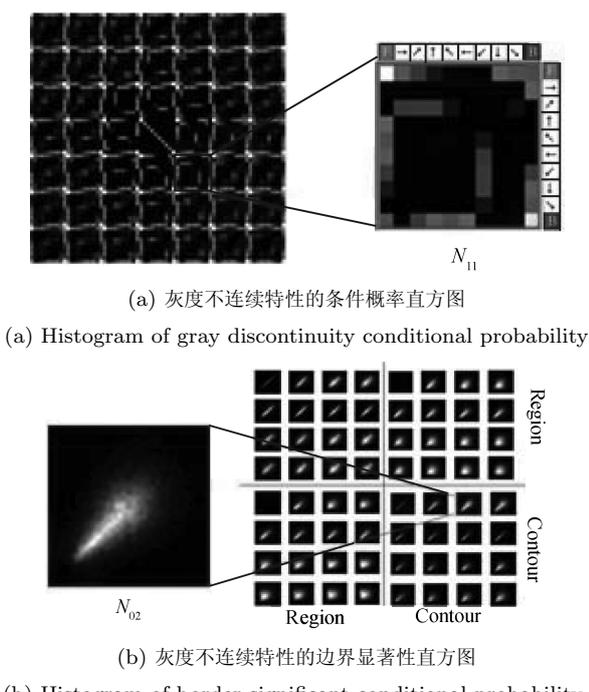


图 4 灰度不连续特性的条件概率与边界显著性直方图

Fig. 4 Histograms of gray discontinuity conditional probability and border significant conditional probability

### 3 实验结果

本文数据全部来自 Berkeley 图像分割数据集. 该数据集提供了在 200 幅图像上的约 1000 个人工标注. 由于 Berkeley 数据集的图像内容保罗万象, 而且大多关于自然场景、风光、人物或动植物, 因此 Berkeley 数据集能够很好地反映自然轮廓线的一般统计特征.

本文使用 Malik 的边缘检测算法对 V1 区的功能进行建模<sup>[8]</sup>. 对 V4 区的预注意模型, 我们在自动分割方式中使用谱残差模型<sup>[19]</sup> 进行建模, 在半自动分割方法中使用了用户的笔画.

首先对一部分 Berkeley 分割数据的测试图片进行了自动前景分割的测试, 以预注意机制的显著性<sup>[19]</sup> 作为赋值为背景的数据代价, 然后运行 Markov 模型直至迭代收敛, 并将本文自动前景分割模型与基于谱分割的分割算法进行比较. 选用的谱分割算法是多尺度的正规割算法<sup>[27]</sup>, 由于该算法无法控制聚类的数量, 直接将聚类数量设置为 2, 将导致不正确的分割. 因此, 首先将图像分割为 3040 个小块, 并对每个小块按照它们的显著性聚为两类, 最终确定场景的分割. 自动前景分割算法的结果如图 5 所示, 分界线为深色. 不难看出, 本算法不仅具有更平滑的边缘, 同时也更好地突出了图像的前景. 这是由于在谱分割算法仅具有将图像像素分离成几个聚类的能力, 但并不能从这些聚类中选择出前景. 此外, 谱分割算法对形状的控制仅限于 MRF 模型的拓扑结构, 而该结构所保证的轮廓线统计特征却与自然轮廓统计性特征略有不同, 这也导致了与本文模型相比其算法在一些图像上的边缘不够自然.

本模型也能够处理半自动图像分割的任务. 本文设计了一个简单的用户交互系统, 在 Berkeley 分割数据的测试图片上进行半自动前景分割的测试: 当用户打开该交互系统后, 该系统提交给用户一系列的图像, 并由用户在图像上描绘前景、背景笔画. 系统自动对当前图像进行边缘检测和前景分割, 并将结果返回到屏幕上. 用户可以在上一次计算的基础上添加或修改笔画, 从而修正之前的分割结果. 为了将本文的半自动前景分割算法与以前的算法作比较, 以同样的输入进行测试. 该方法是首先在图像上进行 Grabcut, 然后通过 Lazy snapping 方法进行边缘细化.

半自动前景分割的实验结果如图 6 所示, 深色粗线为用户笔画, 细线为分割结果. 由于本文的算法考虑到了边界平滑的因素, 因此最终运算的结果比半自动前景分割算法具有更为平滑的边缘.

### 4 结论

本文提出了一个前景分割的生理模型. 不同于以往由图像像素灰度引导的分割算法, 本模型具有很好的鲁棒性, 边界较为平滑. 神经生理实验发现, 视皮层分割机制也是主要受到边缘检测结果的引导, 这意味着本文模型具有比以往模型更好的生理基础, 因此可以作为一个基础计算模型便于后续的认知研究的展开. 同时, 本文使用了 Markov 随机场模型对前景分割机制进行建模: 通过在 Markov 随机场模型的编码中嵌入轮廓线的方向性, 成功地实现了基于 Gestalt 知觉组织原则的约束. 此外, 为了确定 Markov 随机场模型的具体形式, 本文对 Gestalt 知觉组织原理的统计表现进行的定量研究.



图 5 自动前景分割 (每组图像从上至下依次为: 原始图像、显著性图谱、正规割、本文模型)

Fig. 5 Automatic foreground segmentation (Each column from top to bottom are original image, significant patterns, regular segmentation, and our segmentation.)

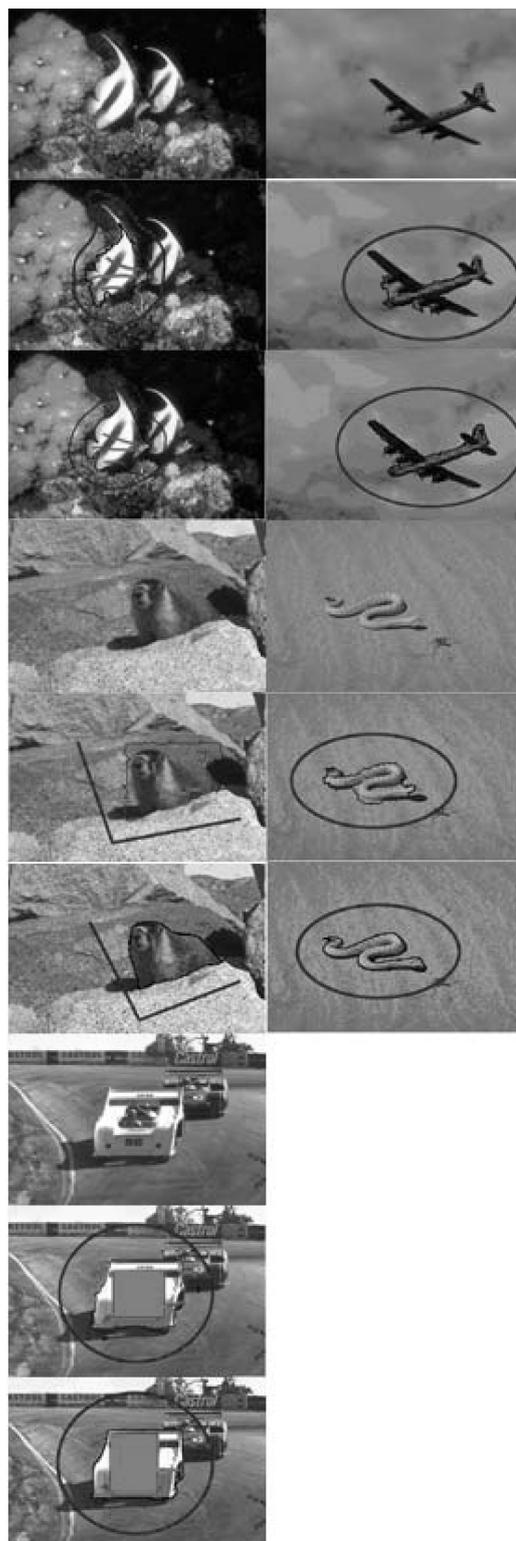


图 6 半自动前景分割 (每组图像从上至下依次为: 原始图像、基于 Grabcut + Lazy snapping 的前景分割、本文模型)

Fig. 6 Semi-automatic foreground segmentation (Each column from top to bottom are original image, the Grabcut and Lazy snapping based foreground segmentation, and our segmentation.)

与大多数现有的分割方法不同, 本文研究的主要目的是给出一个生理和计算上都可行的前景分割模型, 为以后的神经视觉模型研究奠定基础. 然而, 由于本文算法从工程计算的角度也有现实意义, 因此有必要进一步分析本文与其他计算机视觉中的前景分割算法的联系与区别.

1) 基于谱分割的方法<sup>[27]</sup> 在计算流程上与本文的模型有很大不同. 从计算原理而言, 它更多地考虑邻域与邻域或是像素与像素之间的距离, 却几乎没有考虑轮廓线的特征. 这导致了谱分割方法有时无法产生平滑的分割. 此外, 由于谱方法只将数据投影到一个低维欧氏空间, 因此必须事先知道图像中对象的个数才可能得到较准确的聚类. 错误的聚类数量估计常常导致产生错误的分割结果. 这也是谱分割算法常常导致过度分割的原因. 由于本文模型使用用户笔画或预注意模型引导了图像的分割, 因此较谱分割方法而言更偏好对用户而言“自然”的分割结果.

2) 基于高斯混合模型的 Markov 随机场<sup>[26]</sup> 使用的是和本文相同的 Markov 随机场结构. 但同谱分割方法一样, Markov 随机场方法没有考虑到边界的作用. 因此以往的 Markov 随机场模型很难把 Gestalt 知觉组织原则嵌入进去. 这导致了以往的 GMM (Gaussian mixture model) 方法产生的边界不够平滑的问题.

3) 基于学习的图像分割模型: 在基于学习的图像分割模型中, 最早探索了 Gestalt 知觉组织原则的统计表现. 然而, 这些学习模型大多只是在过滤伪边缘的基础上对 Gestalt 模型进行推理, 但对由于边缘检测算子造成的边缘漏检和定位不准确的问题则无能为力. Tu 等提出的 MRF 算法尽管不存在上述问题<sup>[28]</sup>, 但他们提出的模型在计算上很难快速优化, 必须用 MCMC (Markov chain Monte Carlo) 方法采样才能获得, 训练时间较慢, 因此, 也不适用于实时运算的求解.

本文模型为前景分割问题后续的神经计算模型研究给出了一个基本的计算框架, 但还有若干特征暂未包括在目前模型中: 包括反馈连接、快速颜色分割系统以及高级区域的参与.

反馈作用是视皮层分割机制的重要组成部分, 然而它在前景分割中发挥着怎样的作用尚不清楚. 反馈可能的作用之一可能在于引导注意力所关心区域的边缘检测, 或是提高边缘密集区域边缘检测的分辨率. 在本文的计算框架下, 反馈连接可能的作用之一是激活更细致尺度下部分神经元的边缘分割作用. 这样, 视皮层可以把有限的计算资源集中在对前景分割最有价值的位置. 在后续工作中, 可以考虑通过一个多尺度、有注意迁移机制的 Markov 模型实

现该目的.

神经科学家倾向于认为在前景分离的任务中, 存在着一个快速的颜色分割系统对分割起到引导作用<sup>[29]</sup>. 可以认为, 该系统的主要作用和反馈连接相似, 都是为了引导 Markov 模型把计算资源集中到有价值的部分. 快速颜色分割系统的作用可能在于为图像建立子像素级的表示, 从而增加 Markov 模型中的信息带宽, 提高系统计算的速度. 然而, 如果解决在快速颜色分割系统中的定位误差和欠分割问题还是一个需要研究的问题.

在前景分割机制中, 大脑的 LOC (Lateral occipital complex) 区域发挥着重要的作用, 不同于底层区域, LOC 区域具有平移、旋转、仿射、对比度等多种不变性<sup>[30]</sup>. 正是 LOC 区域所具有的这种识别不变性支持了 Gestalt 知觉组织原则中另一重要定律: 对称律的实现. 在本文框架下, 对称律可以通过附加一个形状匹配系统 (例如文献 [31] 中的算法) 获得. 对应点之间一旦得到匹配, 它们之间的形状相似性约束被激活, 从而可以从残缺的局部边界中估计出完整的边界. 此外, 更高级区域 (如颞叶) 担负着识别物体轮廓的功能. 在轮廓识别与匹配的基础上, 物体类别知识也可以参与到分割任务中来, 并通过类似的机制形成与物体类别相一致的错觉轮廓.

附录 目标函数的非子模块性

不失一般性, 考虑一维的情况. 由于前景和背景垂直相邻的概率接近 0, 因此对应的概率应接近无穷大. 同样地, 左边像素为背景而右边像素为向下的轮廓线的概率也为 0. 以此类推, 在图 7 中标记为  $c_{max}$  的元素对应的概率都为 0. 考虑当前所有像素都标记为背景时将标签扩张到  $B, CU$  的情况, 考虑此时图 7 中右边图中的实线连接上的代价矩阵及其对应子矩阵. 由于有:

$$PB, B < P(B, CU) \tag{A1}$$

因此

$$a = -\log_2 P(B, B) > -\log_2 P(B, CU) = b \tag{A2}$$

此时, 代价矩阵的主对角线上的元素之和大于次对角线的代价之和. 因此代价矩阵是非子模块性的, 因此不能用 Graph-cut 算法求解<sup>[21]</sup>.

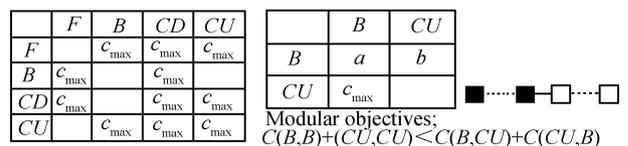


图 7 水平方向上相邻单元的平滑代价矩阵 (F: 前景; B: 背景; CD: 切向向下的轮廓线; CU: 切向上轮廓线;  $C_{max}$ : 最大代价)

Fig. 7 Smooth price matrix in horizontal direction for adjacent units (F shows the foreground; B shows the background; CD shows the contour which cut-direction is down; CU shows the contour which cut-direction is up; and  $C_{max}$  shows the largest price.)

## References

- 1 Mazza V, Turotto M, Carlo Umiltà. Foreground-background segmentation and attention: a change blindness study. *Psychological Research*, 2005, **69**(3): 201–210
- 2 Tommasia L, Vallortigarab G. Figure ground segregation modulates perceived direction of ambiguous moving gratings and plaids. *Vision Research*, 1998, **39**(4): 777–787
- 3 Qiu F T, Sugihara T, von der Heydt R. Figure-ground mechanisms provide structure for selective attention. *Nature Neuroscience*, 2007, **10**(11): 1492–1499
- 4 von Helmholtz H. *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss, 1867. 450
- 5 Ehrenfels C V. On Gestalt-qualities. *Psychological Review*, 1937, **44**(6): 521–524
- 6 Wertheimer M. Laws of organization in perceptual forms. *A Source Book of Gestalt Psychology*. London: Routledge and Kegan Paul, 1938. 71–88
- 7 von der Heydt R, Peterhans E. Mechanisms of contour perception in monkey visual cortex. *Journal of Neuroscience*, 1989, **9**(5): 1731–1748
- 8 Martin D R, Fowlkes C C, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(5): 530–549
- 9 Martin D R, Fowlkes C C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 416–423
- 10 Ren X F, Malik J. A probabilistic multi-scale model for contour completion based on image statistics. In: Proceedings of the 7th European Conference on Computer Vision. Berlin, German: Springer, 2006. 312–327
- 11 Elder J H, Goldberg R M. Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, 2002, **2**(4): 324–353
- 12 Sokolov E N. The problem of Gestalt in neurobiology. *Neuroscience and Behavior Physiology*, 1997, **27**(4): 323–332
- 13 Li Z P. Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, 2000, **13**(1): 25–50
- 14 Ehrenstein W H, Spillmann L, Sarris V. Gestalt issues in modern neuro-science. *Axiomathes*, 2003, **13**(3-4): 433–458
- 15 Grossberg S, Mingolla E, Williamson J. Synthetic aperture radar processing by a multiple scale neural system for boundary and surface representation. *Neural Networks*, 1995, **8**(7-8): 1005–1028
- 16 Kokkinos I, Deriche R, Faugeras O, Maragos P. Computational analysis and learning for a biologically motivated model of boundary detection. *Neurocomputing*, 2008, **71**(10-12): 1798–1812
- 17 Li Y, Sun J, Tang C K, Shum H Y. Lazy snapping. *ACM Transaction on Graphics*, 2004, **23**(3): 303–308
- 18 Rother C, Minka T, Blake A, Kolmogorov V. Cosegmentation of image pairs by histogram matching-incorporating a global constraint into MRFs. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2006. 993–1000
- 19 Hou X D, Zhang L Q. Saliency detection: a spectral residual approach. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2007. 1–8
- 20 Hammersley J M, Clifford P. *Markov Field on Finite Graphs and Lattices*. Oxford: Oxford University, 1971
- 21 Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(2): 147–159
- 22 Rao R P N. Neural models of Bayesian belief propagation. *The Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge: MIT Express, 2006. 235–270
- 23 Popple A V. Context effects on texture border localization bias. *Vision Research*, 2003, **43**(7): 739–745
- 24 Kolmogorov V. Convergent tree-reweighted message passing for energy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(10): 1568–1583
- 25 Kovacs I, Julesz B. A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences of the United States of America*, 1993, **90**(16): 7495–7497
- 26 Wainwright M J, Simoncelli E P. Scale mixtures of Gaussians and the statistics of natural images. *Advances in Neural Information Processing Systems*. Cambridge: MIT Press, 2000. 855–861
- 27 Shi J B, Malik J. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 888–905
- 28 Tu Z W, Zhu S C. Image segmentation by data-driven Markov chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 657–673
- 29 Moller P, Hurlbert A C. Psychophysical evidence for fast region-based segmentation processes in motion and color. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, **93**(14): 7421–7426
- 30 Norcia A. Imaging the time-course of figure-ground segmentation. *Journal of Vision*, 2007, **7**(15): 13
- 31 Shekhovtsov A, Kovtun I, Hlavac V. Efficient MRF deformation model for non-rigid image matching. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2007. 1–6



汤慧旋 复旦大学计算机科学与工程系硕士研究生。主要研究方向为计算机视觉研究。E-mail: 0124143@fudan.edu.cn (TANG Hui-Xuan Master student in the Department of Computer Science, Fudan University. Her main research interest is computer vision.)



危辉 复旦大学计算机系教授，博士。1998年于北京航空航天大学计算机系获得博士学位，2000年浙江大学计算机系博士后出站。主要研究方向为人工智能与认知科学研究。本文通信作者。E-mail: weihui@fudan.edu.cn

(WEI Hui Professor in the Department of Computer Science, Fudan University. He received his Ph. D. degree from Beijing University of Aeronautics and Astronautics in 1998 and finished post-doctor work in the Department of Computer Science, Zhejiang University in 2000. His research interest covers artificial intelligence and cognitive science research. Corresponding author of this paper.)