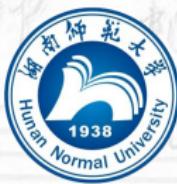


用于交互式图像分割的 自适应表观分离

指导老师：李乔良教授

答辩人：彭自立



湖南师范大学 数学与统计学院

二〇一八年十一月

目录

1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢



1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢

- 论文的研究背景与意义
- 图像分割国内外研究动态
- 本文研究工作与组织结构

HUNAN NORMAL UNIVERSITY



理论意义

- 图像分割是进行图像分析与理解的重要先导步骤之一，它决定最终的图像分析和图像理解的结果和质量。二元图像分割的目标是把图像像素分割成两类：前景和背景。
- 本人提出的自适应表观分离模型，在分割过程中融入了对图像内容高层信息的理解，通过交互式分割获得更精准的分割结果，大大降低了图像分析与理解的难度。
- 此外，本人的方法充分利用了现有的最优化算法、数学工具并做了改进，促进了图像分割研究的新发展。

◀ back



实际意义

- 分割结果可用于在不丢失有用信息的前提下进行数据的压缩和传输，同时由于分割后的图像与噪声的关系减弱，也具有一定的降噪功能；
- 便于进行图像的语义检索；
- 便于进行3D重建和可视化；
- 便于对目标物体的尺寸、体积或者容积进行测量；
- 有助于达到特定的分析目的或满足其他实际需求，如医学分析、目标跟踪与识别等.

◀ back



发展历程和分类

图像分割是从完全人工分割向半自动分割再到全自动分割逐步发展的。早期的图像分割技术倾向于区域分裂与归并，近二十年的图像分割算法通常是优化某个目标函数。

通过对交互式图像分割方法的思路和特点进行深入分析与归纳总结，本文将其划分为如下5大类：

- 基于归一化切割的方法[Shi and Malik, 2000, Li and Chen, 2015];
- 基于图割的方法[Boykov and Jolly, 2001, Gorelick et al., 2017];
- 基于随机游走的方法[Grady, 2006, Dong et al., 2016];
- 基于区域的方法[Tao et al., 2007, Ning et al., 2010];
- 基于边的分割方法[Kass et al., 1988, Li et al., 2010];

◀ back



本文研究工作及创新点

- ① 提出了一种新颖的用于交互式图像分割的表观分离模型。该模型通过结合图像前景-背景的颜色特征差异和对应的测地线距离差异信息，减少甚至消除了来自背景的杂乱干扰，获得良好的初始软分割结果，为实现最终准确的分割迈出了一大步。
- ② 提出结合表观分离模型与Dense CRF框架的精准图像分割算法。本文算法可根据待分割图像的特性，对每个图像自适应地设置并微调Dense CRF框架中的相关参数，将所提出的表观分离模型应用于该框架，经过一次均值场推断即可获得较准确的分割结果，且降低了用户涂鸦交互工作量。
- ③ 提出结合表观分离、表观交叠惩罚与Graph-Cut框架的快速图像分割算法。将本文提出的表观分离模型和表观交叠信息合理融入到高效的图割框架中，提出合理的表观交叠惩罚权重方案，缓解了OneCut算法倾向于出现孤立点的问题，提高了其分割精度、降低了用户的涂鸦交互工作量。



绪论

本文研究工作与组织结构

本文组织结构

全文共分为5章：

- 1 绪论：介绍了论文的研究背景与意义、图像分割算法国内外研究动态、本文的主要研究工作与内容的安排。
- 2 用于交互式图像分割的基础框架：图模型、稠密条件随机场中的高效均值场推断框架、图割框架。
- 3 基于Dense CRF框架的自适应表观分离交互式图像分割算法。
- 4 使用测地表观交叠图割的交互式图像分割算法。
- 5 结语：对本文的研究工作进行了总结与展望。

◀ back



1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠分割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢

● 图模型

● Dense CRF高效推断框架

● 图割框架



图像与图

将图像建模成无向图后，为了实现图像分割，我们可引入一个标签集合 L 以指示分割状态（如某个标签 $l_i = 1$ 对应前景目标，或标签 $l_i = 0$ 对应背景）。这时，图 G 中的顶点 $v \in V$ 需与某个随机变量 $l \in L$ 相关联，由此，对图像 I 的分割问题即描述为了对图 G 中的顶点求标签的问题。通常而言，邻近的图像像素（可能非直接邻接）很可能具有相同的标签，对应图中相邻的结点的标签也倾向于相同。因此进一步地，假设该图模型满足马尔科夫属性：

$$\Pr(l_i | L_{V \setminus \{v_i\}}) = \Pr(l_i | L_{\mathcal{N}(i)})$$

那么图像分割就可基于贝叶斯理论建模为在马尔科夫随机场（MRF）中求最优标签的问题。有了输入图像和用户提供的交互信息，就可以通过最大后验概率（maximum a posteriori, MAP）方法求得图中余下顶点的标签估计值。



图像与图

基于马尔科夫随机场，根据贝叶斯规则与MAP估计，图像分割的目标函数可表示为：

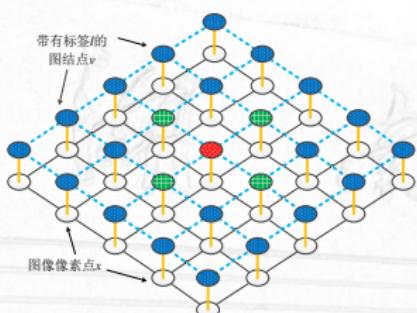


Figure 2.1: 一个 5×5 大小的图像相应的4-邻域马尔科夫随机场的图模型示例。

$$\begin{aligned}
 \hat{l}_{1\dots N} &= \arg \max_{l_{1\dots N}} \Pr(l_{1\dots N} | x_{1\dots N}) \\
 &= \arg \max_{l_{1\dots N}} \frac{\Pr(l_{1\dots N}, x_{1\dots N})}{\Pr(x_{1\dots N})} \\
 &= \arg \max_{l_{1\dots N}} \frac{\prod_{i=1}^N \Pr(x_i | l_i) \Pr(l_{1\dots N})}{\Pr(x_{1\dots N})} \\
 &= \arg \max_{l_{1\dots N}} \prod_{i=1}^N \Pr(x_i | l_i) \Pr(l_{1\dots N}) \\
 &= \arg \max_{l_{1\dots N}} \sum_{i=1}^N [\log \Pr(x_i | l_i) + \log \Pr(l_{1\dots N})]
 \end{aligned}$$

图像与图

由马尔科夫属性和吉布斯分布，根据Hammersley - Clifford定理，先验概率 $\Pr(l)$ 的负对数似然可表示为成对相互作用的基团势函数 E_p 之和：

$$-\log \Pr(l_{1\dots N}) \propto \sum_{(i,j) \in E} E_p(i,j), \quad (2.2)$$

在图像分割中，势 E_p 称为边界项或平滑项，表示的是相邻标签之间非一致性的惩罚（如图2.1中所示的蓝色虚线）。与之相应的，条件概率 $\Pr(x_i|l_i)$ 的负对数似然记为势 E_d ，它常被称为区域项或数据项，反应的是分配某个标签的代价（如图2.1中所示的橙色实线）：

$$E_d(i) = -\log \Pr(x_i|l_i). \quad (2.3)$$



图像与图

由等式2.1至等式2.3可知，图像分割总的势能函数为区域项与边界项之和：

$$E(l_{1\dots N}) = \sum_{i \in V} E_d(i) + \sum_{(i,j) \in E} E_p(i,j). \quad (2.4)$$

综上，通过MAP估计最优标签进行图像分割，即为求使得势能函数 $E(l_{1\dots N})$ 最小的标签集合：

$$\hat{l}_{1\dots N} = \arg \min_{l_{1\dots N}} \sum_{i \in V} E_d(i) + \sum_{(i,j) \in E} E_p(i,j). \quad (2.5)$$

[◀ back](#)



条件随机场

MRF关注的是随机变量的联合概率分布 $\text{Pr}(\mathbf{l}, \mathbf{x})$ ，而CRF关注的是随机变量的条件概率分布 $\text{Pr}(\mathbf{l}|\mathbf{x})$ 。在MRF中，经典的贝叶斯模型假设标签的先验分布 $\text{Pr}(\mathbf{l})$ 与观测值 \mathbf{x} 独立：

$$\text{Pr}(\mathbf{l}|\mathbf{x}) \propto \text{Pr}(\mathbf{x}|\mathbf{l}) \text{Pr}(\mathbf{l}), \quad (2.6)$$

而现实情况可能需要根据观测值来调整先验模型。从本质上来说，条件随机场是给定了观察值集合的马尔科夫随机场，所以，它的条件分布形式除了多了观察集合以外，与马尔科夫随机场的分布形式是一样的。



条件随机场

条件随机场用数学符号描述为：给定观察数据 \mathbf{x} ，如果每个标签 l_i 都满足马尔科夫性：

$$\Pr(l_i|\mathbf{x}, L_{V \setminus \{v_i\}}) = \Pr(l_i|\mathbf{x}, L_{\mathcal{N}(i)}) \quad (2.7)$$

则标签集合 \mathbf{l} 是一个条件随机场。因为在某些情况下，等式2.6中的贝叶斯规则可能不再满足，所以对 $\Pr(\mathbf{l}|\mathbf{x})$ 不做显式分解有时处理起来更方便。由马尔科夫随机场与吉布斯分布的等价性，在使用条件随机场模型时，可直接将该模型的负对数似然写作数据项和平滑项之和：

$$\begin{aligned} E(\mathbf{l}|\mathbf{x}) &= E_d(\mathbf{l}|\mathbf{x}) + E_p(\mathbf{l}|\mathbf{x}) \\ &= \sum_{i \in V} E_d(l_i|\mathbf{x}) + \sum_{(i,j) \in E} E_p(l_i, l_j|\mathbf{x}) . \end{aligned} \quad (2.8)$$



均值场推断理论

在均值场近似中，近似分布 $Qr(\cdot)$ 一般选择成因子相乘的形式，即：

$$Qr(\mathbf{x}) = \prod_{i \in V} q_i(x_i), \quad (2.9)$$

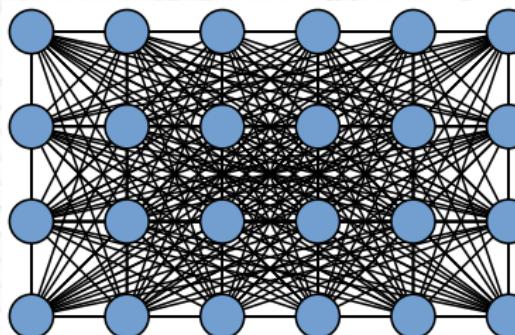
记 $\mathbf{q} = \{q_i(x_i)\}$ ，称为伪边缘，每一个因子都遵循： $q_i(x_i) \geq 0, \forall i, x_i$ ，且 $\sum_{x_i} q_i(x_i) = 1, \forall i$ 。这就意味着，只要确定了 $Qr^*(\mathbf{x})$ ，目标 $\mathbf{x} = (x_1, \dots, x_N)$ 的最大后验概率估计就可以通过选择每一项最大的近似 $\hat{x}_i = \arg \max_{x_i} q^*(x_i)$ 来对整体 Pr 进行近似。在信息论与机器学习实践中，相似性通常采用Kullback-Leibler散度（KL-散度）进行度量。近似分布 Qr 相对于分布 Pr 的KL-散度定义为：

$$D_{KL}(Qr || Pr) = \sum_{\mathbf{x}} Qr(\mathbf{x}) \log \left(\frac{Qr(\mathbf{x})}{Pr(\mathbf{x})} \right), \quad (2.10)$$



DENSE CRF中的高效均值场推断

Krähenbühl和Koltun在2011年提出了在稠密连接的条件随机场(Dense CRF)中的均值场推断，能够高效获得非常精准的图像分割结果。这主要归功于以下两个原因：(1) 像素相关性的长程建模；(2) 利用高斯卷积的近似实现具有高斯边势的Dense CRF的快速推断。



DENSE CRF中的高效均值场推断

Dense CRF的吉布斯能量定义：

$$E(x|I, \theta) = \sum_i \varphi_i(x_i|I, \theta) + \sum_{i < j} \varphi_{ij}(x_i, x_j|I, \theta) \quad (2.11)$$

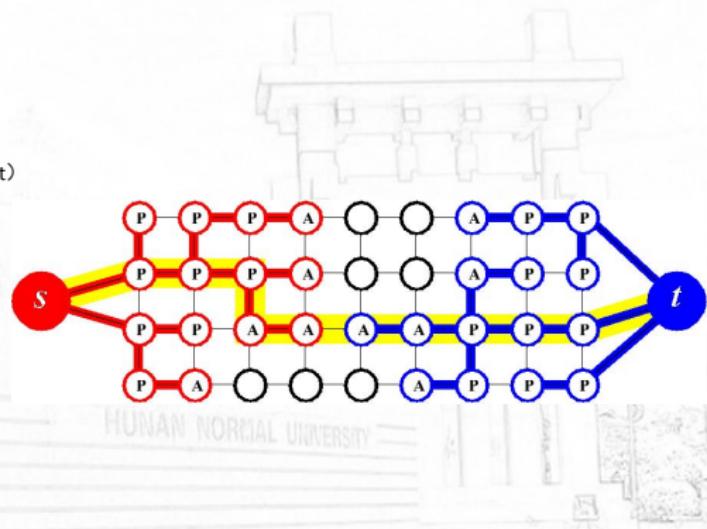
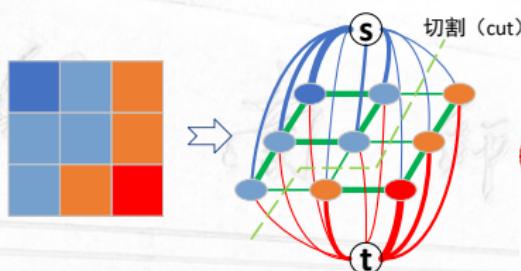
其中 $x \in \{B, F\}$ 为与模型 θ 相关的标签， B 表示背景， F 表示前景， i 和 j 为像素的索引号（1 至 N ）， N 为图像 I 中的像素数。 φ_i 为一元项，衡量像素 i 的标签分配 x_i 代价； φ_{ij} 为成对项，对相邻的相似像素分配不同标签加入惩罚。

一元项可以取任意形式，成对项必须为带权高斯核的线性组合。一旦满足这样的形式，Dense CRF 框架使得我们能够在均值场近似框架中利用高斯滤波技术获得来自快速最大后验边缘（MPM）推断的解。

◀ back



最小割/最大流与BOYKOV-KOLMOGOROV算法



◀ back



1 絮论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠切割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢

- 引言
- 相关工作
- 基于Dense CRF的自适应表观分离算法
- 实验结果与分析
- 本章小结

DENSE CRF中的均值场推断

简要介绍了Krähenbühl和Koltun在2011年提出的在稠密连接的条件随机场（Dense CRF）中的均值场推断，能够高效获得非常精准的图像分割结果。

我们的表观分离模型目前以该Dense CRF框架为基础，总领本章的主要工作。

◀ back



GRABCUT、ONECUT、DENSECUT等

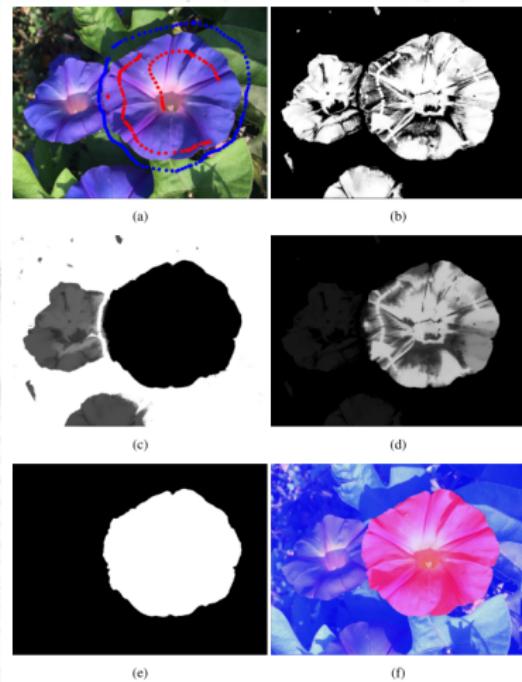
对相关交互式分割技术进行了概括总结，并指出经典的GrabCut、OneCut[Tang et al., 2013]、DenseCut[Cheng et al., 2015]等交互式分割算法存在的不足。

◀ back



动机

本章提出的算法试图解决上述三个典型算法存在的问题，这也是图像分割领域常常遇到的难题：当前景与背景颜色模型相近或杂乱时，如何以尽量少的用户交互工作量，通过高效的算法实现准确的分割。相关分割说明与示例如右图，后面的部分会对图中的细节作详细介绍。



所提出方法的理论基础

本章的方法是基于稠密条件随机场框架中的高效滤波的推断技术。为灵活调整各子项的相对重要性，现将稠密条件随机场的吉布斯能量改写为：

$$E(\mathbf{x}|\mathbf{I}, \theta) = \lambda \sum_i \varphi_i(x_i|\mathbf{I}, \theta) + \sum_{i < j} \varphi_{ij}(x_i, x_j|\mathbf{I}, \theta), \quad (3.1)$$

一旦条件随机场的表达式符合上述形式，[Krähenbühl and Koltun, 2013]的工作可使得本章在均值场近似框架中利用高斯滤波技术，从快速最大后验边际推断中获得优化解。

◀ back



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

所提出的方法 |

对于一元项，本章算法将其定义为：

$$\varphi_i(x_i) = -\log P(x_i), \quad (3.2)$$

其中 $P(x_i)$ 不再像 DenseCut 中那样简单地使用朴素的前景/背景颜色模型：

$$Color(x_i) = \frac{P(\Theta_{x_i}, I_i)}{P(\Theta_F, I_i) + P(\Theta_B, I_i)}, \quad (3.3)$$

而是使用颜色模型之间的差异 $C(x_i)$ 以及测地线距离之间的差异 $G(x_i)$ ，即本章所提出的表观分离模型 ($P(x_i = F)$) 的定义如下，相似的情形 $P(x_i = B)$ 可以简单地进行相应定义)：

$$P(x_i = F) = (C(x_i = F) + G(x_i = F) \times 8) \times 0.12, \quad (3.4)$$

$$c(x_i = F) = \frac{P(\Theta_F, I_i) - P(\Theta_B, I_i)}{P(\Theta_F, I_i) + P(\Theta_B, I_i)}, \quad (3.5)$$



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

所提出的方法 II

$$C(x_i = F) = \begin{cases} c(x_i = F) & \text{if } c(x_i = F) > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (3.6)$$

$$Geo(x_i = F) = \frac{D_F(pix_i) - D_B(pix_i)}{D_F(pix_i) + D_B(pix_i)}, \quad (3.7)$$

$$G(x_i = F) = \begin{cases} -Geo(x_i = F) & \text{if } Geo(x_i = F) < 0 \\ 0 & \text{otherwise} \end{cases}. \quad (3.8)$$

$P(x_i)$ 之所以定义成那样的形式，是为了将其大致限制在0到1之间，使其成为像素*i*属于表观模型的一个有效概率。 $P(\Theta_l, I_i) \in (0, 1)$ 表示的是某个颜色 I_i 属于颜色模型 Θ_l 的概率密度值，可使用高斯混合模型对其进行描述，其中标签 $l \in \{F, B\}$ 。这样， $c(x_i = F)$ 即表示前景颜色模型和背景颜色模型之间的差异。



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

所提出的方法 III

$D_l(pix_i)$ 为测地距离项，它度量的是从像素 pix_i 到最近的种子像素 F 或 B 的测地线距离。它可通过经典的Dijkstra算法计算得到。这样， $Geo(x_i = F)$ 即表示到前景与到背景的测地线距离之间的差异。本质上说，测地线距离是在颜色模型的梯度图上的最短带权路径。

本算法使用高效估计得到的前景/背景颜色模型之间的差异来近似客观的表观交叠，如式3.5和3.6所示的 $C(x_i)$ 。

对于成对项，本章追随[Krähenbühl and Koltun, 2013]、DenseCut中的工作，将其建模为关于颜色特征和位置信息的高斯函数，形如：

$$\varphi_{ij}(x_i, x_j) = u(x_i, x_j)(w_1 K_1(\mathbf{f}_i, \mathbf{f}_j) + w_2 K_2(\mathbf{f}_i, \mathbf{f}_j)) , \quad (3.9)$$

$$K_1(\mathbf{f}_i, \mathbf{f}_j) = \exp\left(-\frac{{dist}_{ij}^2}{\theta_d^2} - \frac{{\Delta I}_{ij}^2}{\theta_{II}^2}\right) , \quad (3.10)$$



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

所提出的方法 IV

$$K_2(f_i, f_j) = \exp\left(-\frac{\Delta I_{ij}^2}{\theta_{I2}^2}\right), \quad (3.11)$$

其中 $u(x_i, x_j)$ 是用于反应标签兼容性的函数，本章算法将其设置为 Potts 模型： $u(x_i, x_j) = 1_{[x_i \neq x_j]}$ 。 K_1 和 K_2 分别为定义在特征 f_i 和 f_j 上的高斯核函数； w_1 和 w_2 为对应核函数的权重因子； $dist_{ij}$ 、 ΔI_{ij} 分别表示位于 i 和 j 处的像素之间的距离、颜色差异； θ_d 、 θ_{I1} 和 θ_{I2} 用于调整平滑度、相似度以及紧致度。等式 3.9 至等式 3.11 的含义与 DenseCut 算法中的表述是一致的。在此，本算法舍弃了 DenseCut 算法中的第三个核函数，原因是它仍然反应的是颜色的相似度，且仅用作微调，而本算法通过合理设置相应系数完全可将其取代。

◀ back



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

相关参数设置 |

由于朴素颜色模型 $Color(x_i)$ 本质上扮演着简单贝叶斯分类器的角色，所以我们可以估计分类的误差：

$$\varepsilon = \frac{1}{2} \left[\frac{\sum_{x_i \in F} (1 - Color(x_i))}{|\Omega_F|} + \frac{\sum_{x_i \in B} (1 - Color(x_i))}{|\Omega_B|} \right]. \quad (3.12)$$

当该误差非常小时 ($\varepsilon \approx 0$)，说明该分类器是比较可靠的，那么有必要分配大的权重给表观分离模型；当误差 ε 过大（比如由于杂乱且相近的颜色模型导致误差 $\varepsilon \geq 0.5$ ），说明该分类器无法准确反映出前景与背景之间的颜色差异，那么应该给表观分离模型分配较低的权重甚至是零权重；分类误差 ε 会随着颜色模型的不明确性的增加而增大，因此，根据该误差的大小给表观分离模型分配适当小的权重才是明智之举。



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

相关参数设置 II

这样，可为三个权重参数引入一个置信度项 κ :

$$\kappa = \begin{cases} 0 & \text{if } \varepsilon \geq 0.5 \\ 1 - 2\varepsilon & \text{otherwise} \end{cases} \quad (3.13)$$

根据上述分析与实验经验，本章设置Dense CRF框架中一元项权重系数为 $\lambda = \kappa$ ；表观核函数的权重系数设置为 $w_1 = 55 \times (1 - \kappa)$ ，用于权衡具有相似颜色的相邻像素分配相同的标签；平滑核函数的权重系数设置为 $w_2 = 0.8 \times \kappa$ ，用于权衡移除孤立点的规模。按照上述方式设置后，表观分离模型、平滑度和相似度之间的相对权重就会根据逐个图像的内部信息自动调整，这样也就增强了本章所提出的算法的自适应性和灵活性。



└ 基于Dense CRF框架的自适应表观分离交互式分割算法

└ 基于Dense CRF的自适应表观分离

相关参数设置 III

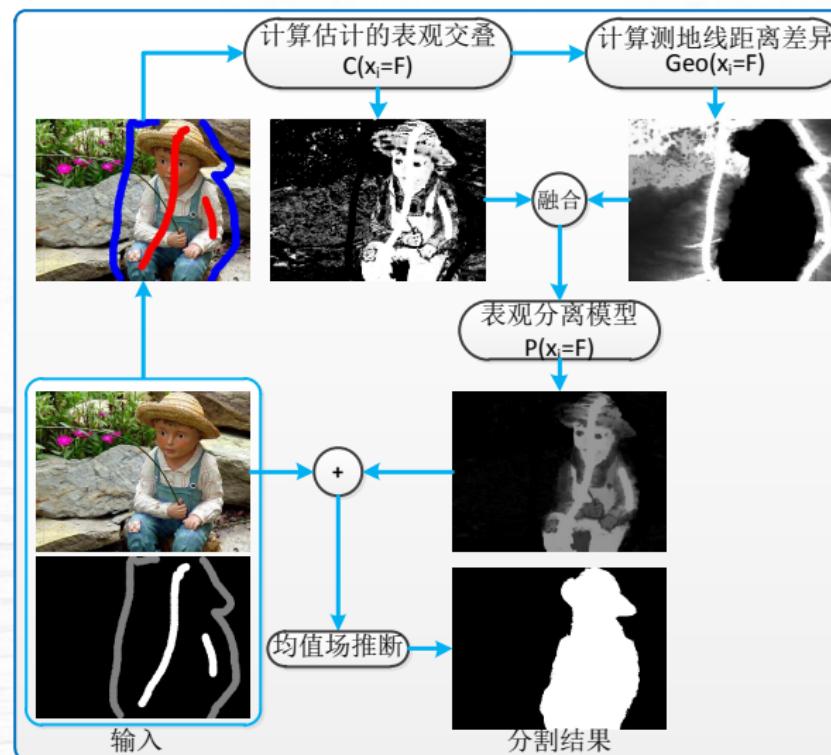
由于现实世界中的目标和场景通常是比较复杂的，将 θ_d 、 θ_{I1} 和 θ_{I2} 设置为固定值显然无法反映出图像本身的特性，而且可能会过于依赖于特定的图像数据集。因此，本章并不像DenseCut的工作那样根据数据集的统计特性设置为固定值，而是根据每个图像数据按照如下方式自动微调： $\theta_d = 0.093\sigma$, $\theta_{I1} = 1.8\sigma$, $\theta_{I2} = 1.5\sigma \times (1.1 - \kappa)$ ，其中 σ 为待分割图像平均颜色差异 ΔI_{ij}^2 的平方根。

HUNAN NORMAL UNIVERSITY

◀ back

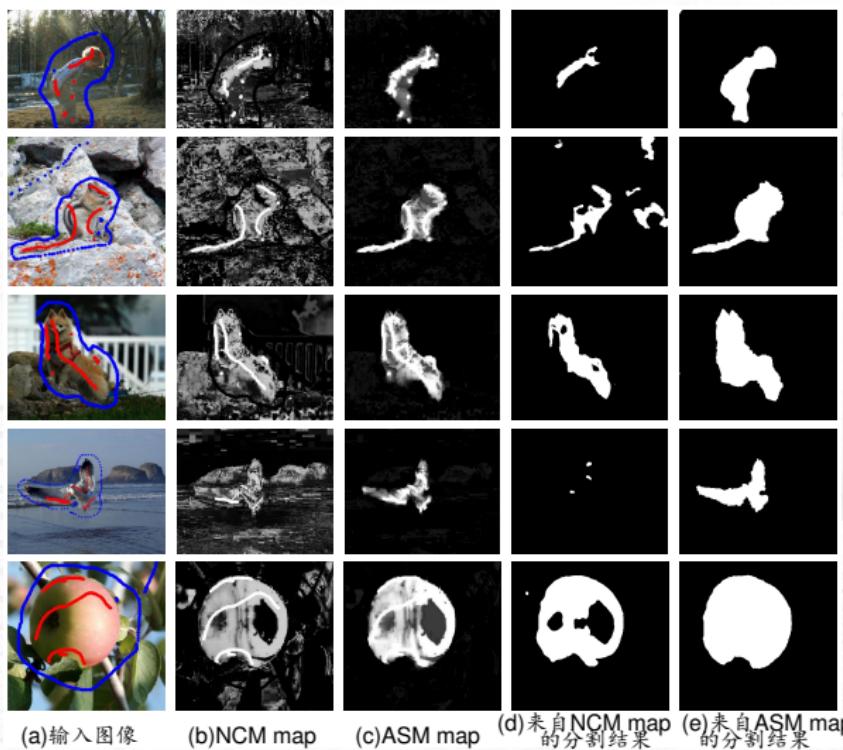


算法的主要框架和实现

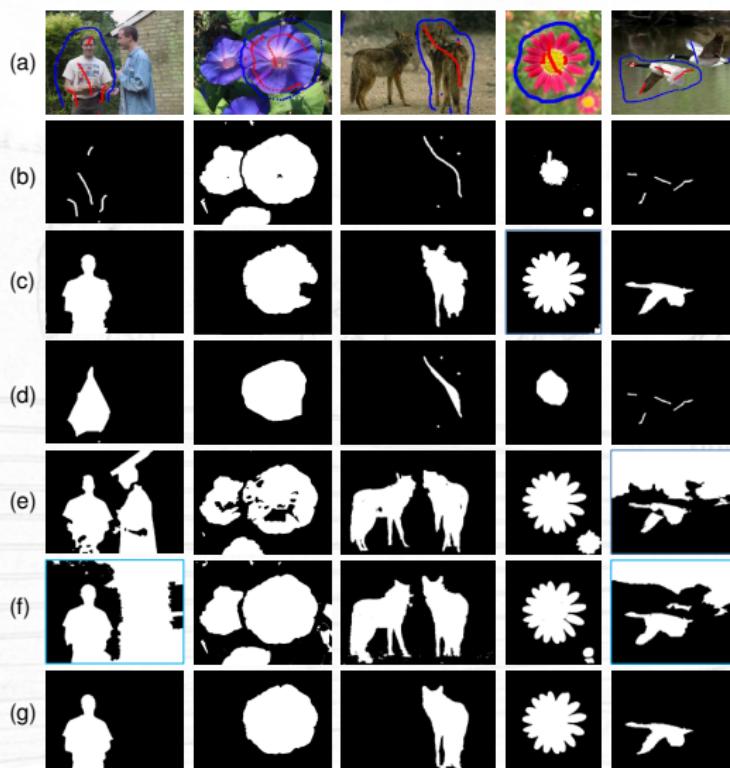


◀ back

表观分离模型 VS. 朴素颜色模型



减少用户涂鸦交互工作量



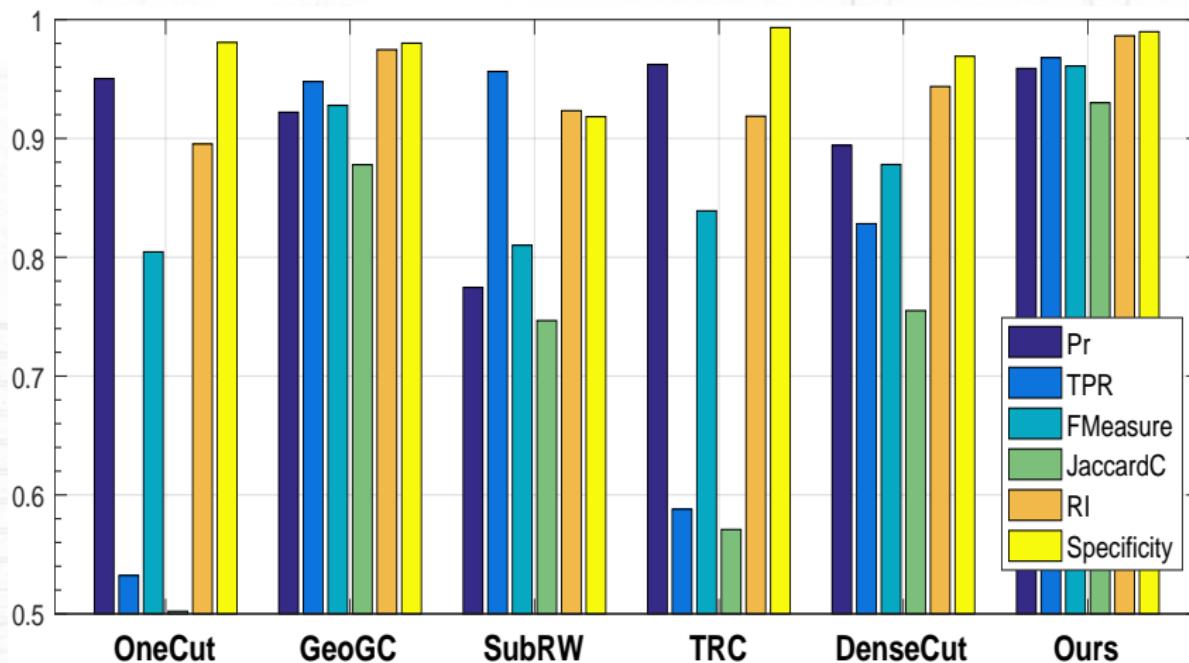
实例分割示例。

(a) 行为标注了前景和背景涂鸦的输入图像；
 (b)-(g) 行分别为OneCut、GeoGC、TRC、DenseCut、SubRW和本算法的分割结果

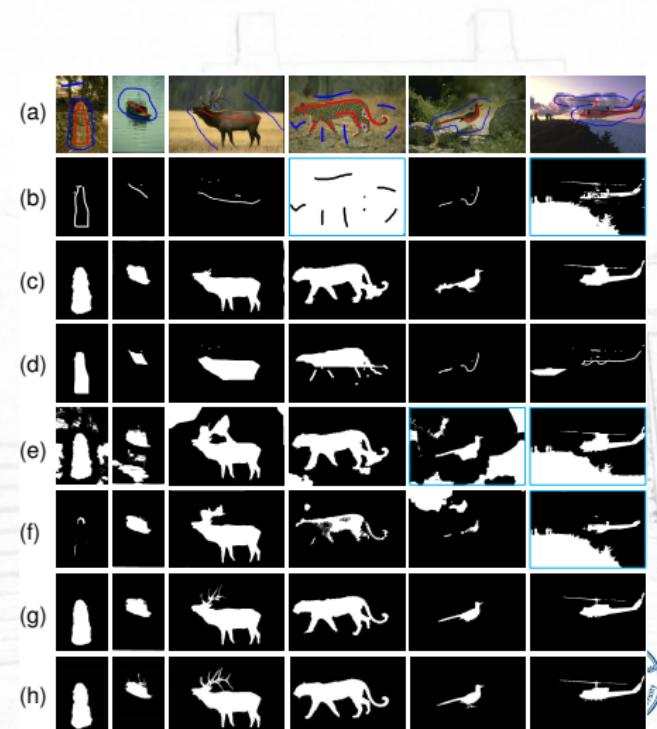
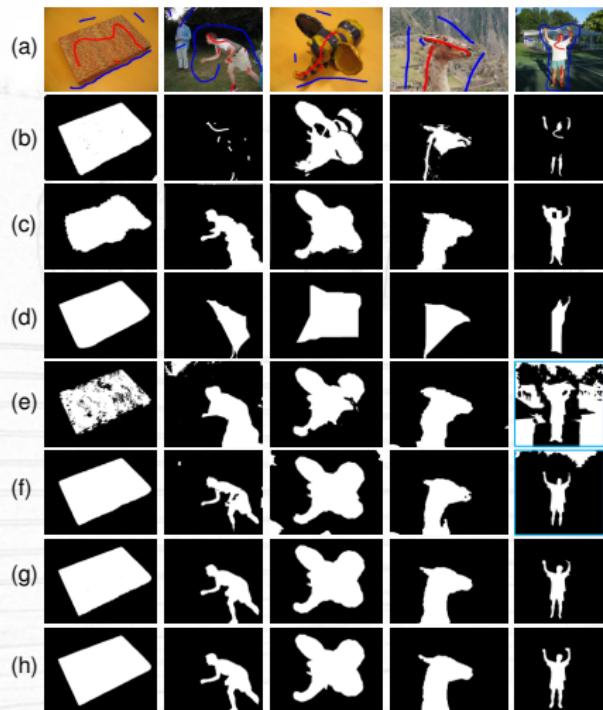


◀ back

定性和定量比较



定性和定量比较



算法时间分析

Table 1: TRC、本章提出的方法（Ours）、DenseCut、OneCut、GeoGC以及SubRW等方法在三个测试数据集上的计算时间比较

		TRC	Ours	DenseCut	OneCut	GeoGC	SubRW
MSRA数据集	平均	10.243	0.525	0.345	0.719	0.257	2.145
	最大	59.440	1.848	0.511	5.432	0.930	12.912
	最小	0.541	0.120	0.122	0.228	0.104	0.385
GrabCut数据集	平均	37.503	1.777	0.695	1.808	0.470	3.742
	最大	182.695	4.670	0.859	8.050	0.880	5.851
	最小	2.767	0.462	0.294	0.294	0.148	1.471
BSD500数据集	平均	17.619	0.720	0.438	1.715	0.351	3.342
	最大	288.037	0.908	0.542	5.065	0.502	9.448
	最小	7.870	0.431	0.371	0.443	0.195	1.543

◀ back



本章小节

本章提出了用于交互式图像分割的表观分离模型，并将其应用到稠密条件随机场框架中，所得到的自适应表观分离交互式分割算法获得了较精准的分割结果。本章所提出的算法具有好的逐个图像的自适应性，不仅降低了对种子点位置的特定要求、减轻了用户交互的工作量，而且不需要复杂的参数学习过程、摆脱了对特定数据集的依赖。基于稠密条件随机场框架，通过使用本章提出的表观分离模型，仅经过一次均值场推断即可获得精确的分割结果，因此，与高效的OneCut算法一样，本章所提出的方法是诸如GrabCut、TRC等迭代优化技术的高效替代方案。与OneCut中所使用的表观交叠惩罚相比，本章所提出的算法几乎没有产生大量孤立点的倾向。而本算法的主要不足之处在于，对于某些颇为复杂的图像，要获得精准的分割结果可能需要基于前面的分割结果继续加入交互才能实现。

◀ back



1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢

- 引言
- 相关工作
- 测地表观交叠图割算法
- 实验与讨论
- 本章小结



引言

在本文第一章中已经介绍了图像分割问题的研究背景与重要意义，还指出了交互式图像分割相对于全自动分割的优势所在，在此不再赘述。

在图像处理领域，Boykov等人利用最小割/最大流（min-cut/max-flow）图优化算法，在2001年提出了一种在N-维图像中优化边界能量和区域能量的技术[Boykov and Jolly, 2001]。此外，他们还提出了交互式图像分割方法的图割（graph-cut）框架，且获得了良好的分割结果，这在很大程度上激励着有关计算机视觉和图像处理中基于图割的研究和应用的向前发展。

◀ back



相关研究工作简介

传统的有关图割的交互式分割方法：

- GeoGC算法[Price et al., 2010]
- GrabCut算法[Rother et al., 2004]
- OneCut算法[Tang et al., 2013]
- TRC算法[Gorelick et al., 2017]
- RGBD图像线索选择图割算法[Feng et al., 2016]

与深度学习相结合的、涉及图割的方法：

- 联合卷积神经网络和上下文迁移的方法[Wang et al., 2018]
- 其他将深度学习与图割进行结合的尝试：[Wu et al., 2018, Mahapatra, 2017, Ullah et al., 2018]等

◀ back



动机

TRC算法效率并不高，而且对于那种本身存在孔洞的或者具有非凸形状的目标，它的分割结果也不准确。此外，它需要多次迭代优化才能收敛到能量最小，特别是对于那种纹理复杂的图像，所需的迭代次数可能会更多。GrabCut算法和TRC算法都属于迭代优化算法，而本章试图寻找一种无需迭代的高效分割方案，这类似于OneCut算法。但是，以画线涂鸦方式进行交互的OneCut算法存在以下几点不足：

- (1) 对用户标注的种子点位置过于敏感；
- (2) 存在孤立点问题；
- (3) 仅根据用户提供的种子点交互信息很难直接决定究竟应该分配多大的惩罚给表观交叠项。

GeoGC算法对于只有少量种子点的情况以及目标边界模糊的情况，它无法完全准确地分割出目标，这无疑会加重用户的交互工作量。



◀ back

算法设计思路 |

本章提出的测地表观交叠算法将测地线距离约束、表观交叠信息合理融入到了图割优化框架中：

$$E(L) = \lambda \cdot \sum_{x_i \in I} R_i(L_i) + \sum_{(x_i, x_j) \in N} B_{i,j}(L_i, L_j). \quad (4.1)$$

为了分别权衡每一个子项的相对重要性，本章将4.1式中的区域项表示为带权重系数的各子项的和：

$$R_l(x_i) = s_l(x_i) + \alpha_1 M_l(x_i) + \alpha_2 G_l(x_i), \quad (4.2)$$

其中， x_i 表示图像中的一个像素； α_1 和 α_2 分别用于指定子项 $M_l(x_i)$ 和子项 $G_l(x_i)$ 的重要性；4.2式中每一个子项的含义与GeoGC算法中的表示是一致的： $s_l(x_i)$ 表示来自用户涂鸦的硬性约束项、 $M_l(x_i)$ 表



算法设计思路 II

示来自全局颜色特征的颜色模型项、 $G_l(x_i)$ 表示基于测地线距离的测地模型项[Price et al., 2010]。

$$s_l(x_i) = \begin{cases} \infty & \text{if } x_i \in \Omega_l \\ 0 & \text{otherwise} \end{cases}, \quad (4.3)$$

$$M_l(x_i) = P_l(C(x_i)), \quad (4.4)$$

$$G_l(x_i) = \frac{D_l(x_i)}{D_F(x_i) + D_B(x_i)}, \quad (4.5)$$

其中， Ω_l 是由用户通过涂鸦方式指定的种子点集合，标签 $l \in \{F, B\}$ ； $C(x_i) = c$ 指的是在像素点 x_i 处的颜色强度值； $P_l(C(x_i))$ 为



└ 使用测地表观交叠图割的交互式图像分割算法

└ 测地表观交叠图割算法

算法设计思路 III

在像素点 x_i 处标签为 l 的颜色模型值，可通过归一化 Ω_l 的颜色概率密度函数（PDF） $Pr(c|l)$ 计算得到，即：

$$P_l(c) = \frac{Pr(c|l)}{Pr(c|F) + Pr(c|B)}, \quad (4.6)$$

颜色概率密度函数可通过快速高斯变换（FGT）[Yang et al., 2003]高效求得。 $D_l(x_i)$ 表示的是测地线距离图，它定义为从像素 x_i 到种子点 s （前景种子点： F 或背景种子点： B ）的最短测地线距离 $d_l(s, x_i)$ ，表达式形如：

$$D_l(x_i) = \min_{s \in \Omega_l} d_l(s, x_i), \quad (4.7)$$

$$d_l(x_i, x_j) = \min_{\Re_{x_i, x_j}} \int_0^1 \left| W_l(\Re_{x_i, x_j}(p)) \cdot \dot{\Re}_{x_i, x_j}(p) \right| dp, \quad (4.8)$$



└ 使用测地表观交叠图割的交互式图像分割算法

└ 测地表观交叠图割算法

算法设计思路 IV

其中 \mathfrak{M}_{x_i, x_j} 表示的是在加权图 $W_l(x_i)$ 上的从 x_i 到 x_j 的路径，参数为 $p \in [0, 1]$ ，表示该点是否在路径上。本质上，给定标签 l ，加权图 $W_l(x_i)$ 其实就是颜色模型 $P_l(c)$ 的梯度图，也就是说， $W_l(x_i) = \nabla P_l(c)$ 。

本章所提出的算法中的边界项是通过组合使用 OneCut 算法中的表观交叠惩罚 [Tang et al., 2013] 以及常用的平滑项，即：

$$B(x_i, x_j) = -\beta \left\| \theta^L - \theta^{\bar{L}} \right\|_{L_1} + |\partial S|, \quad (4.9)$$

其中 $L \subset \Omega$ 表示的是一个分割，且 $\bar{L} = \Omega \setminus L$ ， θ^L 和 $\theta^{\bar{L}}$ 是分别通过非归一化的前景和背景颜色直方图所表示的表观模型， $\|\cdot\|_{L_1}$ 表示的是计算 L_1 范数， β 为近似表观交叠的惩罚因子；等式 4.9 中的第二项是常用的颜色对比度敏感的平滑项，即：

$$|\partial S| = \sum w_{x_i, x_j} |l_{x_i} - l_{x_j}|, \quad (4.10)$$



算法设计思路 V

$$w_{x_i, x_j} = \frac{1}{d} \exp \frac{-\Delta I^2}{2\sigma^2}, \quad (4.11)$$

其中， l_{x_i} 是为像素 x_i 分配的标签， ΔI 表示像素 x_i 和像素 x_j 之间的颜色差异， σ^2 表示整个图像上的平均 ΔI^2 ， d 表示像素 x_i 与像素 x_j 之间的距离。

根据GeoGC、OneCut以及文献[Kolmogorov and Zabih, 2004]等的工作，本章提出的算法的目标能量函数是次模的，因此它可利用图割框架高效实现能量最小化。

◀ back



└ 使用测地表观交叠图割的交互式图像分割算法

└ 测地表观交叠图割算法

相关调参工作 |

表观交叠的惩罚因子 β :

根据所估计的表观模型的模棱两可程度相应地调整 β 值。

引入一个误差项 ρ :

$$\rho = \frac{1}{2} \left[\frac{\sum_{x \in F} P_B(C(x))}{|\Omega_F|} + \frac{\sum_{x \in B} P_F(C(x))}{|\Omega_B|} \right]. \quad (4.12)$$

当该误差项 $\rho \approx 0$ 时，表示建立的颜色模型是比较可靠的，那么与之对应的表观交叠就可能较小，此时可以分配一个大的权重给表观交叠项；当表观模型变得不那么清晰明确（即颜色表观模型变得非常凌乱、表观交叠严重）时，误差项 ρ 也会随之增大，那么，给近似的表观交叠较低的权重才是更合理的方式：

$$\begin{aligned} \beta_{t+1} &= \begin{cases} 0.25 \times \beta_t & \text{if } \rho > 0.9 \\ (1 - \rho) \times \beta_t & \text{otherwise} \end{cases}, \\ \beta_0 &= 1.4. \end{aligned} \quad (4.13)$$



相关调参工作 II

颜色模型和测地线模型的约束 α :

有了反应表观的“镜子” ρ , 相关权重系数 α 的值就可以随着表观的杂乱和交叠严重程度的变化而联动调整。

$$\kappa = \begin{cases} 1 - 2\rho & \text{if } \rho < 0.5 \\ 0 & \text{otherwise} \end{cases}, \quad (4.14)$$

$$\alpha_1 = 80 \times (1 - \kappa), \quad \alpha_2 = 800 \times \kappa. \quad (4.15)$$

此处的 κ 可看作是所估计的表观模型的置信度。

颜色仓的个数:

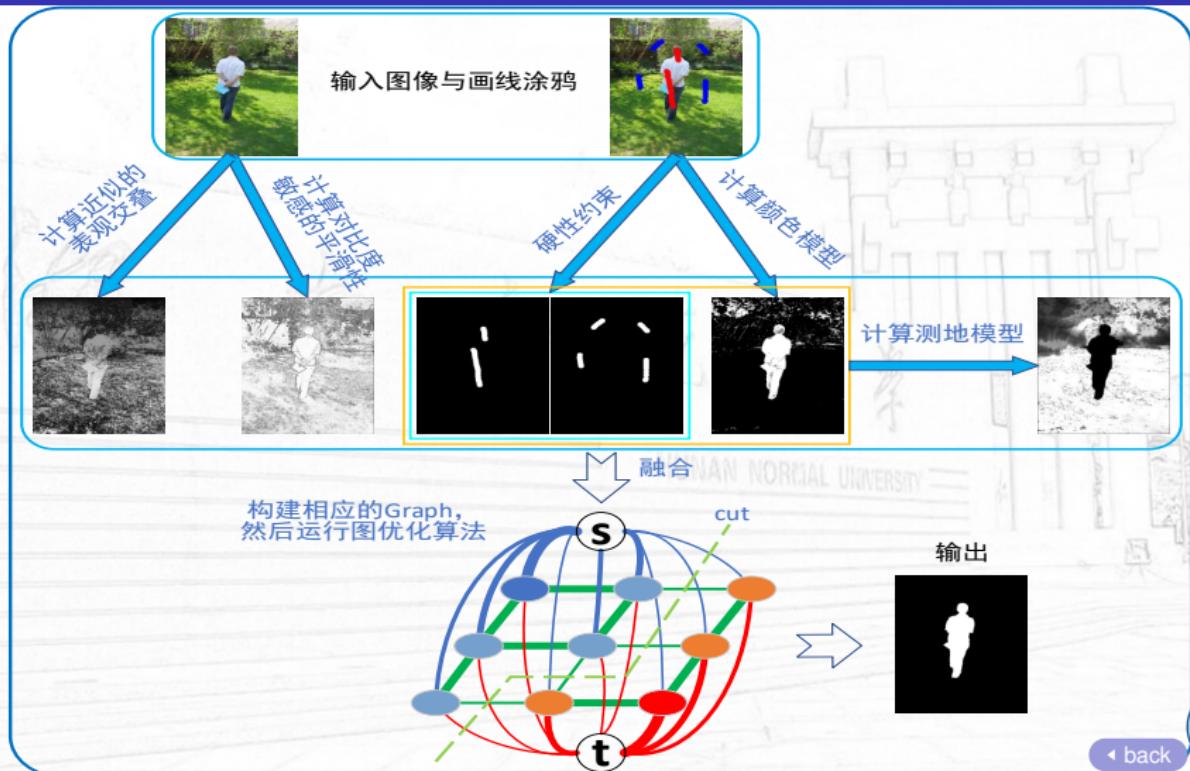
对于低分辨率（分辨率在400*300以下）的待分割图像设置为128；中等大小分辨率（分辨率在400*300至640*480之间）的图像（也是本章中用于实验的大多图像的大小）设置为64；高分辨率（分辨率在640*480以上）的图像则设置为16。



使用测地表观交叠图割的交互式图像分割算法

测地表观交叠图割算法

算法流程图



与ONECUT算法和GEOGC算法的关系 |

本章所提出的算法重点关注的是解决OneCut算法易出现孤立点的问题以及降低其用户交互工作量。OneCut算法实质上考虑的是用颜色直方图反映前景和背景表观的差异，然而颜色直方图反映的是整幅图像的全局特征，它是图像中像素的各种颜色在整个图像中所占的比例的表示，并不关心每种颜色所处的空间位置信息，所以它是没办法准确全面描述图像中的具体目标的。这也就是OneCut算法易出现孤立点问题以及加入后续的用户交互可能会破坏已有的分割结果的根源所在。基于此，本算法与OneCut算法的最大不同之处在于，通过加入反映了空间位置信息的测地线距离，从根源上对OneCut算法加以改进。其次，对表观交叠惩罚权重的设置进行了更合理的考虑。

◀ back



与ONECUT算法和GEOGC算法的关系 II

GeoGC算法对原有的测地线分割算法进行了改进，降低了算法对种子点位置的敏感度；它也对解决传统图割算法存在的短切问题做了尝试。因此，GeoGC算法具有提高分割准确度、降低用户交互工作量的优点。虽然本章提出的算法与GeoGC算法均在图割框架中加入了测地线距离信息，但本算法还加入了全局颜色特征、对区域项中的子项权重做了更细致的调整，此外，本算法采用了稍复杂的但更能反映出目标边界特征的边界项。

综上所述，本章所提出的算法结合了OneCut算法和GeoGC算法各自的优势，同时避免了它们各自存在的缺陷，达到了提高算法精度、降低用户交互工作量的目的。由于还做了相关参数的合理调整，本算法在效率上与这两个算法基本相当甚至更高效。

◀ back



人工合成彩色纹理图上的实验

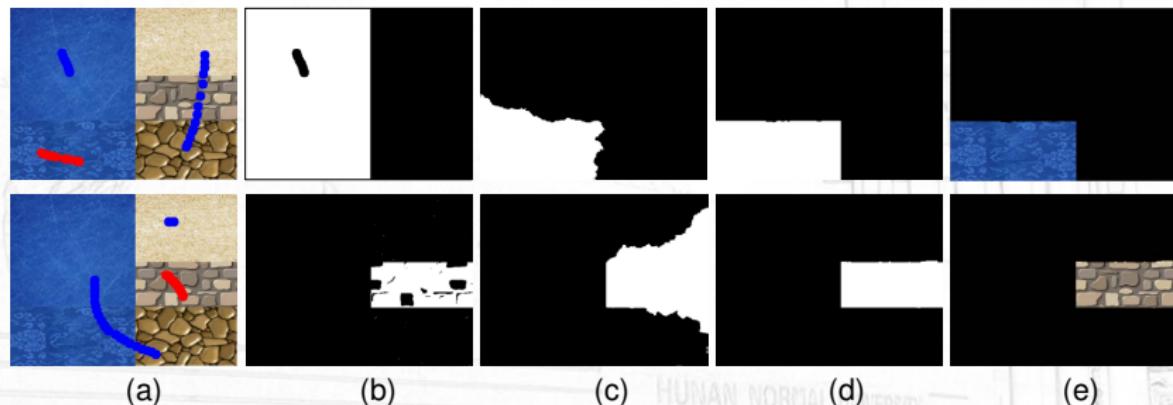
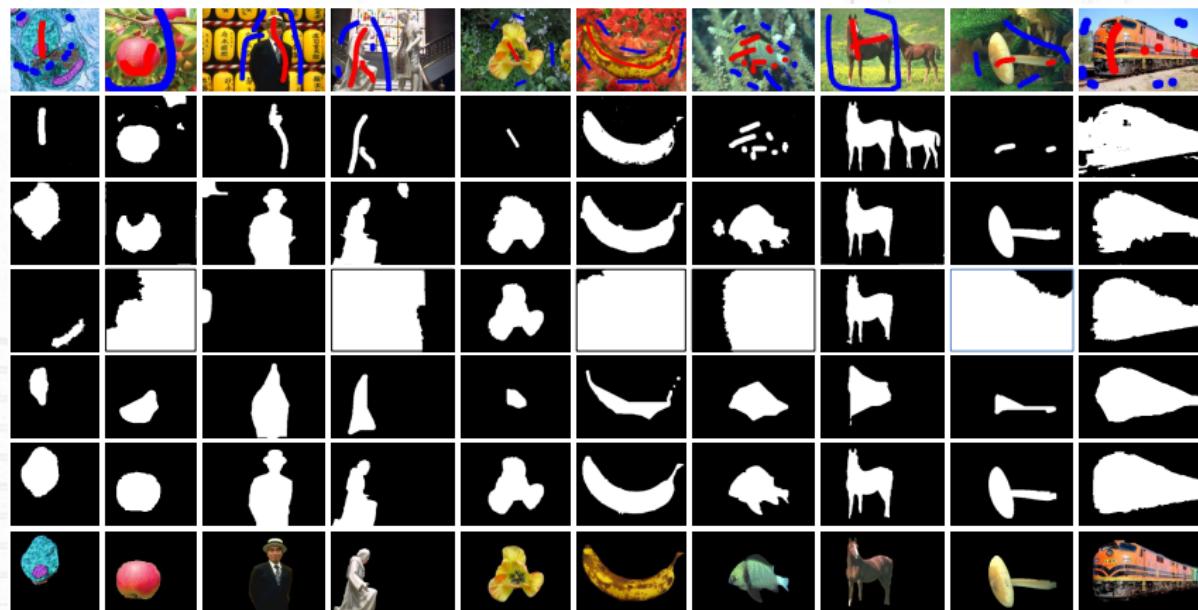


Figure 4.1: 本章所提算法、OneCut和GeoGC方法对特定设计的人工合成彩色纹理图像进行分割的比较，上下两行分别表示了分割任务1和任务2的比较情况。**(a)**列为由5块不同类型的彩色纹理合成的具有挑战性的人工图像。此处，红色涂鸦标示为前景，而用蓝色标记表示图像的背景部分；**(b)**至**(d)**列分别展示了在提供相同的种子点后由OneCut、GeoGC和本方法分割出的前景目标的掩膜；**(e)**列显示了用本方法得到的**(d)**中的掩膜遮罩在原纹理图上的结果。

◀ back



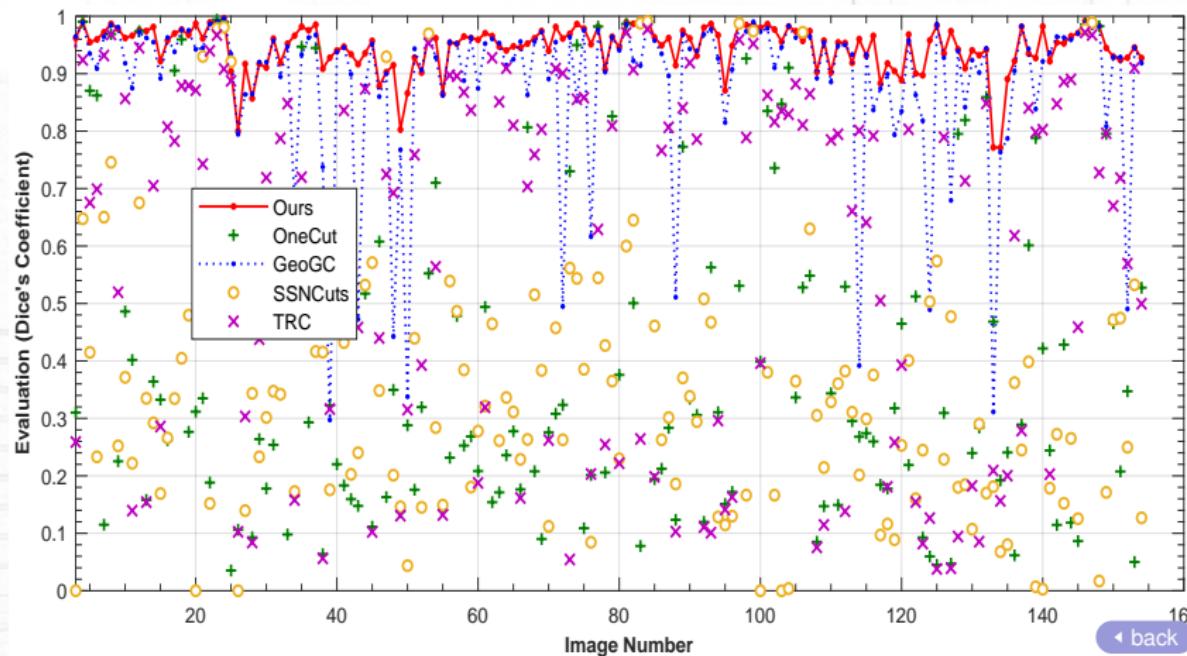
分割结果的定性和定量比较 |



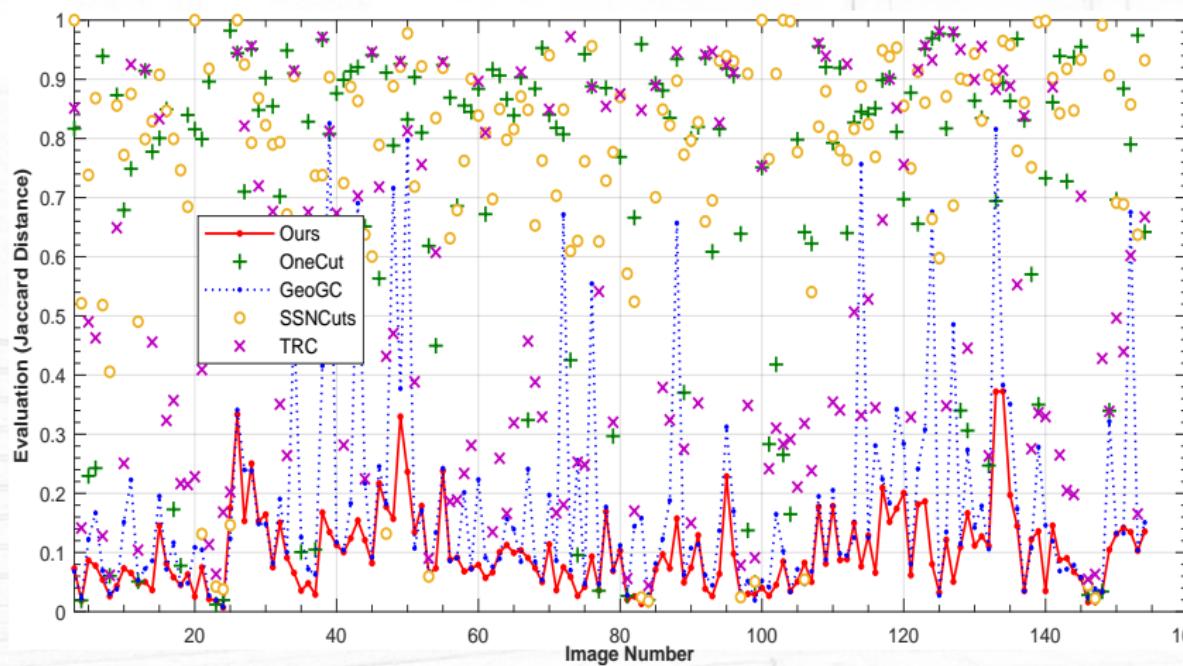
◀ back



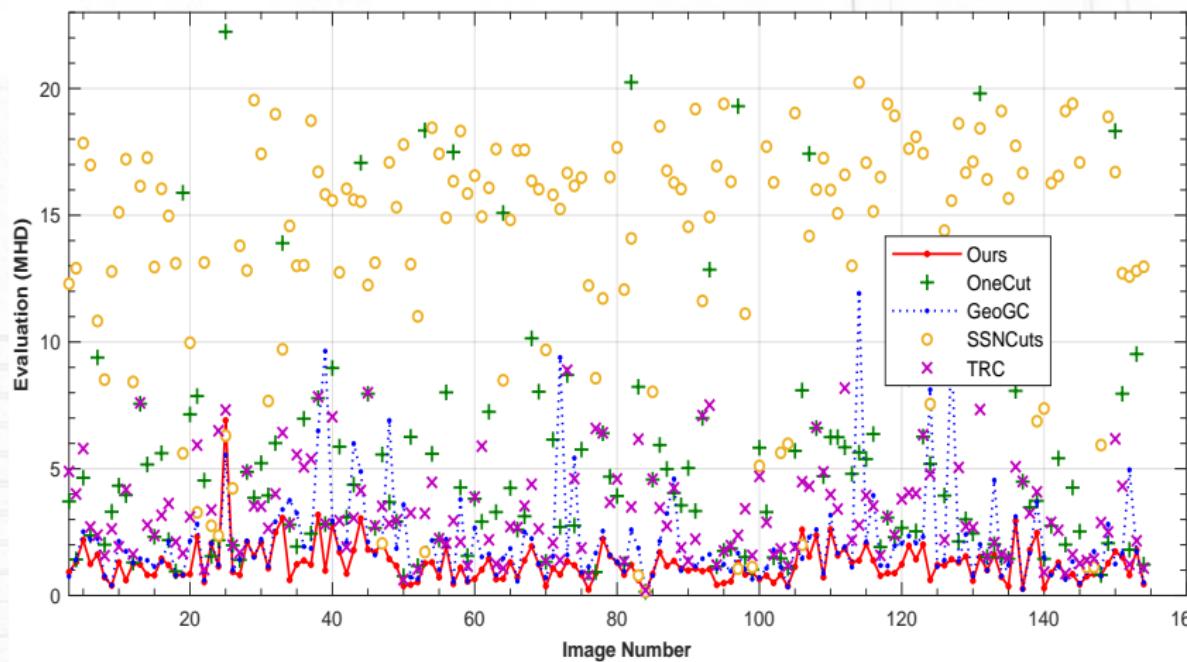
分割结果的定性和定量比较 II



分割结果的定性和定量比较 III



分割结果的定性和定量比较 IV



分割结果的定性和定量比较 V

各个算法的平均错误率(MER)

方法	最小MER	最大MER	平均MER
Ours	0.08%	11.93%	1.57%
OneCut	0.13%	84.89%	15.81%
GeoGC	0.11%	32.19%	3.70%
SSNCuts	0.11%	91.05%	58.54%
TRC	0.28%	36.81%	8.33%

各个算法的分割运行时间

方法	最短时间	最长时间	平均时间
Ours	0.141	2.032	0.587
OneCut	0.233	16.019	1.914
GeoGC	0.104	1.671	0.384
SSNCuts	0.169	182.060	15.819
TRC	0.477	130.345	19.134

◀ back



本章小结

本章提出了一个使用测地表观交叠图割的交互式图像分割算法，该算法融合了测地距离信息、表观交叠信息和边信息到图割优化框架中。本算法并不是对GeoGC和OneCut方法的简单组合，而是试图充分利用测地线距离知识和丰富的颜色信息，并且基于表观交叠所反映的不明确性来权衡这两种信息的重要程度，其中的表观交叠是根据来自用户的涂鸦交互信息以及图像本身的特征推断出来的。在大多分割任务中，本算法比GeoGC和OneCut算法表现更好。实验也验证了，本算法相对于OneCut、SSNCuts以及TRC算法，在分割速度和精度上，实现了较大幅度的提升；与GeoGC相比，本算法不但以与其相当的效率取得了更精准的分割结果，而且进一步降低了用户的交互工作量。本算法最大的不足之处在于，要获得对细长目标的精准分割可能依赖于种子点的位置。

◀ back



1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢

- 本文研究工作总结
- 研究展望



本文研究工作总结

本文以交互式的分割所追求的目标作为出发点和落脚点，力图做到以较小的用户交互工作量实现尽可能高效、精准的分割。本文提出的自适应表观分离算法，在分割过程中融入了对图像内容高层信息的理解，通过交互式分割获得更精准的分割结果，大大降低了图像分析与理解的难度。此外，本文的方法充分利用了现有的最优化算法、数学工具并根据具体的应用条件做了改进，促进了图像分割研究的新发展。本文主要的研究工作与成果总结如下：

- ① 提出了一种实用的用于交互式图像分割的表观分离模型。
- ② 针对DenseCut算法存在的不足，提出了基于Dense CRF框架的自适应表观分离算法。
- ③ 针对OneCut算法分割结果倾向于产生大量孤立点，提出了一个使用测地表观交叠切割的交互式图像分割算法。



扩展和改进计划

随着科技的发展和进步，图像数据包含的内容愈加丰富，图片的尺寸也越来越大，图像分割的实际应用场景也越来越丰富，这些给图像分割技术的发展提出了新的要求。针对本文有限的研究结果及其中存在的一些不足，本人计划从以下几个方面继续扩展和改进现有工作：

- ① 探索在用户交互过程中，为获得精准分割进行种子点选取的规律。这样，用户只要遵循该规律进行交互，即可达到进一步降低用户交互工作量的目的。
- ② 对二元分割工作进行扩展，扩展到多标签的多类分割任务中。
- ③ 探索软边界技术，实现对细长目标的分割、进一步降低对种子点位置的敏感程度。
- ④ 将交互式的分割技术与深度学习进行结合，提高自动化水平的同时保证分割的完整度与准确性。



◀ back

└ 主要参考资料

1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢



部分参考文献列表 |

-  Boykov, Y. Y. and Jolly, M. P. (2001).
Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images.
In *Proceedings Eighth IEEE International Conference on Computer Vision, ICCV 2001*, volume 1, pages 105–112
vol.1, Vancouver, BC, Canada.
-  Cheng, M., Prisacariu, V. A., Zheng, S., Torr, P. H. S., and Rother, C. (2015).
Densecut: Densely connected crfs for realtime grabcut.
Comput. Graph. Forum, 34(7):193–201.
-  Dong, X., Shen, J., Shao, L., and Gool, L. J. V. (2016).
Sub-markov random walk for image segmentation.
IEEE Trans. Image Processing, 25(2):516–527.
-  Feng, J., Price, B. L., Cohen, S., and Chang, S. (2016).
Interactive segmentation on RGBD images via cue selection.
In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pages 156–164, Las Vegas,
NV, USA. IEEE Computer Society.
-  Gorelick, L., Veksler, O., Boykov, Y., and Nieuwenhuis, C. (2017).
Convexity shape prior for binary segmentation.
IEEE Trans. Pattern Anal. Mach. Intell., 39(2):258–271.
-  Grady, L. (2006).
Random walks for image segmentation.
IEEE Trans. Pattern Anal. Mach. Intell., 28(11):1768–1783.



部分参考文献列表 II



Kass, M., Witkin, A., and Terzopoulos, D. (1988).

Snakes: Active contour models.

International Journal of Computer Vision, 1(4):321–331.



Kolmogorov, V. and Zabih, R. (2004).

What energy functions can be minimized via graph cuts?

IEEE Trans. Pattern Anal. Mach. Intell., 26(2):147–159.



Krähenbühl, P. and Koltun, V. (2013).

Parameter learning and convergent inference for dense random fields.

In *Proceedings of the 30th International Conference on Machine Learning, ICML 2013*, volume 28 of *JMLR Workshop and Conference Proceedings*, pages 513–521, Atlanta, GA, USA. JMLR.org.



Li, C., Xu, C., Gui, C., and Fox, M. D. (2010).

Distance regularized level set evolution and its application to image segmentation.

IEEE Trans. Image Processing, 19(12):3243–3254.



Li, Z. and Chen, J. (2015).

Superpixel segmentation using linear spectral clustering.

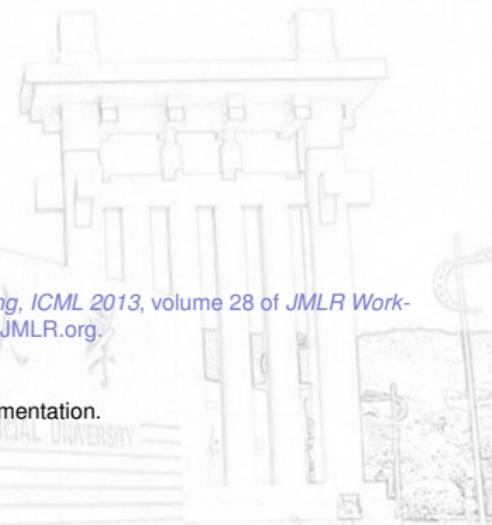
In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1356–1363, Boston, MA, USA.



Mahapatra, D. (2017).

Semi-supervised learning and graph cuts for consensus based medical image segmentation.

Pattern Recognition, 63:700–709.



部分参考文献列表 III



Ning, J., Zhang, L., Zhang, D., and Wu, C. (2010).

Interactive image segmentation by maximal similarity based region merging.
Pattern Recognition, 43(2):445–456.



Price, B. L., Morse, B. S., and Cohen, S. (2010).

Geodesic graph cut for interactive image segmentation.

In *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, pages 3161–3168, San Francisco, CA, USA. IEEE Computer Society.



Rother, C., Kolmogorov, V., and Blake, A. (2004).

"grabcut": Interactive foreground extraction using iterated graph cuts.

ACM Trans. Graph., 23(3):309–314.



Shi, J. and Malik, J. (2000).

Normalized cuts and image segmentation.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):888–905.



Tang, M., Gorelick, L., Veksler, O., and Boykov, Y. (2013).

Grabcut in one cut.

In *IEEE International Conference on Computer Vision, ICCV 2013*, pages 1769–1776, Sydney, Australia. IEEE Computer Society.



部分参考文献列表 IV



Tao, W., Jin, H., and Zhang, Y. (2007).

Color image segmentation based on mean shift and normalized cuts.

IEEE Transactions on Systems Man and Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man and Cybernetics Society, 37(5):1382–1389.



Ullah, M., Iltaf, A., Hou, Q., Ali, F., and Liu, C. (2018).

A foreground extraction approach using convolutional neural network with graph cut.

In *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, pages 40–44.



Wang, Q., Gao, J., and Yuan, Y. (2018).

A joint convolutional neural networks and context transfer for street scenes labeling.

IEEE Trans. Intelligent Transportation Systems, 19(5):1457–1470.



Wu, X., Zhong, Z., Buatti, J. M., and Bai, J. (2018).

Multi-scale segmentation using deep graph cuts: Robust lung tumor delineation in MVCBCT.

In *15th IEEE International Symposium on Biomedical Imaging, ISBI 2018, Washington, DC, USA, April 4-7, 2018*, pages 514–518. IEEE.



Yang, C., Duraiswami, R., Gumerov, N. A., and Davis, L. S. (2003).

Improved fast gauss transform and efficient kernel density estimation.

In *9th IEEE International Conference on Computer Vision (ICCV 2003)*, pages 464–471, Nice, France. IEEE Computer Society.

◀ back



1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢



阶段性成果

已发表或完成的学术论文：

- [1] Peng Zi-Li and Li Qiao-Liang, Adaptive Appearance Separation for Interactive Image Segmentation Based on Dense CRF. IET Image Processing, 2018. (SCI, 录用.)
- [2] Peng Zi-Li, Qu Shao-Jun and Li Qiao-Liang. Interactive image segmentation using geodesic appearance overlap graph cut. (Submitted)
- [3] Qu Shao-Jun, Tan Huang, Peng Zi-Li and Li Qiao-Liang. Image Segmentation Based on Hochbaum's Pseudo-flow Algorithm and Improved OneCut. (Submitted)

参与的科研项目：

国家自然科学基金面上项目（NSFC No. 11471002），图像分割中若干图论问题的研究，2015-1-1~2018-12-31.



◀ back

1 绪论

2 用于交互式图像分割的基础框架

3 基于DenseCRF的自适应表观分离

4 测地表观交叠图割

5 结语

6 主要参考资料

7 阶段性成果

8 致谢



致谢

本论文是在我的导师李乔良教授的深切关怀和悉心指导下完成的。衷心感谢我的导师李乔良教授对我在科研上精心指导；在职业选择上细心规划；更在为人处世上为我树立榜样。李老师对我方方面面的指导，让我受益良多，在此谨向李老师致以诚挚的谢意和崇高的敬意。

感谢实验室中参与课题讨论的文有为老师、陈明老师，以及邱相炎、李金龙等同学。尤其感谢在实验过程中为我解惑的瞿绍军博士和谭煌博士。感谢为我一遍又一遍修改论文的妻子，论文中的语法错误与笔误的修正离不开她的支持与帮助。另外，还要感谢我的父母和其他帮助过我的老师和同学。

最后，向百忙之中参与审阅、评议本论文的各位老师，向参与本人论文答辩的各位老师表示由衷的感谢！

◀ back



谢谢大家！

欢迎批评指正！

