

一种基于双层框架的仿射类图像抠像方法

姚桂林^{1,2} 赵志杰^{1,2} 苏晓东^{1,2} 辛海涛^{1,2} 胡文^{1,2} 秦相林¹

摘要 仿射类抠像方法主要分为 KNN (K-nearest neighbor) 类和 Matting Laplacian 类方法, 本文结合这 2 种方法的优点提出了一种基于仿射类的多层次抠像方法. 其中, 第一层为绝对像素的划分层次或预处理层次, 采用了基于 KNN 类简单权重与相对远距离的搜索方法, 并结合初始 Trimap 未知区域大小无关的方式; 第二层为混合像素的计算层次或最终抠像层次, 充分利用了第一层计算获得的剩余混合像素的宽度, 自适应地调整 Matting Laplacian 中的颜色线性模型所构成颜色近邻的核宽度. 每个层次均按图像的全局颜色重叠程度相应调整合理的搜索范围. 本文的实验具备以下特点: 1) 预处理层次之后采用了若干典型的后续抠像方法, 以展现本文方法相比于其他预处理方法对后续抠像操作步骤的优越性和兼容性; 2) 最终抠像层次引入了若干其他抠像方法, 以验证本文抠像方法的优越性. 实验表明, 相比于其他单层次的仿射类方法, 无论对于计算绝对像素还是混合像素, 本文方法都可以大幅提升计算结果的准确率.

关键词 图像抠像, 仿射类抠像, Matting Laplacian, KNN 搜索, 颜色线性模型

引用格式 姚桂林, 赵志杰, 苏晓东, 辛海涛, 胡文, 秦相林. 一种基于双层框架的仿射类图像抠像方法. 自动化学报, 2021, 47(1): 209–223

DOI 10.16383/j.aas.c180356



开放科学 (资源服务) 标识码 (OSID):

A Hierarchical Framework on Affinity Based Image Matting

YAO Gui-Lin^{1,2} ZHAO Zhi-Jie^{1,2} SU Xiao-Dong^{1,2} XIN Hai-Tao^{1,2} HU Wen^{1,2} QIN Xiang-Lin¹

Abstract Affinity based image matting methods can be categorized into KNN (K-nearest neighbor) based matting and matting Laplacian based matting, and this paper raises a hierarchical framework on affinity based matting according to the analyses of the advantages of these two popular affinity based image matting methods. The first opaque pixel classification layer, also named as pre-processing layer, employs a relatively far searching fashion based on simple weights in KNN and is spatial irrelevant to the unknown region of the initial Trimap. The second mixed pixel computation layer, also named as final matting layer, adaptively adjusts the kernel size of the color line model in matting Laplacian according to the remaining size of the unknown region. Each layer adjusts proper searching range adaptively according to the overlapping degree between global foreground and background colors. The following distinctions are provided in the experiments. First, several representative matting algorithms are processed after the first layer to show the superiority and compatibility of our pre-processing method over other pre-processing methods. Second, several alternative matting methods are also processed after the first layer to show the superiority of our final matting method over other matting methods. Experimental results show that our approach can greatly raise the solving precisions for both opaque and mixed pixels.

Key words Image matting, affinity based matting, matting Laplacian, KNN (K-nearest neighbor) searching, color line model

Citation Yao Gui-Lin, Zhao Zhi-Jie, Su Xiao-Dong, Xin Hai-Tao, Hu Wen, Qin Xiang-Lin. A hierarchical framework on affinity based image matting. *Acta Automatica Sinica*, 2021, 47(1): 209–223

收稿日期 2018-05-28 录用日期 2019-01-02

Manuscript received May 28, 2018; accepted January 2, 2019
黑龙江省自然科学基金 (F2018021, LH2019F044), 哈尔滨商业大学校级科研项目 (18XN021, 2016TD001), 哈尔滨商业大学青年创新人才支持计划 (2019CX02), 黑龙江省哲学社会科学研究规划项目 (18GLB029) 资助

Supported by Heilongjiang Provincial Natural Science Foundation (F2018021, LH2019F044), Technological Project of Harbin University of Commerce (18XN021, 2016TD001), Young Creative Talents Support Project of Harbin University of Commerce (2019CX02), Heilongjiang Provincial Philosophy and Social Science Planning Project (18GLB029)

本文责任编辑 刘跃虎

Recommended by Associate Editor LIU Yue-Hu

1. 哈尔滨商业大学计算机与信息工程学院 哈尔滨 150028 2. 黑龙

数字图像的抠像与合成问题是在虚拟现实图像处理领域的两种经典问题, 目的是从一幅合成图像中将前景物体从背景中分离出来, 并合成到一幅新背景图像中. 对于合成问题, 对于在数字图像 I 中的某个位置 $i = (x, y)$, 给定前景图像的颜色 F_i 、背景图像颜色 B_i 以及一个透明度标量 $\alpha_i \in [0, 1]$,

江省电子商务与信息处理重点实验室 哈尔滨 150028

1. School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028 2. Heilongjiang Provincial Key Laboratory of Electronic Commerce and Information Processing, Harbin 150028

其观察到的颜色 I_i 能够用如下抠像公式表示为

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i \quad (1)$$

如果 $\alpha_i = 0$ 或 $\alpha_i = 1$, 将点 i 称作绝对前景或者绝对背景, 并统称为绝对像素. 对于其余 $0 < \alpha_i < 1$ 的点, 称为混合像素.

抠像问题为合成问题的逆问题, 同时也是欠约束问题: 给出各点的合成像素 I_i , 需要求出未知量 F_i 、 B_i 与 α_i . 很多一般背景抠像问题都需要用户提供手工输入的辅助三分图 Trimap, 如图 1(a) 和 (b) 所示, 它包括已知绝对前景区域 Ω_F 和已知绝对背景区域 Ω_B , 剩余部分为未知区域. 于是, 依据已知像素, 并根据抠像公式及某些先验条件, 可以估计未知区域内的每个像素的 $\{\mathbf{F}, \alpha\}$ 值. 基于 Trimap 的图像抠像算法通常分为如下 2 种方式: 1) 基于仿射的抠像方法^[1-7]. 该方法并不孤立地计算每个点, 而是考虑像素间的相关性, 利用间接、迭代的方式递归地求出未知区域中各个点的 α 值. 2) 基于采样的抠像方法^[8-17]. 该方法对每个未知点独立从已知区域中选取样本, 并采用逐对样本的计算方式, 而并不考虑各个像素之间的联系.

本文主要讨论仿射类方法^[18]. 虽然相比仿射类方法, 采样类方法的采样位置和采样方式变化较多, 而且获取样本方式非常直接, 然而, 仿射类方法具有以下优势. 1) 仿射类方法充分强调了像素间的相关性, 该方式也可视为由已知区域向未知区域缓慢、渐进的计算, 因此在最终 α 结果的平滑性以及带给或用户的视觉感受上要明显好于采样类方法. 2) 采用远距离搜索的仿射类方法对绝对点的计算效果较好, 而在事实上, 未知区域内的大多数点为绝对像素, 因此该类方法总体准确率上有较为明显的优势. 本文在第 1 节首先阐述了仿射类方法的 2 种分类方式以及它们优点和缺点, 在第 2 节中针对 KNN 类的特点讨论了绝对像素划分 (即预处理) 方式, 在第 3 节中针对 Matting Laplacian 类的特点讨论了剩余混合像素的计算, 并在第 4 节中讨论了这 2 类方法各自

的实验结果. 最后第 5、6 节讨论了仿射方法目前存在的问题并得出本文的结论. 本文基于仿射方法的双层次抠像框架如图 1 所示. 需要说明的是, 由于仿射类方法与采样类方法、乃至其他类如特殊类^[19]、深度学习类^[20-22] 等方法采用完全不同的计算机制, 而且本文主要研究前者, 因此无论在理论研究和实验数据上, 本文并未引入任何其他类方法, 而且也不将结果传至 α 评估系统^[23] 中, 以保持仿射类方法的整体一致性.

1 仿射类抠像方法概述

假设每个未知像素的 α 值是它的 K 邻域像素 α 值 $\alpha_1, \alpha_2, \dots, \alpha_K$ 的线性组合, 表示为

$$\alpha_i = w_1 \alpha_1 + \dots + w_j \alpha_j + \dots + w_K \alpha_K \quad (2)$$

实际求解中, 所有未知点 i 的 α_i 可通过求解如下大型稀疏线性方程得到

$$(L + \theta D) \alpha = \theta \mathbf{b} \quad (3)$$

其中, L 为长与宽均为图像像素总数的稀疏方阵, 未知点 i 所在行对应 K 邻域的系数记为 w_j , $j = 1, \dots, K$, D 为对角阵, 在已知点所对应的对角线元素的位置为 1, 未知点为 0, θ 为某个较大的数. \mathbf{b} 为长度等于像素总数的列向量, 前景 Ω_F 处位置为 1, 背景未知区域 Ω_U 位置为 0. 上式可采用共轭梯度法求解.

1.1 仿射类方法的分类

各仿射类方法、亦即稀疏矩阵 L 的构造方式之间的主要区别是每个未知点 i 的若干近邻像素位置的远近以及相应权值 w_j 的构造方式. 按是否采用抠像公式以及在抠像中的实际效果, 仿射类方法可分为不采用抠像公式的简单权重类方法以及采用抠像公式的 Matting Laplacian 类方法.

1.1.1 简单权重类方法

该类方法如图 2 左栏所示, 其特点是不采用抠

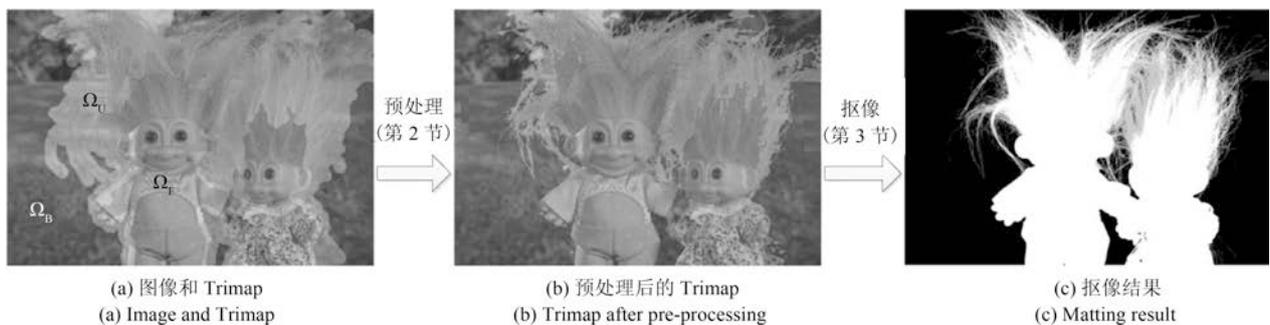


图 1 图像抠像问题的基本输入输出和本文的双层次抠像结构

Fig. 1 Input and output of image matting and the hierarchical framework of our method

像公式, 直接采用颜色差异的方法确定权重. 具体的, 当前点 i 与各近邻点 j 颜色差异的某个负函数即为各个权重, 即与 j 的颜色越接近, 权值越大, 二者的 α 越相似. 起初的 Random Walk 算法^[6] 仅采用 8 邻域处最近的搜索范围, 但无论在绝对像素还是混合像素, 其计算效果均很不理想. Nonlocal 算法^[3] 对邻域空间进行了扩展, 采用以每个未知点为中心、半径为 r 的方形搜索寻找 K 个颜色最相近的点. 为减少计算量, 采样半径 r 通常不大. KNN 算法^[4] 的搜索距离则更远, 需要寻找每个点在颜色约束和空间距离约束的特征下最近的 K 个点.

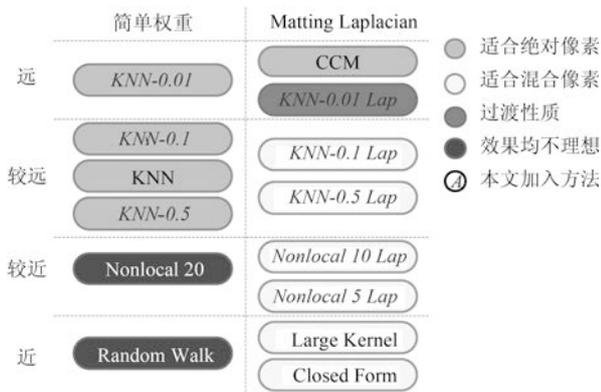


图 2 仿射类方法的 2 种分类方式对应的算法及适用范围 (其中 Lap 表明在原简单权重方法的基础上采用 Matting Laplacian)

Fig. 2 Two types of affinity based matting method and the corresponding algorithms and application scopes (where Lap denotes the application of matting Laplacian based on simple weight methods)

根据式 (2), 若 α_i 或其 K 个近邻像素的 α_j 均接近于 0 或 1, 即它们均为绝对前景或背景, 此时的权值 w_j 显然几乎无法构成影响. 因此, 在 KNN 算法中讨论的权值是否采用 $\exp(-x)$ 还是 $1-x$ 的问题则显得不大重要. 但若 α_i 和 α_j 为小数, 即它们均为混合点, 则权 w_j 非常重要. 然而, 由于简单权重模式下的 w_j 的产生并不采用抠像公式, 而真实值 α^{true} 是由抠像公式算出, 因此此时的结果肯定不理想. 由此可得出的结论是: 该类方法适合于计算绝对像素, 而不是混合像素.

另一方面, 搜索范围较近的 Random Walk 和 Nonlocal 等方法, 需要经过多次传递才能到达已知区域, 因此较易造成误差累积. 相反, 搜索范围较广的 KNN 算法可以通过较少次数的传递或直接搜索到达已知区域, 会产生较少误差累积, 更有利于绝对前景和背景的计算. 这与采样类方法中直接访问已知区域的方式较为相似.

绝对像素计算步骤也同时在文献 [24] 中进行了深入讨论. 据统计, 在评估系统^[23] 提供的所有训练

图像中, 大型号与小型号 Trimap 中的绝对点数量大大超过了混合点的数量, 由此可以看出该步骤的重要性. 进一步, 该步骤必须单独采用一个独立的、不采用抠像公式的步骤, 使其区别于普通的混合像素计算步骤, 而简单权重类方法则完全符合这种绝对像素计算准则. 目前, 绝对像素的计算的算法仅出现在采样类计算方式中, 但局部类方法^[13-16, 19] 容易遗漏较远处的样本, 误划分率偏高, 而全局类方法^[17] 易受全局前景和背景颜色重叠影响, 误划分率较高. 由此可以看出, 目前的绝对像素划分方法仍然受到全局与局部采样方式相互平衡的困扰.

1.1.2 Matting Laplacian 类方法

该类方法如图 2 右栏所示, 它假设在某“局部集合” S 内, 任何点 i 的前景和背景颜色 F_i 和 B_i 都是两个固定颜色 $\{F_1, F_2\}$ 和 $\{B_1, B_2\}$ 的线性组合, 称为颜色线性模型 (Color line model). 结合抠像公式并经推导, 结论为在某个集合 S 内的 $(i, j) \in S$, 点 i 与 j 产生的权重为

$$-\frac{1}{|S|} \left(1 + (\mathbf{x}_i - \boldsymbol{\mu}_S)^T \left(\Sigma_S + \frac{\varepsilon}{|S|} X_3 \right)^{-1} (\mathbf{x}_j - \boldsymbol{\mu}_S) \right) \quad (4)$$

其中, \mathbf{x}_i 与 \mathbf{x}_j 分别为点 i 与 j 的三维颜色向量, X_3 为 3×3 单位阵, Σ_S 、 $\boldsymbol{\mu}_S$ 分别为 S 内的 3×3 协方差矩阵与 3×1 的均值向量, ε 为一个较小的数. 未知点 i 与 j 的权重为所有包含二者的集合 S 所产生的权重之和. 显然, 由于采用了抠像公式, 该类方法非常适合于混合点的计算, 也是目前为止, 仿射类方法中对混合点最重要的计算方式.

该类方法的关键问题是相关集合 S 如何选择. Closed-Form 算法^[1] 提出将集合 S 定义为较小的局部窗口, 对于宽度较窄的前景硬边界效果较好, 然而该类小窗口无法充分涵盖较宽的混合区域或前景空洞等实例的颜色信息. Large Kernel 算法^[2] 改进了基于空间局部窗口的思想, 提出了核宽度 r 与不同未知局部区域的宽度成比例的方式. 然而事实上, 核宽度很大程度上依赖于真实混合区域的宽度信息, 而未知区域仅为用户的粗略手工输入, 尤其输入为稀疏的线条不能充分表达混合区域的实际信息, 现实中还会存在很多偏差. CCM 算法^[5] 采用了 KNN 方式中的颜色近似的方法选取集合 S , 取代了空间方式的局部窗口. 由于该方法搜索范围过大并引入大量已知样本, 在实际中, 它对绝对像素计算的效果较好, 但对于混合像素的计算也受到全局绝对样本影响, 出现过于二值化的现象.

CNN Matting 方法^[20] 提到, KNN 算法与 Closed Form 算法之间的区别并不直接, 由此采用基于深度学习方式对二者进行融合. 然而经本文

讨论至此, 它们的区别已经非常明朗, 即 KNN 类方法适用于绝对像素的划分, 而 Matting Laplacian 类 (即 Closed Form 类) 适用于混合像素的计算. 本文从仿射类方法自身入手, 据各类方法自身的优点和缺点做出相互的补充, 避免了大量训练数据以及长时间训练等问题.

图 3 显示了 5 种典型算法在 α 评价系统^[23] 的 52 幅训练图像上 (27 幅训练图像中的大、小 Trimap 且不包含 GT16), 各 α^{true} 区间的平均 MSE 比较 (不包含 Random Walk), 显然, KNN 在 0 ~ 0.25 与 0.95 ~ 1 区间较好, 印证了它非常擅长于绝对像素划分, 而较为平滑的背景使得在绝对背景划分上更加优秀. CCM 仅在 0 ~ 0.05 区间较好, 说明它受前景和背景重叠的影响非常大. Closed Form 与 Large Kernel 大致 0.25 ~ 0.95 间较好, 印证了它们擅长混合像素的计算. 同时, 近距离 Nonlocal 算法在各区间中均不理想.

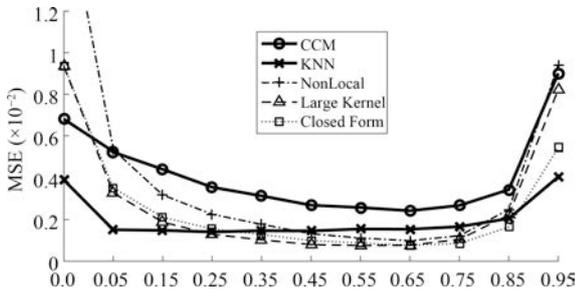


图 3 各 α^{true} 区间中 5 种传统仿射类算法 α 结果的平均 MSE 比较 (其中 x 轴坐标中的 0.0 表示 0.0 ~ 0.05 区间等)
Fig. 3 MSE comparison on five traditional affinity based matting algorithms in each α^{true} interval (where 0.0 in x -label denotes the range of 0.0 ~ 0.05, etc)

1.2 本文方法的提出

事实上, KNN 类算法中 K 近邻像素的远近的选取, 对该类方法的绝对像素划分结果会产生重要影响. 在前景与背景的整体重合度不高的前提下, 选择更远距离的搜索范围, 不仅会为当前未知点提供更多相似的已知点, 而且在 Trimap 真实样本距离未知点较远时 (如未知区域较大、或者前景存在空洞) 计算效果更好. 但如果较远处的前景与背景颜色存在较大重合, 则会产生计算错误, 此时需要采用较小的搜索范围. 因此, 所选取的近邻像素的远近也是一个重要的平衡. 然而, KNN 采用固定的远距离与近距离相结合的方式, 无法根据空间上的颜色变化情况自适应调整.

基于上述分析, 本文在第 2 节新引入了 KNN 搜索上的 3 种范围 (如图 2 左栏所示), 并依据全局重叠与局部重叠程度, 相应采用其中 2 种层次相结合的方式, 以期对大部分绝对像素进行划分, 并进

一步缩小未知区域的宽度. 同时, 由于无法保证初始 Trimap 与混合点之间的关系, 该层次采用与初始 Trimap 未知区域大小无关的方式, 这与 Large Kernel 算法不同.

如前文所述, 混合像素计算过程中的 Matting Laplacian 类方法中的核宽度、或者搜索空间的大小, 需要与真实未知区域的宽度相一致. 若搜索范围过小则无法充分涵知区域的信息, 造成 CNN 算法^[20] 提到的“过平滑”, 而搜索范围过大会导致混合像素的绝对像素化, 如 CCM. 由于在第 2 节中, 相当多的绝对像素已被划分, 且混合像素区域的大致轮廓也已确定, 因此在此基础上, 第 3 节引入了 KNN 及 Nonlocal 搜索上的 5 种范围 (如图 2 右栏所示), 并采用结合未知区域宽度的 Matting Laplacian 方法, 计算最终剩余混合像素.

2 绝对像素划分

本节讨论基于 KNN 搜索方式的绝对像素划分. 第 2.1 节提出了搜索范围因子的概念, 并定义 3 种搜索范围下的直接点差异度, 为后续章节打下基础; 第 2.2 节提出了图像重合概念, 将图像划分为低重合度图像和高重合度图像, 前者按 2.3 节计算 α , 后者按第 2.4 节计算 α . 最后按第 2.5 节的截取方式得到预处理后的 Trimap.

2.1 搜索范围因子与直接点差异度

首先, 类似于 KNN 算法, 定义像素 i 的特征矢量 $\mathbf{x}_i^s(f) = [r_i, g_i, b_i, sx_i(f), sy_i(f)]$ 为 i 的 3 个通道的颜色值, $\{r_i, g_i, b_i\}$ 为点 i 的空间距离特征, 表述为

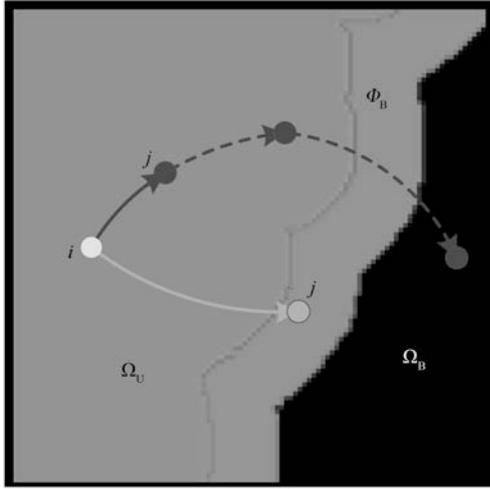
$$\begin{cases} sx_i(f) = f \cdot \frac{x_i}{(x_i^2 + y_i^2)^{\frac{1}{2}}} \\ sy_i(f) = f \cdot \frac{y_i}{(x_i^2 + y_i^2)^{\frac{1}{2}}} \end{cases} \quad (5)$$

其中, x_i 和 y_i 为 i 的空间横坐标与纵坐标, f 为控制空间距离远近的重要参数. f 越大, 采集的像素越近, 反之则越远. 在 KNN 算法中, f 的取值为 1 和 0.01 之间的融合, 其最终效果近似于 0.5, 如图 2 所示.

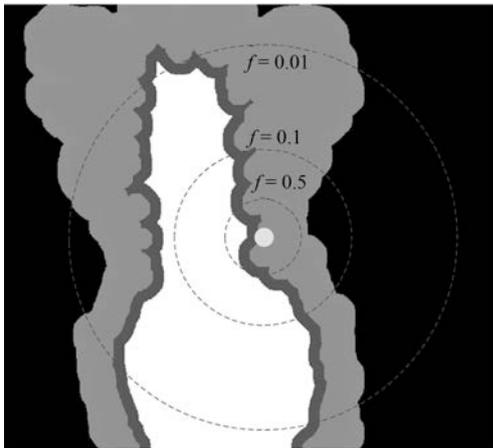
前文曾提到, 搜集像素位置的远近是 KNN 类方法的决定因素之一, 该问题非常依赖于前景与背景颜色的重合度. 然而, 传统的仿射类方法并不是直接访问已知区域的, 直接获取每个点的前景和背景信息. 因此, 本文提出一种类似于采样方式的、直接访问已知区域的方法, 用来模拟和引导每个未知点间接搜索距离的方式. 这里需要利用 FLANN (Fast library approximate nearest neighbors) 方法, 在相关特征 $\mathbf{x}_i(f)$ 下, 靠近未知区域的 5 个像素宽度的已知前景区域 Φ_F 和背景区域 Φ_B 位置 (简记为 F

与 B), 对每个像素 i 分别寻找最近的前景和背景点 j 各 $K = 10$ 个, 将它们称为直接点, 如图 4(a) 下方的 j 点所示 (以背景为例). 在 \mathbf{x}_i 为 \mathbf{x}_i^s 的前 3 个空间无关特征分量、即颜色特征下, 计算未知点 i 与上述前景与背景直接采样点对于 K 个直接点的平均颜色差异 $d_i(f, F)$ 与 $d_i(f, B)$, 之后在给定距离参数 f 之下, 计算点 i 的前景与背景颜色的差异系数:

$$DF_i(f) = \frac{\max(d_i(f, F), d_i(f, B))^{\theta_1}}{\min(d_i(f, F), d_i(f, B))} \quad (6)$$



(a) 间接搜索与直接搜索
(a) Direct searching and indirect searching



(b) 3 种 f 下的模拟搜索范围
(b) Simulated searching ranges for 3 f values

图 4 以背景为例, 未知点的各种搜索方式与搜索范围

Fig. 4 Various kinds of searching manners and searching ranges for an unknown pixel

其中, θ_1 的取值在 1 至 2 之间时, 前景与背景的区别效果最好, 这里取经验值 $\theta_1 = 1.5$. $DF_i(f)$ 越高, 表明在距离参数 f 下, 点 i 的直接前景与背景颜色差异性越大, 重合性越低, 该距离参数 f 就比较理想; 否则, 该距离参数就不理想.

对每个未知点 i , 计算近、中、远 3 个级别的差异度 $DF_i(0.5)$, $DF_i(0.1)$ 和 $DF_i(0.01)$, 并为后面的计算所采用. 这 3 个级别的近似搜索范围如图 4(b) 所示.

2.2 像素采集远近及图像的重合度判别

按文献 [25] 的理论, 局部 (或近距离) 采样方式应作为基本方式, 而全局 (或远距离) 采样方式是局部的一种补充. 而这对于仿射类方法同样有效. 考虑到各图像中前中前景与背景的重叠程度不尽相同, 采集像素远与近的范围也不相同. 首先, 在未知区域 Ω_U 内、靠近已知前景和背景区域的位置, 分别定义 1 个像素宽度为“学习区域” (简记为 F 与 B), 首先在距离参数 f 之下, 计算这 2 个区域到各自已知前景和背景的平均颜色差异 $d_L(f, F)$ 与 $d_L(f, B)$,

$$rt_1 = \max(d_L(0.5, F) - d_L(0.01, F), d_L(0.5, B) - d_L(0.01, B)) \quad (7)$$

rt_1 越小表明近处像素点已经足够判别绝对像素点; 否则, 需要采集远处像素来实现.

其次, 在特定集合 \mathcal{M} 下定义如下准则:

$$rt_2(\mathcal{M}) = \sum_{i \in \mathcal{M}} \frac{\delta(Cr)}{\sum_{i \in \mathcal{M}} \delta(\widetilde{Cr})} \quad (8)$$

其中, 判定条件 $Cr = (DF_i(0.5) < DF_i(0.01))$, $DF_i(f)$ 根据式 (6) 定义, \widetilde{Cr} 为对 Cr 求反, δ 是判别函数. 如果 $DF_i(0.5)$ 大于 $DF_i(0.01)$, 表明 i 点近处的前景和背景的区别度好于远处. 而比率 rt_2 越小则表明幅图像, 较远处的重合度较高, 计算误差也会较大, 需要保守的采集近处像素. 对整幅图像定义如下准则 T_1 :

$$T_1 = rt_1 \cdot rt_2(M_1)^{\theta_2} \cdot rt_2(M_2)^{\theta_3} \quad (9)$$

其中, $M_1 = \{i | i \in \Omega_U \wedge DF_i(0.1) < 1\}$, $M_2 = \{i | i \in \Omega_U \wedge DF_i(0.1) \geq 1\}$, $\theta_2 = 0.7$, $\theta_3 = 0.4$. 之后, 若 $T_1 \leq 0.7 \times 10^{-4}$, 则该图像为前景与背景重合度较高的图像; 否则, 若 $T_1 \geq 1.2 \times 10^{-4}$, 则该图为重合度较低的图像. 若 T_1 在上述二值之间, 则进一步令 $T_2 = rt_1 \cdot rt_2(M_1)^{\theta_2}$, 若 $T_2 < 0.8 \times 10^{-4}$, 则该图为重合度较高的图像, 否则为重合度较低的图像. 上述参数均根据经验取值.

最后, 设定低度重合的图像采用 2 个远距离级别的 f 为 0.1 和 0.01, 而高度重合的图像采用 2 个近距离的 0.5 和 0.1. 每种重合度形式均采用近距离搜索范围为主、远距离为辅的搜索形式.

2.3 低重合度图像

对每个未知点 i , 计算如下比例:

$$\begin{cases} rt_i(O) = \frac{d_i(0.1,O)}{d_i(0.01,O)}, O = F, B \\ rt_i = \max(rt_i(F), rt_i(B)) \end{cases} \quad (10)$$

其中, d_i 为在第 2.1 节中求得的 i 与直接点颜色差异. 定义如下 $f = 0.01$ 处的判别准则

$$rt_i > \theta_4 \wedge DF_i(0.01) > \theta_5 \quad (11)$$

经第 1 节分析可知, 背景通常较为平滑, 而前景变化相对较大, 因此, 若 $rt_i(F) > rt_i(B)$, 则 $\theta_4 = 5$, 否则 $\theta_4 = 1.5$, 同时对上述情况均取 $\theta_5 = 1$, 上述参数取值表明对前景的约束要高一些. 上式成立时表明: 1) 该点在 $f = 0.01$ 下与直接前景点或背景点的差异较在 $f = 0.1$ 时小; 2) 该点在 $f = 0.01$ 下对应直接点的前景与背景之间差异较大. 图 5 的第 2、3 列显示了 DF 的例子.

如前文所述, 该方法需要利用直接点的前景与背景的重合程度, 模拟 KNN 类方法对间接点的采集远近. 具体的, 利用 FLANN 方法, 对每个未知点 i 在空间特征 $\mathbf{x}_i^s(0.1)$ 、 $\mathbf{x}_i^s(0.01)$ 下, 在 Ω_U 处以及靠近它的 40 个像素宽度的已知前景和背景区域位置 (比直接点区域要宽), 寻找与 i 最近的点 j 各 K 个, 称为间接点, 如图 4(a) 上方的 j 点所示. 于是, 若满足式 (11) 对应的约束条件, 则未知点 i 选

取远处 $f = 0.01$ 的 K 个间接点, 否则选取中距离处 $f = 0.1$ 的. 最后, 对于像素 i 搜索到的 K 个像素, 令它们同时对应了式 (2) 中的 K 个位置, 颜色特征为 \mathbf{x}_j , 则该类方法的对应式 (2) 下的权值为 $w_j = 1 - \|\mathbf{x}_i - \mathbf{x}_j\|/3$. 采用共轭梯度法求解所 α 值, 记为 $\alpha^{\text{pre-our}}$.

为了比较, 本文也实现了单独运行的中距离搜索方法 KNN-0.1 以及远距离搜索方法 KNN-0.01. 图 5 同时显示了本文方法能够结合 KNN-0.1 与 KNN-0.01 的优点, 并避免它们的缺点的例子 (包括局部与全局图像). 其中, 前 2 行显示了 2 个前景存在空洞、且前景与背景差异相对较大的图像, 此时第 2、3 列显示了 KNN-0.01 比 KNN-0.1 在较多关键点中的 DF 要大, 本文结果会利用这种优势来覆盖远处的已知背景像素, 并避免在近距离搜索中容易遗漏的像素, 使得更多已知点可以被划分出来. 第 3 行显示了一幅前景与背景重合较大的全局图像, 此时 KNN-0.1 与 KNN-0.01 的 DF 大都较小, 而第 4 行的局部图像显示了它们的 DF 仍然较为相似, 此时对变化过大的前景像素进行过量采集时会产生“过搜索”. 在这 2 种情况下, 本文方法采用相对保守的近距离搜索, 避免了这些情况产生的误划分.

2.4 高重合度图像

由于高重合度图像在直接点处前景与背景的重

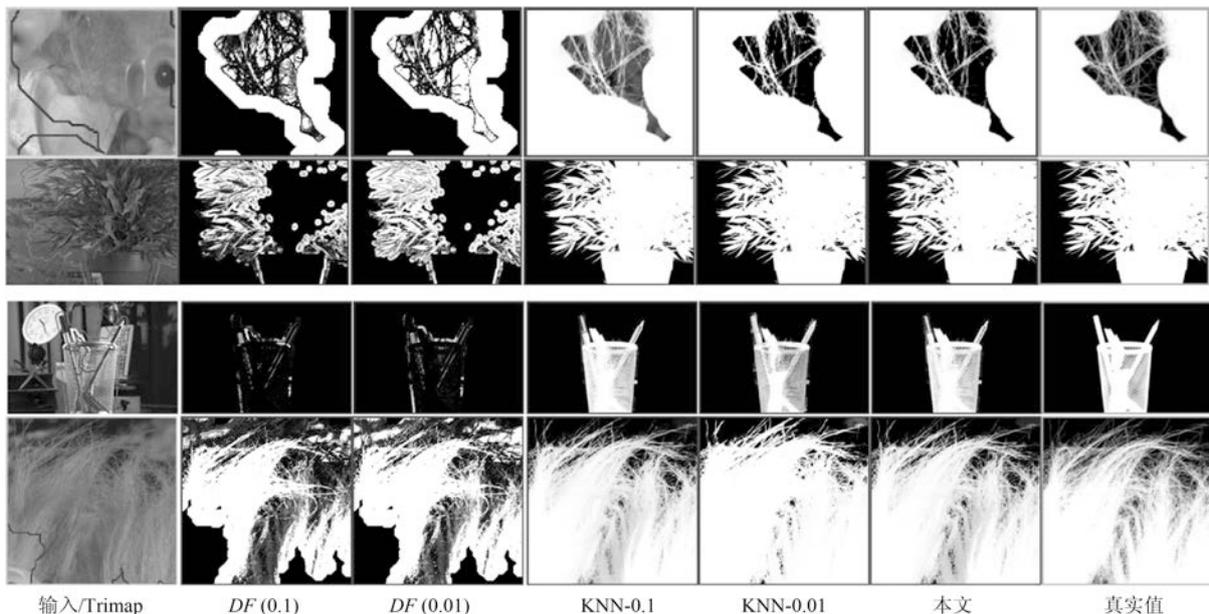


图 5 全局低重合度图像中本文预处理方法与其他 2 种单一搜索方式的结果比较 (前 2 列与后 2 列分别显示了远距离与近距离搜索方法较好的例子, 图像中的线条表示前景与背景边界)

Fig. 5 Comparison of pre-processing results between our method and two unary searching methods in low global overlapped cases (where the first and last two rows show good results in methods with long and short searching ranges respectively, and the lines in the input images show known foreground and background boundaries)

合度也较高, 无法完全提供二者的准确信息, 因此需要利用间接点的方式. 类似于第 2.3 节, 首先, 利用 FLANN 方法, 在上节定义的对每个未知点 i , 在空间特征 $\mathbf{x}_i^s(0.5)$ 、 $\mathbf{x}_i^s(0.1)$ 下, 寻找最近的 K 个间接点 j . 之后在 f 下计算每个未知点 i 与 K 个间接点的平均差异, 记为 $d_i(f)$, 并定义如下 $f = 0.1$ 处的判别准则

$$\frac{d_i(0.5)}{d_i(0.1)} \geq \theta_6 \wedge DF_i(0.1) \geq \theta_7 \quad (12)$$

其中, $\theta_6 = 10$, $\theta_7 = 1$, 上述参数均根据经验取值. 上式成立时表明: 1) 该点在 $f = 0.1$ 下与邻域间接点的差异较在 $f = 0.5$ 时小很多; 2) 该点 $f = 0.1$ 下对应直接点的前景与背景之间差异较大. 于是, 若满足式 (12) 对应的约束条件, 则未知点 i 选取较远处 $f = 0.1$ 的 K 个间接像素, 否则选取 $f = 0.5$ 处的. 求解 $\alpha^{\text{pre-our}}$ 过程仍采用共轭梯度法. 为了比较, 我们也实现了单独运行的近距离方法 KNN-0.5. 图 6 的 2 个例子 (包括局部和全局图像) 显示了前景与背景在远距离搜索 KNN-0.1 上存在较大重合的图像, 此时 KNN-0.5 较 KNN-0.1 在 DF 的很多像素上要大一些, 本文则倾向于采用 KNN-0.5 来避免这类重合所造成的计算错误. 其中第 2 个例子是测试图像, 没有真实值, 因此实际上我们引入了评价系统^[23]上排名最高算法的结果图作为近似真实值.

最后, 一种特例是该图所有未知点与前景和背景的颜色差异都较大, 该例在正常计算时误差较大, 此时该例完全简化为 KNN-0.1, 该平均颜色差异 \bar{d} 计算为

$$\bar{d} = \sum_{i \in \Omega_U} \frac{\min(d_i(0.01, F), d_i(0.01, B))}{n_U} \quad (13)$$

其中, 直接像素差异 $d_i(0.01, O)$ 在第 2.1 节定义, n_U 为未知点的个数, 该约束条件为 $\bar{d} \geq 0.003$.

2.5 Trimap 的更新

经之前的分析可知, KNN 类方法对绝对像素的计算较好, 因此需要将 $\alpha^{\text{pre-our}}$ 的绝对像素作为已解决的像素, 并且将混合像素留给 Matting Laplacian 方法计算. 首先, 对于未知像素 i , 成为新已知前景 Ω'_F 与新已知背景 Ω'_B 需要满足的条件为

$$\begin{cases} \Omega'_F = \{i | \alpha_i^{\text{pre-our}} \geq 0.95\} \\ \Omega'_B = \{i | \alpha_i^{\text{pre-our}} \leq 0.05\} \end{cases} \quad (14)$$

之后, 对得到的新前景 Ω'_F 与新背景 Ω'_B 区域进行形态学开操作, 以消除它们细小且孤立的结构. 最后, $\alpha^{\text{pre-our}}$ 中去除 Ω'_F 与 Ω'_B 得到新的混合区域 Ω'_U , 上述 3 个新区域构成了一个新 Trimap.

3 混合像素计算

图 7 显示了经过前节得到的新 Trimap 的 6 个例子, 可以明显看出, 大部分绝对点都已被计算为已知点, 而且剩余未知区域也可以基本描述出混合像素的分布情况, 包括变化迅速的实心物体边缘 (硬边界)、毛发式边界 (软边界)、半透明物体等, 这对于核宽度的选择或采用其他处理方式都是极大的帮助. 同时, 也很好解决了 Large Kernel 算法中提出的仅依据初始 Trimap 中未知区域大小的缺陷.

依据前文的理论, 混合像素的计算需要首先产生第 1.2.2 节中的集合 S , 并利用 Matting Laplacian 的方式生成权重. 不同的是, 本文并不采用 Closed Form 和 Large Kernel 算法中的基于空间连续的搜索方式, 而是采用颜色特征相匹配的 Non-local、KNN 及 CCM 的搜索方式, 该方式可以扩大搜索范围, 并减少颜色线性模型带来的计算误差. 这里每个未知点产生集合 S 的原则是: 一方面, 该集合要尽可能引入已知前景和背景区域, 以消除在未知区域中传递次数过多而造成的累积误差, 另一方

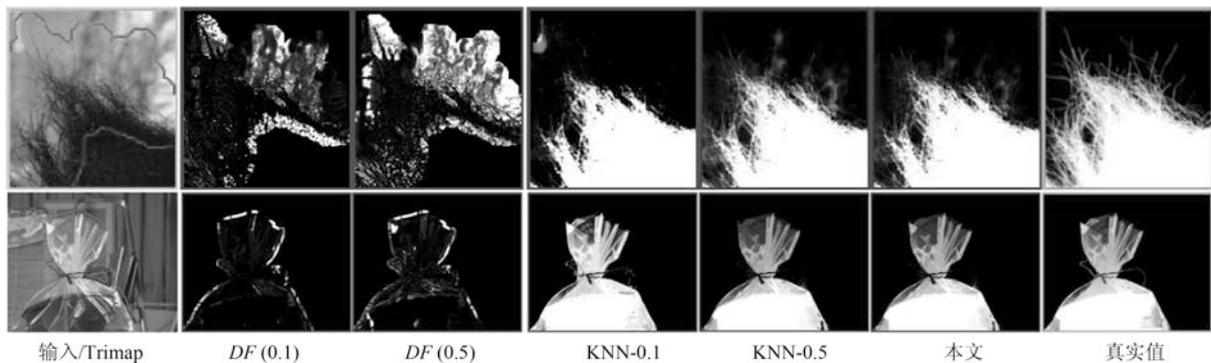


图 6 全局高重合度图像中本文预处理方法与其他 2 种单一搜索方式的结果比较

Fig. 6 Comparison of pre-processing results between our method and two unary searching methods in high global overlapped cases

面, 获取的已知像素不能过多, 以减少颜色过度重合造成的计算误差.

具体的, 在新 Trimap 下, 首先计算每个未知点 i 与新已知前景和背景的最近空间距离, 分别记为 $D_i(F)$, $D_i(B)$, 并求出二者的平均距离, 记为 D_i , 之后利用该参数调整 Nonlocal 方法中的方形搜索半径 r . 进一步, 若 r 较大, 逐点方形搜索的方式所花开销太大, 这时需采用 KNN 或 CCM 算法中的 FLANN 方法. 此时仍然遵循第 2 节的多层次搜索原则, 分为 $f = 0.5$ 、 0.1 、 0.01 的 3 级搜索方式. 此外, 本节依然采用与第 2 节类似的、计算整幅图像重合度的方式, 以此作为搜索范围的依据. 在新 Trimap 下, 对每个未知点 i 获取经第 2.1 节计算的 3 个级别的差异度 $DF_i(0.5)$ 、 $DF_i(0.1)$ 和 $DF_i(0.01)$, 按第 2.2 节重新计算已知前景和背景的平均颜色差异 $d_L(f, F)$ 与 $d_L(f, B)$, 并按式 (9) 重新计算整幅图像的重叠度 T_1 , 若 $T_1 < 2.0$ 则为高重合度图像, 否则为低重合度图像.

根据每个未知点 i 与已知区域的空间距离, 具体分为以下 3 种常规情形, 如图 7(a)~(c) 所示:

1) “硬边界”, 需要满足至少如下 2 种情形之一. a) $D_i \leq 8$; b) 整幅图像是高重合度图像. 2 个局部的例子如图 7(a) 所示. 此时对未知点 i 采用 Nonlocal 的方形搜索方法, 而搜索窗口要尽量覆盖到已知前景和已知背景区域, 但也要避免窗口过宽. 因此在半径为 r 下, 按单纯的颜色特征 x 搜索 $K = 10$

个与之最近邻的像素. 半径 r 计算为

$$r = \min(5, \max(D_i(F), D_i(B))) \quad (15)$$

2) “软边界”, 需要满足 $8 \leq D_i < 40$. 该情形代表较多毛发的边缘, 2 个局部的例子如图 7(b) 所示. 类似于第 2.4 节, 利用 FLANN 方法, 在整幅图像中对每个未知点 i , 在 $f = 0.5$ 、 $f = 0.1$ 时寻找最近的 K 个间接像素. 并定义 $f = 0.1$ 处的判别准则为 $DF_i(0.1) \geq 0.1$. 若满足该约束条件, 则未知点 i 选取较远处 $f = 0.1$ 的 K 个间接像素, 否则选取较远处 $f = 0.5$ 的.

3) 长毛发边缘与前景空洞, 需要满足 $D_i \geq 40$, 表明该点与已知背景距离很大, 2 个局部的例子如图 7(c) 所示. 类似于第 2.4 节, 对 i 在 $f = 0.1$ 、 $f = 0.01$ 时, 寻找最近的 K 个间接像素, 并定义 $f = 0.01$ 处的判别准则为 $DF_i(0.01) \geq 0.1$, 若满足该约束条件, 则未知点 i 选取较远处 $f = 0.01$ 的 K 个间接像素, 否则选取较远处 $f = 0.1$ 的.

有时, 整幅图像未知区域的空间距离或颜色差异较大, 分为以下特例, 如图 7(d)~(f) 所示, 此时以上常规方法并不奏效, 需独立于上述情形之外单独考虑.

4) 存在大量与已知区域差异较大的点, 或者所谓的“突兀点”, 此时需要满足 $\bar{d} > 0.003$, \bar{d} 为式 (13) 定义的未知区域的平均颜色差异. 一个全局的例子如图 7(d) 所示, 其中旗子部分与前景和背景的

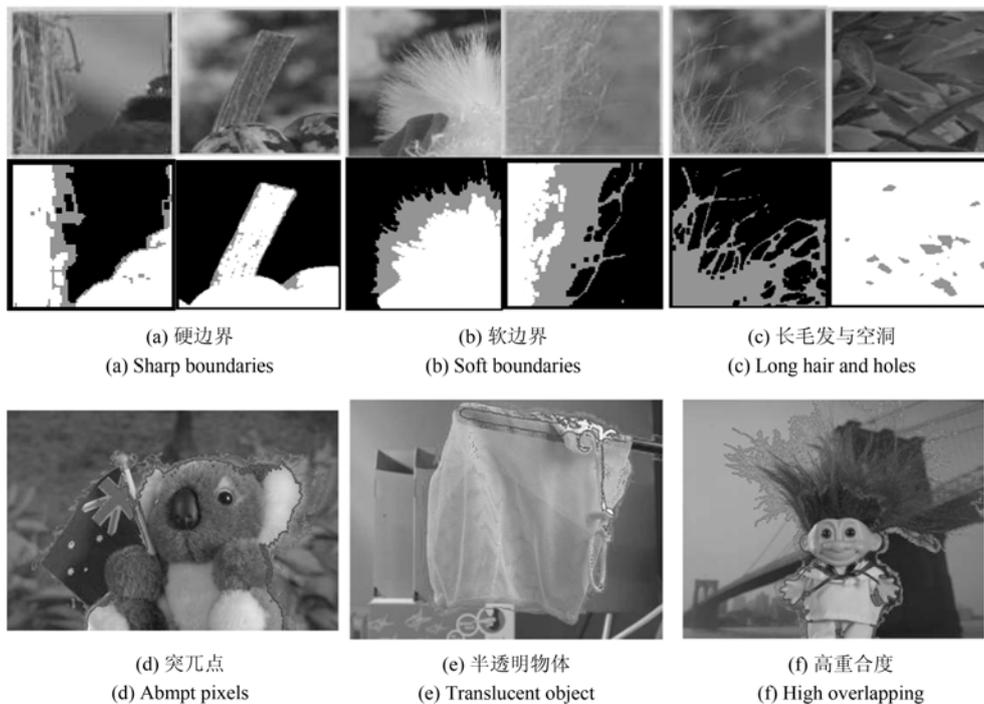


图 7 预处理后的各种 Trimap 情况

Fig. 7 Various kinds of Trimap after pre-processing

颜色差异都很大. 此时在实践中将式 (15) 中 \min 函数内的 \max 也同样变为 \min 可以使得最终抠像效果大为改善.

5) 存在大量半透明像素的前景物体, 如塑料袋、网袋、杯子等. 此时需要满足 $\bar{D} > 25$, \bar{D} 为所有 D_i 的平均值, 且不满足条件 4) 的情形. 在大量像素需要远距离搜索的情况下, 与 CCM 类似, 此时 Matting Laplacian 会经常失效, 又分为以下 2 种情形: a) 该图为低重合度图像, 一个全局的例子如图 7(e) 所示. 此时不做后续处理, 直接采用预处理后的结果. b) 该图为高重合度图像, 一个全局的例子如图 7(f) 所示. 此时意味着前面的预处理步骤很可能也不准确, 这里直接采用 KNN 算法下的 HSV 色彩区间的结果.

最后, 对于上述前 4 种情形搜索到的 K 个像素, 按第 1.1.2 节形成 Matting Laplacian 的集合 S , 并按式 (4) 为像素 i 提供 K 个未知的权值贡献, 采用共轭梯度法求出最终解 α^{lap} . 要注意的是, 此时式 (3) 中的列向量 \mathbf{b} , 对应所有已知点位置为其实际的 α 值, 即新已知前景 Ω'_F 处位置的范围为 $[\alpha^{\text{upper}}, 1]$, 新已知背景 Ω'_B 处位置的范围为 $[0, \alpha^{\text{lower}}]$, 其余未知点位置为 0. 求解中, 若遇到共轭梯度法不收敛的情况, 则直接采用预处理后的结果即可.

4 实验

为了体现本文方法对于 Trimap 大小的兼容性, 对于评价系统^[23]中的 27 幅训练图像, 除了大型号与小型号的 Trimap 之外, 我们另外为它们创建了“巨大”型号 Trimap, 它们通过将大型号 Trimap 未知区域膨胀 10 个像素、并保留面积小于 500 的前景和背景区域得到.

4.1 绝对像素划分

本文选择了 9 种仿射类的预处理方法, 包括 5 种已存在的算法 (不包含 Random Walk), 3 种 KNN 类算法 KNN-0.5、KNN-0.1、KNN-0.01, 以及本文方法. 其中, Large Kernel 算法仅简单地将 Matting Laplacian 的核宽度设为 2, Nonlocal 算法的搜索半径为 20, 搜索 10 个最相似的像素. 对各方法中获得的 α 进行预处理方式均按式 (14) 进行截取. 为了比较预处理方法的有效性, 未经预处理的方法也被引入.

4.1.1 预处理划分代价的比较

与传统意义上的预处理评价准则不同, 由于在本文预处理步骤中的已知区域会引入一些半透明点,

因此在某种划分下, 将未知点 i 的代价函数定义为

$$\mathcal{L}_i = \delta(i \in \Omega_U \wedge i^{\text{true}} \notin \Omega'_U) + \frac{|\alpha_i - \alpha_i^{\text{true}}|}{\Delta_\alpha \cdot \delta(i \notin \Omega'_U)} \quad (16)$$

其中, Ω'_U 为新未知区域, i^{true} 为 i 点在真实结果 α_i^{true} 下的划分方式, δ 为判别函数, $\Delta_\alpha = 0.1$. 该式加号左侧为漏划分代价, 右侧为误划分代价, 这与 High Resolution 方法^[17]中误划分代价的简单 2 倍不同.

表 1 显示了上述 10 种预处理方法 (包含无预处理) 在 3 种型号 Trimap 下的总体代价比较, 本文算法在各种不同型号的 Trimap 下, 划分代价均为 10 种预处理方法的最低. 其他方法中, 远距离方法 (如 CCM、KNN-0.01) 在巨大型号 Trimap 下代价较低, 近距离搜索方法 (如 KNN-0.5) 在小型号 Trimap 下代价较低, 过近距离搜索方法 (如 Nonlocal、Closed Form、Large Kernel) 在各种型号的 Trimap 下代价均较高. 搜索范围适中的 KNN-0.1 各种代价也较低, 而 KNN 算法本身效果与 KNN-0.5 较为相似.

表 1 9 种预处理方法与未预处理方法在 3 种型号 Trimap 之下对所有训练图像的代价值之和的比较

Table 1 Comparison on sum cost of all the training images within 9 pre-processing methods and no pre-processing method over three types of Trimap

预处理方法	代价值 ($\times 10^{-3}$) & Trimap 型号		
	巨大	大	小
无预处理	36.6	22.4	13.7
Large Kernel	14.6	11.0	8.3
Nonlocal	17.3	10.2	6.8
Closed Form	16.4	10.4	6.6
KNN-0.5	16.3	9.4	5.5
KNN	14.6	9.1	5.8
KNN-0.01	11.4	8.5	5.9
CCM	10.3	8.4	6.3
KNN-0.1	9.8	6.3	4.7
本文方法	9.3	6.3	4.6

图 8 显示了从图 5 和图 6 中获取的 5 个局部图像, 对它们采用 10 种预处理方法 (包含无预处理) 的预划分代价. 同样, 本文算法的划分代价基本为 10 种预处理方法的最低.

4.1.2 与后续抠像算法的联合比较

预处理方法不仅需要单独比较它们的漏划分和误划分率, 同时注意到, 预处理步骤和抠像步骤终究是一个整体, 预处理步骤的成败需要最终的抠像结果反映出来, 因此需要适当引入一些后

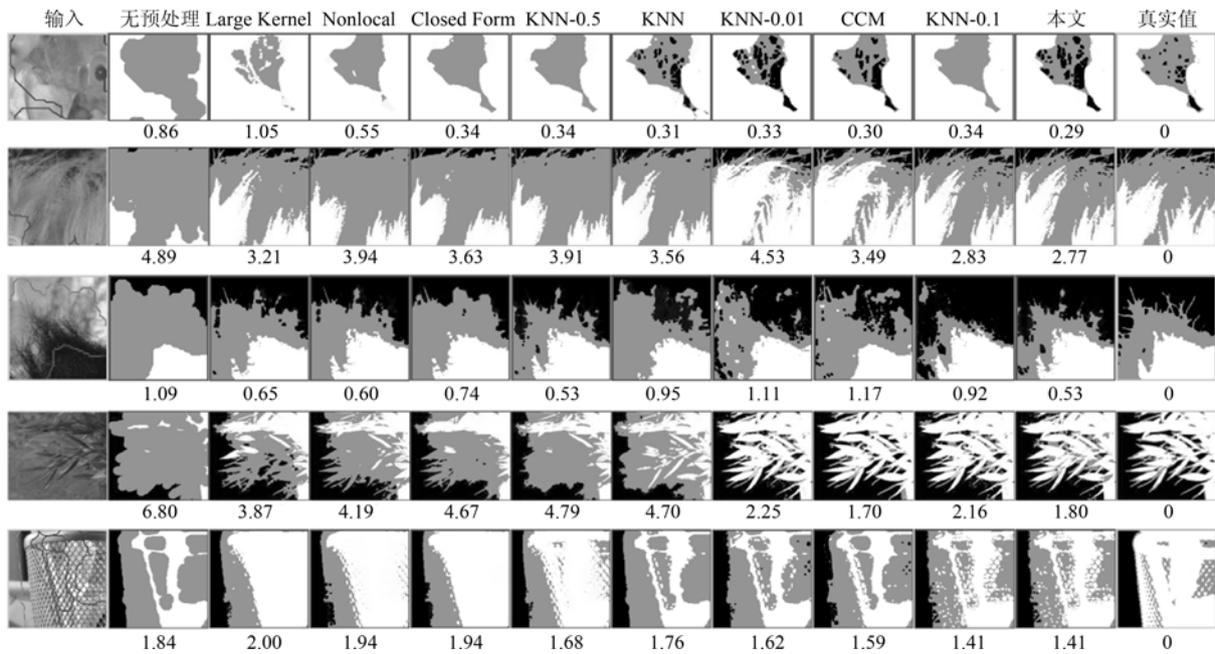


图 8 9 种预处理方法与无预处理方法在一些局部图像中的代价比较 ($\times 10^{-4}$)

Fig. 8 Comparison on cost of several local images in 9 pre-processing methods and no pre-processing method ($\times 10^{-4}$)

续的抠像方法加以验证. 本文选择了 8 种典型的后续抠像算法, 其中 4 种仿射类算法 (Closed Form、Nonlocal、KNN、CCM), 3 种采样类算法 (Robust、Global、KL-D-Sparse), 1 种特殊算法 (Sparse Coded). 由于在预处理步骤中同时会引入一些半透明已知点, 它们也会为后续抠像步骤产生作用, 因此需要对这些抠像算法做出稍许改进. 1) 仿射类方法按第 3 节方式初始化并求解; 2) 采样类方法, 假设每个未知点 i 的真实前景颜色 F_i 及背景颜色 B_i 与它的前景样本点 j 和背景样本点 k 的实际颜色相同, 利用抠像式 (1) 并推导可得

$$\begin{cases} F_i = \frac{(1-\alpha_k)F_j - (1-\alpha_j)B_k}{\alpha_j - \alpha_k} \\ B_i = \frac{\alpha_j B_k - \alpha_k F_j}{\alpha_j - \alpha_k} \end{cases} \quad (17)$$

其中, F_j 及 B_k 为前景和背景样本点实际显示的颜色, α_j 及 α_k 为前景和背景样本点的 α 值. 之后 α_i 通过如下一般求解公式得到

$$\alpha_i = \frac{(I_i - B_i)(F_i - B_i)}{\|F_i - B_i\|^2} \quad (18)$$

3) Sparse Coded 算法, 对于未知点 i , 设其字典 D 中 p 点的透明度值为 α_p , 编码系数为 β_p , 于是有

$$\alpha_i = \sum_{p \in D} \alpha_p \beta_p \quad (19)$$

图 9 显示了在巨型、大型、小型的 Trimap 下 10 种预处理方法针对上述 8 种抠像算法的比较, 因

此每种型号下共有 80 个结果, 每个结果 (即图 9 中的每个点) 是由 27 个训练图像结果的平均 MSE 计算得到. 显然, 本文算法在各种不同型号的 Trimap 下, 对 8 种抠像算法的平均排名均位列 10 种预处理方法的第一位, 在每种单一抠像算法下的排名同样位居前列. 与上节类似, 远距离方法 (CCM、KNN-0.01)、较近距离搜索方法 (KNN-0.5) 分别在巨大型号和在小型号 Trimap 下抠像效果较好, 而过近距离搜索方法 (Nonlocal、Closed Form、Large Kernel) 在各种型号的 Trimap 下的抠像效果都较差. 搜索范围适中的 KNN-0.1 排名也比较靠前, KNN 算法抠像效果与 KNN-0.5 较为相似.

与上节不同, 无预处理方法的排名并不在最后, 说明一些预处理中误差过大的误划分对整体抠像结果影响非常大. 而且, 无预处理方法随着 Trimap 未知区域的变小而效果变好, 说明预处理的主要功能还是对于未知区域内的绝对点的划分. 另一个有趣的现象是“自预处理”操作, 即用一种算法本身进行截取, 之后再采用该算法本身做出最终抠像. 结果表明, 这种方式并不能为该算法带来实质性的提升. 例如如图 9(a) 中, 若采用 Closed Form 进行预处理并采用 Closed Form 进行抠像, 效果比单纯采用 Closed Form 抠像要差. 从而进一步说明预处理与抠像步骤进行分离、并采用不同方法的必要性.

4.2 混合像素计算

在第 4.1 节本文预处理结果的基础上, 为了与本文混合像素的计算方法相比较, 本文采用了其他

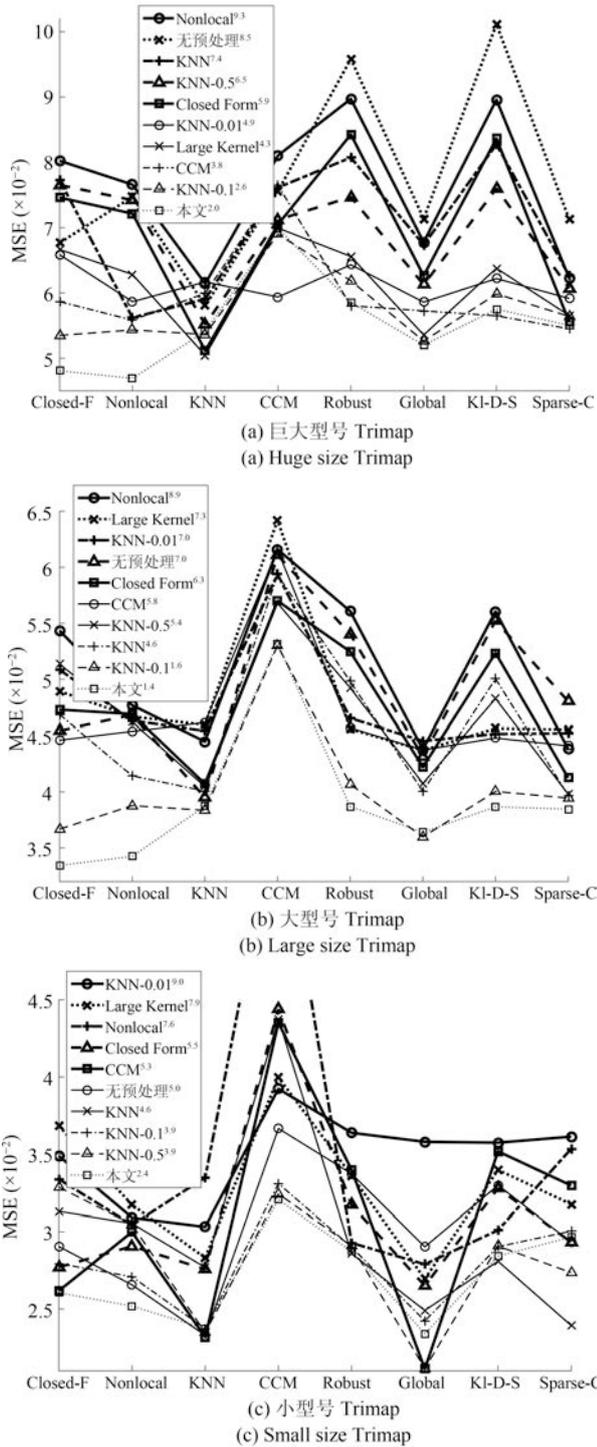


图9 3种型号的 Trimap 下, 对所有训练图像抠像结果的平均 MSE 比较, 格式为: 算法名称^{排名}, 排名为在 8 种后续抠像算法下, 每种预处理方法在 10 种预处理方法中的平均排名

Fig. 9 Average MSE Comparisons on the matting results of all the training images over three types of Trimap, where the format is [algorithm name]^{rank} and the rank denotes the average rank for each of the pre-processing methods out of 10 over the 8 matting methods

8 种单一混合像素的计算方法, 其中包括搜索半径 5、10、20 的 Nonlocal 方法, 将式 (15) 中 min 函数分别采用 min 和 max 时的方法, KNN-0.5 和 KNN-0.1 方法, 以及 Closed Form 方法 (即 Nonlocal 的搜索半径为 1). 另外, 第 4.1 节处理后的预处理结果 (即第 2 节中的 $\alpha^{\text{pre-our}}$) 也被直接引入 (记为无后处理). 注意, 这里各方法与第 4.1 节的区别是它们均采用 Matting Laplacian 取代了简单权重.

图 10 显示了在图 7(a)~(c) 中的 6 个局部图像中采用上述 10 种计算方法 (包括无处理) 的 MSE 比较. 显然无论在硬边界还是软边界, 本文方法的计算效果均为最好. 同时注意并未采用 Matting Laplacian 的无预处理方法, 该结果虽然也较为不错, 但相比上述方法中的一些结果要差, 这验证了前文提出的需要进一步引入 Matting Laplacian 进行混合像素计算的基本理论. 在其他方法中, 对于未知区域很窄的第 1、2 个例子, Closed Form、半径为 5 的 Nonlocal 等近距离方法尚且不错, 但在第 3、4 个例子中随着未知区域的扩大, 一些半径为 10、20 的 Nonlocal 方法变好, 而至最后未知区域非常宽的第 5、6 例子时, 固定搜索半径的方法已基本失效. 然而, 由于 KNN-0.5 及 KNN-0.1 的搜索范围并不完全受距离因子的局限, 可以根据颜色变化情况, 灵活和动态选择较远处的像素, 因此它们的效果始终较好.

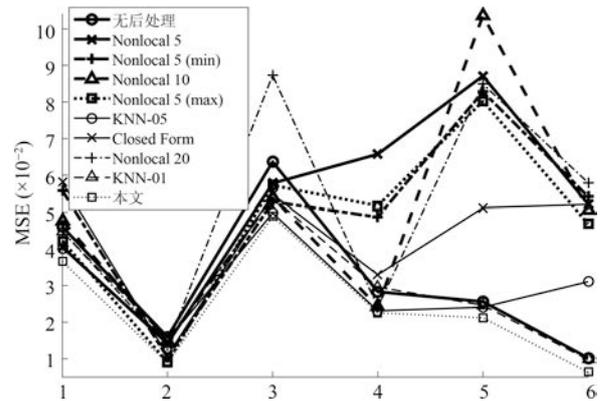


图 10 对于图 7 的前 6 个局部图像及预处理后的 Trimap, 采用 10 种方法进行抠像计算后的 MSE 比较, 其中前 2 个例子为硬边界, 中间 2 个例子为软边界, 后 2 个例子为长毛发边缘与前景空洞, 且这些例子中未知区域的宽度也逐渐增大

Fig.10 MSE comparison on matting results of the first 6 local images in Fig. 7 for 10 matting methods, where the first, median, and last 2 cases are hard boundaries, soft boundaries, and long hair edges and foreground holes respectively, in which the sizes of unknown regions gradually enlarge

为了显示本文方法的综合计算能力, 与上节类似, 上述 10 种方法在 3 种型号 Trimap 下的

27 幅训练图像的 α 结果的平均 MSE 比较均显示在图 11 中. 同时, 为了显示 α 结果分别对于绝对像素和混合像素的处理能力, 与图 3 类似, 将各 α^{true} 划分为若干均匀区间, 并将绝对像素区间 ($0 \sim 0.15$ 、 $0.85 \sim 1$) 与混合像素区间 ($0.15 \sim 0.85$) 分别显示.

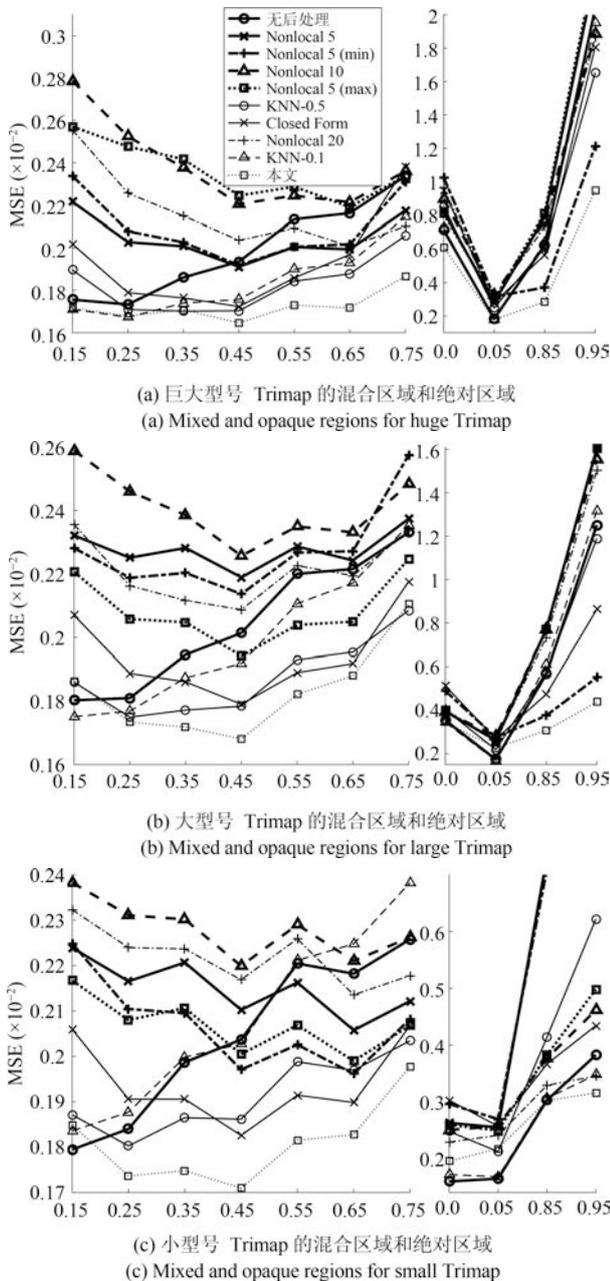


图 11 3 种型号的 Trimap 下, 在各 α 区间上, 10 种混合像素计算方法对所有训练图像的抠像结果的平均 MSE 比较, 其中 x 坐标轴中的 0.15 表示 $0.15 \sim 0.25$ 区间等

Fig. 11 Average MSE comparisons on matting results of all the training images for 10 matting methods over 3 types of Trimap in each α range, where 0.15 in x -label denotes the range of $0.15 \sim 0.25$, etc.

显然, 相对其他处理方法, 本文方法除了在混合像素上计算精确外, 对剩余绝对像素的计算仍然具有明显优势. 在其他方法中, 由于大宽度未知区域的例子偏多, 且占据误差计算的主导地位, 因此固定搜索宽度的 Nonlocal 方法始终不理想, 而搜索范围相对灵活的 KNN-0.5 及 KNN-0.1 方法则要好很多. 一个特殊的例子是搜索半径为 1、且效果同样不错的 Closed Form 方法, 此时仍然可以理解为 Matting Laplacian 在较小范围内更为有效, 因此搜索范围的远近也是一个微妙的平衡.

4.3 算法的整体评价

最后, 我们将本文中的绝对像素划分步骤与混合像素计算步骤合并为一个整体算法, 并与 9 种已经出现的仿射类算法 (或变形) 做出比较. 其中采用简单权重方法的有: 搜索半径 20、10、5 的 Nonlocal 方法, KNN-0.01、KNN-0.1、KNN-0.5、KNN 方法; 采用 Matting Laplacian 方法的有: CCM 和 Closed Form 方法. 需要注意的是, 上述单层次方法是可以与本文双层次方法相比较的, 主要由于这些单层次的方法也可以视为双层次方法, 即首先对自身做预处理 (先运行一次该算法并做出 $\alpha \geq 0.95$ 和 $\alpha \leq 0.05$ 的绝对像素划分)、再用自身做后处理 (在预处理基础上再运行一次该算法). 如前文所述, 这些算法按双层次方法运行 2 次, 和只运行该算法一次的效果相近, 因此可以用一次运行的上述算法模拟 2 层次运行结果.

表 2 显示了上述 10 种方法在 3 种型号的 Trimap 下的平均 MSE 总体排名和分别排名, 本文方法对整体和各 Trimap 型号的效果均为最优.

表 2 10 种仿射类方法最终抠像结果的 MSE 比较
Table 2 Final MSE comparison on matting results for 10 affinity based matting methods

算法	MSE($\times 10^{-2}$) & Trimap 型号		
	巨大	大	小
Nonlocal 5	13.9	8.8	5.2
Nonlocal 10	10.5	6.7	4.0
CCM	7.6	6.1	3.8
KNN-0.01	7.5	5.6	2.4
Nonlocal 20	7.5	4.7	3.0
Closed Form	6.8	4.5	2.6
KNN-0.5	6.1	3.8	2.4
KNN	5.8	4.0	2.3
KNN-0.1	4.9	3.6	2.4
本文方法	3.2	2.5	2.2

图 12 显示了各算法对于抠像效果的视觉比较, 其中前 2 个例子 (前 4 行) 显示了 2 幅训练图像. 其

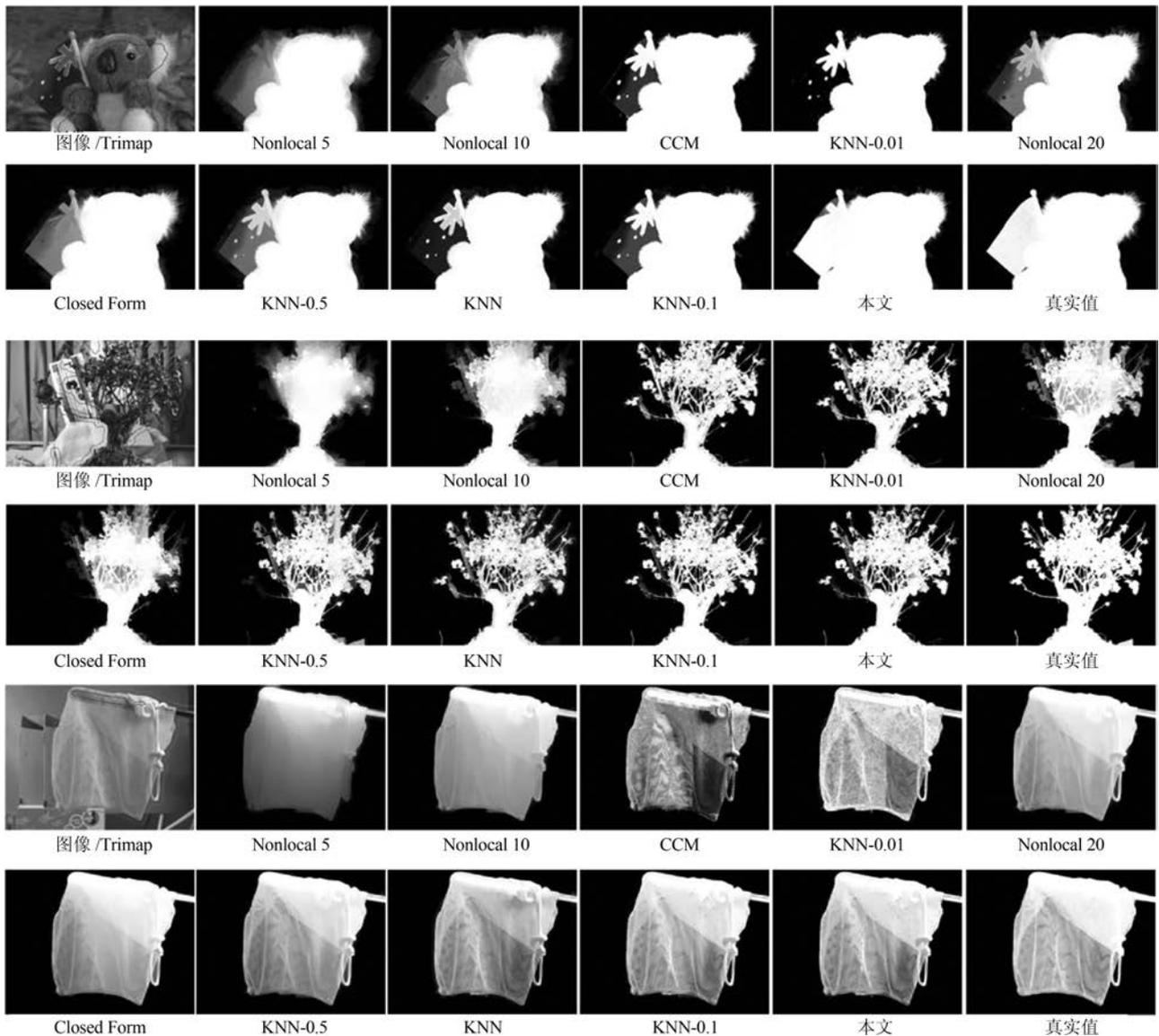


图 12 10 种仿射类抠像方法的 3 个实例

Fig. 12 3 Cases of matting results for 10 affinity based matting methods

中第一幅来自于图 7 (d), 国旗部分与前景和背景部分的颜色差异均非常大, 而第 2 幅的前景与 Trimap 的形状均非常复杂. 本文方法对这 2 个例子均能处理的非常好. 第 3 个例子 (后 2 行) 显示了 1 幅测试图像, 它们经第一步预处理后的未知区域仍然非常大, 且无法通过普通 Matting Laplacian 计算得到, 本文方法能够获得最好的结果.

5 讨论与后续工作

1) 离散与连续的 KNN 类搜索方式. 本文对 KNN 类算法中的 3 种由远及近的搜索范围因子 0.01、0.1、0.05, 是 3 种离散化的初级方式. 而如何根据前景背景的重合情况, 采用连续函数方式进行

搜索, 需要后续工作的进一步研究.

2) 简单权重与 Matting Laplacian 的适用范围. 如图 2 所示, 简单权重方法较适用在中远距离范围, 而 Matting Laplacian 方法较适用在中近距离搜索范围, 但在二者的共同区域、即中距离或偏远的范围上, 应采用何种权重的计算方式仍显得非常模糊, 例如 KNN-0.01、KNN-0.1、Nonlocal 20 等方法, 对于它们确切的应用范围仍需要进一步实验来验证.

3) 初始 Trimap 未知区域大小的问题. 介于本文在第一步预处理中对于 Trimap 输入无关的设计模式, 并未考虑初始 Trimap 未知区域的形状和大小, 这对于未知区域较宽的实例非常有效. 然而事实上, 当初始 Trimap 的未知区域非常小时 (如评价系统中的小号 Trimap), 也可以适当考虑采用更加保

守的搜索方式以减少误划分, 这种方式的抠像结果需要进一步实验验证。

4) 仿射类方法的局限性. 尽管本文在预处理步骤中采用了直接像素, 然而采样类方法的迭代最终需要间接的方式求解, 与采样类方法相比, 仍然缺乏前景和背景的直接信息, 在某些场合中容易造成间接误差的累积, 而且无法像采样类方法在算法结束后对单个点进行观察与调试, 若效果不佳, 只能尝试改进一些算法参数, 并重新运行整个程序, 这也影响到了该类算法的后续分析和改进. 因此类似于后处理方法^[26-28], 本文方法仍然可以与采样方法结合。

5) 对于大规模图像和视频数据的抠像应用, 算法效率非常重要, 今后的工作将考虑采用多核 CPU、众核 GPU 等并行技术^[29-30] 来加速所提出的算法, 以便处理大数据量的图像和视频。

6 结论

本文采用了“绝对像素划分—混合像素计算”的双层次计算方式求解仿射类图像抠像问题. 其中, 第一层预处理步骤采用了基于 KNN 类简单权重的多种搜索距离的联合方式, 第二层后续抠像步骤采用了基于 Matting Laplacian 的多种搜索距离的联合方式, 以充分发挥这两种传统抠像方法的特点. 同时, 本文针对全局与局部前景与背景的重合度, 相应采用不同的搜索层次, 并采用了以局部近距离搜索为基础、全局搜索为补充的策略. 在实验中, 本文同时引入了若干中间层次的抠像方法与本文方法进行比较, 以验证本文方法对上述 2 个层次的抠像结果的有效性. 最后, 本文对后续探索方向进行了展望, 以推进下一步的工作。

致谢

感谢赵世杰教授、徐耀群教授和白世贞教授在本文写作过程中的支持和帮助。

References

- Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(2): 228–242
- He K M, Sun J, Tang X O. Fast matting using large kernel matting Laplacian matrices. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 2165–2172
- Lee P, Wu Y. Nonlocal matting. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 2193–2200
- Chen Q F, Li D Z Y, Tang C K. KNN matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(9): 2175–2188
- Shi Y F, Au O C, Pang J H, Tang K T, Sun W X, Zhang H, et al. Color clustering matting. In: Proceedings of the 2013 IEEE International Conference on Multimedia and Expo. San Jose, USA: IEEE, 2013. 1–6
- Grady L, Schiwietz T, Aharon S, Westermann R. Random walks for interactive alpha matting. In: Proceedings of the 2005 International Conference on Visualization, Imaging and Image Processing. Benidorm, Spain: ACTA, 2005. 423–429
- Zhang Z P, Zhu Q S, Xie Y Q. Learning based alpha matting using support vector regression. In: Proceedings of the 2012 19th IEEE International Conference on Image Processing, Orlando, USA: IEEE, 2012. 2109–2112
- Chuang Y Y, Curless B, Salesin D H, Szeliski R. A Bayesian approach to digital matting. In: Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition. Kauai, USA: IEEE, 2001. 264–271
- Wang J, Cohen M F. An iterative optimization approach for unified image segmentation and matting. In: Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, China: IEEE, 2005. 936–943
- Wang J, Cohen M F. Optimized color sampling for robust matting. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA: IEEE, 2007. 1–8
- Rhemann C, Rother C, Gelautz M. Improving color modeling for alpha matting. In: Proceedings of British Machine Vision Conference. Leeds, UK: BMVA, 2008. 1155–1164
- He K M, Rhemann C, Rother C, Tang X O, Sun J. A global sampling method for alpha matting. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE, 2011. 2049–2056
- Shahrian E, Rajan D, Price B, Cohen S. Improving image matting using comprehensive sampling sets. In: Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013. 636–643
- Varnousfaderani E S, Rajan D. Weighted color and texture sample selection for image matting. *IEEE Transactions on Image Processing*, 2013, **22**(11): 4260–4270
- Karacan L, Erdem A, Erdem E. Image Matting with KL-Divergence Based Sparse Sampling. In: Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile: IEEE, 2015. 424–432
- Gastal E S L, Oliveira M M. Shared sampling for real-time alpha matting. *Computer Graphics Forum*, 2010, **29**(2): 575–584
- Rhemann C, Rother C, Rav-Acha A, Sharp T. High resolution matting via interactive trimap segmentation. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA: IEEE, 2008. 1–8
- Yao Gui-Lin, Yao Hong-Xun. A Review on Affinity Based Algorithms in Image Matting. *Journal of Computer-Aided Design & Computer Graphics*, 2016, **28**(4): 677–692 (姚桂林, 姚鸿勋. 基于仿射方法的图像抠像算法综述. *计算机辅助设计与图形学学报*, 2016, **28**(4): 677–692)
- Johnson J, Varnousfaderani E S, Cholakkal H, Rajan D. Sparse Coding for Alpha Matting. *IEEE Transactions on Image Processing*, 2016, **25**(7): 3032–3043
- Cho D, Tai Y W, Kweon I S. Deep convolutional neural network for natural image matting using initial alpha mattes. *IEEE Transactions on Image Processing*, 2019, **28**(3): 1054–1067

- 21 Shen X Y, Tao X, Gao H Y, Zhou C, Jia J Y. Deep automatic portrait matting. In: Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016. 92–107
- 22 Xu N, Price B, Cohen S, Huang T. Deep image matting. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017. 311–320
- 23 Rhemann C, Rother C, Wang J, Gelautz M, Kohli P, Rott P. A perceptually motivated online benchmark for image matting. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009. 1826–1833
- 24 Yao G L. A survey on pre-processing in image matting. *Journal of Computer Science and Technology*, 2017, **32**(1): 122–138
- 25 Yao G L, Zhao Z J, Liu S H. A comprehensive survey on sampling based image matting. *Computer Graphics Forum*. 2017, **36**(8): 613–628
- 26 Chen X, He F Z, Yu H P. A matting method based on full feature coverage. *Multimedia Tools and Applications*, 2018: 1–29
- 27 Yao Gui-Lin, Liu Shao-Hui, Li Min, Yao Hong-Xun, Hu Wen. Applications and analyses on post processing in image matting. *Acta Electronica Sinica*, 2017, **45**(3): 719–729 (姚桂林, 刘绍辉, 李敏, 姚鸿勋, 胡文. 后处理在数字抠像中的应用与解析. *电子学报*, 2017, **45**(3): 719–729)
- 28 Zhang Zhan-Peng, Zhu Qing-Song, Xie Yao-Qin. The latest research progress on digital matting. *Acta Automatica Sinica*, 2012, **38**(10): 1571–1584 (张展鹏, 朱青松, 谢耀钦. 数字抠像的最新研究进展. *自动化学报*, 2012, **38**(10): 1571–1584)
- 29 Zhou Y, He F Z, Hou N, Qiu Y M. Parallel ant colony optimization on multi-core SIMD CPUs. *Future Generation Computer Systems*, 2018, **79**: 473–487
- 30 Zhou Y, He F Z, Qiu Y M. Dynamic strategy based parallel ant colony optimization on GPUs for TSPs. *Science China Information Sciences*, 2017, **60**(6): 260–262



姚桂林 哈尔滨商业大学副教授. 主要研究方向为图像处理, 人工智能.
E-mail: glyao@hrbcu.edu.cn
(YAO Gui-Lin Associate professor at Harbin University of Commerce. His research interest covers image processing and artificial intelligence.)



赵志杰 哈尔滨商业大学教授. 主要研究方向为图像处理, 智能信息处理.
E-mail: zhaozj@hrbcu.edu.cn
(ZHAO Zhi-Jie Professor at Harbin University of Commerce. His research interest covers image processing and intelligent information processing.)



苏晓东 哈尔滨商业大学教授. 主要研究方向为人工智能, 智能信息处理. 本文通信作者. E-mail: suxd@hrbcu.edu.cn
(SU Xiao-Dong Professor at Harbin University of Commerce. His research interest covers artificial intelligence and intelligent information processing. Corresponding author of this paper.)



辛海涛 哈尔滨商业大学教授. 主要研究方向为智能信息处理, 大数据分析.
E-mail: xht@hrbcu.edu.cn
(XIN Hai-Tao Professor at Harbin University of Commerce. His research interest covers intelligent information processing and big data analysis.)



胡文 哈尔滨商业大学教授. 主要研究方向为商业智能, 嵌入式技术, 社会计算.
E-mail: huw@hrbcu.edu.cn
(HU Wen Professor at Harbin University of Commerce. His research interest covers business intelligence, embedded technology and social computing.)



秦相林 哈尔滨商业大学教授. 主要研究方向为图像处理, 计算机网络.
E-mail: qxl@hrbcu.edu.cn
(QIN Xiang-Lin Professor at Harbin University of Commerce. His research interest covers image processing and computer network.)