

基于纹理聚类的抠图算法

阳伟*, 甘涛, 兰刚

(电子科技大学 电子工程学院, 成都 610054)

(*通信作者电子邮箱 yy0838@126.com)

摘要:针对当图像纹理比较丰富时抠图难的问题,基于 K 近邻(KNN)抠图算法提出了一种纹理聚类抠图(TCM)方法。该方法首先提取出纹理特征;然后用该纹理特征与颜色和位置特征一起构造新的特征空间;接下来在该特征空间上聚类近邻像素以构造Laplacian抠图矩阵;最后利用闭形解求解不透明度。在基准数据集上的实验结果表明,该方法的总排名有显著提升,对于纹理丰富的图像取得了比较好的抠图效果。

关键词:纹理聚类抠图; K 近邻抠图; Laplacian矩阵; 闭形解

中图分类号: TP391.413 **文献标志码:** A

Texture clustering matting algorithm

YANG Wei*, GAN Tao, LAN Gang

(School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu Sichuan 610054, China)

Abstract: To solve the problem that traditional matting methods do not perform well in highly textured regions, a Texture Clustering Matting (TCM) algorithm based on K -Nearest Neighbor (KNN) matting was proposed. First, the texture features were extracted. Second, a new feature space which contained color, position and texture information was constructed. Third, the matting Laplacian matrix was constructed by clustering neighbors in the new feature space. Last, the opacity was solved by using the closed-form solution. The experiments on benchmark datasets indicate that the overall ranking of the proposed method is significantly improved, which achieves relatively leading matting effect for highly textured image.

Key words: Texture Clustering Matting (TCM); K -Nearest Neighbor (KNN) matting; Laplacian matrix; closed-form solution

0 引言

在图像或视频编辑中,常常会用到背景替换技术,把图像中感兴趣的前景目标精确地提取出来,再与其他背景合成一幅新的图像。该技术即大家熟知的数字抠图(Digital matting)技术。要完成该工作,需要估计出图像的前景 F 、背景 B 和不透明度 α 。用 z 表示像素下标,那么每个像素 I_z 可以用前景色 F_z 、背景色 B_z 和不透明度 α_z 线性合成。合成公式表示如下:

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z \quad (1)$$

其中 $\alpha_z \in [0, 1]$ 。当 $\alpha_z = 1$ 时,表示该像素为前景;当 $\alpha_z = 0$ 时,表示该像素为背景;当 $0 < \alpha_z < 1$ 时,表示该像素为前景和背景的混叠。对于彩色图像,式(1)中有3个方程和7个未知变量,这在数学上是一个不适定(ill-posed)问题。所以,这就需要提供先验假设或通过用户交互提供的额外信息来为抠图增加约束条件,再行求解。

根据如何利用图像统计特性和近邻像素的相关性,自然图像抠图可以大致分为基于颜色采样和基于像素相似性传播两类技术。基于颜色采样的方法是:假设局部颜色平滑,采样颜色相近的前景 F 和背景 B ,利用合成公式估计未知区域像素的 α 。在基于像素相似性传播的抠图方法中,从颜色、空间位置等方面描述近邻像素的相似性,利用这种相关性求约束解。

在基于颜色采样的方法中,贝叶斯抠图(Bayesian Matting, BM)^[1]把抠图问题转变成求解最大后验概率问题,从数学角度对求解问题作了详细证明,是近十年来重要的参考和改进对象。但是由于该算法存在计算量大、误差迭代累

计的问题,当前景和背景颜色稍微接近时,效果往往不佳;鲁棒抠图(Robust matting)^[2]为了提高采样的效果,添加了信任系数,可以采样到距离更远、种类更丰富的样本,但是当未知区域比较宽时,采样的效果并不明显;最近提出来的颜色和纹理加权抠图(Weighted color and texture matting)算法^[3]创新地考虑了图像的纹理特征,通过添加纹理信息提高采样的可靠性,取得了比较好的抠图效果。

在基于像素相似性传播的方法中,泊松抠图(Poisson Matting, PM)^[4]假设图像的前景和背景平滑过渡,在图像的亮度梯度场上求解一个泊松方程,利用狄里克雷边界条件求解。在纹理比较复杂的图像中,前景和背景平滑过渡的假设被打破,抠图效果不佳。闭形式抠图(Closed-Form Matting, CFM)^[5-6]假设原始图像在一窗范围内,前景色 F 和背景色 B 都可以分别表示成某两种颜色的线性组合,即Color line模型,通过最小化二次代价函数来求解闭形解。文中还推导出了拉普拉斯抠图矩阵,这些重要的理论成果为后续的研究奠定了基础,被广泛地采用;非局部抠图(Nonlocal matting)^[7]利用非局部准则(nonlocal principle)^[8]重新构建了拉普拉斯矩阵,使用聚类算法聚类近邻相似成为可能; K 近邻(K -Nearest Neighbor, KNN)抠图^[9]利用 K 近邻算法,对图像特征聚类,构建相似矩阵,再求闭形解,在当时取得了比较领先的抠图效果。

上述的方法存在两个问题:第一,当前景和背景区域的颜色接近或色阶重叠时,采集的样本不能对目标像素作一个准确的表示,利用这些样本来估计 α 的效果不佳;第二,在纹理密集的区域,硬边(strong edges)降低了近邻像素相关性描述

收稿日期:2013-04-23;修回日期:2013-06-07。

作者简介:阳伟(1986-),男,四川宜宾人,硕士,主要研究方向:图像处理、模式识别;甘涛(1975-),男,四川成都人,副教授,博士,主要研究方向:图像压缩、数字音视频技术、多媒体通信;兰刚(1970-),男,四川成都人,讲师,硕士,主要研究方向:多媒体通信。

的准确度,使得 α 不能够可靠地从已知区域向未知区域传播。目前的自然图像抠图方法趋向于多种技术的融合。比如,颜色采样法和相似性传播法的融合,局部处理和非局部处理的融合,颜色特征、纹理特征和几何形态的融合等。另外,基于机器学习的方法也是数字抠图的新尝试。它通过一个机器学习的过程,降低算法对抠图问题的假设程度,从模式识别的角度求解。

本文针对图像纹理比较丰富时抠图难的问题,基于 KNN 抠图算法,提出了一种纹理聚类抠图(Texture Clustering Matting, TCM)算法。该方法把图像的纹理也作为一种特征,同颜色特征一起重新描述近邻像素的相关性,极大地提高了近邻像素相关性描述的准确度。在基准数据集上的实验结果证明了本文算法能有效地提取出纹理丰富区域的前景目标。

1 KNN 抠图算法

KNN 抠图算法的主要步骤如下:

1) 提取特征向量。

利用原始图像的颜色和位置信息提取向量特征:

$$X(i) = (\cos(h), \sin(h), s, v, x, y)_i \quad (2)$$

其中: h, s, v 分别为 HSV 颜色空间的各个分量; (x, y) 为像素的位置坐标。

2) 利用 K 近邻算法^[10] 聚类近邻像素,构造拉普拉斯矩阵。

利用内核函数:

$$k(i, j) = 1 - \|X(i) - X(j)\| / C \quad (3)$$

其中: $X(i)$ 为特征向量; C 为权值调节系数,它保证 $k(i, j)$ 的取值在 0 到 1 之间。把相似性权值 $k(\cdot)$ 代入相似矩阵^[7],求得拉普拉斯抠图矩阵:

$$L = D - A \quad (4)$$

其中 D 为对角矩阵, $D_{ii} = \sum_j A_{ij}$ 。

3) 闭形式求解。

加入用户约束信息(三分图),求得闭形解:

$$\alpha = (L + \lambda M)^{-1}(\lambda v) \quad (5)$$

其中: M 为对角矩阵,表征了用户对已知像素区域的标记; v 为向量,表征了用户对前景像素区域的标记; λ 为约束系数(例如 $\lambda = 1000$)。本文通过添加纹理特征,重新提取特征向量,提出了一种基于纹理聚类的抠图算法。

2 纹理聚类抠图算法

传统的抠图方法只利用了图像的颜色信息,忽略了纹理也是重要的视觉特征。文献[3]提取出了纹理,与颜色特征一起作为采样条件,有效地采样到前景/背景颜色样本对。该文提出的 TCM 算法,在颜色和位置特征的基础上,添加了纹理特征。把这些特征混合在一起,能够更有效地聚类近邻像素,对目标像素作更精确的表示。TCM 流程如图 1 所示。

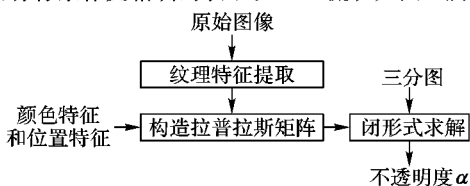


图 1 纹理聚类抠图过程

2.1 纹理特征提取

利用小波分解的多尺度描述纹理,得到高维的纹理特征。利用主成分分析(Principal Components Analysis, PCA)和线性判别分析(Linear Discriminant Analysis, LDA)法两次降维,最终不仅降低了维数,而且增加了前景和背景的分度。流程

如图 2 所示。

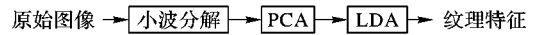


图 2 纹理特征提取过程

2.1.1 基于小波分解的纹理描述

小波分解,取出纹理。采用二层 Haar 小波分解,得到四个子图 A, H, V, D , 它们分别表示近视图像、水平、垂直和对角纹理。每层小波的四个子图以 3×3 的子窗求均值,并求得低频信号 A 在小窗内的梯度和方差,用以表征颜色的变化。最后把这些子图的大小按 bicubic 插值法调整到原图大小。纹理特征可以表示为:

$$T = \{A_{(l,c)}^{grad}, A_{(l,c)}^{var}, A_{(l,c)}, H_{(l,c)}, V_{(l,c)}, D_{(l,c)}\} \quad (6)$$

其中: c 表示颜色通道; $l = 1, 2$ 表示小波尺度。

2.1.2 PCA 降维

利用 PCA 算法进行第一次降维。以保持原始数据的绝大部分信息为前提,把原始高维的数据投影到一个低维的特征空间。

PCA 算法过程如下:

- 1) 输入图像纹理特征 $T_{N \times 36}$, N 为图像像素个数, 36 为特征维数。
 - 2) 求协方差矩阵。对 36 个特征求均值,对所有的样本 $T_{N \times 36}$, 都减去对应的均值,从而得到协方差矩阵。
 - 3) 计算 C 的特征值,把特征值降序排列。
 - 4) 计算出最大的前 d 个特征值对应的特征向量,构成特征向量矩阵 $U_{36 \times d}$ 。
 - 5) 将原始数据在 $U_{36 \times d}$ 上投影,获得主特征 $P_{N \times d}$ 。
- 对式(6)中 36 维纹理特征 $T_{N \times 36}$ 采用 PCA, 留下至少 96% 的信息,得到降维后的纹理特征 $P_{N \times d}$ 。流程如图 3 所示。

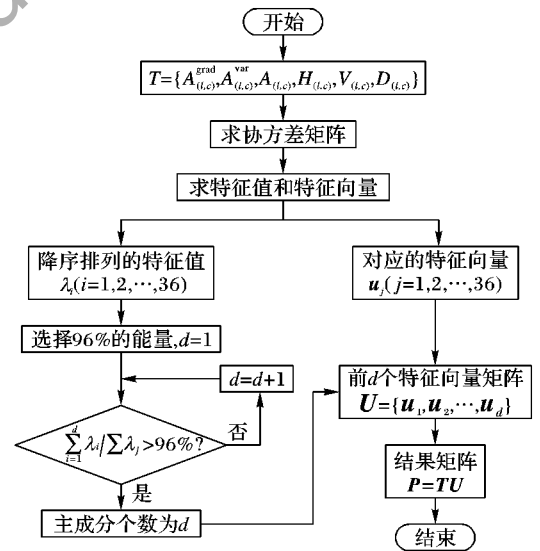


图 3 PCA 算法流程

原始图像的特征维数由 36 维下降到 d 维, d 取值在 5 ~ 7, 还需要利用线性判别分析法二次降维,增加前景和背景纹理的分度。

2.1.3 LDA 降维

利用 LDA 算法进行第二次降维。寻找一种投影空间,把前景和背景区域的纹理区分开来。对前景 F 和背景 B 区域像素的纹理分别用 K 均值法分类。把分类信息和纹理作为输入进行 LDA 分析,取前 3 维投影,这样,原始图像的特征维数由 36 维下降到 3 维。其中,聚类的类数可以由赤池信息准则(Akaike Information Criterion, AIC)估计得到。

LDA 算法过程如下:

- 1) 利用 k -means 算法,分别对输入前景和背景的纹理

$P_{N \times d}$ 分类;

- 2) 计算类间离散度矩阵 S_b 和类内离散度矩阵 S_w ;
- 3) 计算矩阵 $S_w^{-1} S_b$ 的特征值;
- 4) 计算出前 3 个特征值对应的特征向量, 构成投影矩阵

$U_{d \times 3}$;

5) 将纹理 $P_{N \times d}$ 在 $U_{d \times 3}$ 上投影, 纹理维数由 d 下降到了 3 维。

LDA 算法流程如图 4 所示。

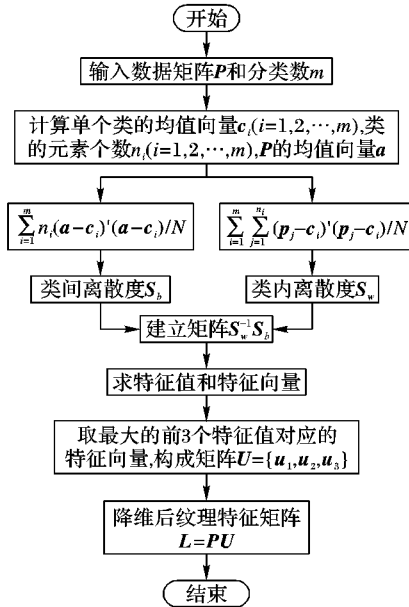


图 4 LDA 算法流程

对 GT01 进行小波分解, 采用 PCA、LDA 相结合的办法提取出纹理特征, 量化到 $[0, 255]$, 显示成图像与原图比较, 如图 5 所示。原始图像红色通道如图 5(a), 前景和背景的色阶有很大部分重叠; 纹理特征提取后的第一维纹理(图 5(b))明显把前景和背景的色阶区分开来, 第二维纹理(图 5(c))凸显了前景目标的纹理特征。

2.2 构造新特征向量求解

提取出纹理特征以后, 与颜色特征、空间位置特征一起构

造新的特征。图像的颜色空间有 RGB 或 HSV, 但是 HSV 颜色空间与人们的视觉感知接近, 通常效果较好。像素 i 的特征向量定义为:

$$X(i) = \begin{cases} (r, g, b, v, t_1, t_2, t_3, x, y)_i, & \text{RGB 空间} \\ (\cos(h), \sin(h), s, v, t_1, t_2, t_3, x, y)_i, & \text{HSV 空间} \end{cases} \quad (7)$$

其中: r, g, b, h, s, v 分别为 RGB 和 HSV 颜色空间的各个分量; t_1, t_2, t_3 为提取的纹理特征; (x, y) 为像素的位置坐标。与式(2)相比, 该式增加了 3 维纹理特征 t_1, t_2, t_3 。

利用式(7)的替代式(2), 引用 KNN 抠图方法, 求得不透明度 α 。



图 5 纹理特征

3 实验结果分析

本文在基准数据集^[11]上进行实验。数据集包含 8 幅测试图和训练图。这 8 幅图像几乎囊括了纹理比较丰富的毛发、图像的前景目标和背景的颜色色阶重叠、半透明物体、有硬边的叶子等特殊情况。它们按照透明度由强到弱排序为: Net、Plastic bag、Doll、Troll、Donkey、Elephant、Pineapple、Plant。用户交互采用了三分图方式, 每幅图像对应 3 张三分图: small、large 和 user。small 图的提示信息最强, 制作最精细; large 图的提示信息最弱, 制作最粗糙; user 图介于两者之间。

利用 TCM 算法得到抠图结果后, 在 Rhemann 等建立了一个网上测试平台^[12]上进行了算法比较实验。平台上的标准抠图结果(Ground truth)由三角抠图法生成。把实验结果与 Ground truth 进行比较, 得到一个客观的评价。

本实验在 Matlab 平台上实现, PC 系统配置为 AMD Athlon II X2 250 处理器, 主频 3.00 GHz, 4 GB 内存。实验用测试数据共 8 幅图片, 分别在 3 种三分图测试下测试, 选取相似或比较典型算法作为比较。如 KNN 算法、Closed-Form 算法、Bayesian 算法和 Poisson 算法。结果如表 1 所示。

表 1 不同算法的绝对误差和均方误差排名

算法	Sum of Absolute Differences				Mean Squared Errors			
	overall rank	avg. small rank	avg. large rank	avg. user rank	overall rank	avg. small rank	avg. large rank	avg. user rank
TCM	9.6	12.9	9.9	6.1	9.3	12.6	8.9	6.4
KNN matting	12.8	14.1	13.6	10.5	10.3	11.5	10.5	8.8
Closed-Form matting	13.8	13.0	12.0	16.3	13.6	12.3	12.9	15.8
Bayesian matting	26.9	27.0	27.5	26.1	26.7	26.8	27.3	26.0
Poisson matting	28.8	29.0	28.6	28.8	28.9	29.0	28.8	28.9

从表 1 可以看出, 该文的 TCM 算法在加入纹理特征以后, 抠图质量得到了显著的提高。其中, 绝对误差和(Sum of Absolute Difference, SAD)的平均总排名由 KNN 算法的 12.8 上升到了 9.6; 均方误差(Mean Squared Error, MSE)的平均总排名由 KNN 算法的 10.3 上升到了 9.3。以 large 三分图作为输入, 这 8 幅图像的绝对误差和均方误差比较如图 6 所示(这 8 幅图像的顺序为: Troll、Doll、Donkey、Elephant、Plant、Pineapple、Plastic bag、Net)。

从图 6 可以看出, 该文的算法的改善效果明显, 提高了前景目标提取的精确度。图 7 和图 8 为以三分图 small 为输入的 α 比较。

在图 7~8 中, (a)~(g) 分别为: 原始图像(Input)、三分

图(Trimap)、Poisson Matting (PM)、Bayesian Matting (BM)、Closed-Form Matting (CFM)、KNN matting 和本文算法(TCM)。图 9 至 12 为图 Troll 和 Pineapple 在以三分图 large 和 user 为输入时的 α 比较。

从图 9~12 可以清晰地看出, 本文的算法取得了最好的效果。

4 结语

纹理聚类抠图是一种基于像素相似性传播的方法。纹理扮演了重要的角色, 使得近邻像素的相关性能够准确地表示, α 可靠地从已知区域向未知区域的传播, 弥补了 KNN 抠图算法的缺陷, 取得了较好的抠图效果。在今后的工作中, 可以考

虑添加几何形态等局部区域特征,进一步提高前景目标提取的精确度。

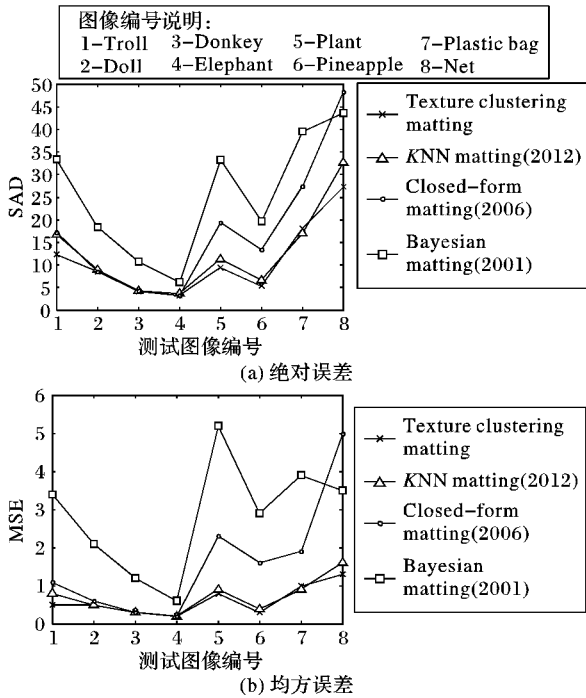


图 6 多种算法抠图结果的绝对误差和均方误差比较

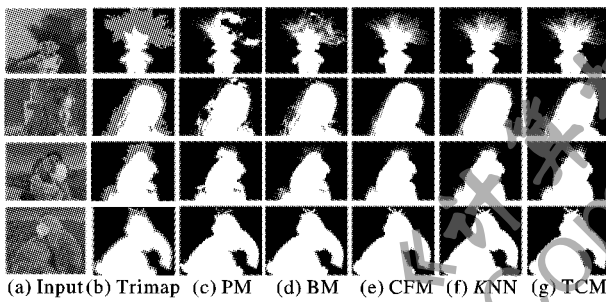


图 7 前 4 幅图抠图结果

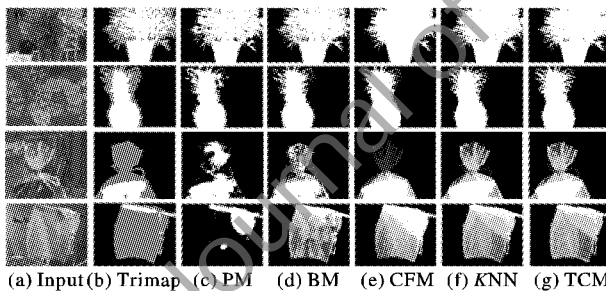


图 8 后 4 幅图抠图结果

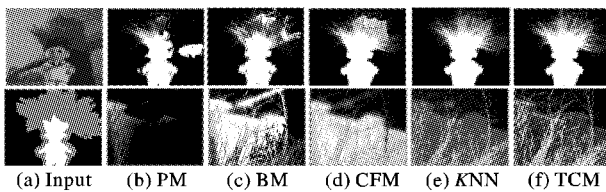


图 9 Troll 比较 (large 输入)

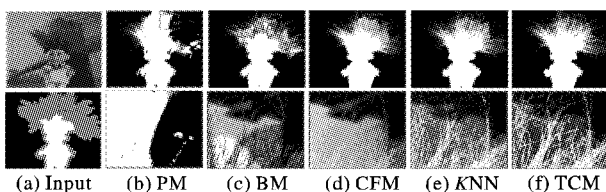


图 10 Troll 比较 (user 输入)

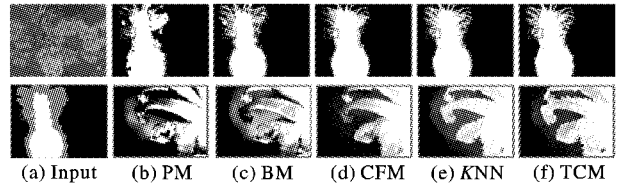


图 11 Pineapple 比较 (large 输入)

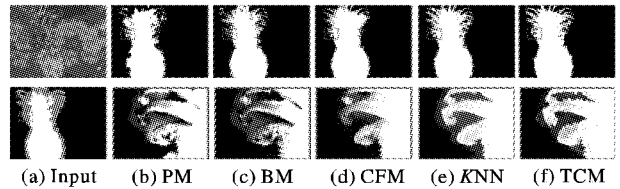


图 12 Pineapple 比较 (user 输入)

参考文献:

- [1] CHUANG Y Y, CURLESS B, SALESIN D H, *et al.* A Bayesian approach to digital matting [C]// CVPR 2001: Proceedings of the 2001 IEEE Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2001: 264-271.
- [2] WANG J, COHEN M F. Optimized color sampling for robust matting [C]// Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2007: 1-8.
- [3] SHAHRIAN E, RAJAN D. Weighted color and texture sample selection for image matting [C]// Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2012: 718-725.
- [4] SUN J, JIA J, TANG C K, *et al.* Poisson matting [J]. ACM Transactions on Graphics, 2004, 23(3): 315-321.
- [5] LEVIN A, LISCHINSKI D, WEISS Y. A closed-form solution to natural image matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 228-242.
- [6] LEVIN A, RAV-ACHA A, LISCHINSKI D. Spectral matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(10): 1699-1712.
- [7] LEE P, WU Y. Nonlocal matting [C]// CVPR'11: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2011: 2193-2200.
- [8] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising [C] // CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2005: 60-65.
- [9] CHEN Q F, LI D, TANG C-K. KNN matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(9): 2175-2188.
- [10] MUJA M, LOWE D G. Fast approximate nearest neighbors with automatic algorithm configuration [C]// Proceedings of the Fourth International Conference on Computer Vision Theory and Applications. [S. l.]: INSTICC Press, 2009: 331-340.
- [11] RHEMANN C, ROTHER C, WANG J, *et al.* A perceptually motivated online benchmark for image matting [C] // CVPR 2009: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2009: 1826-1833.
- [12] RHEMANN C, ROTHER C, WANG J, *et al.* Alpha matting evaluation website [EB/OL]. [2013-03-04]. <http://www.alphamatting.com>.