

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **11**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年11月 第17卷 第11期(总第199期)

目次

综述

数字图像合成技术综述 吴昊, 徐丹(1333)

图像处理和编码

基于群稀疏的结构化字典学习 郭景峰, 李贤(1347)

SSIM 度量虚拟视点绘制失真的深度图帧内编码 喻莉, 张军涛, 邓慧萍, 向森, 周鹏, 左雯, 王宁(1353)

统计量移位的鲁棒无损图像信息隐藏 李晓博, 周诠(1359)

伪造图像典型篡改操作的检测 左菊仙, 刘本永(1367)

图像分析和识别

融合灰度和 SURF 特征的红外目标跟踪 范新南, 丁朋华, 刘俊定, 张学武(1376)

海面温度栅格图的锋面提取与矢量化 崔雪森, 周为峰, 王栋, 张胜茂(1384)

交通场景中车辆的运动检测与阴影消除 王彬, 冯远静, 郭海峰, 张贵军(1391)

基于随机点积图的图像标注改善算法 孙登第, 罗斌, 郭玉堂(1400)

图像理解和计算机视觉

有监督子空间建模和稀疏表示的场景分类 段菲, 章毓晋(1409)

对立色 LBP 模型的目标跟踪 张炯, 宁纪锋, 颜永丰, 于伟(1418)

计算机图形学

联合骨架与边界特征的平面形状分解…………… 蒋建国, 周丹凤, 郝世杰, 郭艳蓉, 詹曙(1425)

屏幕空间自适应的地形 Tessellation 绘制…………… 张兵强, 张立民, 艾祖亮, 张建廷(1431)

遥感图像处理

SAR 图像稀疏优化滤波…………… 杨萌, 张弓(1439)

分段线性动态矩匹配条带去除…………… 秦雁, 邓孺孺, 何颖清, 陈蕾, 陈启东(1444)

基于 Harris 角点和 SIFT 描述符的高分辨率遥感影像匹配算法…………… 陈梦婷, 闫冬梅, 王刚(1453)

第八届图像图形技术与应用学术会议征文通知…………… (1460)

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 11 期

2012 年 11 月 16 日出版

主管单位 中国科学院

主 办 中国科学院遥感应用研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文

编辑出版 《中国图象图形学报》编辑出版委员会

北京 9718 信箱 邮编 100101
电子信箱:jig@irsa.ac.cn
电话:010-64807995 010-82614429
网 址:www.cjig.cn

印刷装订 北京北林印刷厂

广告经营许可证 京朝工商广字第 0346 号

总 发 行 北京报刊发行局

订 购 全国各地邮局

国外发行 中国国际图书贸易总公司
(中国国际书店)
(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing Application,
CAS China Society of Image and Graphics
Institute of Applied Physics and Computational
Mathematics

Chief editor LI Xiaowen

Editor, Publisher Editorial and Publishing Board
of Journal of Image and Graphics
(P. O. Box 9718, Beijing 100101, China)
E-mail:jig@irsa.ac.cn

Distributed by Beijing Bureau for Distribution of Newspapers
and Journals

Domestic All Local Post Offices in China

Foreign China International Book Trading Corporation
(P. O. Box 399, Beijing 100044, China)

Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2012)11-1333-14
论文引用格式: 吴昊,徐丹. 数字图像合成技术综述[J]. 中国图象图形学报,2012,17(11):1333-1346.

数字图像合成技术综述

吴昊, 徐丹

云南大学计算机科学与工程系, 昆明 650091

摘要: 数字图像合成一直是图像处理中的研究热点,在图片编辑,平面设计,电影特效等领域有着广泛的应用。从原图像中准确地提取目标物体并将其无缝地合成到新背景下是图像合成的基本目标。按所使用的关键技术分类,现有的数字图像合成技术可分为基于 α 分量的图像合成,基于梯度场的图像合成和基于多分辨率模型的图像合成。首先详述了3类方法中的典型算法,并从合成质量、鲁棒性、运算效率等方面进行分析比较,然后对新的图像合成应用方式进行了扩展介绍,最后总结了现有图像合成方法的普遍局限性,并探讨了图像合成今后面临的挑战和发展方向。

关键词: 图像合成;抠图;泊松克隆;无缝合成

Survey of digital image compositing

Wu Hao, Xu Dan

Department of Computer Science and Engineering, Yunnan University, Kunming 650091, China

Abstract: Digital image compositing is receiving increasing attention from the image processing research community, and it has been used in many applications such as photo editing, graphic design, and for visual effects in movies. The essential of image compositing is how to extract foreground object from a given image and composite it to a new background seamlessly. In this paper, we give an overview over the different approaches of digital image compositing, which are categorized as matting-based image compositing, gradient-based image compositing, and multi-resolution-based image compositing. The representative methods in each category are stated briefly, and then we compare and analyze their performance according to quality of compositing, robustness and computation time. Next, several state-of-art applications of image compositing are introduced as extension of this survey. Finally, the limitations, the research challenges and the future directions of image compositing are discussed.

Key words: image compositing; matting; Poisson cloning; seamless compositing

0 引言

在图像处理领域,数字图像合成(以下简称图像合成)技术一直是一个重要的研究热点,其研究价值来源于两个事实,第一、获取图像的代价是不相同的,图像合成可减少获取特定图像时所受的时间、

空间、花费上的制约;第二、图像合成可以产生新的图像内容,而有些内容甚至无法直接从现实世界获得。对图像合成的研究已经有很长历史,从早期简单的图片编辑到现在复杂精致的电影特效,随处可见图像合成技术的身影,与其他图像处理技术的结合也为图像合成扩展出了新的应用方式。

图像合成通常是指在输入图像(原图像)中提

收稿日期:2011-07-26;修回日期:2012-04-17

基金项目:国家自然科学基金项目(61163019);教育部博士点基金项目(20095301110006)

第一作者简介:吴昊(1982—),男,云南大学通信与信息系统专业博士研究生,主要研究方向为图像分割与合成,非真实感图形绘制。

E-mail: WuHao19820311@163.com

取感兴趣的物体或区域,通过对该区域的颜色、纹理、亮度、噪声等多个方面的调整,使其无缝地融合到新的图像背景中。图像合成技术可以按照多种方式分类:根据用户的交互程度可以分为全自动图像合成和半自动图像合成;根据合成结果可以分为颜色无缝的图像合成,边界无缝的图像合成以及纹理无缝的图像合成;按照关键技术则可以分为基于 α 分量的图像合成,基于梯度场的图像合成以及基于多分辨率模型的图像合成。虽然分类方式多样,但对图像合成的研究基本上包含两个问题,一是如何从原图像中准确高效地提取出用户感兴趣的合成区域;二是如何将提取出的合成区域无缝地合成到目标图像中,使其在颜色、纹理、噪声等多个特征上与目标图像一致,同时又能最大限度地保留合成区域原有特征。值得注意的是,虽然第一个问题与图像分割或抠图(Matting)的目标相似,但合成区域并非只限于图像的前景,它可能包含图像前景,也可能只是前景的一部分,换言之,合成区域可以是任意的。对于第二个问题,在最大限度地保留合成区域原有特征时,应根据其应用决定,既可保留部分特征,也可保留全部特征。图像合成技术不仅应用前景广阔,而且具有重要的学术价值,对图像合成的研究不仅有助于理解图像的构成,了解人眼感知图像的要素,还为其他图像编辑技术提供相关理论和基础。因此,图像合成技术一直是图像处理和计算视觉中的一个研究热点。本文整理了现有的典型图像合成方法,按其涉及的关键技术进行分类,详细介绍了各类方法中的代表算法,并对它们的合成质量、鲁棒性、运算效率等方面进行分析比较。同时,对图像合成技术较新的应用方式进行了简要介绍,并展望了图像合成技术的未来发展趋势。

1 基于 α 分量的图像合成

假设图像的前景为 $F(z)$, 背景为 $B(z)$, 并已知前景的不透明程度为 $\alpha(z)$, 则基于 α 分量的图像合成可通过前景和背景的线性组合实现,即

$$I(z) = \alpha(z)F(z) + (1 - \alpha(z))B(z) \quad (1)$$

式中, $I(z)$ ($z = (x, y)$) 表示合成后的图像。式(1)最早用于绘制前景到任意背景时减少锯齿现象^[1], 现已成为图像合成和抠图的基本公式之一。如果已知目标区域 $F(z)$ 及其对应的 $\alpha(z)$, 则可直接用式(1)将图像合成到新背景下,但难点是如何从一

幅已知图像中准确地提取出前景及 α 分量。在用 3 个通道表示的彩色图像中,该问题需要用 3 个方程求解 7 个未知量,是一个具有不确定解的“病态(ill-posed)”问题^[2]。值得注意的是,当基于 α 分量进行图像合成时,抠图和图像合成构成了一对互逆问题(图 1)。如果限定抠图的结果用于图像合成,则基于 α 分量的图像合成方法实际上即为基于抠图的图像合成。本文将按颜色采样方法和相似性度量方法对基于 α 分量的图像合成予以归类介绍。

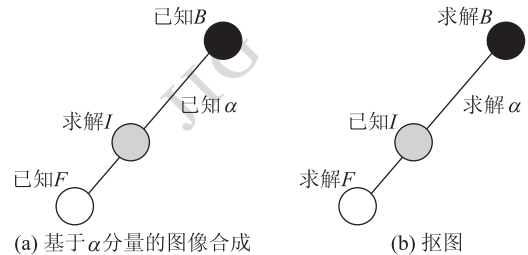


图 1 图像合成与抠图的关系

Fig. 1 The relation between image compositing and matting

1.1 基于颜色采样的抠图与合成

一般认为,抠图是非限定问题,在缺乏先验知识的情况下是无法直接求解的。Smith 等人^[3]在图像获取过程中使用了常量颜色背景,以此减少方程中的未知数个数。这种限制背景颜色的抠图技术通常称为“蓝屏抠图”,虽然该方法可以获得精确的结果,但难以应用到自然图像的抠图中。为弱化“蓝屏抠图”中的限制条件,使其能应用到更一般的自然图像中,许多抠图方法都要依赖于用户交互和对图像统计信息的先验假设,而基于颜色采样的抠图正是一类基于图像统计特性的方法,它通常要求用户指定一个三元图(Trimap),即确定的前景 Ω_F , 确定的背景 Ω_B , 以及未知区域 Ω_U (图 2)。使用三元图不但减少了求解范围(只求解在 Ω_U 中的像素),更重要的是能通过 Ω_F, Ω_B 提供必要的先验信息。正如前文所述,虽然抠图是非限制性问题的,但相邻像素间的强相关性却可用于降低求解问题的难度,从



图 2 三元图

Fig. 2 Trimap

对图像特性的统计上来看,具有相似颜色的相邻像素通常也具有相似的不透明度 α ,在这样的局部平滑性假设基础上,求解未知像素所需的样本可以从邻近的 Ω_F, Ω_B 中采集,并且样本颜色接近于未知像素真实的前景色 $F(z)$ 和背景色 $B(z)$,进而可以从样本得到 $F(z)$ 和 $B(z)$ 的近似估计。一旦确定了 $F(z)$ 和 $B(z)$,那么 α 值则可以由式(1)直接得到。

1.1.1 参数化采样方法

Ruzon 等人^[4]提出了参数化采样的抠图算法,该方法隐式假设未知区域 Ω_U 是前景边界周围的窄带,在窄带中的骨架线上,以单位间隔设置锚点(anchor points)。采样过程和 α 值的估计按以下步骤完成:

1) 在每个锚点定义一个局部窗口,每个窗口应覆盖部分前景像素,部分背景像素,以及部分未知像素,各个窗口的覆盖范围互不重叠(图 3(a))。

2) 分别采样各局部窗口中的前景像素和背景像素,然后在 CIE-Lab 空间中对这些样本值聚类并估计相应的高斯分布。

3) 假设未知像素的观察颜色 $I(z)$ 也服从高斯混合分布,则该分布由配对的前景高斯分布和背景高斯分布插值得到, α 值是使得观察颜色 $I(z)$ 概率最大化的线性插值系数(图 3(b))。

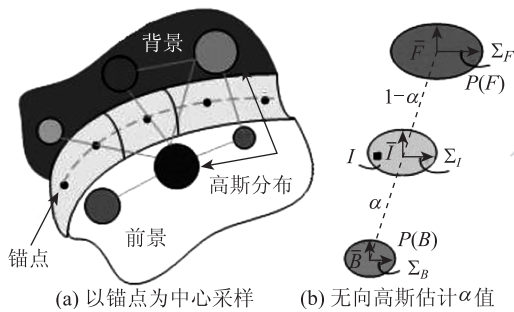


图 3 Ruzon^[4]的颜色采样方式和 α 值估计过程
Fig. 3 Color sampling and α estimating used by Ruzon^[4]

当图像颜色变化很平滑且用户给出的 Trimap 较精确时,该方法能获得较好的结果,但它也存在一些问题。首先,选取适当的锚点位置来使窗口互不重叠很困难。其次,每个未知像素的 α 值被独立计算,这可能导致抠图结果不连续。最后,当图像存在复杂的纹理时,用低阶的参数统计模型(无向高斯)为样本颜色建模会产生较大的拟合错误。

在上述算法基础上,Chuang 等人^[5]提出了基于

贝叶斯框架的抠图模型,虽然该方法仍用高斯混合模型为前景和背景颜色建模,但在采样方法和 α 估计算法上有所改进,主要体现在:1) 采样窗口不再定义在一定间隔的锚点上,而是定义在连续的区域中,并分别从前景和背景的边界逐渐向内滑动(图 4(a));2) 除前景像素和背景像素外,已经计算过的未知区域的像素同样被采样;3) 将抠图问题表示为后验概率,并使用最大化后验概率(MAP)技术和有向高斯来估计 α 值(图 4(b))。

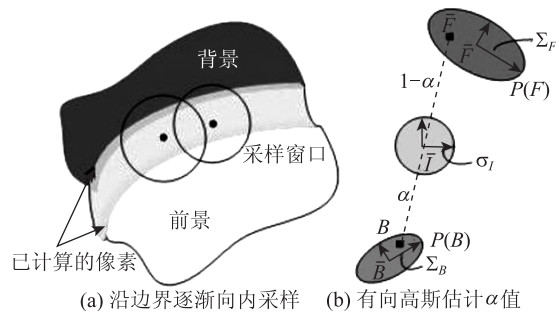


图 4 贝叶斯抠图^[5]的颜色采样方式和 α 值估计过程
Fig. 4 Sampling and α estimating by Bayesian matting^[5]

相较文献[4]的算法,贝叶斯抠图具有以下优点:1) 在连续域中定义采样窗口时使抠图结果更为平滑;2) 将计算过的未知像素作为样本可以拟合出各向异性的高斯分布,增强了算法的鲁棒性;3) α 值的估计有更严谨的推导过程,最大化后验概率技术也易于扩展和改进^[6]。但前景和背景颜色的建模依然采用了高斯混合模型,这导致前景和背景的颜色分布相似时,估计出的 α 值有大量噪点。此外,贝叶斯抠图的结果依然依赖于 Trimap 的精细程度。

1.1.2 非参数化采样方法

当给定图像的颜色分布满足假设时,参数化采样方法能得到准确和平滑的 α 值,但如果颜色分布并非高斯型时,参数估计将会有较大错误。对于这种情况,部分学者采用了非参数化模型来处理颜色样本。在 Berman 等提出的 Knockout^[7]算法中,未知像素 $I(z)$ 的前景颜色 $F(z)$ 由邻近的前景颜色样本加权平均得到,每个颜色样本的权重由未知像素与样本像素的空间距离确定。对于背景估计,首先用同样的方式求得 $B'(z)$,再根据 $I(z)$ 与 $F(z)$ 的相对位置调整 $B'(z)$ 来得到背景颜色 $B(z)$ 。则在不同通道的 α 值为

$$\alpha_R = \frac{I_k(z) - B_k(z)}{F_R(z) - B_k(z)} \quad (2)$$

式中, $k=r, g, b$, 最终的 α 值由这 3 个分量加权平均得到, 其权重正比于对应的颜色通道上前景与背景的距离。

Wang 等人^[8]把 α 值看作一个随机变量, 并以累加直方图的方式用前景和背景样本来估计其概率密度函数(PDF)。累加得到的直方图可以被认为未知像素 α 值的 PDF 的离散近似。在该直方图的基础上, 局部一致且最优的 α 估计将通过置信度传播(belief propagation)^[9]算法得到。上述算法并未验证样本对的有效性, 特别当图像含有复杂的纹理模式或用户提供的 Trimap 较为粗糙时, 有效样本对往往只占全部样本对的小部分, 为了寻找有效的样本对。文献[10]提出了一个可信度(confidence value)模型来评价样本对的质量, 该模型认为, “好”的样本对应具有以下 3 个特点: 1) 样本对的线性组合要接近未知像素的观察色; 2) 在颜色空间中, 前景样本和背景样本的差异应该足够大; 3) 当未知像素的颜色接近前景样本或背景样本时, 该像素更有可能完全属于前景或背景。通过计算样本的可信度, 不仅增强了样本颜色变化较大时计算 α 值的鲁棒性, 也提高了其计算速度。

基于颜色采样的抠图是一类有效的自然图像抠图方法, 它将未知像素周围的前景和背景像素作为样本采集, 并利用样本的统计信息和像素间的相关性来估计 α 值。当用户提供的 Trimap 较精致且输入图像含有平滑区域时, 基于颜色采样的抠图算法通常能获得较准确的结果, 反之, 如果 Trimap 很粗糙或图像含有大量强烈变化的纹理时, 算法的局部平滑性假设就难以得到保证, 从而导致抠图质量难以令人满意。

1.2 基于相似性度量的抠图与合成

错误地分类样本是基于颜色采样方法的重要缺陷, 图像纹理模式越复杂, 该缺陷越明显。为了避免该问题, 部分学者不再使用颜色样本, 而是通过定义相邻像素间的相似性来获取图像统计特性。由于像素间的相似性通常定义在一个很小的范围内(如 3×3 邻域), 局部平滑性假设更容易得到满足, 算法也更具鲁棒性。

1.2.1 基于颜色距离的相似性定义

在基于相似性度量的抠图算法中, 一种简单的相邻像素间的相似性可定义为

$$w_{ij} = \exp \left\{ - \frac{\| \mathbf{I}(i) - \mathbf{I}(j) \|^2}{\sigma^2} \right\} \quad (3)$$

式中, σ 是自由参数, 可以由用户指定。在此基础上, Grady 等人^[11]提出了一个基于随机游走模型^[12]的 α 估值算法, 该方法认为 α 值是从未知像素的位置起, 一个随机行者(random walker)在达到背景像素以前首先到达前景像素的概率, 而式(3)被用于计算随机行者在相邻像素间的转移概率。但是, 直接在 RGB 颜色空间中计算式(3)得到的概率并不可靠^[13], 因此文献[11]使用局部保持投影(LPP)技术^[14]对 RGB 颜色空间进行投影, 投影后的相似性函数为

$$w_{ij} = \exp \left\{ - \frac{(\mathbf{I}(i) - \mathbf{I}(j))^T \mathbf{Q} (\mathbf{I}(i) - \mathbf{I}(j))}{\sigma^2} \right\} \quad (4)$$

式中, \mathbf{Q} 由一组特征向量构成。实验结果表明, 经 LPP 投影后的相似性函数比直接使用式(4)能更好地区别前景边界与背景边界, 而且可以通过 GPU (graphic process unit) 来提高运算速度。该方法不仅有效地将随机过程的相关理论应用到抠图中, 更重要的是, 它为探讨在什么颜色空间中度量相似性更好提供了新的思路。

像素间的相似性函数通常建立在颜色的欧氏距离上, 但它们之间还存在着空间距离和连通性的关系。为了有效地度量未知像素与前景和背景的关系, Bai 等人^[15]用测地距离(geodesic distance)来计算 α 值。测地距是两点间测地线的长度(图 5), 在由图像构成的图(Graphic)上, 它是两点间的最短路径, 在这条路径上, 相邻两点间的距离可表示为颜色差的函数, 颜色越相似, 则距离越小。

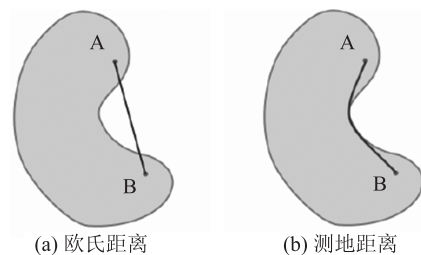


图 5 测地距离示意图

Fig. 5 The illustration of geodesic distance

当图像的前景或背景内部有较强边缘时, 直接用颜色差函数来定义测地距离会在这些边缘产生不正确的 α 的变化(在前景或背景内部区域, α 值应保持不变), 这种情况在图像颜色变化较大时十分明显, 为了缓解该问题, 他们在计算测地距离时并未直接使用颜色差的函数, 而是先计算各像素属于前景

的概率 $P_F(z)$ (通过快速核密度估计算法^[16]得到), 然后用相邻像素间 $P_F(z)$ 的变化来定义测地距离, 形式为

$$d(z_1, z_2) = \min \int_0^1 W C_{z_1, z_2}(p) dp \quad (5)$$

式中, $C_{z_1, z_2}(p)$ 是连接像素 z_1, z_2 的参数化路径(形参 $p \in [0, 1]$), W 即上述相邻像素间 $P_F(z)$ 的变化, 可定义为 $W = \nabla P_F(z)$, 在估计 α 值时, 该方法综合了两个因素: 1) 测地距离, 用于测量像素与已知前景或已知背景的接近程度; 2) 属于前景的概率, 用于度量未知像素的颜色有多少可能是属于前景。从实验结果看, 基于测地距离抠图具有以下优点: 1) 通过在相邻像素间定义合适的相似函数, 算法能从少量用户给定的种子像素中得到高质量的抠图结果, 减少了对精细 Trimap 的依赖; 2) 通过快速距离变换算法^[17], 测地距离的计算能在线性时间内完成, 所需的存储容量度也很小, 这对于将其扩展到视频上十分重要。但因为使用了种子像素来估计分布模型, 当前景和背景的颜色分布有较大重叠区域时 $P_F(z)$ 很难被正确估计, 进而影响最终的抠图结果。图 6 示意了基于测地距离的抠图, 图 6(a) 是输入图像, 用户分别以黑线和白线指定前景和背景, 图 6(b) 和图 6(c) 分别表示到前景和背景的测地距离, 黑色表示近距离, 白色表示远距离, 图 6(d) 是抠图结果。

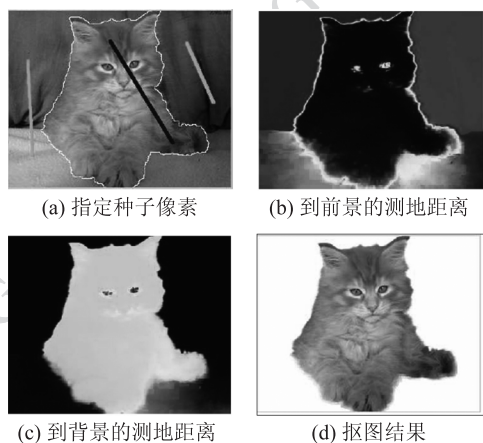


图 6 基于测地距离的抠图^[15]

Fig. 6 Geodesic distance-based matting^[15]

Zheng 等人^[18]提出了一个与测地距离^[15]相似的算法, 使用绒连通度(FC)^[19]来度量像素间的关系, 相较测地距离, FC 并未将路径中所有相邻像素的连接强度(相似性)相加, 而是代之以寻找整条路

径中最小的连接强度。这种方式能更好地捕捉到图像中的长绒毛体结构, 如发丝, 但也使其对噪声非常敏感。此外, 当不能正确估计前景和背景的颜色分布时, 该方法也不如测地距离稳定。

1.2.2 基于梯度的相似性定义

另一种定义相似性的方法是泊松抠图^[20], 与直接估计单个像素的 α 值不同, 泊松抠图根据 α 值的变化率来定义相似性。假设图像的前景和背景颜色都是平缓变化的, 对式(1)两边同时取偏导可得

$$\nabla \alpha(z) \approx \frac{1}{F(z) - B(z)} \nabla I(z) \quad (6)$$

式中, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$, 各像素的 α 值可通过求解最小化问题得到, 即

$$\alpha^* = \arg \min_{\alpha} \iint_{z \in \Omega_U} \left\| \nabla \alpha(z) - \frac{1}{F(z) - B(z)} \nabla I(z) \right\|^2 dz \quad (7)$$

式中, Ω_U 是 Trimap 中的未知区域, 由式(7)可知, 每个像素的 α 值都受相邻像素的 α 值的约束, 因此需要以全局最优的方式同时估计出所有未知像素的 α 值。虽然在梯度场上定义相似性使抠图结果很平滑, 但该方法的问题是 α 值的准确性依赖于 $F(z)$ 和 $B(z)$ 的估计精度。在具有复杂场景的图像上, 取空间距离最近的前景样本和背景样本来估计 $F(z)$ 和 $B(z)$ 并不准确, 从而导致估计出的 α 值有较大错误。为了缓解该问题, 用户可以进一步指定需要细化的局部区域 Ω_L (图 7(b)), 并通过图像滤波和求解局部泊松方程来改善结果, 但这也同时增加了交互量和使用难度。此外, 该方法所做的假设过于严格, 这对大多数图像而言是难以满足的。

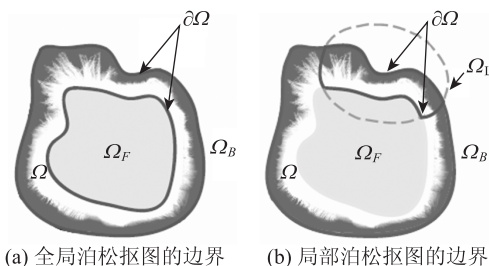


图 7 泊松抠图^[20]

Fig. 7 Poisson matting^[20]

为避免对颜色分布估计的依赖, Levin 等人^[21]假设未知像素的 α 值与其邻域范围内的像素颜色存在线性关系, 并由该假设推导出关于 α 的二次代

价函数,在该代价函数中,定义相邻像素相似性的方法与式(3)类似,不过它并未使用全局尺度参数 σ ,而是估计图像局部的均值和方差,以局部自适应方式提高算法的精度。

频谱抠图(spectral matting)^[22]不再用式(1)作为基本合成公式,而是将其一般化为 N 个图层的线性组合,即

$$I(z) = \sum_{k=1}^N \alpha^k(z) F^k(z) \quad (8)$$

式中, $F^k(z)$ 表示第 k 个图层, $\alpha^k(z)$ 表示对应图层的抠图分量。通过将频谱分割^[23-24]技术与抠图进行类比,文献[22]推导出抠图分量可以通过求解拉普拉斯抠图矩阵^[25]的最小特征值向量获得(矩阵元素表示相邻像素的相似程度),进而由这些抠图分量构建最终的抠图结果。

1.2.3 相似性度量与颜色采样的比较

在含有复杂场景的图像中,基于相似性度量的抠图算法比基于颜色样本的方法有更好的鲁棒性,其原因在于相似性通常定义在一个很小的局部窗口内,如相邻像素间,这使算法在复杂图像上能更好地保证局部平滑性假设,从而稳定性更好。基于相似性度量算法的另一个优势是全局地估计 α 值,更容易获得平滑的结果。但这类算法也存在以下不足:1)求解过程是从已知像素扩展到未知像素,之前估值上的微小错误会被累加并传播到后续计算中,最终产生较大的估值错误;2) α 值与前景 $F(z)$ 不是同时估计,而是分别进行。基于相似性度量的方法通常是通过定义相似性来计算 α ,然后再基于估计出的 α 值来计算前景 $F(z)$,而基于样本的方法则是同时考虑两者,并联合求取一个最优解。因此,当前景和背景的颜色分布重叠部分较少时,基于颜色样本的方法能得到更准确的结果。此外,Trimap的精细程度对基于颜色样本的抠图方法影响更大。

1.3 小结

基于 α 分量的图像合成是一类常用且有效的方法,其难点在于高精度地从任意图像中提取出前景及其对应的混色权重 α 。当新背景的颜色、纹理或风格与前景接近时,使用 α 值进行简单的线性混合能产生真实、自然的合成效果。如果用户希望的合成区域恰是原图像的前景时,基于 α 分量进行图像合成优势更为明显,加之其速度上的优势,使它被广泛应用于影视产业中。但如果待合成区域只是原图前景的一部分,或者背景与前景在颜色上存在较

大的视觉反差,只依靠 α 值合成会产生明显的边界,背景与前景的颜色过渡也很不自然。为消除合成边界,并使合成图像在颜色、对比度等多个特征上更具一致性,合成过程需要更复杂的模型,而非简单地将前景和背景进行线性组合。但无论如何, α 分量确是图像合成中的重要因素,即使不直接用于合成计算,它也常用于对合成区域进行预处理,或细化复杂边界附近的合成效果。

2 基于梯度场的图像合成

该类方法的焦点是如何将目标区域无缝地合成到新的背景中,其难点在于既要使合成区域与背景在颜色、纹理等多个方面一致,又要最大限度地保持合成区域的原有特征。泊松克隆(Poisson cloning)^[26]较好地平衡了这个问题,它将合成过程建模为带引导场的membrane插值,而用户仅仅需要提供合成区域的边界。如果将合成后的像素值设为未知函数 $f(z)$,图像合成的membrane插值等价于最小化问题

$$f^* = \arg \min_f \iint_{z \in \Omega_U} \|\nabla f(z) - v(z)\|^2 dz \quad (9)$$

式中, $v(z)$ 是图像的引导场, Ω_U 是图像的合成区域,在用式(9)合成得到的图像中,颜色变化的程度会尽可能地接近引导场 v ,如果将合成区域的引导场设置为原图像的梯度场,即 $v(z) = \nabla g_s(z)$,则合成图像能很好地保持原图像的局部颜色变化强弱的关系(梯度特征)。为了进一步使合成区域在边界与背景无缝地混合,在合成边界上有

$$f(z) = g_t(z), z \in \partial\Omega_U \quad (10)$$

式中, $g_t(z)$ 为目标图像。图8(d)是泊松克隆的合成效果,它在边界是无缝的。

泊松克隆具有许多优点:1)提取合成区域的方法简单,用户只需在原图像上大致地给出待合成区域,并不需要精确地跟踪物体的轮廓线(图8(a));2)可以保证合成图像在颜色上的无缝性,将合成区域的边界值设置为目标图像的边界值,使得合成图像在边界上不存在明显的变化,而membrane插值进一步保证了颜色平滑地传播到整个合成区域中;3)有多种高效的泊松方程算法可用于求解式(9)^[27],且能通过GPU进一步加速^[28]。但泊松克隆并不能保证合成图像在纹理上是无缝的。此外,当边界穿过颜色变化较大的背景区域时,很容易导致渗效应(bleeding effect)而使合成图像模糊。

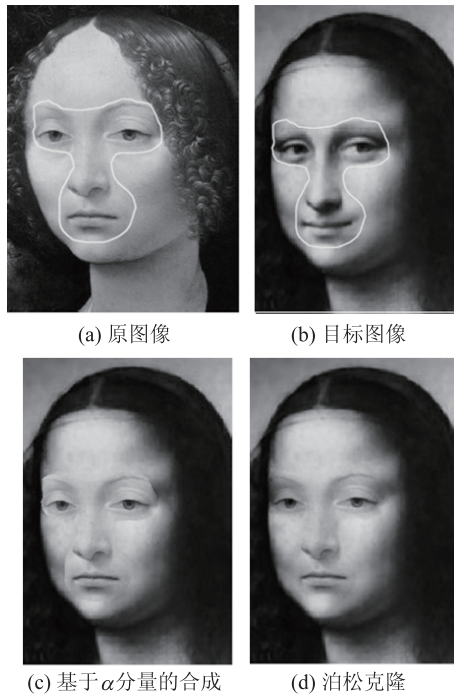


图 8 基于 α 分量的合成及泊松克隆^[26]

Fig. 8 Matting-based compositing and Poisson cloning^[26]

2.1 边界条件的优化

文献[29]进一步研究了泊松克隆中边界条件对合成结果的影响,并指出当原图像与目标图像在边界上的颜色之差为常数时, f^* 有最优值。因此,图像的合成效果可以通过优化边界来改善。最优边界可按下述主要步骤得到:

- 1) 基于 Grab-Cut 算法^[13]得到合成物体的轮廓 $\partial\Omega_{obj}$,并与用户指定的边界 $\partial\Omega_0$ 构成求解区域 Ω (图 9(a));
- 2) 连接 $\partial\Omega_{obj}$ 、 $\partial\Omega_0$ 上距离最近两点作一切割线 (Cut),将其两端的像素隔开;
- 3) 从切割线一端的像素开始,用最短路程^[30]算法在求解区域 Ω 中寻找到达切割线另一端像素的

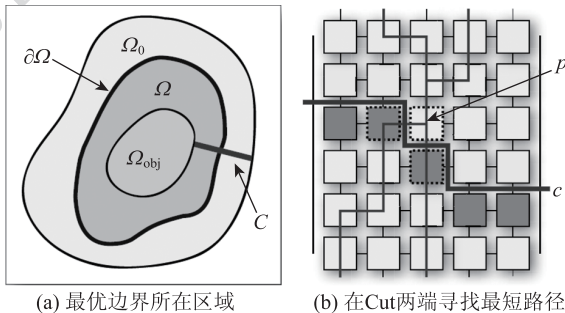


图 9 优化合成边界^[29]

Fig. 9 Boundary condition optimization^[29]

最短路程(图 9(b)),路径长度定义在每个像素上,并由以下目标函数确定,其中 k 为常数。

$$l(z) = (g_s(z) - g_t(z) - k)^2 \quad (11)$$

用最优边界求解泊松方程较大地改善了合成图像的质量,但由于将合成边界限制在物体的轮廓边界 $\partial\Omega_{obj}$ 和用户指定的边界 $\partial\Omega_0$ 之间,待合成区域总是包含整个前景物体,这使该方法难以实现前景物体任意部分的合成。此外,通过上述步骤得到的最优边界只是局部最优解,在 Ω 中仍可能存在一个能量变化更小的全局最优解。

文献[31]也采用了类似的边界优化方法,但轮廓边界 $\partial\Omega_{obj}$ 是基于泊松抠图^[20]得到,而且在寻找最优边界时,并不只是将 $\partial\Omega_{obj}$ 作为初始条件,还将 $\partial\Omega_0$,以及求解区域 Ω 中的平均颜色也作为新的初始值,通过比较从 3 个不同初始值中得到的局部最优边界,选取其中最小的作为求解泊松方程的边界条件。该方法实质上是用多个初始值对局部最优解进行“扰动”,以期能更逼近全局最优解。虽然进一步优化边界条件使合成质量有所提高,但该方法也未能实现前景物体任意部分的合成。

为了解决该问题,文献[32]不再使用封闭边界 $\partial\Omega$ 来确定合成区域,而是首先由用户指定图像中需要合成的部分,进而通过 closed-form 抠图^[21]确定合成区域。该方法不仅能实现图像任意部分的合成,而且交互方式更为简单。此外,他们的实验结果表明,通过迭代法求解泊松方程时,较少的迭代次数产生的解虽然不精确,但用这些解却可以生成有视觉意义的合成效果。

2.2 额外的限制条件

虽然优化边界可以使合成图像的局部变化特性得到保持,但可能会导致合成区域的整体亮度或色调与新背景一致,从而失去合成区域原有的色调。为此,Guo 等人^[33]提出了一种基于用户标记的方法,在保证合成图像的边界无缝性和局部变化特性时,尽可能地保持原图像的色调。该方法需要用户标记原图像中的特定像素以保证合成后被标记像素的颜色值不会改变,像素标记为合成加入了新的硬限制条件(除边界条件外),使泊松方程的解能保留原图像的颜色特性,避免合成区域的色调在新背景下失真。但额外的限制条件也可能使被标记像素周围有较大的拟合错误,从而导致合成区域内部也存在“渗色效果”。

图像粗糙化^[34]为减少因拟合错误带来的混色效

应提供了有效的方法,他们以双边滤波函数为核函数,将非线性的图像结构粗糙化为线性表示。虽然转化过程中会丢失图像细节,但粗糙化后的图像具有同质的特性,能有效减少求解线性系统时产生的拟合错误。在此基础上,Wu 等人^[35]进一步探讨了将泊松克

隆^[26]投影到线性空间求解的方法,他们通过对直接投影法、梯度投影法,以及梯度和边界投影法的比较,证明了先将图像的梯度场和边界条件转化为线性表示,再求解泊松方程能有效地减少边界附近的拟合错误,从而消除泊松克隆中的“渗色效果”(图 10(c))。

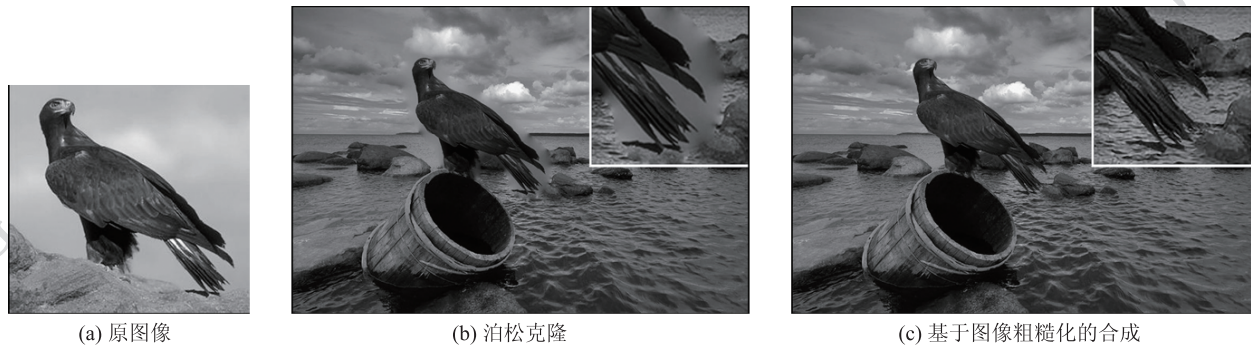


图 10 基于图像粗糙化消除“渗色效果”^[35]

Fig. 10 Eliminating bleed effect based on image coarsening^[35]

2.3 小结

当边界能量变化较小时,基于梯度的图像合成方法十分有效,它不仅能做到在边界和颜色上的无缝合成,而且在合成图像上最大限度地保留了原图像的局部对比度。虽然可以通过优化边界和增加其他限制条件进一步优化合成效果,但如果原图像的纹理模式与目标图像相差太大时,该类方法难以消除边界上纹理之间的差异,平滑的前景图像合成到高噪声的背景下也不自然。

3 基于多分辨率模型的图像合成

为使合成图像在纹理和噪声上具有一致性,分析并利用图像在不同尺度下的特性是一类有效的方法。Burt 等人^[36]将图像合成视作两幅图像的线性混合,混合权重由样条函数决定。样条函数的拟合步骤如下:

- 1) 分别构建原图像 A 和目标图像 B 的拉普拉斯金字塔 L_A, L_B ;
- 2) 由 L_A, L_B 构建合成图像的拉普拉斯金字塔 L_S, L_S 在图像中线上的值为 L_A 和 L_B 的均值;
- 3) 扩展并叠加金字塔各层的 L_S , 生成最终的合成图像 S 。

该方法在离合成边界较远的地方混合低频分量,而在离边界较近的区域则计算局部均值来得到细节,因此在效果上与泊松克隆类似。但在金子塔

的每层进行合成时会使用到来自上一层的原图像和目标图像的信息,这导致距离较远的像素也参与到混合中,从而产生不期望的模糊效果。尽管合成效果并不突出,但它提出了两个重要的结论:1) 通过样条函数无缝地合成图像时,图像的带宽应大致为一个倍频程(octave);2) 等价的合成权重函数(样条函数)可在不同分辨率下混合叠加相应的子带得到。图 11 表示以金字塔各层的等价权重函数来拟合插值样条函数。

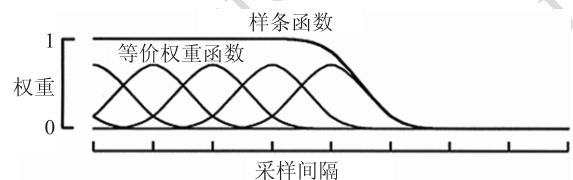


图 11 多尺度下的等价权重函数拟合样条函数^[36]

Fig. 11 Fitting spline function for image compositing by multi-scale weight function^[36]

另一个基于多分辨率的图像合成的有效方法是先对金子塔分解后的子带进行直方图匹配,然后由这些子带重构出合成图像。但直接对子带进行直方图匹配会导致图像边界周围有明显的光晕(halo effect),合成痕迹十分明显。Sunkavalli^[37]提出了平滑直方图匹配来解决上述问题(图 12),其算法的改进主要体现在:

- 1) 使用过采样的金字塔构建方法^[38]。在该金

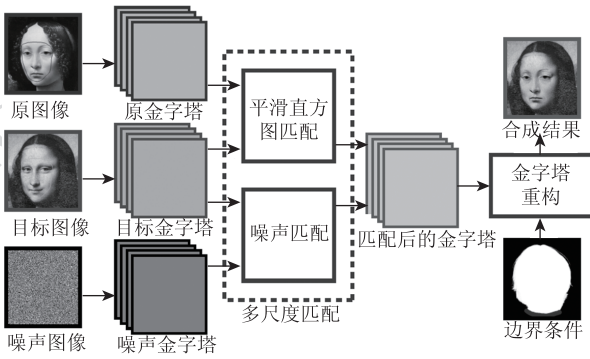


图 12 基于多尺度下平滑直方图匹配的图像合成框架^[37]

Fig. 12 The multi-scale image harmonization framework^[37]

字塔中,滤波后的子带并不进行下采样,金字塔各层与原图像具有相同尺寸,在该金字塔中进行子带的直方图匹配能有效减少合成图像的人造痕迹。

2)将直方图匹配建模为控制增益改变的过程,从而通过调节增益来控制直方图匹配的效果。

3)限制过大的增益值,避免图像的合成痕迹被放大。该步骤可通过设置最大增益值实现。

4)不直接使用增益控制直方图匹配,而是为每层金字塔计算一个活动图 A_k^s 以平滑各层中的子带系数。

$$A_k^s = N(\sigma) * \sum_{i \in lev(k)} |B_i^s| \quad (12)$$

式中, $N(\sigma)$ 为高斯滤波器, $lev(k)$ 为第 k 层金字塔的所有子带的下标集合。计算活动图的目的在于通过累计各层子带幅值和平滑滤波,得到一个在空间上局部平滑的幅值系数,使其不会过度改变子带系数的形状而影响合成效果。平滑直方图匹配后的子带可以通过求解与分解滤波器数相同的线性方程组重构为合成图像,通过将泊松克隆或 α 抠图引入方程组中,还可以进一步限制边界条件。该方法由于在不同频带对图像进行平滑直方图匹配,使目标图像的纹理和对比度特征能较好地转移到原图像中,即使两幅图像在颜色、纹理上相差较大,它也能获得较好的合成效果。如果将噪声图像分离出来(使用高通滤波器)并进行平滑直方图匹配,该方法还能使合成区域的噪声模式与目标图像相似,但其求解过程复杂,运算速度较慢,并且不适用于含有非随机纹理的图像。

4 合成算法的性能比较

对图像合成的运算速度可以通过统计不同算法

在不同尺寸图像上的运算时间得到,而合成质量的客观评价则会随测度模型的不同而产生变化。如在评价基于 α 分量的图像合成方法时,合成质量主要取决于提取 α 分量的准确性,因此量化评价通常建立在算法估计得到的 α 分量与该图像真实 α 分量间的差异的基础上。文献[2]在小规模测试集上(图 13)对不同抠图算法进行了比较,其使用的测度标准为均方误差(MSE),作为参考标准的真实 α 分量则由“蓝屏抠图”^[3]得到。在同一图像上,如果记某抠图方法在 n 个不同的 Trimap 下得到的 MSE 为 $\{e_1, e_2, \dots, e_n\}$,那么最小均方误差 $\min(e_i)$ 将作为测量算法精确度的标准,而最大均方误差 $\max(e_i)$ 将作为测量算法鲁棒性的标准。本文也使用了相同的测试集,并分别对各算法的 $\min(e_i)$ 和 $\max(e_i)$ 进行排名,各算法在 6 幅图像上的平均排名如图 14 所示,排名等级越小,算法的精确度和鲁棒性越高(秩和统计)。

相较于基于 α 分量的图像,基于梯度场或基于多分辨率模型的方法除需要保证合成边界的无缝性外,还需要尽可能地保持合成区域原有的部分特征,如局部对比度、纹理方式等,但合成算法在优化合成区域的过程中并非总能完全保持这些特征。因此,保持合成区域原有特征的能力是评价合成算法的一个重要标准。本文比较了基于梯度场和基于多分辨率模型的合成算法在保持局部对比度上的能力,并将其作为评价合成算法的测度。选取 6 幅图像(图 8、图 10、图 15)作为测试集,评价函数为

$$E_i(I, O) = \sum_{x \in \Omega} |\nabla I(x) - \nabla O_i(x)| \quad (13)$$

式中, $I(x)$ 为原图像, $O_i(x)$ 为用第 i 种算法得到的合成图像, Ω 为合成区域。式(13)通过计算原图像与合成图像梯度的 SAD(绝对差之和)来评价不同合成算法保持局部对比度的能力,其结果如表 1 所示。从表 1 可以看出,除泊松克隆^[26]和均值坐标^[39]外,其他合成方法能有效地减少梯度的 SAD,因此在合成区域能更好地保持原有的局部对比度。图 15(d)在不同合成方法下的 SAD 值非常接近,这是因为原图像和目标图像在合成边界上均是平滑变化,已接近最优边界,故以任意方法合成后均能很好地保持合成区域原有的梯度。需要强调的是,带标记的泊松克隆^[33]的 SAD 值还取决于用户标记的像素数量以及被标记像素是否有利于算法在合成区域对颜色平滑的插值,并非标记数越多, SAD 值越小。因此,用户往往需要多次修改标记才能取得理想的合成结果。

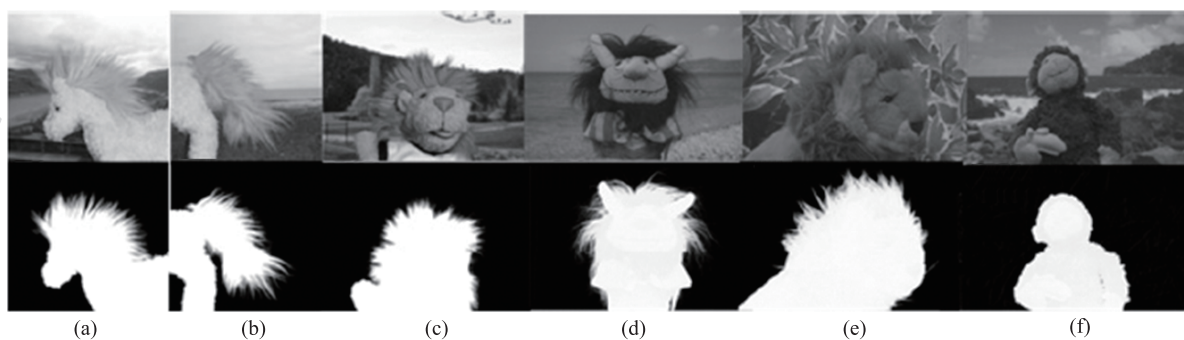


图 13 评价抠图的测试图像以及对应的真实 α 值^[2]

Fig. 13 Test images and corresponding ground-truth mattes used to evaluate matting algorithms^[2]

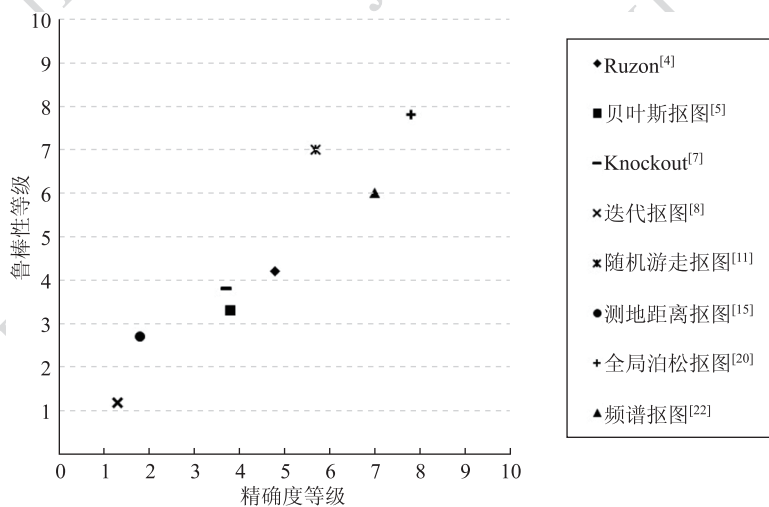


图 14 基于 MSE 测度的秩和统计结果

Fig. 14 Rank-sum results based on MSE metric



图 15 评价合成算法的部分测试图像以及对应的泊松克隆

Fig. 15 Test images used to evaluate compositing algorithms and corresponding Poisson cloning results

表 1 合成区域中梯度的 SAD
Table 1 The SAD of gradient in area of compositing

	图 8	图 10	图 15(a)	图 15(b)	图 15(c)	图 15(d)
泊松克隆 ^[26]	1 245.6	2 122.2	1 457.3	1 912.9	3 245.8	111.3
边界优化 ^[29]	121.2	225.7	146.4	176.2	245.0	110.0
多初始值边界优化 ^[31]	121.1	221.6	143.4	161.0	234.6	110.0
带标记的泊松克隆 ^[33]	678.0	743.5	531.6	645.2	845.6	110.9
粗糙化的图像合成 ^[35]	123.0	226.5	152.3	181.5	265.6	110.1
平滑直方图匹配 ^[37]	174.2	288.4	226.8	234.6	375.3	132.2
均值坐标 ^[39]	1 278.2	2 263.4	1 571.2	2 012.7	3 445.6	112.4

误差函数确实是一种有效的评价图像合成的方法,但并非与人眼的主观感受具有一致性。Rhemann 等人^[40]研究发现,将抠图结果用于图像合成时,低误差的算法却不一定产生主观满意度高的结果。图 16 为两种抠图算法的结果及其对应的 SAD,但根据用户的主观评价,88% 的人认为图 16(b)(e)的效果更好,而这两幅图像的 SAD 却较大。

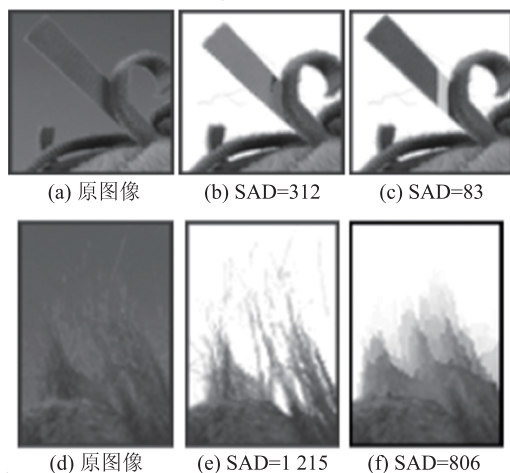


图 16 SAD 值与主观感知不符的抠图结果^[40]

Fig. 16 An instance that SAD metric is not corresponding to the perception of eyes in matting^[40]

该结果表明 SAD 或 MSE 测度并未考虑人眼的感知因素,因此文献[40]建立了基于感知的误差测度模型,其核心主要有两点:1)用连通误差来评价抠图结果中前景的连通性;2)用梯度误差来测量抠图结果中的过度平滑或不连续程度。从实验结果来看,虽然基于感知模型的测度方法与用户的主观评价更接近,但该方法需要严格挑选测试集,模型参数

需要从大量用户的选择中学习,作为评价图像合成的标准仍需改进。但不可否认,将感知模型引入到图像合成的评价体系中将是今后研究图像合成技术的一个重要分支。

5 应用扩展

随着图像处理技术不断的发展,除上述方法外,越来越多的新技术被引入到图像合成中,丰富和拓展了图像合成的应用方式。如在合成迷彩图像(camouflage image)时^[41],被嵌入目标图像的物体的纹理要尽可能与周围一致,以隐藏自身原来的纹理特征来加大感知难度。同时嵌入物体又要留下一定的线索,使观察者经过一定时间的努力后,能通过人眼的连接搜索(conjunction search)^[42]机制感知到它。这两个互相矛盾的目的难以直接用前面章节中介绍的方法实现,因而文献[41]借助亮度赋值和纹理合成来调节嵌入物体的浸入度和突出度,其过程如下:

1)按亮度值对原图像(含有待嵌入物体)和目标图像进行量化,根据结果将其分割为多个连通区域。

2)调整原图像各连通区域的亮度分布,使其介于原始值和目标图像的亮度分布之间,相应的目标函数通过构造一个图(graph)求解,其中图的节点是前景和背景的连通区域。该步骤也可称为对原图像进行亮度赋值。

3)生成纹理库,纹理库中的每一项目标为图像的纹理片(texture patch)及其对应的亮度片。

4)根据亮度赋值的结果在纹理库中寻找最匹配的亮度片,并用其对应的纹理片对原图像进行纹

理合成。该步骤与类比^[43]方法相似。

5) 通过背景变形以及添加额外的连通区域进一步细化合成结果。

当原图像的对比度较高时,该方法能合成出高质量的迷彩图像(图 17(c)),而且将浸入度和突出度参数化后用户可根据需要调节嵌入物体的识别难度。

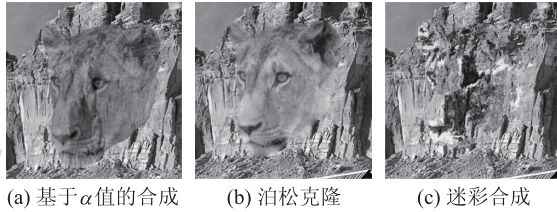


图 17 合成迷彩图像^[41]

Fig. 17 Camouflage image compositing^[41]

另一个图像合成的有趣应用是数字上妆(digital make-up)^[44-45],其目的是通过对给定的上妆样本图像的分析,将化妆效果合成到未上妆的目标图像中。文献[45]结合了多种合成方法实现这一效果,其首先将样本图像和目标图像中脸部的各个区域变形对齐,使其构成对应关系。然后将对各区域分别使用不同的合成方法,如在眼睛和嘴唇部分使用线性混色,对皮肤区域则先用带权最小均方算子^[46]分解出结构层和细节层,再将样本图像的细节层按一定权重与目标图像的细节层混合。最后通过泊松克隆^[26]将参考图像的阴影和高光部分合成到目标图像上。由于采用了分区域合成的策略,即使目标图像和样本图像在色调和风格上差异较大,该方法也能合成出较好的上妆效果。但其也存在两个限制,一是两幅图像必须是正脸对齐;二是合成是在亮度通道上进行,颜色只是直接赋值,因而部分区

域的合成效果会不自然。

图像合成算法往往涉及对线性系统的求解,其运算量大,难以实现高分辨率图像的实时处理。为此,Agarwala 等人^[47]用自适应四叉树细分计算域,并从中重构出规模较小的线性系统加快运算速度。Farbman 等人^[39]在减少合成算法的时间复杂性上做了有效的尝试,他们通过在合成区域上从边界向内插值,将求解泊松方程简化为均值坐标(MVC)插值^[48],有效地减少了运算时间。为了进一步提高 MVC 插值速度,该方法使用了两个优化策略,第一、在合成区域内建立自适应网格,并只对网格顶点进行 MVC 插值,然后再对网格顶点线性插值得到网格内部各像素的值;第二、对边界使用等级化采样,即用于计算每个插值点的边界样本数各不相同,采样数与该插值点到边界的距离成反比。在这两个优化策略下,算法的时间复杂度从 $O(mn)$ 减少到 $O(\sqrt{n})$, m 为边界上的像素数, n 为合成区域中的像素数。此外, MVC 插值是一个高度并行化的过程,可很方便地利用 GPU 进行加速。通过利用图形硬件和改进算法复杂度,该方法不仅能实现高分辨率图像的快速合成,对于中等分辨率的视频也能做到实时合成。这比建立并求解 3D 泊松方程来实现视频合成^[49]在速度上有很大提高,但由于没有考虑帧间的连贯性, MVC 插值合成只适用于内容连续的视频。另外,将合成区域限定为矩形虽然简化了计算,但会留有明显的合成边界,当合成区域的纹理与目标视频帧的纹理相差较大时,矩形合成边界会更加突出。图 18 为视频合成中的两帧,图 18(a) 中矩形区域被无缝地合成到图 18(b) 的对应区域。

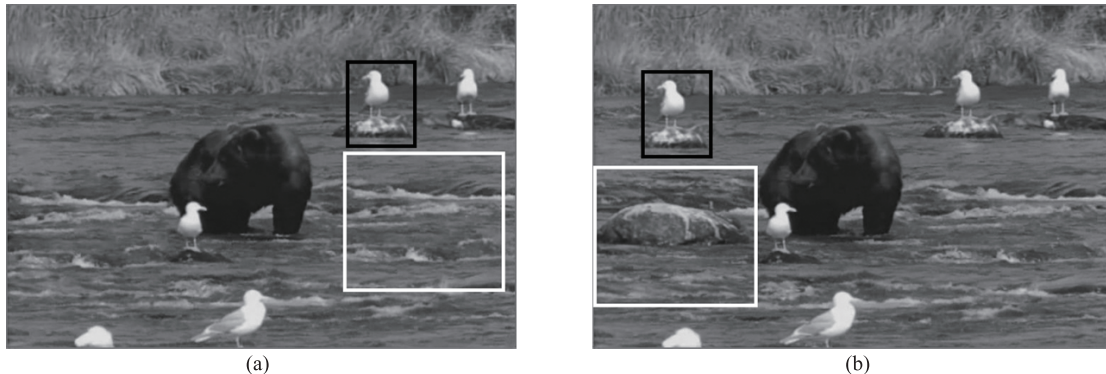


图 18 基于均值坐标的视频合成^[39]

Fig. 18 Video compositing based on mean-value coordinates^[39]

6 结 语

图像合成一直是图像处理领域的研究热点。近年来,随着计算机硬件性能的提高以及对人眼感知特性的不断深入研究,图像合成无论在理论算法还是应用领域都取得了很大的进步。在提取前景目标时,除继续探求高精度的分割外,适应多种图像类型的鲁棒算法也越来越受到重视^[10],不再单纯强调分割质量,注重计算速度与分割效果平衡的实时交互算法^[50]已成为主流研究方向之一。在将前景物体合成到目标图像时,也不再只是实现某一图像特征的无缝合成,而是更强调合成图像在多特征和全局上的无缝性^[37]。除了传统的度量方式外,复杂的感知模型在评估合成图像的质量时越来越接近主观评价。随着理论算法的深入,图像合成的应用领域也在不断扩展,除追求以假乱真的真实感图像合成外,生成特殊风格或特殊用途的非真实感图像合成技术也已展现出其独特魅力。借助可编程图形硬件,图像合成技术在实时应用领域(如视频合成)得到了进一步发展,甚至在图像处理软件中,图像合成功能^[51]已成为其必不可少的模块之一。

虽然图像合成技术已经被广泛应用于日常生活之中,但其潜在的应用方式仍有待于开发,而且现有的合成技术仍存在许多问题,如仍需要较多的用户交互,原图像和目标图像的纹理模式不能相差太大,合成速度受限于图像的分辨率等。更为重要的是,现有的方法主要关注对比度、颜色、纹理等语法层次上的无缝合成,内容上或语义级别的图像合成^[52]仍缺少深入的理论研究,甚至可以认为,图像合成技术还处于起步阶段。对于研究者而言,一个在合成效果、鲁棒性、运算效率、自动性、甚至意义匹配上均能取得高质量结果的方法依然具有很大的挑战性,也仍然值得继续深入研究。

参考文献 (References)

[1] Porter T, Duff T. Compositing digital images [C] // Proceedings of the 11th annual conference on Computer graphics and interactive techniques. New York, USA: ACM, 1984: 253-259.

[2] Wang J, Cohen M F. Image and video matting: a survey [J]. Foundations and Trends in Computer Graphics and Vision, 2007, 3(2): 97-175.

[3] Smith A, Blinn J. Blue screen matting [C] // Proceedings of

ACM SIGGRAPH 1996. New York, USA: ACM, 1996: 259-268.

[4] Ruzon M, Tomasi C. Alpha estimation in natural images [C] // Proceedings of IEEE CVPR 2000. Washington DC, USA: IEEE Computer Society, 2000: 18-25.

[5] Chuang Y, Curless B, Salesin D, et al. A bayesian approach to digital matting [C] // Proceedings of IEEE CVPR 2001. Washington DC, USA: IEEE Computer Society, 2001: 264-271.

[6] Sindeyev M, Konushin V, Vezhnevets V. Improvements of bayesian matting [C] // Proceedings of Graphicon 2007. Moscow, Russia: Moscow State Lomonosov University, 2007: 88-95.

[7] Berman A, Vlahos P, Dadourian A. Comprehensive method for removing from an image the background surrounding a selected object. US, 6135345 [P]. 2000-10-17.

[8] Wang J, Cohen M. An iterative optimization approach for unified image segmentation and matting [C] // Proceedings of IEEE ICCV 2005. Washington DC, USA: IEEE Computer Society, 2005: 936-943.

[9] Weiss Y, Freeman W. On the optimality of solutions of the max-product belief propagation algorithm in arbitrary graphs [J]. IEEE Transaction on Information Theory, 2001, 47(2): 303-308.

[10] Wang J, Cohen M. Optimized color sampling for robust matting [C] // Proceedings of IEEE CVPR 2007. Washington DC, USA: IEEE Computer Society, 2007: 1-8.

[11] Grady L, Schiwietz T. Random walks for interactive alpha-matting [C] // Proceedings of Visualization, Imaging, and Image Processing. Calgary, Alberta, Canada: ACTA, 2005: 423-429.

[12] Grady L. Random walks for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11): 1768-1783.

[13] Rother C, Kolmogorov V, Blake A. Grabcut-interactive foreground extraction using iterated graph cut [C] // Proceedings of ACM SIGGRAPH 2004. New York, USA: ACM, 2004: 309-314.

[14] He X. Locality Preserving Projections [D]. Chicago, USA: University of Chicago, 2005.

[15] Bai X, Sapiro G. A geodesic framework for fast interactive image and video segmentation and matting [C] // Proceedings of IEEE ICCV 2007. Washington DC, USA: IEEE Computer Society, 2007: 16-19.

[16] Yang C, Duraiswami R, Gumerov N, et al. Improved fast gauss transform and efficient kernel density estimation [C] // Proceedings of IEEE ICCV 2003. Washington DC, USA: IEEE Computer Society, 2003: 464-471.

[17] Yatziv L, Bartesaghi A, Sapiro G. O(n) implementation of the fast marching algorithm [J]. Journal of Computational Physics, 2006, 212(2): 393-399.

[18] Zheng Y, Kambhamettu C, Yu J, et al. Fuzzymatte: A computationally efficient scheme for interactive matting [C] // Proceedings of IEEE CVPR 2008. Washington DC, USA: IEEE Computer Society, 2008: 1-8.

[19] Udupa J, Samarasekera S. Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmenta-

- tion[J]. *Graphical Models and Image Processing*, 1996, 58(3): 246-261.
- [20] Sun J, Jia J, Tang C, et al. Poisson matting[C]//*Proceedings of ACM SIGGRAPH 2004*. New York, USA: ACM 2004; 315-321.
- [21] Levin A, Lischinski D, Weiss Y. A closed form solution to natural image matting[C]//*Proceedings of IEEE CVPR 2006*. Washington DC, USA: IEEE Computer Society, 2006; 61-68.
- [22] Levin A, Rav-Acha A, Lischinski D. Spectral matting[C]//*Proceedings of IEEE CVPR 2007*. Washington DC, USA: IEEE Computer Society, 2007; 1-8.
- [23] Shi J, Malik J. Normalized cuts and image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905.
- [24] Yu S, Shi J. Multiclass spectral clustering[C]//*Proceedings of IEEE ICCV 2003*. Washington DC, USA: IEEE Computer Society, 2003; 313-319.
- [25] Grady L, Funka-Lea G. Multi-lable image segmentation for medical application based on graph-theoretic electrical potentials[C]//*Proceedings of Computer Vision and Mathematical Methods in medical and Biomedical Image Analysis*. New York, USA: Springer-Verlag, 2004; 230-245.
- [26] Perez P, Gangnet M, Blake A. Poisson image editing[C]//*Proceedings of ACM SIGGRAPH 2003*. New York, USA: ACM, 2003; 313-318.
- [27] William P. *Numerical Recipes in C++ [M]*. 2nd ed. Beijing: Publishing House of Electronics Industry, 2005; 280-312. [William P. *C++ 数值算法 [M]*. 2 版. 北京: 电子工业出版社, 2005; 280-312].
- [28] Kruger J, Westermann R. Linear algebra operators for gpu implementation of numerical algorithms[J]. *ACM Transactions on Graphics*, 2003, 22(3): 908-916.
- [29] Jiaya J, Sun J, Tang C, et al. Drag-and-drop pasting[C]//*Proceedings of ACM SIGGRAPH 2006*. New York, USA: ACM, 2006; 631-637.
- [30] Goldberg A. A simple shortest path algorithm with linear average time[C]//*Proceedings of the 9th Annual European Symposium on Algorithms*. New York, USA: Springer-Verlag, 2001; 230-241.
- [31] Wu H, Xu D. Improved poisson image editing and its implementation on GPU[C]//*Proceedings of International Conference on Computer-aided Industrial Design & Conceptual Design 2009*. Washington DC, USA: IEEE Computer Society, 2009; 1044-1048.
- [32] Fu X Y, Guo H, Wang Y X, et al. Arbitrary image cloning[C]//*Proceedings of International Conference on Signal Image Technology & Internet Based Systems 2006*. New York, USA: Springer-Verlag, 2006; 644-655.
- [33] Guo D, Sim T. Color me right—seamless image compositing[C]//*Proceedings of the International Conference on Computer Analysis of Images and Patterns 2009*. New York, USA: Springer-Verlag, 2009; 444-451.
- [34] Fattal R, Carroll R, Agrawala M. Edge-based image coarsening[J]. *ACM Transactions on Graphics*, 2009, 29(1): 1-11.
- [35] Wu H, Xu D. Poisson cloning using bilateral image coarsening[C]//*Proceeding of the Second International Conference on Internet Multimedia Computing and Service*. New York, USA: ACM, 2010; 49-52.
- [36] Burt P, Adelson E. A multiresolution spline with application to image mosaics[J]. *ACM Transactions on Graphics*, 1983, 2(4): 217-236.
- [37] Sunkavalli K, Johnson M, Matusik W, et al. Multi-scale image harmonization[C]//*Proceedings of ACM SIGGRAPH 2010*. New York, USA: ACM, 2010, Article No. 125; 1-10.
- [38] Li Y, Sharan L, Adelson H. Compressing and companding high dynamic range images with subband architectures[J]. *ACM Transactions on Graphics*, 2005, 24(3): 836-844.
- [39] Farbman Z, Hoffer G, Lipman Y, et al. Coordinates for instant image cloning[C]//*Proceedings of ACM SIGGRAPH 2009*. New York, USA: ACM, 2009, Article No. 67; 1-9.
- [40] Rhemann C, Rother C, Wang J, et al. A perceptually motivated online benchmark for image matting[C]//*Proceedings of IEEE CVPR 2009*. Washington DC, USA: IEEE Computer Society, 2009; 1826-1833.
- [41] Chu H, Wei H, Niloy J, et al. Camouflage images[C]//*Proceedings of ACM SIGGRAPH 2010*. New York, USA: ACM, 2010, Article No. 51; 1-8.
- [42] Wolfe M. Guided search 2.0: A revised model of visual search[J]. *Psychonomic Bulletin and Review*, 1994, 1(2): 202-238.
- [43] Hertzmann A, Jacobs E, Oliver N, et al. Image analogies[C]//*Proceedings of ACM SIGGRAPH 2001*. New York, USA: ACM, 2001; 327-340.
- [44] Tong S, Tang K, Brown M, et al. Example-based cosmetic transfer[C]//*Proceedings of 15th Pacific Conference on Computer Graphics and Applications*. Washington DC, USA: IEEE Computer Society, 2007; 211-218.
- [45] Guo D, Sim T. Digital face makeup by example[C]//*Proceedings of IEEE CVPR 2009*. Washington DC, USA: IEEE Computer Society, 2009; 73-79.
- [46] Farbman Z, Fattal R, Lischinski D, et al. Edge-preserving decompositions for multi-scale tone and detail manipulation[J]. *ACM Transactions on Graphics*, 2008, 27(3): 1-10.
- [47] Agarwala A. Efficient gradient-domain compositing using quadrees[J]. *ACM Transactions on Graphics*, 2007, 26(3): 1-5.
- [48] Floater M. Mean value coordinates[J]. *Computer Aided Geometric Design*. 2003, 20(1): 19-27.
- [49] Wang H, Raskar R, Ahuja N. Seamless video editing[C]//*Proceedings of ICPR 2004*. Washington DC, USA: IEEE Computer Society, 2004; 858-861.
- [50] Wang J, Agrawala M, Cohen M. Soft scissors: an interactive tool for realtime high quality matting[C]//*Proceedings of ACM SIGGRAPH 2007*. New York, USA: ACM, 2007. Article No. 9; 1-6.
- [51] Georgiev T. Photoshop healing brush: a tool for seamless cloning[C]//*Proceedings of ECCV 2004*. New York, USA: Springer-Verlag, 2004; 1-8.
- [52] Chen T, Cheng M M, Tan P. Sketch2Photo: internet image montage[J]. *ACM Transactions on Graphics*, 2009, 28(5): 1-10.