

## 一种保持颜色的图像合成算法\*

吴昊<sup>+</sup>, 徐丹

(云南大学 计算机科学与工程系, 云南 昆明 650091)

### Color Preserved Image Compositing

WU Hao<sup>+</sup>, XU Dan

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

+ Corresponding author: E-mail: coolbeary@163.com

**Wu H, Xu D. Color preserved image compositing. *Journal of Software*, 2012, 23(Suppl. (2)): 94-104 (in Chinese).** <http://www.jos.org.cn/1000-9825/12029.htm>

**Abstract:** The Gradient-Domain fusion is an effective method for image compositing, but suffers from color distortion when the hue of target image is quietly different from that of source image. The distortion can hardly be eliminated by boundary optimization. This paper proposes a method to correct the artifact of color distortion induced by gradient-domain fusion while keeping the seamless boundary. First, Poisson cloning is applied to luminance component for keeping the local contrast of intensity. Then, for each pixel of source image, a color belief is used to guide color correction and is estimated based on general geodesic distance transform for source image. Next, the color components of each pixel are corrected by minimizing an object function with corresponding color belief, while the object function takes neighborhood colors propagation, the original colors obtained by Poisson cloning are also taken into consideration. The final composites are obtained simply by combining the corrected color and the cloned luminance. Experimental results demonstrate that the proposed method reduces the color distortion of composition effectively while keeping the seamless boundary and requires only few interactions from user. Compared to other gradient-domain based image compositing methods that only optimize boundary conditions, the proposed method keeps both the original color and the relative changes of colors. The use of a soft color belief to compositing provides a smooth transition from foreground to background, which can not be done by hard constraints.

**Key words:** image compositing; Poisson cloning; color correction; color belief; seamless boundary

**摘要:** 基于梯度域的融合是一类有效的图像合成算法,但当目标图像和源图像的色调相差太大时,合成结果往往存在一定程度的颜色失真,且难以通过优化边界条件消除.为此提出了一种能够校正颜色失真并保持无缝边界的图像合成方法,该方法主要分为3个步骤.首先,对亮度分量使用泊松克隆以保持合成图像的局部对比度和无缝边界.其次,估计每个像素的颜色置信度以引导颜色校正,颜色置信度基于通用测地距离估计得到.最后,通过最小化目标函数对合成图像的颜色分量进行校正,目标函数不仅结合了源图像和经过泊松克隆后的图像的颜色信息,还考虑了邻域像素间的颜色扩散,并利用颜色置信度为其各项加权.最终的合成图像通过合并泊松克隆后的亮度分量和校正

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

收稿时间: 2012-05-20; 定稿时间: 2012-09-29

后的颜色分量得到.实验结果表明,通过少量简单、直观的用户交互,所提出的图像合成方法不仅能保持无缝的合成边界,而且能有效地减少合成图像中的颜色失真,在合成颜色差异较大的源图像和目标图像时也具有较高的鲁棒性.与其他仅仅优化边界条件的基于梯度域的合成方法相比,该方法能够更好地保持前景物体的原有颜色和局部对比度,使用颜色置信度而非硬性的限制条件也使得前景能够更平滑地融合到背景中.

**关键词:** 图像合成;泊松克隆;颜色校正;颜色置信度;无缝边界

图像合成通常是指在输入图像(源图像)中提取出感兴趣的区域(region of interest,简称 ROI)并将其无缝地合成到目标图像中.对于理想的图像合成算法而言,“无缝性”是其最重要的目标之一,它要求合成图像在亮度、颜色、纹理、边界、噪声模式等多个方面具有一致性,从而使用户难以感知到其中的合成痕迹.虽然以手工方式逐像素地修改 ROI 可以产生“无缝”的合成图像,但这不仅需要用户有丰富的经验,而且效率低下,非常耗时,不断重复修改的乏味性也可想而知.此外,当面对含有数百万像素的 ROI 时,以手工方式合成图像将变得不切实际,而这种 ROI 在高分辨率图像中相当常见.因此,自动的或只需少量用户交互的高质量合成算法成为了图像合成的重要研究方向,其中,基于图像分割<sup>[1,2]</sup>或抠图<sup>[3,4]</sup>的图像合成是最常用的一类方法.此类方法通常首先从源图像中提取出前景,然后将前景直接覆盖到新的目标图像上(基于二值图像分割的合成),或者先计算表征各像素属于前景程度的 $\alpha$ 值,再根据 $\alpha$ 值将前景与新的背景线性混合(基于抠图的合成).虽然基于 $\alpha$ 值进行图像合成很有效,但求取 $\alpha$ 值的抠图是一个“病态(ill-posed)”问题,为了获得合理的解,通常需要做出额外的限制条件.因此,当假设条件得不到满足时(如前景与背景的颜色分布相似或源图像的 ROI 中含有高对比度的复杂纹理),此类方法难以产生理想的合成结果<sup>[5]</sup>.

另一类图像合成方法则基于对梯度域的操作,如泊松克隆<sup>[6]</sup>,与直接操作像素值的方法不同,泊松克隆通过求解泊松方程,将 ROI 的梯度场拟合为给定的向量场来完成图像合成.如果使用源图像的梯度场作为给定向量场并以对应的目标图像边界上的值为边界条件,泊松克隆不仅能保持合成图像在边界上的无缝性.而且能使合成图像中 ROI 的灰度的局部变化与源图像的变化相一致.基于梯度域操作的图像合成也能很好地保持 ROI 中颜色的对比度,但只要源图像和目标图像在 ROI 的边界上颜色差异较大,该类方法会明显地改变合成区域中的色调,从而导致合成图像中前景物体颜色的失真,其失真程度与源图像和目标图像间的颜色差异成正比.

为解决图像合成中的颜色失真问题,本文提出了一种既能将 ROI 无缝地合成到目标图像中,又能使 ROI 中的前景物体保持原有颜色的图像合成方法.与泊松克隆相比,该方法并未以相同的方式处理图像的 3 个颜色通道,而是首先将图像划分为亮度分量和颜色分量,然后再对其应用泊松克隆,经泊松克隆后的颜色分量会存在失真,为校正颜色失真,本方法首先估计各像素颜色分量的置信度,然后以置信度为权重最小化目标函数以校正颜色失真,校正后的颜色分量将与泊松克隆后的亮度分量重组得到最终的合成结果.除具有泊松克隆的优点以外,本文的方法将估计得到的颜色置信度用于优化颜色校正函数,因此,在合成结果中有效地减少了颜色失真,从而使前景能够平滑、自然地融合到背景中,如图 1 所示,其中,图 1(a)表示用户输入的笔划,鸳鸯身上的笔划指示了用户希望在合成结果中保持原来颜色的大致区域,水域中的笔划指示了无需保持颜色的大致区域.图 1(c)表示泊松克隆,可看到鸳鸯不自然的颜色(颜色失真).图 1(d)表示本文的方法,消除了颜色失真并继续保持无缝边界.

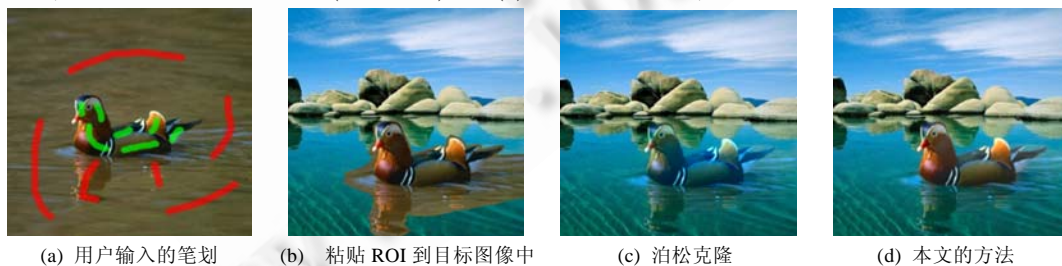


Fig.1 Color preserved image compositing

图 1 保持颜色的图像合成

相比已有的图像合成方法,本文方法的主要贡献和特点有:(1) 一个新的用于校正图像合成中的颜色失真的模型。(2) 一种鲁棒而准确地估计颜色置信度的方法,该颜色置信度可有效用于图像合成中的颜色校正。(3) 一个结合了泊松克隆和颜色校正的图像合成框架,在保证合成图像的边界无缝性和局部对比度的同时,又能合理地保持前景物体的原有颜色。

## 1 相关工作

当 ROI 的边界穿过目标图像中灰度变化强烈的区域(或具有较强梯度值的区域)时,泊松克隆会在该区域产生模糊的光晕,这被称作“渗色效应(bleeding effect)”,“渗色效应”是除颜色失真外泊松克隆存在的另一个重要缺陷.Jiaya 等人<sup>[7]</sup>通过在 ROI 的边界和前景物体边界之间寻找一条最优合成边界来减少“渗色效应”,最优边界削弱了源图像和目标图像在边界上的能量变化,从而减少了在强对比度区域求解泊松方程的错误,最终有效地减少了合成图像中的“渗色效应”。为进一步解决该问题,Tao 等人<sup>[8]</sup>首先减少边界上的旋度,使 ROI 中的梯度场变为近似可积分的,然后以加权积分的方式在调整后的梯度场上重构出合成图像,这不仅减少了求解过程中的错误残差,还能有效地将残差分布到部分复杂纹理区域中,从而使“渗色效应”难以被感知.Chen<sup>[9]</sup>等人通过缝合多幅图像中的 ROI 来实现图像合成,待缝合的多幅图像并不由用户直接给出,而是要求用户先输入待合成物体的草图和对应的意义文本,然后算法自动地在网络上寻找与之最匹配的图像。在合成图像时,该方法首先以过分割算法将 ROI 划分为超像素(一个超像素代表一个分割区域)并估计各超像素在源图像和目标图像上的相似性,然后基于该相似性,对各超像素设定不同的边界条件进行泊松克隆,最后结合抠图将前景物体合成到目标图像中。当 ROI 中的纹理与目标图像相似时,这种混合式的方法能得到高质量的合成结果,但其计算复杂度也较大。

除保持无缝边界和局部对比度之外,使合成图像中 ROI 的纹理和噪声模式与目标图像一致也十分重要.Sunkavalli 等人<sup>[10]</sup>使用多分辨率技术来解决该问题,他们首先在不同尺度下对源图像和目标图像的纹理和噪声进行直方图匹配,然后通过平滑的增益控制来修改每个尺度下的子带系数,从而改善匹配纹理和噪声的直方图时所引起的信号失真现象,最后将抠图和泊松克隆的边界条件相结合,并重构修改后的子带系数得到合成图像。以这种多尺度方法得到的合成图像在对比度、纹理和噪声模式等方面均能与目标图像保持一致,而且具有较强的鲁棒性。但如果 ROI 中的纹理或目标图像的纹理是非随机的,那么该方法将不再适用。

在减少合成算法的时间复杂性上,Farbman 等人<sup>[11]</sup>作了有效的尝试,他们通过在 ROI 上从边界向内插值,将泊松克隆转化为均值坐标(mean-value coordinates,简称 MVC)插值,有效地缩短了运算时间。为了进一步提高 MVC 插值的速度,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(adaptive mesh),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,MVC 插值的计算速度比泊松克隆有了很大提高,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序列也能做到实时合成。但 MVC 插值在本质上是泊松克隆的快速替代算法,因此图像的合成质量并未得到改善。

另一个多 ROI 的合成算法由 Agarwala 等人<sup>[12]</sup>提出,他们基于 graph-cut<sup>[13]</sup>为每一个 ROI 在目标图像中寻找对应的最佳缝合线,为了进一步减少缝合线痕迹,该算法也使用了类似泊松克隆的梯度域合成算法(除将狄利克雷(Dirichlet)边界条件替换为诺伊曼(Neumann)边界条件)。虽然该算法在一定程度上能保持 ROI 原有的颜色,但其难以确保为每个 ROI 都找到最佳缝合线,这导致在部分 ROI 边界周围形成明显的可见缝隙。此外,graph-cut 算法复杂度较高,难以做到实时合成。

上述工作在一定程度上都对泊松克隆进行了扩展和改进,但并未有效地解决其颜色失真问题.Guo 等人<sup>[14]</sup>提出了一种基于用户标记的方法,该方法需要用户标记 ROI 中的特定像素以保证合成后被标记像素的颜色值不会发生改变,这相当于除边界条件外,在标记像素上引入额外的硬性限制条件。该限制条件和适当的像素权重结合在一起可以控制错误残差的分布,进而根据需要控制合成图像对源图像颜色的保留程度,有助于减轻泊松克隆中的颜色失真,但这也可能导致被标记像素周围产生更多的错误残差和“渗色效应”。Lalonde 等人<sup>[15]</sup>预先建立

了一个基于图像的前景物体库,然后从该图像库中检索符合特定特征的前景物体作为 ROI 合成到目标图像中,这些特征包括相机位置、光照、物体的实际大小等.为减小“渗色效应”与颜色失真,他们使用二值混合掩膜(binary blending mask)来协助合成,当掩膜覆盖的区域与前景物体一致时,该区域保留不变,以外的区域则使用带权泊松克隆进行合成,二值混合掩膜实际上也是一种硬性的限制条件,它强制前景物体不参与到合成运算中,其缺陷是当源图像和目标图像的光照条件相差太大时,合成图像会因为前景物体与背景的亮度反差显得不够自然.此外,从有限的前景物体库中检索 ROI 在很大程度上限制了合成的多样性.本文的方法并未使用硬性的限制条件,而是采用了软限制条件(即颜色置信度)来校正颜色,结合闭合形式的颜色校正函数,所提出的方法能够更好地缓解颜色失真,而且不引入多余的“渗色效应”.

## 2 流程概述

本文方法的主要任务是生成具有无缝边界的合成图像,并使其中的前景物体保持适当的原有颜色.该方法可通过 3 个步骤实现,即亮度克隆、估计颜色置信度以及颜色校正.

在亮度克隆阶段,源图像和目标图像将首先被分为亮度分量和颜色分量,然后用泊松克隆将源图像 ROI 中的亮度分量合成到目标图像的亮度分量上,经泊松克隆后的亮度分量能很好地保持 ROI 中的局部对比度,同时形成无缝的合成边界.而对于颜色分量,直接使用泊松克隆会产生颜色失真,特别是当 ROI 中的前景物体的颜色与目标图像相差较大时,颜色失真程度会更加严重.为了获得正确的颜色分量(尤其是前景物体的颜色分量),需要估计颜色的失真程度,换言之,需要知道 ROI 中哪些像素需要保留其原来的颜色以及保留的程度.为此,在估计颜色置信度阶段,本文提出了一种基于通用测地距离<sup>[16]</sup>来估计各像素的颜色需要被保留的程度的方法,估计得到的颜色置信度将用于后续的颜色校正.在最后阶段,通过最小化一个闭合形式的目标函数来校正泊松克隆后的颜色分量,该目标函数主要由两项组成,即颜色保真项和颜色传播项,在上一步中得到的颜色置信度将用于产生该目标函数的权重.最终的合成图像则通过重新合并泊松克隆后的亮度分量和校正后的颜色分量而得到.图 2 给出了本文方法的流程图.

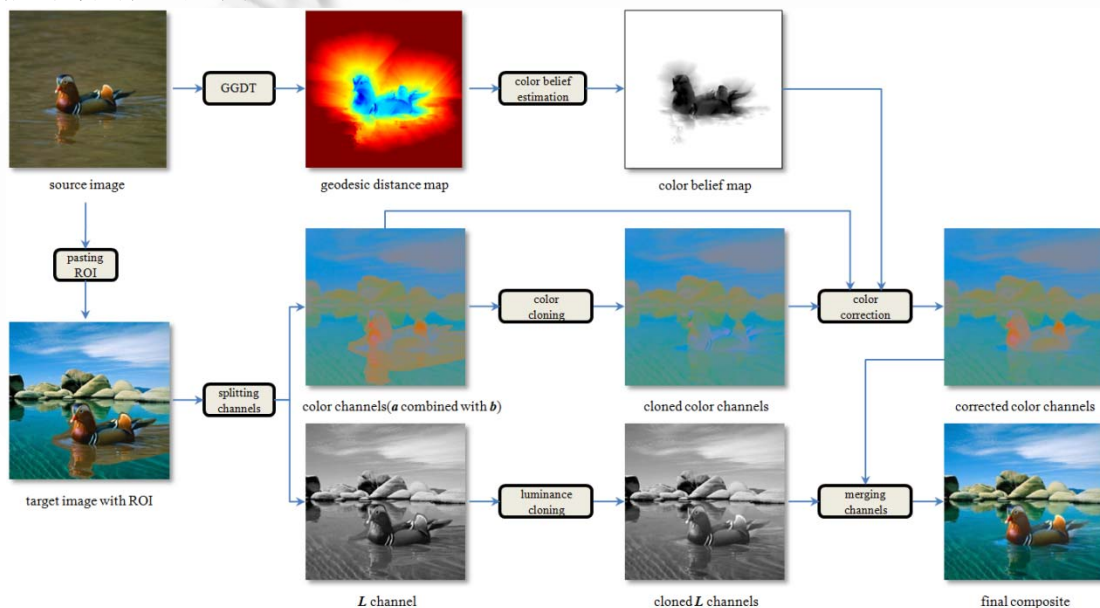


Fig.2 The workflow of our method

图 2 本文方法的流程图

### 3 算法

给定目标图像  $I_t$ 、源图像  $I_s$  以及用户在源图像上指定的 ROI  $\Omega_s$ , 本文的方法无缝地将  $\Omega_s$  合成到目标图像  $I_t$  上, 同时合理地保持  $\Omega_s$  中前景物体的原有颜色及其局部对比度, 下面将详细论述实现该方法的关键步骤。

#### 3.1 亮度克隆

为分别处理亮度分量和颜色分量, 输入图像首先被转换到 CIE-Lab 颜色空间. 源图像  $I_s$  的 3 个颜色通道记为  $L_s, a_s, b_s$ , 目标图像  $I_t$  的颜色通道记为  $L_t, a_t, b_t$ . 在合成  $\Omega_s$  到目标图像时, 应保持其局部对比度, 这是因为相比全局亮度上的缓慢改变, 人类的视觉系统对局部对比度的变化更为敏感<sup>[11]</sup>. 此外, 合成结果在  $\Omega_s$  边界上的值应尽可能地与目标图像一致, 以避免由亮度变化导致明显的合成边界, 当源图像的亮度分量  $L_s$  与目标图像的亮度分量  $L_t$  相差较大时, 合成边界十分明显. 为了有效解决上述两个问题, 本文对亮度分量使用泊松克隆来保证无缝边界和局部对比度, 亮度分量的合成可以表达为如下形式:

$$\arg \min_{L_c} \iint_{\Omega_s} \|\nabla L_c - \nabla L_t\|^2 \quad (1)$$

并使用 Dirichlet 边界条件:

$$L_c|_{\partial\Omega_s} = L_t|_{\partial\Omega_s} \quad (2)$$

式(1)中  $L_c$  为所求泊松克隆后的亮度分量,  $\nabla$  为梯度算子. 式(2)中  $\partial\Omega_s$  表示 ROI 的边界. 虽然可以将泊松克隆应用到颜色分量来获得彩色的合成图像, 但如果  $\Omega_s$  中的色调和目标图像  $I_t$  相差太大, 则合成结果会使  $\Omega_s$  中的色调偏向目标图像的色调(如图 1(c)所示). 泊松克隆本质上是构造一个谐函数, 将源图像和目标图像在  $\partial\Omega_s$  上的差异平滑地插值到  $\Omega_s$  内部, 这是泊松克隆引起颜色失真的内在原因. 为了更清楚地说明对颜色分量直接使用泊松克隆的缺点, 可以假设一种特殊情况, 即边界  $\partial\Omega_s$  上每个像素的颜色分量在源图像和目标图像间的差异为常数  $K_a, K_b$  (下标分别表示 CIE-Lab 颜色空间中的  $a, b$  两个颜色通道). 在该假设下, 泊松克隆等价于对  $\Omega_s$  中的每个像素的颜色分量简单增加常量  $K_a, K_b$ . 当  $K_a, K_b$  很小时, 合成结果的色调与源图像相比并不会会有明显的变化. 随着  $K_a, K_b$  的增加, 合成图像的色调会逐渐偏离源图像, 最终导致在感知上有较大的颜色失真. 虽然优化边界条件可以使边界  $\partial\Omega_s$  上的能量变化接近常数<sup>[7]</sup>, 但能量变化为常数并不能保证该常数值很小, 因此仅仅优化边界条件虽有助于减少“渗色效应”, 但并不能有效地消除颜色失真. 为此, 本文提出了一种校正泊松克隆中颜色失真的方法, 它通过估计颜色置信度来调整颜色分量.

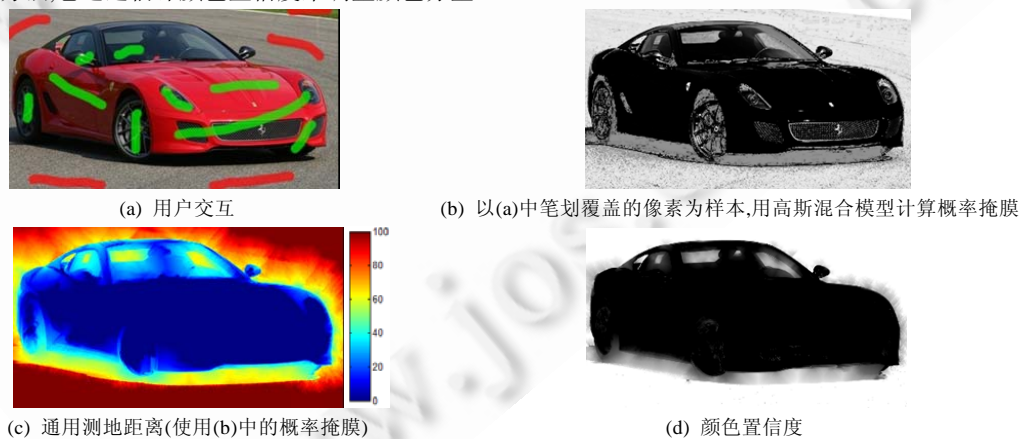


Fig.3 Color belief estimation

图 3 估计颜色置信度

在图 3 中, 图 3(a)表示用户交互, 前景样本像素由车体上笔划所覆盖, 背景样本像素由车周围笔划所覆盖; 图 3(b)表示以图 3(a)中笔划覆盖的像素为样本, 用高斯混合模型计算概率掩膜, 较黑的像素属于背景的概率更低; 图 3(c)表示通用测地距离(使用图 3(b)中的概率掩膜); 图 3(d)表示颜色置信度. 较黑的像素经泊松克隆后会有较

高的颜色失真.

### 3.2 估计颜色置信度

颜色置信度测度了泊松克隆后的颜色分量的可信程度,对于每个像素  $\mathbf{x}$ ,颜色置信度记为  $B(\mathbf{x}) \in [0,1]$ ,在 ROI 中的前景物体上的颜色失真通常比其他区域更容易被感知,因此,颜色置信度应正比于像素  $\mathbf{x}$  到前景物体的距离,即  $B(\mathbf{x}) \propto D(\mathbf{x})$ .换言之,当泊松克隆后颜色失真时,前景物体上的颜色置信度较低,而置信度会随像素到前景的距离增加而增大.此外,颜色置信度在相同特征的区域应是平滑的,从而使颜色分量能自然、平滑地从前景物体过渡到其他区域.因此,本文将颜色置信度建模为对距离函数  $D(\mathbf{x})$  的平滑插值,形式如下:

$$B(\mathbf{x}) = 3 \left( \frac{D(\mathbf{x}) - D_{\min}}{D_{\max} - D_{\min}} \right)^2 - 2 \left( \frac{D(\mathbf{x}) - D_{\min}}{D_{\max} - D_{\min}} \right)^3, \quad \mathbf{x} \in \Omega_s \quad (3)$$

式(3)中,  $D_{\min}$ ,  $D_{\max}$  分别为  $D(\mathbf{x})$  的最小值和最大值.颜色置信度的估计准确度与距离函数  $D(\mathbf{x})$  紧密相关.简单的距离变换,如欧氏距离,并不能用于有效地估计颜色置信度,因为它只度量了像素间的空间距离而并未考虑,如颜色距离等其他因素.本文采用了更有效的距离变换方式,即通用测地距离变换(general geodesic distance transform,简称GGDT)<sup>[16]</sup>来计算  $D(\mathbf{x})$ .除空间距离和颜色距离外,GGDT引入了一个实值掩膜  $M_s(\mathbf{x}) \in [0,1]$  来测度每一个像素属于非前景区域的概率,其形式化的表达如下所示:

$$D_{\text{GGDT}}(\mathbf{x}, M_s, I_s) = \min_{\mathbf{x}' \in \Omega_s} (d(\mathbf{x}, \mathbf{x}') + \nu M_s(\mathbf{x}')) \quad (4)$$

式(4)中,自由参数  $\nu$  调整路径距离与实值掩膜间的相对重要程度.而路径距离  $d(\mathbf{x}, \mathbf{x}')$  可按下式定义:

$$d(\mathbf{x}, \mathbf{x}') = \min_{\Gamma \in P_{\mathbf{x}, \mathbf{x}'}} \int_0^1 \sqrt{\|\Gamma'(s)\|^2 + \gamma^2 (\nabla I_s \cdot \mathbf{u})^2} ds \quad (5)$$

式(5)中,  $P_{\mathbf{x}, \mathbf{x}'}$  是像素  $\mathbf{x}$  到像素  $\mathbf{x}'$  的所有路径的集合,集合中的每条路径都可由参数化函数  $\Gamma(s)$ ,  $s \in [0,1]$  来描述.  $\Gamma'(s) = \frac{\partial \Gamma(s)}{\partial s}$ , 而  $\mathbf{u} = \frac{\Gamma'(s)}{\|\Gamma'(s)\|}$  表示正切于路径  $\Gamma(s)$  的单位向量.图像梯度对空域距离的相对贡献由参数  $\gamma$  决定.与常用的测地距离相比,GGDT 引入了实值掩膜  $M_s(\mathbf{x})$  来将概率分布引入距离变换中,在本文中  $M_s(\mathbf{x})$  由下式得到:

$$M_s(\mathbf{x}) = g \left( \ln \frac{\Pr(I(\mathbf{x}) | \alpha(\mathbf{x}) = \text{BG})}{\Pr(I(\mathbf{x}) | \alpha(\mathbf{x}) = \text{FG})} \right) \quad (6)$$

式(6)中,  $g(t) = \frac{1}{1 + \exp(-t/\sigma)}$ , 文中的实验均设  $\sigma = 4.5$ , 而  $\Pr(I(\mathbf{x}) | \alpha(\mathbf{x}) = \text{BG})$  和  $\Pr(I(\mathbf{x}) | \alpha(\mathbf{x}) = \text{FG})$  分别为像素  $\mathbf{x}$  的背景颜色概率和前景颜色概率,  $\alpha(\mathbf{x})$  是二值标记,用于表示  $\mathbf{x}$  属于背景(BG)或前景(FG).与 Grabcut<sup>[2]</sup> 估计背景和前景的颜色概率的方法类似,本文也将这两个概率分布表示为两个不同的高斯混合模型(Gaussian mixture model,简称GMM),模型的参数则分别从用户笔划所覆盖的样本像素中估计得到(用户提供的笔划分为背景笔划和前景笔划,如图3(a)所示).高斯混合模型对噪声不敏感,在背景的颜色分布和前景的颜色分布近似时有较强的鲁棒性,因此,比使用直方图累积技术<sup>[16]</sup>能更准确地估计背景和前景的颜色概率分布.

分析式(4)可知,GGDT 实际上是在寻找一条从像素  $\mathbf{x}$  到其他像素的能量最小的路径,该能量由路径距离  $d(\mathbf{x}, \mathbf{x}')$  和实值掩膜  $M_s(\mathbf{x})$  构成.换言之,GGDT 是在考虑了空间距离、颜色距离及颜色概率分布下的最短路径.由于概率掩膜  $M_s(\mathbf{x})$  提供了更多像素间关系的信息,GGDT 在图像分割方面已显示出较强的准确性和鲁棒性,此优点也同样体现在对颜色置信度的估计中.在源图像上使用通用测地距离变换后,有  $D(\mathbf{x}) = D_{\text{GGDT}}(\mathbf{x}, M_s, I_s)$ , 这时,各像素的颜色置信度  $B(\mathbf{x})$  可通过式(3)计算得到.求得的  $B(\mathbf{x})$  将用于引导后续的颜色校正过程.图3给出颜色置信度的估计.

### 3.3 颜色校正

该步骤基于颜色置信度来校正由泊松克隆引起的颜色失真,颜色校正过程被建模为最小化校正颜色分量

的目标函数,目标函数由颜色保真项和颜色传播项组成,形式如下:

$$E = E_{cf} + \lambda E_{cp} \quad (7)$$

**颜色保真项  $E_{cf}$ .** 当 ROI 中的像素具有高颜色置信度时,说明泊松克隆后的颜色分量失真程度较小,此时应保留其颜色特征.反之,在低置信度区域泊松克隆后的颜色分量可信度低,应参考源图像的颜色分量进行校正.为方便求解,  $E_{cf}$  以 2 次函数的形式定义.

$$E_{cf} = \sum_{\mathbf{x} \in \Omega_s} \left( (1 - B(\mathbf{x})) \|C_r(\mathbf{x}) - C_s(\mathbf{x})\|^2 + B(\mathbf{x}) \|C_r(\mathbf{x}) - C_p(\mathbf{x})\|^2 \right) \quad (8)$$

式(8)中,源图像的  $C_r(\mathbf{x})$ ,  $C_s(\mathbf{x})$ ,  $C_p(\mathbf{x})$  分别表示待求解的校正后的颜色分量、源图像的颜色分量以及泊松克隆后的颜色分量(颜色分量为  $a, b$  两个颜色通道构成的向量).  $B(\mathbf{x})$  为上一步估计出的颜色置信度.

**颜色传播项  $E_{cp}$ .** 如果像素的颜色置信度处于中间值区域,校正后的颜色应同时具有源图像和泊松克隆后的颜色分量的特性.但简单地按颜色置信度线性混合  $C_s(\mathbf{x})$  和  $C_p(\mathbf{x})$  并不能产生合适的颜色分量,这是因为颜色置信度并未考虑颜色的局部一致性,即颜色在一个小邻域内可能是平滑一致的,而如果邻域跨过边缘,则颜色会有较大的变化.因此,应该考虑各像素邻域内的颜色对其颜色校正的影响,与线性混合相比,更好的方法是将具有极端值颜色置信度(较高或较低的颜色置信度)的像素的颜色分量向中间值颜色置信度的区域传播.换言之,当像素属于 ROI 中的前景与属于背景的程度相近时,  $C_r(\mathbf{x})$  应更多地由其邻域像素的颜色所决定.所以,本文将颜色传播项建模为

$$E_{cp} = \sum_{\mathbf{x} \in \Omega_s} B_1(\mathbf{x}) \left\| C_r(\mathbf{x}) - \sum_{\tilde{\mathbf{x}} \in N(\mathbf{x})} w(\mathbf{x}, \tilde{\mathbf{x}}) C_r(\tilde{\mathbf{x}}) \right\|^2 \quad (9)$$

式(9)中,  $N(\mathbf{x})$  为像素  $\mathbf{x}$  的相邻像素的集合.  $w(\mathbf{x}, \tilde{\mathbf{x}})$  是定义在像素  $\mathbf{x}$  与其相邻像素  $\tilde{\mathbf{x}}$  的相似性上的权重函数,具体定义为考虑了空间距离和颜色距离的仿射函数.

$$w(\mathbf{x}, \tilde{\mathbf{x}}) = \exp\left(\frac{-\|I_s(\mathbf{x}) - I_s(\tilde{\mathbf{x}})\|^2}{\sigma_1^2}\right) \exp\left(\frac{-\|\mathbf{x} - \tilde{\mathbf{x}}\|^2}{\sigma_2^2}\right) \quad (10)$$

式(10)中,  $\sigma_1$  为邻域集合  $N(\mathbf{x})$  中所有像素的颜色值所估计出的标准差,设  $\sigma_2 = 2.5$ ,  $B_1(\mathbf{x})$  用于加权颜色传播项,定义为关于颜色置信度的函数.

$$B_1(\mathbf{x}) = \exp\left(\frac{-|1 - 2B(\mathbf{x})|}{\sigma_3^2}\right) \quad (11)$$

固定式(11)中的  $\sigma_3 = 0.6$ ,显然,当像素的颜色置信度  $B(\mathbf{x})$  为中间值时,  $B_1(\mathbf{x})$  会具有较大的值,从而使颜色传播项在目标函数中占主导地位.本文所使用的颜色传播模型类似于图像上色<sup>[17]</sup>和闭合形式抠图<sup>[18]</sup>的方法,它们都在像素  $\mathbf{x}$  的小范围邻域内建立一个线性函数,并通过在重叠的邻域窗口上全局优化目标函数,使得像素的颜色值向周围其他像素传播,当局部窗口内像素间的线性关系假设成立时,该颜色传播模型能合理地扩散颜色分量,并具有较好的边缘意识性(edge-aware).

在式(7)中,  $\lambda$  控制颜色传播项和颜色保真项之间的相对重要程度,本文设  $\lambda = 1.2$ , 使目标函数偏向以颜色传播为主.由于  $E_{cf}$  和  $E_{cp}$  均是以二次函数的形式定义,目标函数  $E$  的最小化可通过求解一个稀疏线性系统来完成.需要强调的是,在实现算法时,并不需要校正图像中所有像素的颜色,而只需校正 ROI 中的像素,最终校正后的颜色分量  $C(\mathbf{x})$  表示为

$$C(\mathbf{x}) = \begin{cases} \arg \min_{C_r(\mathbf{x})} E, & \mathbf{x} \in \Omega_s \\ C_p(\mathbf{x}), & \text{otherwise} \end{cases} \quad (12)$$

最终的合成图像则通过重组泊松克隆后的亮度分量  $L_c(\mathbf{x})$  和本文方法校正后的颜色分量  $C(\mathbf{x})$  得到.

#### 4 实验结果

通过少量的用户交互,采用本文方法生成的合成图像具有较少的颜色失真和无缝的边界,图 4 对本文方法

和其他合成方法的结果进行了比较.在图 4(b)中,除在 ROI 的边界附近有明显的“渗色效应”外,前景物体(摩托车手)呈现出不自然的颜色,明显地偏离了源图像的色调,存在较严重的颜色失真现象.虽然通过优化边界条件<sup>[7]</sup>可以消除“渗色效应”,但如图 4(c)所示,前景物体仍然存在颜色失真现象.图 4(d)为基于抠图的图像合成,虽然该方法中前景颜色的失真程度较低,但前景周围部分区域的颜色(摩托车以下的区域)并未能自然地融合到目标图像中,而且存在较生硬的合成边界.此外,由于目标图像和源图像的亮度差异较大,导致合成图像中前景的亮度与背景不符.通过对像素标记来附加硬性的限制条件<sup>[14]</sup>虽然在一定程度上减少了颜色失真,但细节部分仍存在颜色失真(图 4(e)中摩托车手的衣服和轮胎部分),而且在轮胎附近有轻微的“渗色效应”(硬限制条件引入的错误残差).从图 4(f)可以看出,本文的方法不仅更合理地保持了前景物体的原有颜色,而且前景到背景的颜色转变也平滑、自然,亮度特性也与目标图像一致.

图 5 比较了源图像中 ROI 的色调与目标图像相差较大时的合成结果,泊松克隆明显地改变了前景物体的色调(偏向不自然的颜色),而本文的方法消除了颜色失真,使前景物体自然地融合到背景中并继续保持无缝的合成边界.图 6 中,本文的方法除减少颜色失真外,对阴影部分和透明区域(车窗)也做出了合适的调整(并非对源图像简单的复制,而是将颜色置信度和颜色传播模型相结合),使其与目标图像保持一致.图 7 则展示了本文的方法不仅能有效地校正颜色失真,而且能合理地调整 ROI 的亮度,使其与新背景一致(图 7(d)中湖水的亮度比在源图像中要暗,这与周围沙漠的亮度相符).

值得注意的是,在所有的示例中,本文的方法仅需要少量的用户笔划,但更多的笔划的确有助于更准确地估计颜色置信度,从而更有效地引导颜色校正.因此,本方法可以通过不断地在细节区域增加笔划以增量式地修改合成结果.

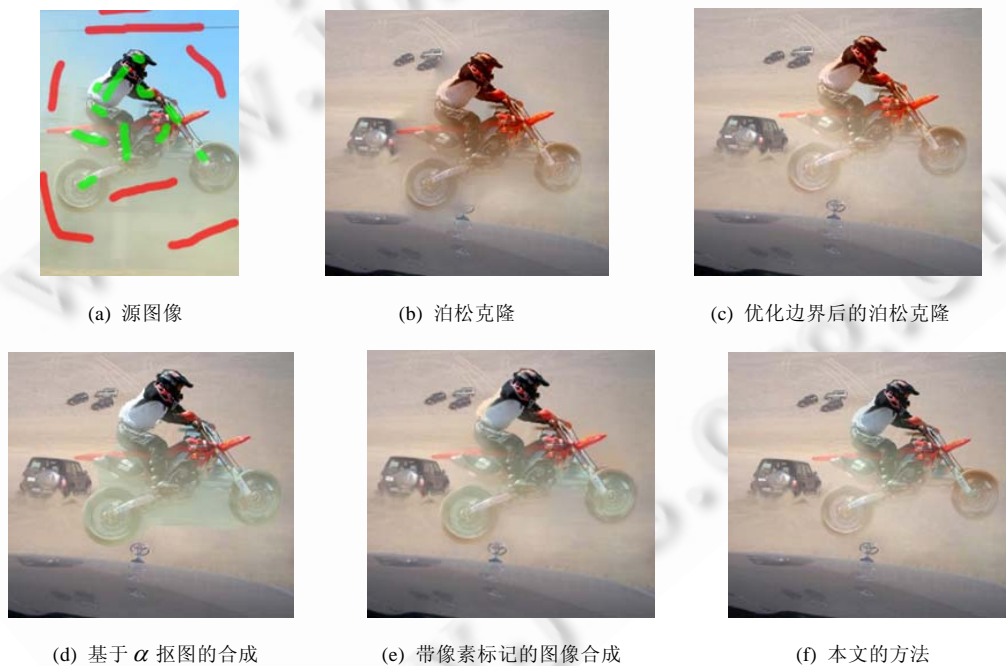


Fig.4 Comparison of our method with other compositing methods

图 4 与其他方法的比较





Fig.5 Example of color correction

图 5 校正颜色失真



Fig.6 Example of shadow and transparent area correction

图 6 阴影和透明区域的校正

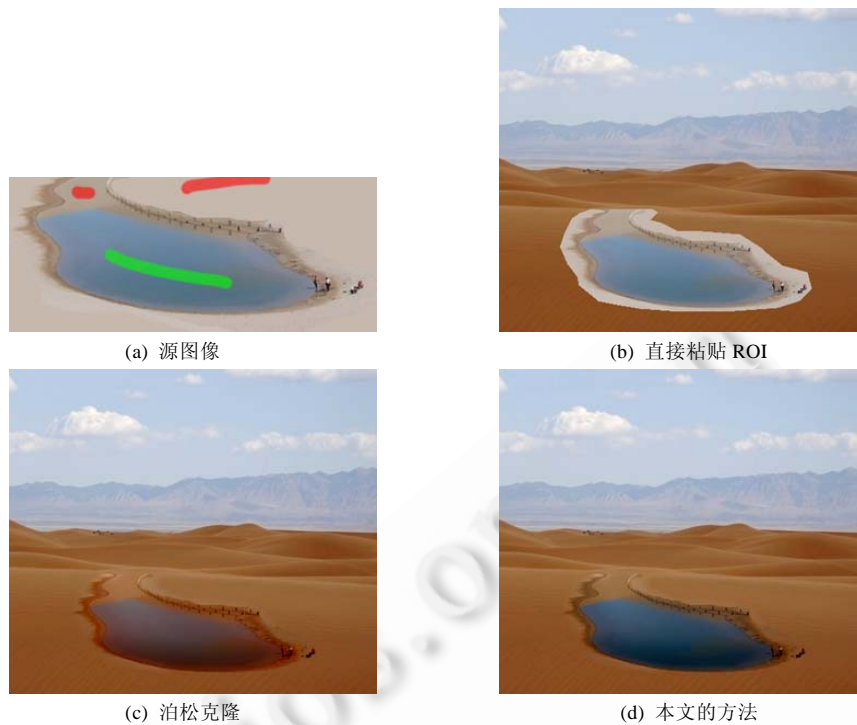


Fig.7 Example of keeping the luminance of foreground be consistent with that of target image

图 7 保持前景亮度与目标图像一致

## 5 结 论

本文提出了一种无缝的图像合成方法,在保持无缝边界和亮度的局部对比度的情况下,使 ROI 中的前景物体能合理地保留其原有颜色.图像合成通常会对各颜色通道进行相同的操作,但本文将图像分为亮度分量和颜色分量,并以不同的方式进行合成.为保持 ROI 中的局部对比度并使合成后的图像在边界上是无缝的,亮度分量直接以泊松克隆的方式合成到目标图像上.直接泊松克隆颜色分量则会导致颜色失真,这种失真在源图像和目标图像的色调相差太大时十分明显,因此,本文提出了通过最小化目标函数来校正颜色失真的方法,用于颜色校正的目标函数由颜色保真项和颜色传播项构成,颜色保真项的作用是从原图像中保留合理的颜色分量,而颜色传播项则负责对邻域像素传播的合理颜色分量.为合理地引导颜色校正过程,本文基于通用测地距离来估计颜色置信度,颜色置信度测度了泊松克隆后的颜色分量的可信程度,能可靠地用于对颜色保真项和颜色传播项加权.实验结果表明,与泊松克隆或其他优化边界的基于梯度域的合成算法相比,本文的合成方法不仅能保持无缝边界,而且能有效地减少 ROI 中前景物体的颜色失真,使前景物体的颜色能自然地过渡到周围其他区域.即使源图像和目标图像的色调相差较大,该方法在保持前景合理的颜色方面也体现出了较强的鲁棒性.在用户交互方面,除指定 ROI 外,只需提供少量的笔划来指定希望保留颜色的部分,非常简单和直观.

但是本文的方法也存在以下不足,第一,当 ROI 中的前景物体到周围区域的边缘不明显时,估计出的颜色置信度误差较大,这导致了周围区域的颜色会错误地渗入到前景物体上.第二,本文方法只试图解决图像合成中的颜色失真问题,不能有效地处理源图像的 ROI 和目标图像在纹理上的不一致性.后续的工作包括对 ROI 的深度估计(用于改变前景物体的尺寸)、匹配 ROI 与目标图像的纹理以及分解图像的光照和反射率等,通过将其引入合成算法中来生成感知上更为真实的合成图像.

## References:

- [1] Li Y, Sun J, Tang C, Shum H. Lazy snapping. In: Hart J, ed. Proc. of the SIGGRAPH. New York: ACM, 2004. 303–308. [doi: 10.1145/1186562.1015719]
- [2] Rother C, Kolmogorov V, Blake A. Grabcut-Interactive foreground extraction using iterated graph cut. In: Hart J, ed. Proc. of the SIGGRAPH. New York: ACM, 2004. 309–314. [doi: 10.1145/1186562.1015720]
- [3] Chuang Y, Curless B, Salesin D, *et al.* A Bayesian approach to digital matting. In: Proc. of the CVPR. Washington: IEEE Computer Society, 2001. 264–271. [doi: 10.1109/CVPR.2001.990970]
- [4] Wang J, Cohen M. Simultaneous matting and compositing. In: Proc. of the CVPR. Washington: IEEE Computer Society, 2007. 1–8. [doi: 10.1109/CVPR.2007.383079]
- [5] Wang J, Cohen M. Image and video matting: A survey. Foundations and Trends in Computer Graphics and Vision, 2007,3(2):1–78. [doi: 10.1561/06000000019]
- [6] Perez P, Gangnet M, Blake A. Poisson image editing. In: Proc. of the SIGGRAPH. New York: ACM, 2003. 313–318. [doi: 10.1145/1201775.882269]
- [7] Jiaya J, Sun J, Tang CK, Shum HY. Drag-and-Drop pasting. In: Proc. of the SIGGRAPH. New York: ACM, 2006. 631–637. [doi: 10.1145/1179352.1141934]
- [8] Tao M, Johnson M, Paris S. Error-Tolerant image compositing. In: Proc. of the ECCV. Berlin: Springer-Verlag, 2010. 31–44. [doi: 10.1007/978-3-642-15549-9\_3]
- [9] Chen T, Cheng MM, Tan P, Shamir A, Hu SM. Sketch2Photo: Internet image montage. In: Proc. of the SIGGRAPH Asia. New York: ACM, 2009. Article No.124. [doi: 10.1145/1661412.1618470]
- [10] Sunkavalli K, Johnson M, Matusik W, Pfister H. Multi-Scale image harmonization. In: Proc. of the SIGGRAPH. New York: ACM, 2010. Article No.125. [doi: 10.1145/1833349.1778862]
- [11] Farbman Z, Hoffer G, Lipman Y, Cohen D, Lischinski D. Coordinates for instant image cloning. In: Proc. of the SIGGRAPH. New York: ACM, 2009. Article No.67. [doi: 10.1145/1576246.1531373]
- [12] Agarwala A, Dontcheva M, Agrawala M, Drucker S, Colbum A, Curless B, Salesin D, Cohen M. Interactive digital photomontage. In: Hart J, ed. Proc. of the SIGGRAPH. New York: ACM, 2004. 294–302. [doi: 10.1145/1186562.1015718]
- [13] Boykov Y, Jolly M. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Proc. of the ICCV. Washington: IEEE Computer Society, 2001. 105–112. [doi: 10.1109/ICCV.2001.937505]
- [14] Guo D, Sim T. Color me right-seamless image compositing. In: Proc. of the Int'l Conf. of Computer Analysis of Images and Patterns. New York: Springer-Verlag, 2009. 444–451. [doi: 10.1007/978-3-642-03767-2\_54]
- [15] Lalonde JF, Hoiem D, Efors A, Rother C, Winn J, Criminisi A. Photo clip art. In: Proc. of the SIGGRAPH. New York: ACM, 2007. Article No.3. [doi: 10.1145/1275808.1276381]
- [16] Criminisi A, Sharp T, Rother C, Pérez P. Geodesic image and video editing. ACM Trans. on Graphics, 2010,29(5):1–15. [doi: 10.1145/1857907.1857910]
- [17] Levin A, Lischinski D, Weiss Y. Colorization using optimization. In: Hart J, ed. Proc. of the SIGGRAPH. New York: ACM, 2004. 689–694. [doi: 10.1145/1186562.1015780]
- [18] Levin A, Lischinski D, Weiss Y. A closed form solution to natural image matting. In: Proc. of the CVPR. Washington: IEEE Computer Society, 2006. 61–68. [doi: 10.1109/CVPR.2006.18]



吴昊(1982—),男,云南昆明人,博士生,主要研究领域为图像和视频处理,非真实感图形绘制.



徐丹(1968—),女,博士,教授,博士生导师,CCF会员,主要研究领域为基于图像的建模与绘制,非真实感图形绘制,生物特征识别技术.