

山东大学

毕业论文(设计)

论文(设计)题目:

《一种基于 Poisson 方法新的图像融合方法》

姓 名 胡永涛
学 号 20061301049
学 院 计算机科学与技术
专 业 计算机科学与技术
年 级 2006
指导教师 屠长河

2010 年 3 月 5 日

目录

摘要.....	I
Abstract.....	II
第一章 导论.....	1
1.1 背景.....	1
1.2 问题及意义.....	2
1.3 论文结构.....	2
第二章 相关工作.....	3
2.1 感兴趣区域（或对象）的提取方面的工作.....	3
2.1.1 魔术棒方法.....	4
2.1.2 智能剪刀方法 ^[3]	4
2.1.3 Bayes 方法.....	4
2.1.4 Knockout 方法 ^[7]	4
2.1.5 Graph Cut 方法 ^[8-9]	4
2.1.6 GrabCut 方法.....	5
2.2 Poisson 图像融合方面的工作.....	5
第三章 解决方法.....	6
3.1 算法第一步：利用 GrabCut 算法扣取出要融合的对象.....	8
3.1.1 采用 graph cut 进行图像分割.....	8
3.1.1.1 图像分割.....	8
3.1.1.2 能量最小化法分割.....	9
3.1.2 GrabCut 分割算法.....	10
3.1.2.1 颜色数据模型.....	10
3.1.2.2 迭代能量最小化分割.....	11
3.1.2.3 用户交互和不完整三分图.....	13
3.1.3 透明性.....	13

3.1.3.1 边界扣取技术.....	13
3.1.3.2 前景的确定.....	15
3.2 算法第二步：利用扣取出的要融合的对象得到融合时所需的对象对应融合蒙板.....	17
3.3 算法第三步：利用 Poisson 融合算法进行融合。.....	17
3.3.1 引导下的调整的 Poisson 求解法.....	17
3.3.2 无缝隙克隆.....	20
第四章 结论.....	23
4.1 主要工作.....	23
4.2 本文得到的结论.....	25
4.3 与原先的 Poisson 融合方法进行对比.....	27
4.3.1 原先的 Poisson 融合方法得到的结果.....	27
4.3.2 改进后的 Poisson 融合方法得到的结果.....	28
4.3.4 对比.....	29
4.4 进一步研究的方向.....	30
4.4.1 算法第二步：获得对象的融合蒙板方面.....	30
4.4.2 算法第三步：进行 Poisson 融合方面.....	30
4.4.3 对源图像的大小要求方面.....	31
参考文献.....	32
致谢.....	34
附录.....	Error! Bookmark not defined.

摘要

一副图像胜于千言万语，但是往往在现实中，人们对已经存在的图像不太满意，图像融合技术应运而生。其目的是先将目标物体从一幅源图像中分离出来，然后将目标物体嵌入到目标图像中。现存的主流技术有以下三种：一种是 Adobe Photoshop 中使用的技术（尚未公开）；第二种是通过构造拉普拉斯金字塔进行插值运算来融合两幅图像；第三种是 Poisson 图像编辑技术^[2]，其已成为近几年诸多融合算法的基础。虽然泊松编辑的方法拥有简洁的表达式，且便于理解和计算，但是，这种算法的最大弱点是它需依赖于用户手动来提取图像边缘，这要求用户即便对于很复杂的图像也要仔细地画出边缘，从而给用户带来了极大的不便。

本文提出了一种基于 Poisson 方法新的图像融合方法，力求使用户从繁琐的手动提取图像边缘的任务中解脱出来，并提高融合的质量。该方法可以分为三个主要过程：第一步，应用 GrabCut 方法^[1]从待融合的图像中提取出要进行融合的区域或对象（感兴趣的区域或对象）；第二步，根据提取出来的要融合的区域或对象进而可以得到其对应的精确轮廓；第三步，采用 Poisson 融合方法^[2]以及充分利用已得到的要进行融合区域或对象的精确轮廓进行图像融合。本文提到的新方法的核心仍是 Poisson 融合方法，因此本身就具有 Poisson 融合方法固有的能够正确地处理光照等效果的固有优点，而且更重要的是，由于对图像先进行了 GrabCut 提取和得到精确轮廓等步骤，很好地克服了 Poisson 融合方法中由于融合轮廓过大造成的融合边界区域的不理想效果，较好地处理纹理和颜色等效果，能够产生较为满意的图像融合结果。

关键字：图像融合 GrabCut 抠图方法 精确轮廓 Poisson 融合方法

Abstract

An image is said to be worth thousands of words. However, in reality, people are not always satisfied with existed images. That's why image blending comes to its life. Its goal is first to extract target objects from source images, and then to embed them to the target image. Nowadays, there are three popular methods as follows: one is the method used in Adobe Photoshop (not yet been published); the second is to blend two images through constructing Laplacian pyramid and doing interpolation; the third is 'Poisson Image Editing', which has been the basis of many blending algorithms in recent years. Although 'Poisson Image Editing' has concise expressions and is easy to understand and compute, its biggest weakness is that it has to rely on the extracted contour by users manually. Because it needs that users can draw the contour even for some complex images, it really gives users a great deal of inconvenience.

This paper presents a new way based on 'Poisson Image Editing' to do image blending. Its goal is to make users free of the tedious task of manually extracting images' contours and improve the quality of integration. This new way of image blending consists of three main processes. Firstly, I will extract the to-be blended areas or objects (areas or objects interested) from the to-be blended images. Secondly, I can obtain the exact contours of the to-be blended areas or objects based on the first step. Thirdly, I will apply Poisson blending to do image blending based on the exact contours of the to-be blended areas or objects. The core of this new way of image blending is still Poisson blending, so it itself has the advantage of handling illumination changes between images which is the immanent advantage of Poisson blending. What's more, and what is important is that, doing GrabCut and getting the exact contours first overcome the bad results of the blending boundary areas, because in traditional Poisson blending, the blending contours are too big. This in all makes the new way of image blending can better handle texture or color differences and produce preferable pleased results.

Keywords: image blending Grabcut exact contour Poisson blending

第一章 导论

1.1 背景

图像融合技术在图像编辑领域处于核心地位，研究的问题是如何将 2 个或多个图像中感兴趣的区域（或对象）进行融合，同时综合考虑到图像相互之间的光照、纹理、颜色等因素。本文提出了一种基于 Poisson 方法新的图像融合方法。

最大限度的提取各自信道中的有利信息，最后综合成高质量的图像，以提高图像信息的利用率、改善计算机解译精度和可靠性、提升原始图像的空间分辨率和光谱分辨率，利于监测。

高效的图像融合方法可以根据需要综合处理多源通道的信息，发展系统对目标探测识别的可靠性和系统的自动化程度。其目的是将单一信源的针对同一目标的不同时间的图像信息或不同类传感器所提供的同一目标的信息加以融合，消除多传感器信息之间存在的冗余信息和矛盾，并通过一系列图像处理技术和方法（在图像融合步骤中都有体现），增强影像中信息的透明度，改善解译图像的精确度、可靠性和使用率，以形成对目标的清晰、完整、准确的描述，为军事应用和非军事应用提供更准确的指导。

诸多方面的优点使得图像融合在医学、遥感、计算机视觉、气象预报和军事目标识别等方面的应用潜力得到了更加充分认识和重视，尤其在军事应用方面。图像融合被认为是克服目前某些难点的技术方向。在航天、航空多种运载平台上，各种传感器所获得的大量光谱遥感图像数据，为目标信息的高效提取提供了良好的处理对象。同时，图像融合也是当代战争中的高端技术和亮点。人们在实践中将图像融合理论和技术得到进一步的认识、发展和提高，并迅速逐步转向实时性图像处理技术。

一般情况下，图像融合由低到高分三个层次：数据级融合、特征级融合、决策级融合。数据级融合也称像素级融合，是指直接对传感器采集来得数据进行处理而获得融合图像的过程，它是高层次图像融合的基础，也是目前图像融合研究的重点之一。这种融合的优点是保持尽可能多得现场原始数据，提供其它融合层次所不能提供的细微信息。

像素级融合中有空间域算法和变换域算法，空间域算法中又有多种融合规则方法，如逻辑滤波法，灰度加权平均法，对比调制法等；变换域中又有金字塔分解融合法，小波变换法。其中的小波变换是当前最重要，最常用的方法。

在特征级融合中，保证不同图像包含信息的特征，如红外光对于对象热量的表征，可见光对于对象亮度的表征等等。

决策级融合主要在于主观的要求，同样也有一些规则，如贝叶斯法，D-S 证据法和表决法等。

融合算法常结合图像的平均值、熵值、标准偏差、平均梯度；平均梯度反映了图像中的微小细节反差与纹理变化特征，同时也反映了图像的清晰度。目前对图像融合存在两个问题：最佳小波基函数的选取和最佳小波分解层数的选取。

针对融合后的图像进行算法评价也是不可或缺的。融合图像的评价参数通常分为基于信息量的评价参数（如：信息熵、交叉熵、联合熵等）、基于统计特征的评价参数（如：均值，中值、标准差、均方差等）、基于相关性的评价参数（如：偏差指数、相关系数等）和基于梯度值的评价参数（如：平均梯度等）。针对实时性系统，完成图像融合全过程的实际消耗时间也是一个关键的参数。

1.2 问题及意义

本文研究的问题正是图像融合，即如何将 2 个或多个图像中感兴趣的区域（或对象）进行融合，同时综合考虑到图像相互之间的纹理、光照、颜色等因素，得到高质量目标图像。

现实存在的图像总是满足不了人们对图像的各种要求的，而图像融合正是可以满足这种要求，使图像的总类大大增加，满足人们各种各样的要求，而且图像本身也具有极好的质量，这种质量体现在纹理、光照、颜色等各个方面。

1.3 论文结构

本章主要介绍本文的背景知识，问题以及其意义，下章介绍涉及到的相关工作，在第三章中给出了问题的解决方法，算法的具体过程。第四章中给出了结论，并以实例验证了算法的可行性，最后也给出了进一步研究的方向。

第二章 相关工作

本文提出的基于 Poisson 方法新的图像融合方法可以分为三个主要过程：第一步，应用 GrabCut 方法从待融合的图像中提取出要进行融合的区域或对象（感兴趣的区域或对象）；第二步，根据提取出来的要融合的区域或对象进而可以得到其对应的精确轮廓；第三步，采用 Poisson 融合方法以及充分利用已得到的要进行融合区域或对象的精确轮廓进行图像融合。

在该部分中本文也将针对这两个部分分别阐述其相关工作：

2.1 感兴趣区域（或对象）的提取方面的相关工作

目前已经有很多种对感兴趣区域（或对象）的提取方法，即是所谓的抠图方法，概括起来有以下六种：魔术棒方法、智能剪刀方法、Bayes 方法、Knockout 方法、Graph Cut 方法和本文中所采用的 GrabCut 方法，图 1 给出了分别采用这几种抠图方法时的抠图效果比较（注：本文只讨论用于前景提取时的情况），同时给出了获得相应结果所需要的用户交互量。从图中可以清楚地看到，GrabCut 方法无论是在用户输入的简单性上，还是在最终结果的质量上都优于其他 5 种方法，这也正是本文采用 GrabCut 方法的原因。

下面对各个方法进行简单说明：

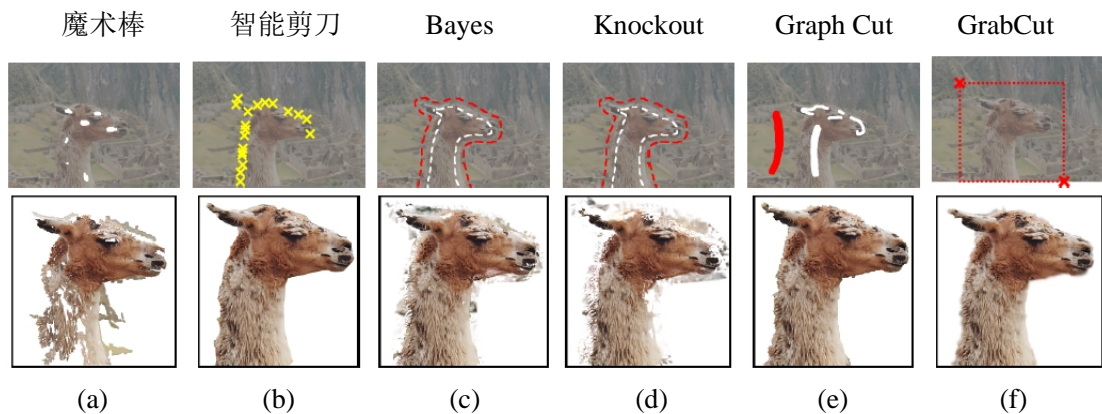


图 1：几种抠图方法的比较。最上面一行是各个方法所需要的用户交互量：白色的画笔/套索(前景)，红色的画笔/套索(背景)，黄色的画笔/套索(边界)。下面的一行是各个方法得到的抠图最终结果。

2.1.1 魔术棒方法

开始时用户指定一点或一个区域，应用颜色信息计算出一定容差范围内的相连像素的区域。尽管用户所需要的交互量是很直观的，但是找到合适的容差并不容易，有时根本就找不到。由于前景和背景的像素的颜色分布存在着大量的重叠现象，这种方法一般得不到很好的结果。

2.1.2 智能剪刀方法^[3]

这种方法允许用户通过鼠标粗略地沿着对象的边界来选择一个“最小耗费轮廓”。当鼠标移动时，将会显示出鼠标光标处到上一次“种子点”的最小耗费路径。如何计算出来的路径不理想，就需要用户指定额外的“种子点”。这个方法的主要缺点很明显：对于高纹理（或者无纹理）区域存在很多可选的“极小”路径。因此，为了得到满意的结果，通常需要用户进行大量（本例中为 19）的交互。活跃轮廓模型是一种自动改进剪刀的方法^[4]。

2.1.3 Bayes 方法

基于论文^[5]中讨论的思想，应用模型轮廓的分布来获得完整的 alpha matte 值^[6]。用户指定一个“三分图” $T = \{T_B, T_U, T_F\}$ ，其中标示出了背景和前景区域 T_B 和 T_F ，计算 alpha 值将针对剩余的区域 T_U 进行。这种方法能够得到高质量的抠图，但是必须满足以下条件：区域 T_U 不是很大且背景/前景的轮廓分布能够很好地进行分离。为了得到内部和外部的路径，需要用户大量的交互。

2.1.4 Knockout 方法^[7]

像 Bayes 方法一样都是基于用户定义的三分图，是 Photoshop 的一个专门的插件，按照这种方法得到的结果也和 Bayes 方法类似，有时质量更加不理想^[6]。

2.1.5 Graph Cut 方法^[8-9]

和 Bayes 方法相似，都包括三分图和概率轮廓模型，这是一个功能强大的优化技术，当前景和背景的轮廓分布不能很好的分离时也能得到比较理想的结果。此外，Graph Cut 技术也能用于图像合成^[10]。

2.1.6 GrabCut 方法

在解决方法一章节中会进行详细介绍。

2.2 Poisson 图像融合方面的工作

Poisson 方程式在计算机视觉方面应用很广泛。它很自然地成为了诸多变分问题的必不可少的条件。特别地，在图像编辑应用方面，主要牵涉到三种相关的工作，下面分别阐述。

通过将高动态变化（HDR）图像的梯度区域按照非线性方式重新组织，能够产生一个不再是梯度区域的向量区域^[10]。将这个向量区域的散度作为 Poisson 方程式的右部部分，并且在诺伊曼（Neumann）边界条件下指定新图像在标准方向上到边界的值为 0，通过解决这样一个方程式可以获得一新图像。相比之下，本文所采用的 Poisson 图像融合方法能够应用于图像中任意的区域，而不仅仅只针对于整幅图像。为了达到这一目的，必须用任意轮廓的狄利克雷（Dirichlet）边界条件去替换矩形轮廓的诺伊曼（Neumann）边界条件。更进一步地，可以扩展应用于梯度的非线性操作符的变化范围，包括最大操作符并尽量减少小操作符，这两种操作符都具有很好的编辑功能。

Elder 等人提出了一种系统能够通过图像的边元素（edgels）实现对图像的编辑^[11]。在这样一个系统中，扣除一个对象，其相关联的边也已被删除；添加一个对象，不仅要考虑到其相关联的边，还要考虑到每个边的两边的颜色值。通过将不断调整的颜色值赋给新边集来得到新图像。这一步可以归结于在狄利克雷（Dirichlet）边界条件下给定边附近的颜色值来求解一拉普拉斯（Laplace）方程式（无右部部分的 Poisson 方程式）。编辑的边和相应的颜色值并不总是那么简单。而且，当转换或者转换到轮廓区域过程中，图像的细节会丢失，这将会造成不理想的结果。与基于小波极值表示方式^[12]（完整的，但不太适用于手动编辑）相比，基于稀疏边的表示方式确实不完整。

通过在选定的区域中分离出亮度部分并且在选择边界上用调整后的亮度值替代原先的亮度值，来达到从图像中扣除斑点的目的^[13]。

在图像编辑方面，目前存在两种技术可以获得无缝隙克隆的效果。第一个技

术就是应用 Adobe Photoshop 中的修补刷。但是这个技术还没有公布，因此，不知道这个技术是否也用到了 Poisson 求解法。

第二个技术是 Burt 等人提出的多重求解图像融合方法^[14]。思想是对感兴趣的图像采用多重求解表示方式，也即是拉普拉斯金字塔（Laplacian pyramid）。在每次单独的求解过程中，将源图像的内容和目标图像的新的周围区域进行整合。通过不断加入不同等级的新的融合拉普拉斯金字塔来不断提升最终融合图像的质量。这种技术导致在选择区域边界的局部细节进行平均的区域进行融合，而对于低频率的区域，则在这些边界更远的区域进行融合。这种快速技术能够在目标区域获得一个近似源拉普拉斯的插入项（在拉普拉斯金字塔的第一级），但是本文采用的 Poisson 融合方法是通过求解 Poisson 方程式来处理这个拉普拉斯的插入项的。更重要的是，这种多重求解融合方法在最终的融合图像中通过金字塔的高级层次还要考虑到源图像和目标图像的远距离像素。这种长距离的融合并不会出现在本文采用的 Poisson 融合方法中，而且这也会产生不理想的结果。

最后，尽管本文采用的 Poisson 融合方法提供了一个用户指定的供引导的调整框架，例如，按照无缝隙克隆源图像的形式，已经很多种调整的方法来自动填充这一区域。这些方法中的第一类由喷绘技术构成，其中提出的基于 PDE 的调整方法就是用来实现选定区域边界部分的^[15-16]。

求解的这种 PDE 比 Poisson 方程式复杂的多，而且仅仅适用于无纹理的区域。基于实例的调整方法^[17-19]是一种相对于喷绘技术的可选方法，其中新的图像区域采用多个小区域进行重新组织进行合成。这些方法很大程度上能够处理大区域空洞和边界区域。而且，他们还能够用于引入纹理^[20-21]。

第三章 解决方法

本文所采用的算法流程见图 2，另外图 2 中还给出了例子说明，这一章节中将对算法的三个步骤进行详细说明：



图 2：图像融合算法流程图。

3.1 算法第一步：利用 GrabCut 算法扣取出要融合的对象

较理想的情况是，不在硬性条件“alpha 值必须为 0 或 1”下，一个图像扣取工具应该能够针对一个三分图的整体引用区域 T_o 产生连续的 alpha 值。这样，那些牵涉到诸如烟雾、头发、树木等情况下的问题可以自动地适当地得到解决。然而，解决这种一般图像扣取问题的技术在当前景和背景的颜色分离的比较大时能够处理的很好，而在重叠的情况下更易遭受失败^[5-6]。这也是采用 GrabCut 算法的动力所在。

首先，应用迭代的 graph cut 方法得到一个“强度”分割（3.1.1 和 3.1.2 部分）。随后进行边界抠图操作（3.1.3 部分），其中 alpha 值在强度分割边界附近的狭窄区域处计算得出。最后，除了在边界处的整个的透明性都不是采用 GrabCut 方法进行，而是采用抠图画笔的方法^[6]，实践证明，这样在不存在重叠的区域都能够工作的很好。

这种方法的新颖之处首先在于处理图像分割上。在这点上对 graph cut 机制进行了两处改进：“迭代建立”和“不完整标记法”，改进后使得用户的交互量大大减低，仅仅只需要围绕着感兴趣的对象圈出一个矩形。这种方法的新颖之处还在于提出了一种新的用于边界抠图中计算 alpha 值的机制，这种机制使得调整计算出的 alpha 值用以提高抠图的质量。

3.1.1 采用 graph cut 进行图像分割

首先，先详细描述一下 Boykov 和 Jolly 提出的分割方法，这也是 GrabCut 算法的基础。

3.1.1.1 图像分割

他们在 2001 年的论文^[8]在给定一初始三分图 T 条件下处理单色（黑白）图像的分割。这个图像的灰度值向量为 $z = (z_1, \dots, z_n)$ ，其中 n 是下标。图像的分割被描述成对每个像素的不透明性值向量 $\underline{\alpha} = (\alpha_1, \dots, \alpha_n)$ 。一般来说，有 $0 \leq \alpha_n \leq 1$ ，但对于强度分割而言，则有 $\alpha_n \in \{0, 1\}$ ，其中 0 代表背景，1 代表前景。参数 $\underline{\theta}$ 描述前景和背景的灰度分布，包括灰度值柱状图：

$$\underline{\theta} = \{h(z; \alpha), \alpha = 0, 1\}, \quad (1)$$

一个是前景的，一个是背景的。这两个柱状图是直接从三分图中各自的区域 T_B ， T_F 中的标记像素组合获得的。（柱状图在灰度等级变化范围上总和标准化为 1，即有： $\int_z h(z; \alpha) = 1$ 。）

分割的任务是为了从给定图像的 z 数据和模型 $\underline{\theta}$ 下推出未知的不透明性变量 $\underline{\alpha}$ 。

3.1.1.2 能量最小化法分割

定义一个能量函数 E 使得它的最小值可以反映一个很好的分割，它能同时受到观察到前景和背景灰度等级柱状图的作用，而且它的不透明性是合乎情理的，能够反映对象的可靠度。这可以通过下面形式的 Gibbs（吉布斯）能量获得：

$$E(\underline{\alpha}, \underline{\theta}, z) = U(\underline{\alpha}, \underline{\theta}, z) + V(\underline{\alpha}, z). \quad (2)$$

在给定柱状图模型 $\underline{\theta}$ 条件下，数据 U 用来评估不透明性分布 $\underline{\alpha}$ 对数据 z 的合适程度，定义如下：

$$U(\underline{\alpha}, \underline{\theta}, z) = \sum_n -\log h(z_n; \alpha_n). \quad (3)$$

平滑度量可以写成：

$$V(\underline{\alpha}, z) = \gamma \sum_{(m,n) \in C} dis(m,n)^{-1} [\alpha_m \neq \alpha_n] \exp -\beta (z_m - z_n)^2, \quad (4)$$

其中， $[\phi]$ 取值为 0 或 1，是 ϕ 的指示函数。 C 是临近像素对的集合， $dis(\square)$ 是临近像素的 Euclidean（欧几里得）距离。这个能量能够提高具有相似灰度等级区域的一致性。在实践中，可以通过将像素赋值为临近像素的值来获取好的结果，前提他们是邻接的（水平或垂直的）或者斜对着（8 连通）。当限定条件 $\beta = 0$ 下，平滑度量就是一个简单的熟知的相对由约束 γ 决定的 Ising prior（伊辛先验模型），它能用来提高任一区域的一致性。Boykov、Jolly 等人提出，设定 $\beta > 0$ 将会更有效，因为这样能够放宽高对比度区域的一致性约束^[8]。他们设定 β 值为：

$$\beta = (2 \langle (z_m - z_n)^2 \rangle)^{-1}, \quad (5)$$

其中， $\langle \square \rangle$ 表示对一个图像样本的期望。这样选择 β 能够确保(4)中的幂函数在高对比度和低对比度间适当地变换。约束变量 γ 设定为 50，以此来最优化性能。这也被证明是一个对大容量图像的可取设置^[22]。

到此，能量函数定义完毕，分割可以作为一个全局的最小量来建立：

$$\underline{\alpha} = \arg \min_{\underline{\alpha}} E(\underline{\alpha}, \underline{\theta}). \quad (6)$$

最小化过程是通过采用一个标准的最小化切割算法来进行的^[8,23]。这个算法是强度分割的基础，下一部分中将重点介绍对此的三点改进，以使得能够进行 GrabCut 的新的强度分割算法。第一，单色（黑白）的图像的柱状图模型将被替换成颜色图像的高斯混合模型（GMM）；第二，单次进行的最小化切割算法将会被替换成一个更强大的迭代过程，它能够在建立和参数学习阶段自动切换；第三，对交互用户的要求也大大放宽，允许交互用户采用不完整标记法——用户只需在三分图中指定 T_B ，而这个步骤可以简单的通过划定一个围绕对象的矩形或者套索来完成。

3.1.2 GrabCut 分割算法

这一部分讲述 GrabCut 强度分割算法的新颖部分：迭代建立和不完备标记法。

3.1.2.1 颜色数据模型

现在针对的图像包括在 RGB 颜色空间的像素 z_n 。由于建立不完整颜色空间柱状图不太可行，因此本文采用一种已经存在的用于软分割采用 GMM 的方法^[5-6]。每个 GMM，一个是背景的，一个是前景的，都是一个全协方差的包括 K 个成分的高斯混合（一般取 $K=5$ ）。为了使 GMM 更容易处理，在最优化框架中引入了一个额外的向量 $k = \{k_1, \dots, k_n\}$ ，其中 $k_n \in \{1, \dots, K\}$ ，将这个向量赋值给每个像素的每个单独的 GMM 成分，每个成分要么来源于背景模型或者来源于前景模型（ $\alpha_n = 0, 1$ ）。

(2)中分割的 Gibbs 能量现在变成：

$$E(\underline{\alpha}, k, \underline{\theta}, z) = U(\underline{\alpha}, k, \underline{\theta}, z) + V(\underline{\alpha}, z), \quad (7)$$

同样取决于 GMM 成分变量 k 。考虑到颜色 GMM 模型，数据 U 现在则定义为：

$$U(\underline{\alpha}, k, \underline{\theta}, z) = \sum_n D(\alpha_n, k_n, \underline{\theta}, z_n), \quad (8)$$

其中 $D(\alpha_n, k_n, \underline{\theta}, z_n) = -\log p(z_n | \alpha_n, k_n, \underline{\theta}) - \log \pi(\alpha_n, k_n)$ ， $p(\square)$ 是一个高斯概率分布， $\pi(\square)$ 是混合权重系数，因此（根据约束），有：

$$D(\alpha_n, k_n, \underline{\theta}, z_n) = -\log \pi(\alpha_n, k_n) + \frac{1}{2} \log \det \sum (\alpha_n, k_n) + \frac{1}{2} [z_n - \mu(\alpha_n, k_n)]^T \sum (\alpha_n, k_n)^{-1} [z_n - \mu(\alpha_n, k_n)]. \quad (9)$$

因此，模型的参数现在变成：

$$\underline{\theta} = \{\pi(\alpha, k), \mu(\alpha, k), \sum (\alpha, k), \alpha = 0, 1, k = 1 \dots K\}, \quad (10)$$

例如，权重 π ，中值 μ 和有针对背景和前景分布的具有 $2K$ 个高斯成分的协方差 Σ 。平滑度量 V 相对(4)中定义的基本没变，除了相应的量在计算时采用的是颜色空间下的 Euclidean 距离：

$$V(\underline{\alpha}, z) = \gamma \sum_{(m,n) \in C} [\alpha_n \neq \alpha_m] \exp -\beta \|z_m - z_n\|^2. \quad (11)$$

3.1.2.2 迭代能量最小化分割

GrabCut 中采用新的能量最小化方案是迭代工作的，以替代先前的单次算法^[8]。由于初始三分图中的 T_U 区域新标记的像素是用来调整颜色 GMM 参数 $\underline{\theta}$ ，这样的优点在于能够自动调整不透明性 α 的值。图 3 给出了 GrabCut 系统的主要原理。步骤 1 很简单，只需要简单地对每个像素 n 列出 k_n 个值即可。步骤 2 是一系列的高斯参数建立过程，如下所述：比如说，对于一个给定的前景模型的 GMM 成分 k ，像素的子集 $F(k) = \{z_n : k_n = k \text{ 且 } \alpha_n = 1\}$ 定义下来。中值 $\mu(\alpha, k)$ 和协方差 $\sum (\alpha, k)$ 按照像素值为 $F(k)$ ，权重为 $\pi(\alpha, k) = |F(k)| / \sum_k |F(k)|$ 的样本中值和协方差标准的方式建立，其中 $|S|$ 表示集合 S 的大小。最后，步骤 3 是一个运用最小化切割的全局最优化^[8]。

算法的结构保证了有正确的收敛性质。这是因为步骤 1 到 3 中迭代最小化的每一步都可以看成是整个能量 E 反过来相对于三组变量 $k, \underline{\theta}, \underline{\alpha}$ 的值。因此， E 是单调递减的，演示见图 4。这样，算法就保证了能够收敛到某一个局部的最小值 E 。当 E 显著减少时，能够很容易地检测出来，也能够自动地终止迭代过程。

迭代最小化的实践优势 图 1(e)和(f)的对比能够看出相对于单次的 graph cut^[8]而言，GrabCut 所采用的迭代最小化的额外优势：即能够大大减少用户所需的交互量。这可以从两个方面进行说明：第一，初始化和最优化以后，所需的用户的编辑量大大减少；第二，初始的交互更简单，例如允许用户进行不完整标记，参见下面部分的描述。

初始化

- 用户通过仅提供 T_B 来初始化三分图 T 。前景设置为 $T_F = \phi$ ； $T_U = \overline{T_B}$ ，即是背景区域的补集。
- 当 $n \in T_B$ 下初始化 $\alpha_n = 0$ ；当 $n \in T_U$ 下初始化 $\alpha_n = 1$ 。
- 背景和前景的 GMM 相应地根据集合 $\alpha_n = 0$ 和 $\alpha_n = 1$ 进行初始化。

迭代最小化

1. 将 GMM 成分赋给每个像素：对 T_U 中的每个 n ：

$$k_n := \arg \min_{k_n} D_n(\alpha_n, k_n, \theta, z_n).$$

2. 根据数据 z 学习得到 GMM 参数：

$$\underline{\theta} := \arg \min_{\underline{\theta}} U(\underline{\alpha}, k, \underline{\theta}, z)$$

3. 建立分割：运用最小化分割解决：

$$\min_{\{\alpha_n, n \in T_U\}} \min_k E(\underline{\alpha}, k, \underline{\theta}, z).$$

4. 重复步骤 1，直到收敛停止。
5. 运用边界抠图法（3.1.3 部分）。

用户编辑

- 编辑：对一些像素进行校准，或者校准为 $\alpha_n = 0$ （背景），或者校准为 $\alpha_n = 1$ （前景）；相应地更新三分图。执行一次上面的步骤 3。
- 调整操作：[可选] 执行整个迭代最小化算法。

图 3: GrabCut 迭代图像分割算法

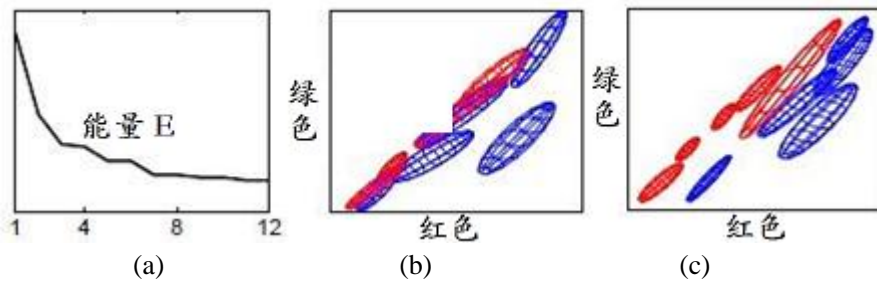


图 4: 对图 1(f)中数据进行的迭代最小化的收敛性。(a)12 次迭代中能量 E 的变化。(b)、(c)分别是在 RGB 颜色空间下的 GMM 的初始情况和收敛后情况。对背景（红色）和前景（蓝色）都使用 $K = 5$ 的混合成分。初始(b)时 GMM 重叠很严重，但是收敛后(c)分离的很好，因为前景/背景的标记已经变的很精确了。

3.1.2.3 用户交互和不完整三分图

不完整三分图。 迭代最小化算法允许用户交互的灵活性。特别地，用户可以使用不完整标记法，即在整个三分图 T 中，用户只需要指定背景区域 T_B ，剩下的 $T_F = 0$ 。根本不需要对前景区域做任何硬性的标记。迭代最小化（图 3）允许在某些像素（前景区域上）进行临时性标记，使得它们能够逐步扣取出来；只有背景标记区域 T_B 被确定，保证以后过程中不再对此区域处理扣取操作。（当然了，一个互补的方案即只指定前景区域也是可行的。）在本文的实现过程中，用户指定一矩形，围绕矩形外部的像素就确定构成了初始的 T_B （图 1(f) 中红色标记）。

进一步的用户编辑。 初始的不完整的用户标记往往不足以能够自动完成整个分割过程，尽管对一些简单的往往就足够了。如果不能够的话，就需要进一步的用户编辑（见图 5）^[8]。通过对像素进行笔刷，使这些像素确定属于前景或者确定属于背景；然后进行图 3 中的步骤 3 最小化过程。注意，只需要对部分错误标记的区域进行笔刷。另外，图 3 中的可选的“调整操作”将会根据用户的编辑更新颜色模型。这能够扩展用户编辑的效果，而这往往是有用的。注意，为了效率，在 Graph Cut 计算出来的最优流能够在用户编辑过程中再次被利用。

3.1.3 透明性

鉴于一个图像扣取工具应该能够产生连续的 alpha 值，本文上面所谈到的强度分割的机制能够被称为“边界扣取技术”，在这种机制中，在强度分割边界狭窄的区域内允许整体透明性。这种机制能够处理出现在光滑对象边界周围的模糊和混合像素问题。这种技术牵涉到以下问题：没有人工痕迹情况下为狭窄区域产生一个 alpha 值，进而恢复前景颜色，使之不会出现来自背景的颜色溢出现象。

3.1.3.1 边界扣取技术

边界扣取技术以一个封闭的轮廓 C 开始，这个轮廓是通过寻找一个与先前迭代强度分割的边界符合的折线获得的。这样，一个新的三分图 $\{T_B, T_U, T_F\}$ 被计算出来，其中 T_U 是通过在轮廓 C 两侧各取 w 个像素（本文取 $w = 6$ ）。目标是计算出 α_n ， $n \in T_U$ ，并且能够在 T_U 内建立一个粗略、健壮模型。这种模型是基于 Mortensen 和 Barrett 等人在 1999 年提出的思想^[3]，但是又进行了 2 点重要的补充：一是保证建立的 α 文件质量的规则；二是在 T_U 内建立 α 的动态规划（DP）

算法。

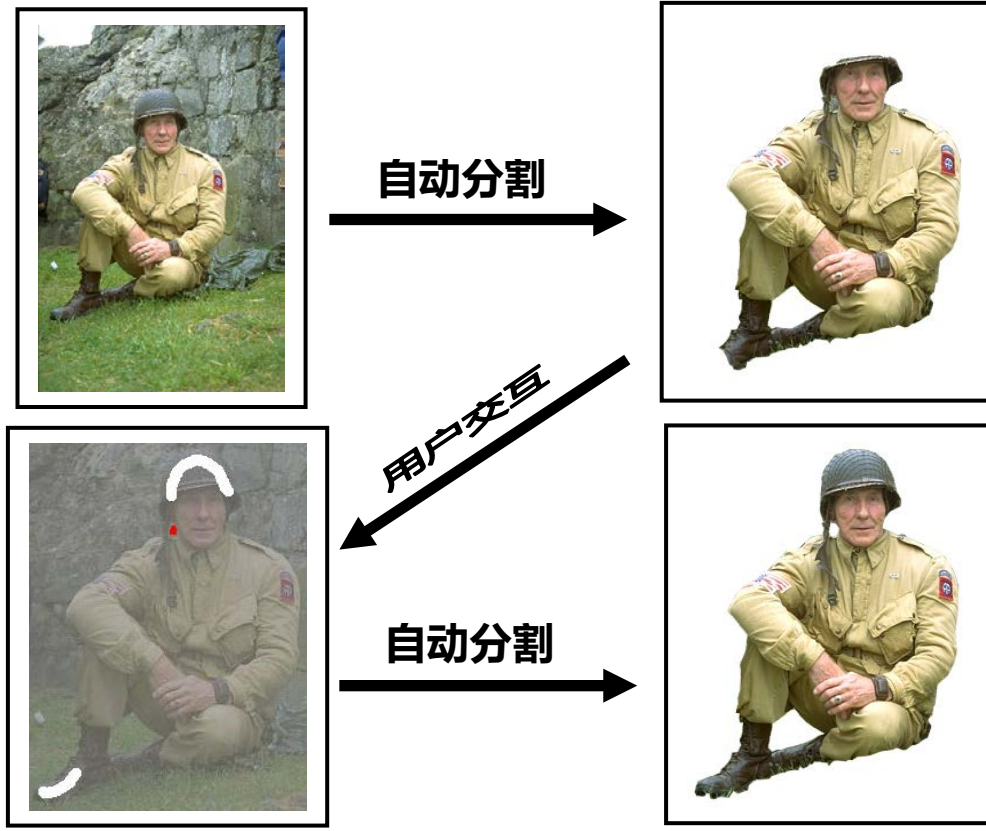


图 5: 用户编辑。在初始用户交互和分割（上一行）之后，需要进一步的用户编辑（图 3）。只需要进行粗略的前景笔刷（白色）和背景笔刷（红色）就足以得到想要的结果（下一行）。

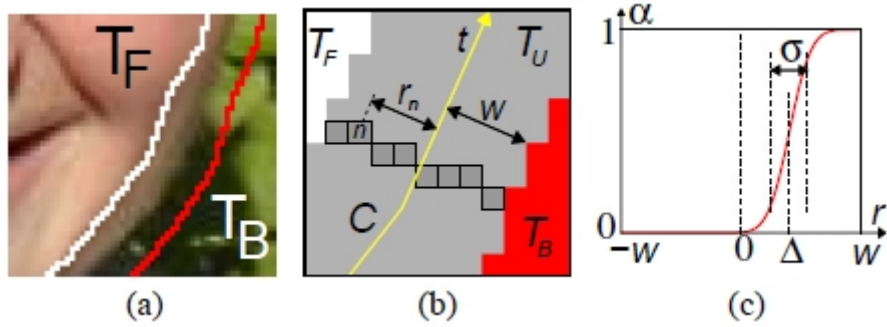


图 6: 边界扣取技术。(a)存在三分图重叠的原始图像。(b)标记出了轮廓参数和距离视图。从强度分割中获得轮廓 C (黄色)。 T_U 内的每个像素被赋值为来自轮廓 C 轮廓参数 t 和距离 r_n 。显示的像素具有共同的 t 值。(c)针对 α 文件的软性的步操作函数 g ，其中，中心 Δ ，宽度 σ 。

设 $t=1, \dots, T$ 是轮廓 C 的参数模型，其中 C 是封闭的曲线， T 是周期。正如图 6(c) 中所示，给每个像素 $n \in T_U$ 都赋以下标 $t(n)$ 。 α 文件按照一个软性的步操作函数 g （图 6(c)）来处理： $\alpha_n = g(r_n; \Delta_{t(n)}, \sigma_{t(n)})$ ，其中， r_n 是一个从像素 n 到轮廓 C

得带符号距离。参数 Δ , σ 决定了 α 文件中相对于 0、1 变换的中心和宽度。同时假设, 具有同样下标 t 的所有像素具有相同的参数 Δ_t , σ_t 。

参数值 $\Delta_1, \sigma_1, \dots, \Delta_T, \sigma_T$ 是通过在 t 上最小化下面的能力函数确定的:

$$E = \sum_{n \in T_U} \bar{D}_n(\alpha_n) + \sum_{t=1}^T \bar{V}(\Delta_t, \sigma_t, \Delta_{t+1}, \sigma_{t+1}) \quad (12)$$

其中, \bar{V} 是一个平滑规则量:

$$\bar{V}(\Delta, \sigma, \Delta', \sigma') = \lambda_1 (\Delta - \Delta')^2 + \lambda_2 (\sigma - \sigma')^2, \quad (13)$$

它的作用就是沿着曲线 C 当 t 增加时使得 α 文件进行平滑地变化 (本文中取 $\lambda_1 = 50$, $\lambda_2 = 10^3$)。对于 DP 计算中, Δ_t 分离到 30 个级别, σ_t 分离到 10 个级别。平滑量 \bar{V} 当文件的数量从 t 到 $t+1$ 变化过程中需要一个训练的过程, 然而, 本文中的规则允许应用距离转换的一个线性时间算法^[24]。如果轮廓 C 是封闭的, 最小化过程不能完全应用单过程的 DP 精确获得, 因此本文中近似地采用两个完整的 DP 过程, 假设第一次过程约定是对 $t = T/2$ 。

数据量定义如下:

$$\bar{D}_n(\alpha_n) = -\log N(z_n; \mu_{t(n)}(\alpha_n), \sum_{t(n)}(\alpha_n)) \quad (14)$$

其中, $N(z; \mu, \Sigma)$ 是一个表示 z 在中值 μ 和协方差 Σ 下的高斯概率密度。(14)中的中值和协方差定义如下^[5]:

$$\begin{aligned} \mu_t(\alpha) &= (1-\alpha)\mu_t(0) + \alpha\mu_t(1) \\ \sum_t(\alpha) &= (1-\alpha)^2 \sum_t(0) + \alpha^2 \sum_t(1). \end{aligned} \quad (15)$$

前景和背景的高斯参数 $\mu_t(\alpha), \sum_t(\alpha), \alpha = 0, 1$ 按照样本的中值和协方差进行确立, 其中对于每个区域 F_t 和 B_t 定义 $F_t = S_t \cap T_F, B_t = S_t \cap T_B$, 其中, S_t 是一个在 t 位置以分割边界 C 为中心的 $L \times L$ 的一个正方形区域 (本文取 $L=14$)。

3.1.3.2 前景的确定

这一步的目的是在不存在来自源图像背景的颜色溢出情况下建立前景像素颜色。这样的颜色溢出在 Bayes 扣取技术中会出现, 因为它所使用的概率函数不能够准确的进行。本文中通过从前景区域 T_F “偷取” 一些像素进行避免这种现象的发生。具体说来, 首先, 通过在像素 $n \in T_U$ 上应用 Bayes 扣取算法获得对景颜色 \bar{f}_n 的建立过程; 然后, 从上面定义的临近像素 $F_{t(n)}$ 中“偷取” 最相似于 \bar{f}_n 的

像素来形成前景颜色 f_n ；最后，可以得到综合应用规则的 α 计算和前景像素偷取技术的边界扣取后的结果，见图 7。



图 7：边界扣取技术相关方法比较。给定图 6 中的边界三分图，由于采用了规则化的 α 文件，GrabCut 能够获得相比于 Knockout 和 Bayes 都好的、清晰的扣取结果。

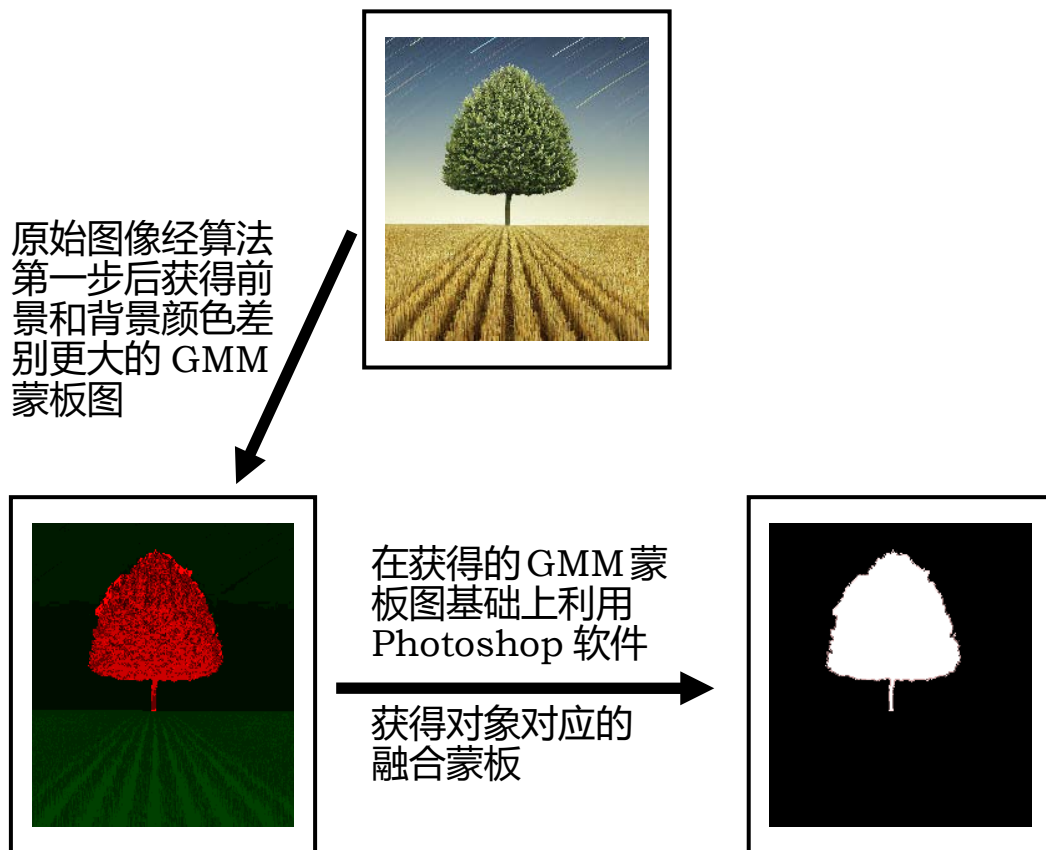


图 8：获得融合蒙板具体流程。

3.2 算法第二步：利用扣取出的要融合的对象得到融合时所需的对象对应融合蒙板

这一步相对来说较简单，是连接算法第一步和第三步的桥梁，只需要根据算法第一步扣取出来的对象进行得到相应得蒙板即可。其中，蒙板定义如下，黑色区域代表要融合的对象，其他区域均为白色区域，且满足蒙板的图像大小和源图像大小相同。

获得融合蒙板具体流程见图 8。

在算法一的程序输出中，除了能够得到扣取出来的对象，还能够得到相对于该对象的 GMM 蒙板（图 8 中左下图），该 GMM 蒙板使得对象和图像中其他区域的颜色差别增大，具体说来，对象为红色，其他区域为不同程度的绿色，这样一来方便了使用 Photoshop 软件获得对象的融合蒙板。实际上，利用 Photoshop 获得对象的融合蒙板一步非常简单，根据本文多次实验，总结得出：只需要在 Photoshop 中载入该 GMM 蒙板，选择魔术棒工具，并设置其容差为 50，点击一下表示对象的红色区域即可选择获得整个对象的区域（注，即使一次点击不足以全部将整个对象选择出来，也只需要在对象上未选择的区域简单进行几次点击即可获得），将其设置为白色，然后反选，将其他区域设置为黑色即可，非常方便。点击表示除对象的其他区域的绿色区域也可以得到同样的效果。

3.3 算法第三步：利用 Poisson 融合算法进行融合。

3.3.1 引导下的调整的 Poisson 求解法

引导下的调整 由于完全可以以单独对每个颜色成分处理的方法去求解调整问题，因此本文只考虑标量的图像函数。图 9 描述了这种标记法：用 R^2 的一个封闭子集 S 表示图像的定义域， Ω 是 S 的一个封闭子集，其边界为 $\partial\Omega$ 。 f^* 是定义在 $S - \Omega$ 区域上的一个已知标量函数， f 是定义在 Ω 上的一个未知的标量函数。用 v 表示定义在 Ω 上的一个向量域。

在 Ω 上 f 对 f^* 最简单的调整时是按照最小化问题的求解定义的膜调整：

$$\min_f \iint_{\Omega} |\nabla f|^2 \text{ 且 } f|_{\partial\Omega} = f^*|_{\partial\Omega}, \quad (16)$$

其中， $\nabla_{\square} = \left[\frac{\partial_{\square}}{\partial x}, \frac{\partial_{\square}}{\partial y} \right]$ 是梯度操作符。最小化必须满足 Euler-Lagrange 等式：

$$\Delta f = 0(\Omega \text{上}) \text{ 且 } f|_{\partial\Omega} = f^*|_{\partial\Omega}, \quad (17)$$

其中, $\Delta_{\square} = \frac{\partial^2_{\square}}{\partial x^2} + \frac{\partial^2_{\square}}{\partial y^2}$ 是 Lagrange 操作符。等式(17)是一个满足 Dirichlet 边界条件的 Laplace 等式。对于图像编辑方面的应用来说, 这个简单的方法会产生不理想的、模糊的调整结果, 这种问题可以通过很多种方法进行克服。一种方法是采用更加复杂的差分等式^[16]。本文采用的方法是调整问题本身, 具体来说, 是通过进一步地引进引导向量形式的约束条件的方式, 见下面的描述。

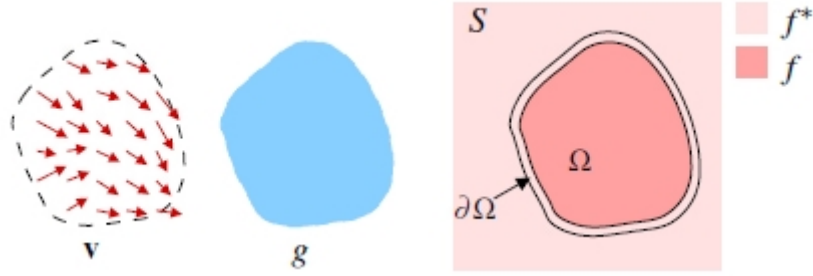


图 9: 引导下的调整的标记法。未知函数 f 在定义域 Ω 上按照向量域 v 的引导进行对目标函数 f^* 的调整, 其中, v 可能是或者不是原函数 g 的梯度域。

引导向量是上面介绍的向量域 v 在最小化问题上的扩展版本:

$$\min_f \iint_{\Omega} |\nabla f - v|^2 \text{ 且 } f|_{\partial\Omega} = f^*|_{\partial\Omega}, \quad (18)$$

它的解是下面 Poisson 等式在 Dirichlet 边界条件下的唯一解:

$$\Delta f = \text{div}v(\Omega \text{上}) \text{ 且 } f|_{\partial\Omega} = f^*|_{\partial\Omega}, \quad (19)$$

其中, $\text{div}v = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$ 是 $v = (u, v)$ 的散度。这是彩色图像 Poisson 编辑法的基本机理:

三个(19)形式的 Poisson 等式在选择的颜色空间上的三个颜色通道上分别单独求解。本文中全部采用的是 RGB 颜色空间, 采用例如 CIE-Lab 同样会得到相似的结果。

当引导向量 v 不适用时, 可以通过另一种理解 Poisson 调整法的方法来解决, 即是在定义 Ω 上定义一个相应的纠正函数 \tilde{f} , 使得 $f = g + \tilde{f}$ 。(19)中的 Poisson 等式变成下面的具有边界条件的 Laplace 等式:

$$\Delta \bar{f} = 0(\Omega \text{上}) \text{ 且 } \bar{f}|_{\partial\Omega} = (f^* - g)|_{\partial\Omega}. \quad (20)$$

因此, 在 Ω 上, 添加的纠正函数 \bar{f} 是源和目标图像在边界 $\partial\Omega$ 上的差值 $(f^* - g)$ 的一个膜调整。这种引导下的调整的特殊性将被用于后面要讲到的无缝隙克隆部分。

分离 Poisson 求解器 (18)式中的变量问题和(19)中的 Dirichlet 边界条件下的关联 Poisson 等式可以通过很多种方法分离求解。

对于分离的图像, 问题可以采用分离像素网格进行自然分离。不失一般性, 本文将对连续的对象以及分离的对象采用相同的标记: S , Ω 现在是定义在无穷分离网格上的有限点集。注意, S 可以包括一副图像的全部像素点或者其子集。对于 S 中的每个像素 p , 用 N_p 表示它在 S 中的 4 连通邻居像素点集, 用 $\langle p, q \rangle$ 表示一像素点对, 且有 $q \in N_p$ 。 Ω 的边界现在为 $\partial\Omega = \{p \in S \setminus \Omega : N_p \cap \Omega \neq \emptyset\}$ 。用 f_p 表示在 p 处的值。目标是计算强度集 $f|_{\Omega} = \{f_p, p \in \Omega\}$ 。

对于定义在任意形状边界上的 Dirichlet 边界条件, 最好直接去分离(18)式中的变量问题, 而对(19)式中的 Poisson 等式则不然。对(18)式进行有限差分分离会产生下面的分离二次最小化问题:

$$\min_{f|_{\Omega}} \sum_{\langle p, q \rangle \cap \Omega \neq \emptyset} (f_p - f_q - v_{pq})^2, \text{ 且对所有的 } p \in \partial\Omega \text{ 存在 } f_p = f_p^*, \quad (21)$$

其中, v_{pq} 是 $v(\frac{p+q}{2})$ 在引导边 $[p, q]$ 上的投影, 例如, $v_{pq} = v(\frac{p+q}{2}) \square \overrightarrow{pq}$ 。它的解满足下面的同步线性等式:

$$|N_p| f_p - \sum_{q \in N_p \cap \Omega} f_q = \sum_{q \in N_p \cap \partial\Omega} f_q^* + \sum_{q \in N_p} v_{pq}. \quad (22)$$

当 Ω 包含 S 边界上的像素点, Ω 扩展到像素网格边界时, 这些像素拥有一个删节版的邻居像素点使得 $|N_p| < 4$ 。注意, 对于 Ω 内部的像素点 p , 即 $N_p \subset \Omega$, (22) 式右边没有边界量, 变成:

$$|N_p| f_p - \sum_{q \in N_p} f_q = \sum_{q \in N_p} v_{pq}. \quad (23)$$

等式(22)构成了一个经典的稀疏(有带的)对称正值系统。由于针对的是任意形状的边界 $\partial\Omega$, 因此必须采用熟知的迭代求解器。本文中的结果或者采用了带连续松弛的 Gauss-Seidel 迭代式, 或者采用了 V-周期多重网格。



图 10: 插入。当将一个具有复杂轮廓的对象插入到一个新的背景下的情况展示了这种方法的强大之处。由于源和目标之间的强烈差异,在这种条件下标准的图像克隆并不适用。

3.3.2 无缝隙克隆

输入梯度 对引导区域 v 的基本选择是直接选自源图像的一个梯度区域。用 g 表示源图像, 调整将在下面的引导式下进行:

$$v = \nabla g, \quad (24)$$

(19)式现在变成:

$$\Delta f = \Delta g(\Omega \text{ 上}) \text{ 且 } f|_{\partial\Omega} = f^*|_{\partial\Omega}. \quad (25)$$

对于数值实现, 连续规则式(24)处理成:

$$v_{pq} = g_p - g_q, \text{ 对所有 } \langle p, q \rangle, \quad (26)$$

将被带入到式(22)中。

这样得到的无缝隙克隆工具确保了源图像和目标图像边界的一致性。从用户输入角度，正如图 10 所示，大部分任务只需要进行简单粗糙的套索选择。

混合梯度 当采用上面讲到的工具时，目标图像在 Ω 内不再保持追踪性。然而，存在想结合 f^* 和 g 的性质的情况，例如想添加一个带孔的对象，或者局部透明的对象到一个纹理的或者杂乱的背景中时的情况。

图 11 给出了这样的例子，例子是在没有负责选择操作基础上将一个文字图层从源图像中扣取出来应用到目标图像中的情况。一种可能的方法是将引导区域 v 定义成一个线性的源和目标结合的梯度区域，但是这种会忽略纹理信息，见图 11。



图 11: 插入带孔对象。(a)是经典方法，基于颜色的选取和 α 蒙板，可能费事，而且经常会残留不理想的光环；(b-c)是无缝隙克隆和经平均化的无缝隙克隆，也都不是很理想；(d)是基于简单套索选取后的混合无缝隙克隆，证明是有效的。

然而，Poisson 方法允许不限制地采用引导区域，这为得到更加具有说服力的效果很大的余地。在 Ω 中的每个点，保留在 f^* 和 g 内的变量强度，并采用下面的引导区域：

$$v(x) = \begin{cases} |\nabla f^*(x)|, & \text{当 } |\nabla f^*(x)| > |\nabla g(x)| \\ g(x), & \text{或者} \end{cases} \text{对所有的 } x \in \Omega, \quad (27)$$

这个引导区域的分离部分为：

$$v_{pq} = \begin{cases} f_p^* - f_q^*, & \text{当 } |f_p^* - f_q^*| > |g_p - g_q| \\ g_p - g_q, & \text{或者} \end{cases} \text{对所有的 } \langle p, q \rangle. \quad (28)$$

图 11 和图 12 演示了此引导区域的效果。

这种混合的无缝隙克隆对于将对象添加到另一个非常相近的目标图像中的情况也是很有用的，图 13 中给出了这样一个例子。

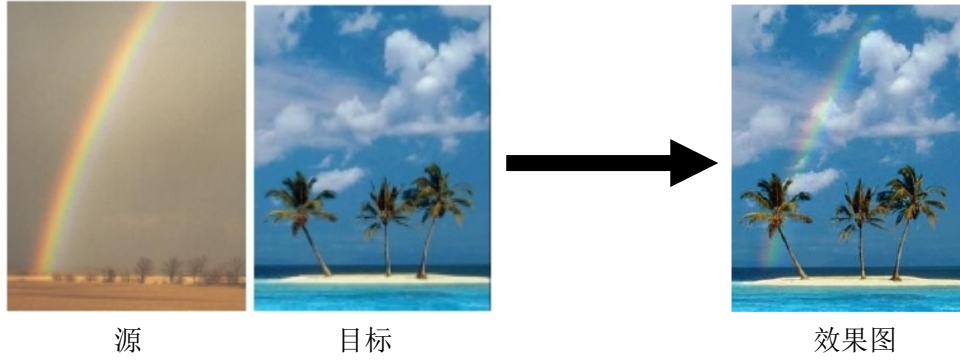


图 12：插入透明对象。



图 13：将对象插入到另一个相近的背景中。在无缝隙克隆下，将会在目标图像中临近选择区域 Ω 的像素延伸到该对象。这种延伸在应用混合梯度作为引导区域情况下是受限的。

第四章 结论

本文提出了一种基于 Poisson 方法新的图像融合方法，并对这种方法进行了细致的描述和实现。

4.1 主要工作

本文提出的基于 Poisson 方法新的图像融合方法可以分为三个主要过程：第一步，应用 GrabCut 方法从待融合的图像中提取出要进行融合的区域或对象（感兴趣的区域或对象）；第二步，根据提取出来的要融合的区域或对象进而可以得到其对应的精确轮廓；第三步，采用 Poisson 融合方法以及充分利用已得到的要进行融合区域或对象的精确轮廓进行图像融合。

下面结合一具体完整的实例进行再一次的阐述：



图 14：背景及融合对象源图像

第一步：应用 GrabCut 方法从待融合的图像中提取出要进行融合的区域或对象（感兴趣的区域或对象）：首先扣取出左图中的足球，得到的足球的对象扣取图以及对应的 GMM 蒙板图见图 15。

同样的方法对同一图像进行处理扣取得到运动员 1（左图中）以及对另一个源图像处理扣取得到运动员 2（右图中）的对象扣取图以及对应的 GMM 蒙板图见图 16。

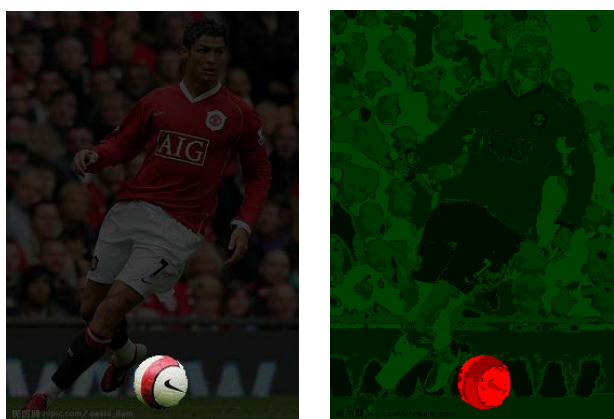


图 15： 足球的对象扣取图以及对应的 GMM 蒙板图



图 16： 两足球运动员的对象扣取图以及对应的 GMM 蒙板图

第二步：根据提取出来的要融合的区域或对象进而可以得到其对应的精确轮廓：为结合第三步的 Poisson 融合，对于这一步，只需要得到各个融合对象的融合蒙板即可，利用前面介绍的方法采用 Photoshop 软件得到各个融合对象的融合蒙版见图 17。



图 17：三融合对象的融合蒙版图

第三步：采用 Poisson 融合方法以及充分利用已得到的要进行融合区域或对象的精确轮廓进行图像融合：待融合的对象是依次融合进背景图像中去的，顺序无所谓，本文采用“先运动员 2，再运动员 1，最后足球”的顺序进行融合的，融合过程记录如下：（见图 18）

4.2 本文得到的结论

本文提出的基于 Poisson 方法新的图像融合方法不是一个全新的方法，但是综合了 GrabCut 和 Poisson 后，使得最终融合的结果兼具两者的优点。本文提到的新方法的核心仍是 Poisson 融合方法，因此本身就具有 Poisson 融合方法固有的能够正确地处理光照等效果的固有优点，而且更重要的是，由于对图像先进行了 GrabCut 提取和得到精确轮廓等步骤，很好地克服了 Poisson 融合方法中由于融合轮廓过大造成的融合边界区域的不理想效果，较好地处理纹理和颜色等效果，能够产生较为满意的图像融合结果。



图 18：融合过程图。上图：融合运动员 2 后；中图：融合运动员 1 后；下图：最终效果。

4.3 与原先的 Poisson 融合方法进行对比

如图 19 所示，要将一个豹子和摩托车融合到目标图像戈壁滩上，

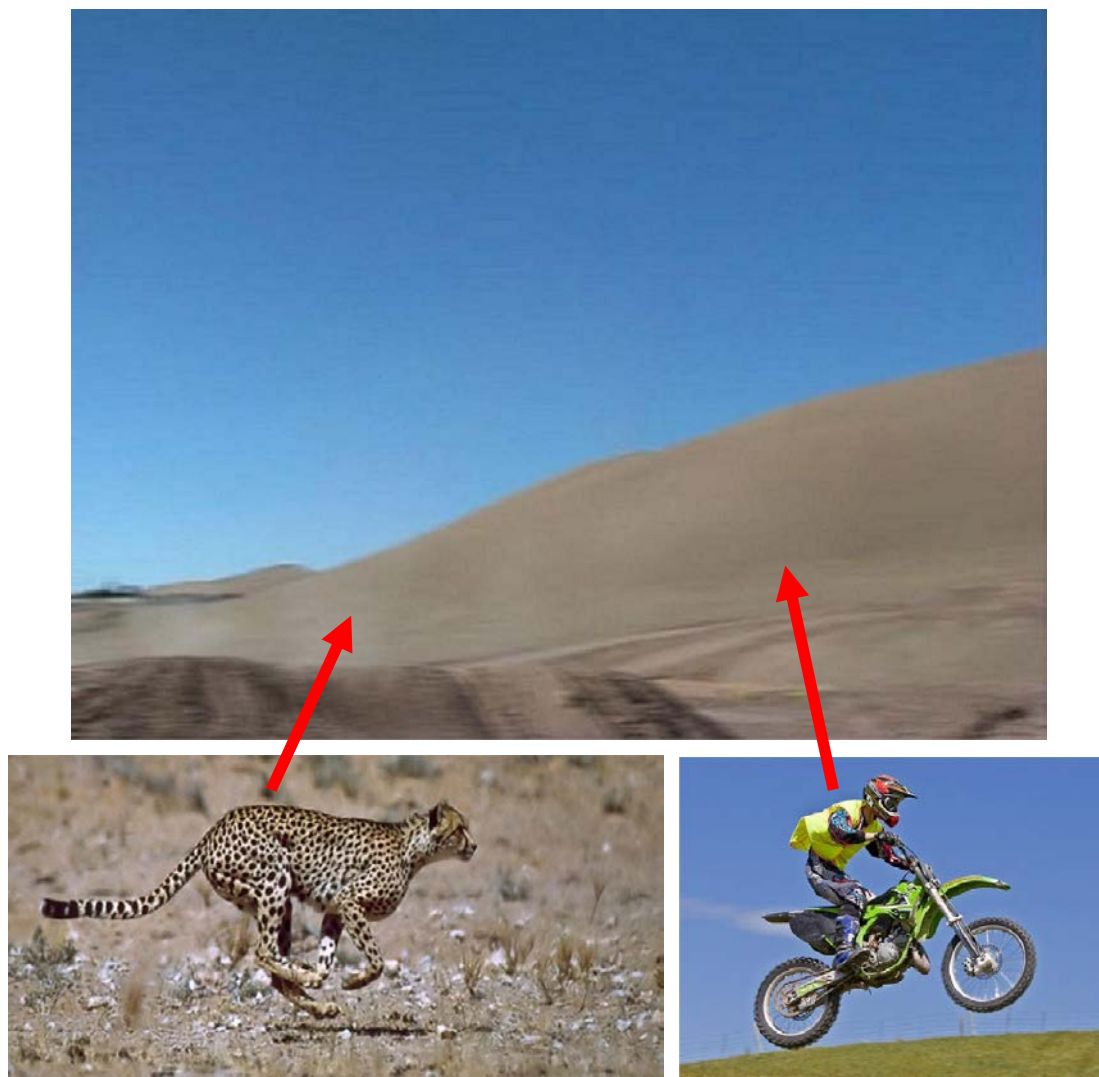


图 19: 背景及融合对象源图像

4.3.1 原先的 Poisson 融合方法得到的结果

图 20 是原先的 Poisson 融合方法用到的融合模板图(上图)和得到的结果(下图)，可以看到原先的 Poisson 融合方法用到的融合模板图由于是用户来指定，不可能很精确。

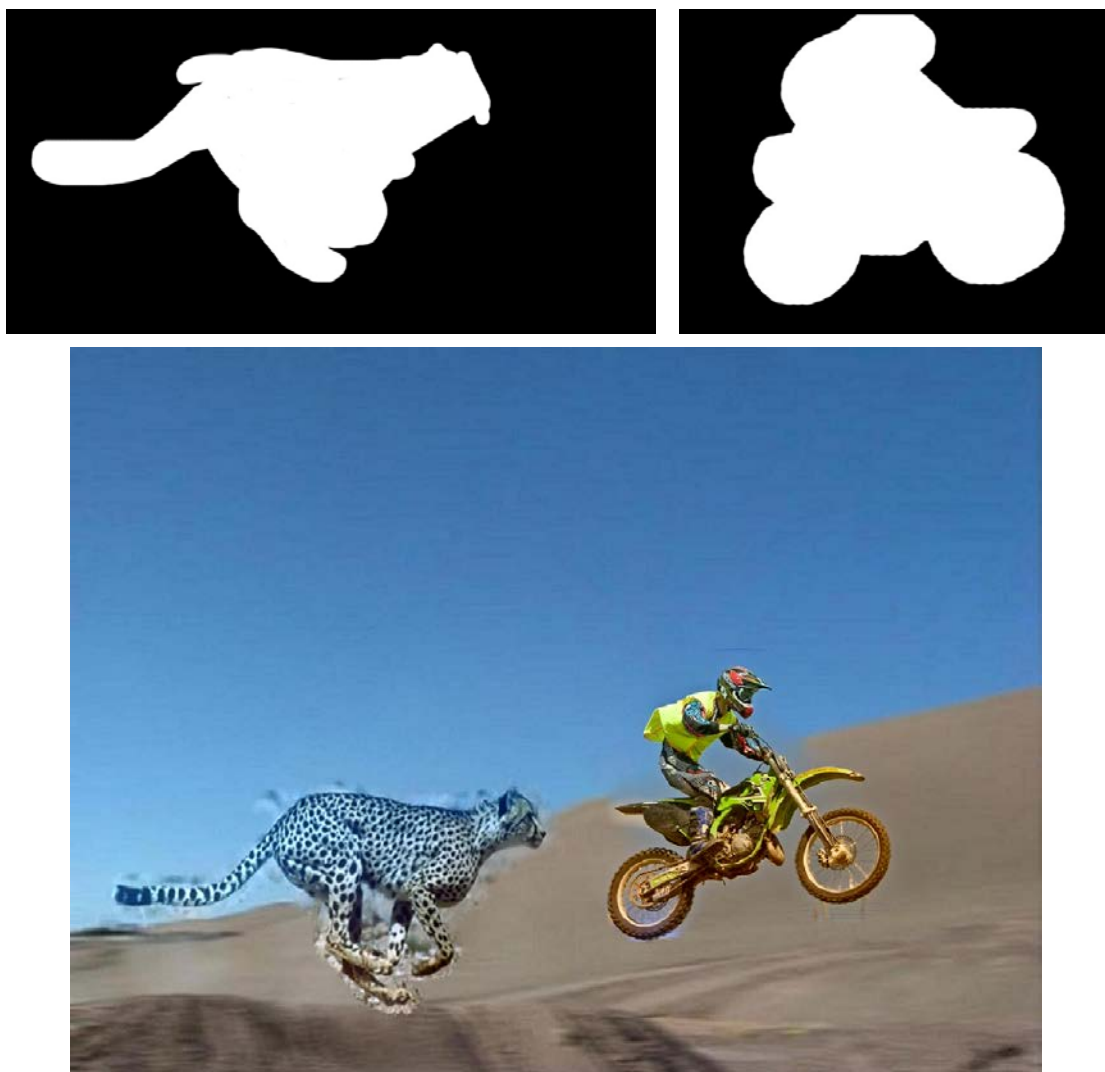


图 20: 原先的 Poisson 融合方法用到的融合模板图（上图）和得到的结果（下图）

4.3.2 改进后的 Poisson 融合方法得到的结果

图 21 是改进后的 Poisson 融合方法用到的融合模板图（上图）和得到的结果（下图）。可以看到，相比于原先的 Poisson 融合方法，改进后的 Poisson 融合方法用到的融合模板图由于使用了 GrabCut 后变得更加精确。为后面的 Poisson 融合方法做好了铺垫。

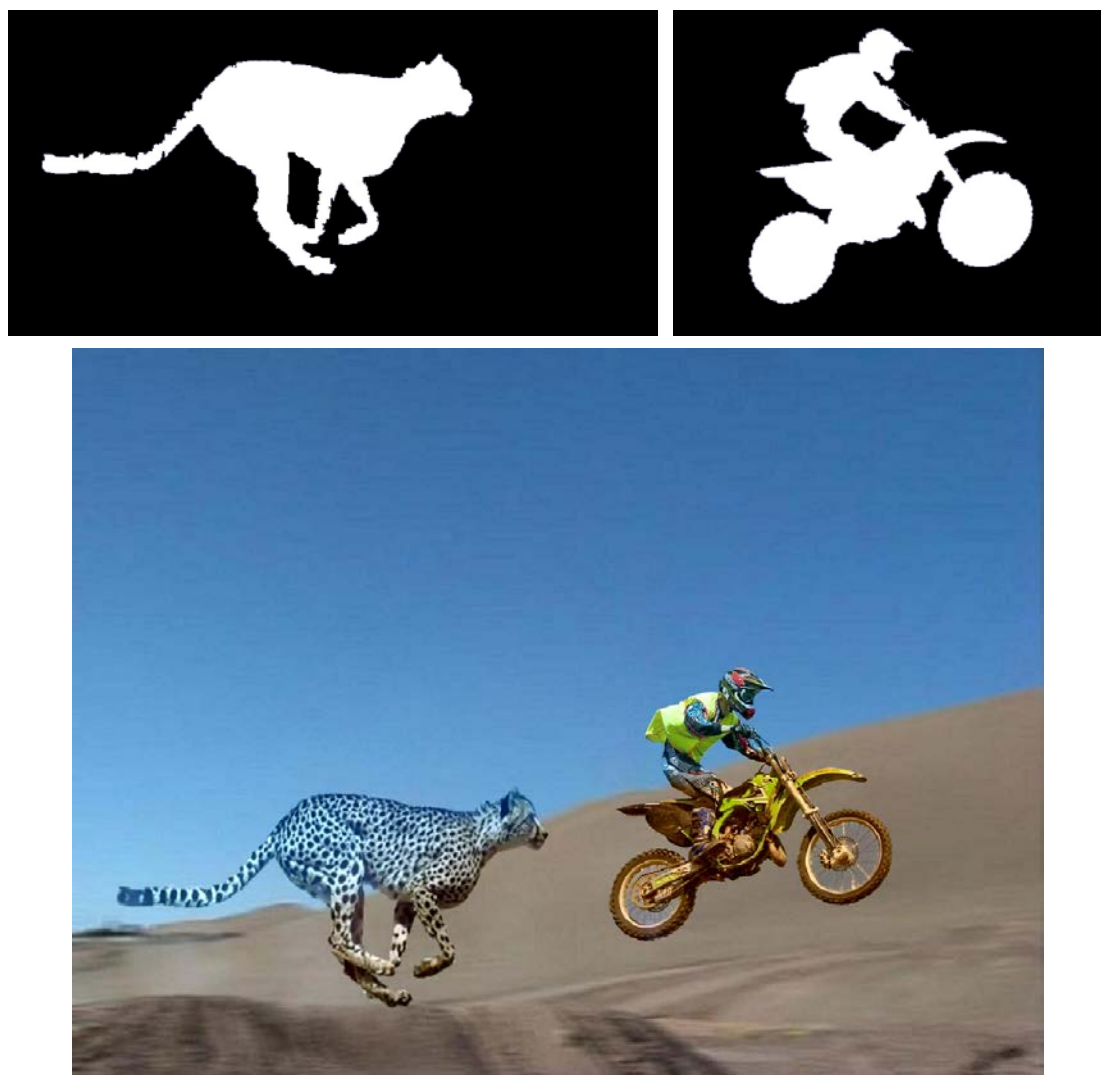
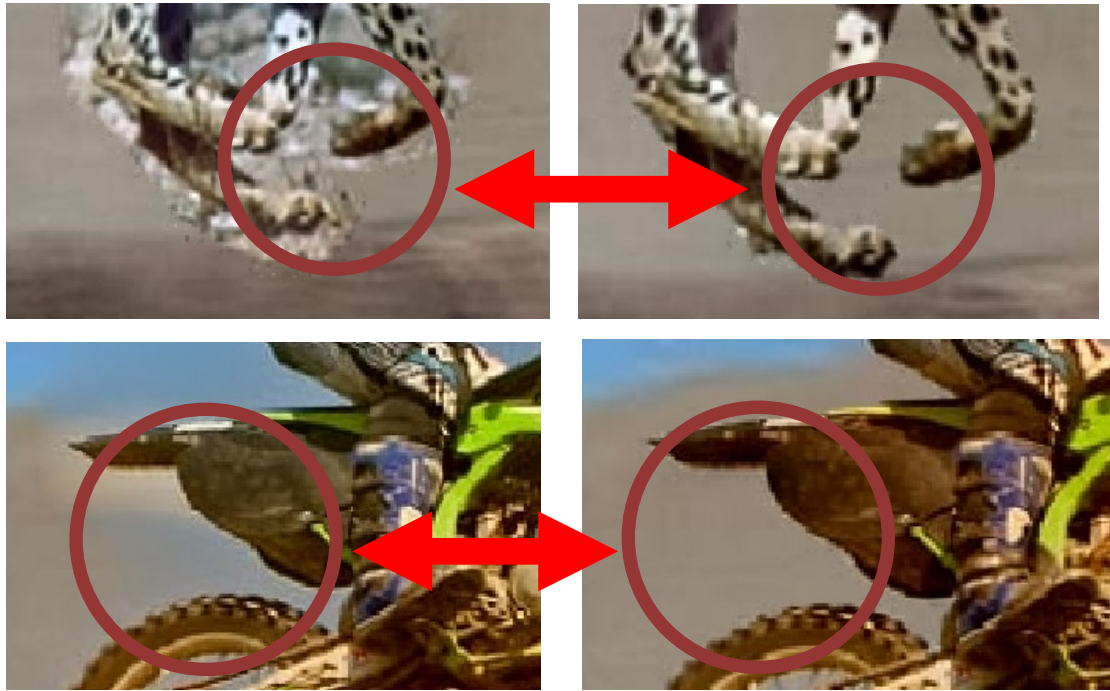


图 21：改进后的 Poisson 融合方法用到的融合模板图（上图）和得到的结果（下图）

4.3.4 对比

图 22 给出了改进后的 Poisson 融合方法和原先的 Poisson 融合方法局部对比，可以看到由于原先的 Poisson 融合方法需要用户自己来指定融合蒙版，这样得到的融合对象的边界不精确，在图像真正的边界和用户指定的不精确边界的区域会造成融合结果很模糊，不理想。而本文采用的改进后的 Poisson 融合方法由于在进行 Poisson 融合之前先采用 GrabCut 获得了融合对象的精确边界，使得原先的造成模糊的可能性降低到最小，从而改善了融合的质量。



原先的 Poisson 融合方法结果

改进后的 Poisson 融合方法结果

图 22: 改进后的 Poisson 融合方法和原先的 Poisson 融合方法局部对比

4.4 进一步研究的方向

当然，由于本人的能力有限，在此只提出了一种综合后的融合方案，虽然效果比较理想，但是有很多值得进一步改进的地方，也是本文今后进一步研究的方向所在，具体来说，有以下几个方面：

4.4.1 算法第二步：获得对象的融合蒙板方面

这一步是利用 Photoshop 软件来做的，虽然比较方便易行，但对于不熟悉该软件的来说是一个问题。今后可以将这一步也用程序实现，并设想能够将本文中提到的 3 步都整合起来，封装为一个完整的可视化程序版本，方便用户的使用。

4.4.2 算法第三步：进行 Poisson 融合方面

在进行 Poisson 融合时，具体到融合对象在背景中的具体位置，需要手动在程序中输入 2 个偏移量，即 x 轴方向偏移量和 y 轴方向偏移量，这使得用户控制融合位置方面显得有点笨拙。今后可以将其改进成鼠标拖拽型的，并支持交互调

节式的，使其更智能化。

4.4.3 对源图像的大小要求方面

本文所用到的所有例子中的融合对象都是经过和背景大小匹配后，即符合大小匹配后的情况，在程序实现方面没有实现对融合对象和背景的自动化匹配一步，而是必须事先手动调整完毕。今后可以将这一步完全实现程序自动化，既简化用户的工作，也有利于方法的推广使用。

参考文献

- [1] ROTHER, C., KOLMOGOROV, V., AND BLAKE, A. 2004. “grabcut”: interactive foreground extraction using iterated graph cuts. SIGGRAPH2004.
- [2] PEREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. SIGGRAPH 2003.
- [3] MORTENSEN, E., AND BARRETT, W. 1995. Intelligent scissors for image composition. Proc. ACM Siggraph, 191–198.
- [4] KASS, M., WITKIN, A., AND TERZOPOULOS, D. 1987. Snakes: Active contour models. In Proc. IEEE Int. Conf. on Computer Vision, 259–268.
- [5] RUZON, M., AND TOMASI, C. 2000. Alpha estimation in natural images. In Proc. IEEE Conf. Comp. Vision and Pattern Recog.
- [6] CHUANG, Y.-Y., CURLESS, B., SALESIN, D., AND SZELISKI, R. 2001. A Bayesian approach to digital matting. In Proc. IEEE Conf. Computer Vision and Pattern Recog., CD-ROM.
- [7] COREL CORPORATION. 2002. Knockout user guide.
- [8] BOYKOV, Y., AND JOLLY, M.-P. 2001. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In Proc. IEEE Int. Conf. on Computer Vision, CD-ROM.
- [9] GREIG, D., PORTEOUS, B., AND SEHEULT, A. 1989. Exact MAP estimation for binary images. J. Roy. Stat. Soc. B. 51, 271–279.
- [10] FATTAL, R., LISCHINSKI, D., AND WERMAN, M. 2002. Gradient Domain High Dynamic Range Compression. ACM Transactions on Graphics 21, 3, 249.256.
- [11] ELDER, J., AND GOLDBERG, R. 2001. Image Editing in the Contour Domain. IEEE Trans. Pattern Anal. Machine Intell. 23, 3, 291.296.
- [12] MALLAT, S., AND ZHONG, S. 1992. Characterization of Signals from Multi-Scale Edges. IEEE Trans. Pattern Anal. Machine Intell. 14, 710. 732.
- [13] LEWIS, J., 2001. Lifting Detail from Darkness. SIGGRAPH 2001 Tech Sketch.
- [14] BURT, P., AND ADELSON, E. 1983. A Multiresolution Spline with Application to Image Mosaics. ACM Transactions on Graphics 2, 4, 217.236.
- [15] BALLESTER, C., BERTALMIO, M., CASELLES, V., SAPIRO, G., AND VERDERA, J. 2001. Filling-in by Joint Interpolation of Vector Fields and Gray Levels. IEEE Trans. Image Processing 10, 8, 1200.1211.
- [16] BERTALMIO, M., SAPIRO, G., CASELLES, V., AND BALLESTER, C. 2000. Image Inpainting. In Proceedings of ACM SIGGRAPH 2000, ACM Press / ACM SIGGRAPH, New-York, E. Fiume, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 417.424.
- [17] BARRET, A., AND CHENEY, A. 2002. Object-Based Image Editing. ACM

- Transactions on Graphics 21, 3, 777.784.
- [18] BORNARD, R., LECAN, E., LABORELLI, L., AND CHENOT, J.-H. 2002. Missing Data Correction in Still Images and Image Sequences. In Proc. ACM International Conference on Multimedia.
- [19] EFROS, A., AND LEUNG, T. 1999. Texture Synthesis by Non-Parametric Sampling. In Proc. Int. Conf. Computer Vision, 1033.1038.
- [20] EFROS, A., AND FREEMAN, W. 2001. Image Quilting for Texture Synthesis and Transfer. In Proceedings of ACM SIGGRAPH 2001, ACM Press / ACM SIGGRAPH, New-York, E. Fiume, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 341.346.
- [21] HERTZMANN, A., JACOBS, C., OLIVER, N., CURLESS, B., AND SALESIN, D. 2001. Image Analogies. In Proceedings of SIGGRAPH 2001, ACM Press / ACM SIGGRAPH, New-York, E. Fiume, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 327.340.
- [22] BLAKE, A., ROTHER, C., BROWN, M., PEREZ, P., AND TORR, P. 2004. Interactive Image Segmentation using an adaptive GMMRF model. In Proc. European Conf. Computer Vision.
- [23] KOLMOGOROV, V., AND ZABIH, R. 2002. What energy functions can be minimized via graph cuts? In Proc. ECCV. CD-ROM.
- [24] RUCKLIDGE, W. J. 1996. Efficient visual recognition using the Hausdorff distance. LNCS. Springer-Verlag, NY.

致谢

从选定题目到现在，经过三个多月的努力，我终于完成了这个本科阶段用时最长、精力最集中、最富挑战性和综合性的项目，在三个月的设计和写作过程中，我的家人、老师、同学和朋友给我充分的支持，而我能顺利的完成了毕业论文写作，更是离不开他们的帮助和建议，这篇论文既总结我的劳动成果，也凝聚了他们的汗水和期望。

同时，论文的完成也标志着大学本科四年生活的结束，这是我人生中最为宝贵的日子之一，大学四年的生活，充满了无穷探索的乐趣、无数激动人心的工作和工作完成后的兴奋和荣誉，这使得我逐渐培养了对科研的兴趣和热忱，也使我确定了从事科研工作的信心和决心。四年的大学生活，没有过多的干扰和影响，远离市区的校区使我们能够躲避城市的喧嚣，对知识的希冀使我们沉下心来刻苦钻研，多样的课程和便捷的互联网保证我们能够接触更多的专业知识和前沿进展，四年里我们以学习为乐，以项目为乐，综合理论学习和实践应用，努力掌握计算机科学的内涵，把握发展的方向，为下一步的科研工作打好了基础。因此，在这里，我向所有给我们提供了这样的学习环境和氛围的人表示感谢！

我要感谢香港大学的导师王文平教授在选题方面对我的建议，并在论文研究的具体内容上给予了我莫大的指导和支持。他在香港那边任务繁重，感谢他还是抽出时间到济南给予我具体的指导和帮助。

我要感谢我的导师屠长河教授，是他在我论文的写作整个过程中给了我细致入微的指导，付出了很多精力。屠教授的工作和科研事务十分繁忙，但是他还是于百忙中对我的日常学习工作进行悉心指导和帮助，使我的三年所学可以融会贯通，并能与社会实践相结合，我被教授渊博的学识、严密的思维方法和闪光的人格魅力以及对学生慈父般的关怀所深深的折服，感动。在我毕业论文的写作过程中，我的导师更是牺牲了很多工作和休息时间，从论文选题到文献处理，再到后期写作修正，多次提出了富有建设性的指导意见，让我的论文在结构上和理论内容上都有了很大的提高。

此外，在论文写作过程中，有很多朋友给我提供了有力的支持，在我进行毕业设计期间，热心的同学们给了我很多的精神鼓励和帮助，没有和他们度过的那些日日夜夜，我的生活是不完整的，因为有了你们才有了我对这段时光的美好追忆。

在这里我要特别感谢同宿舍的五位兄弟，是他们在平时给予了我细致入微的帮助，在我在宿舍做毕业设计时，有时候做到很晚，影响到他们的休息，是他们的包容给予了我宽松的学习环境。

最后，我要向本文中所有参考文献的作者表达深深的谢意，是前辈先人的卓越理论给我们的研究工作带来新的方向。在这里，我还要向百忙中抽出时间对本文进行审阅的各位专家、学者和老师表示深深的感谢。