

电子科技大学
UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

硕士学位论文

MASTER DISSERTATION

(电子科技大学图标)

论文题目: 数字图像抠图算法研究

学科专业: 信息获取与探测技术

指导教师: 李宏亮 教授

作者姓名: 费炳超

学号: 200921020662

分类号_____ 密级_____
UDC_____

学 位 论 文

数字图像抠图算法研究

(题名和副题名)

费炳超

(作者姓名)

指导教师姓名 李宏亮 教授

电子科技大学 成都

(职务、职称、学位、单位名称及地址)

申请学位级别 硕士 专业名称 信息获取与探测技术

论文提交日期 2012. 4. 18 论文答辩日期 2012. 5. 24

学位授予单位和日期

答辩委员会主席

评阅人

年 月 日

注 1 注明《国际十进分类法 UDC》的类号

独 创 性 声 明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

签名：_____

日期： 年 月 日

论文使用授权

本学位论文作者完全了解电子科技大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权电子科技大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后应遵守此规定）

签名：_____

导师签名：_____

日期： 年 月 日

摘 要

数字图像抠图是一种对任意自然图像中指定前景对象进行精确提取的图像处理技术。例如在电影制作中，导演可以利用抠图技术将演员无缝地转移到任意需要的位置。数字抠图技术已经成为视觉效果制作的关键技术，广泛运用在杂志设计，平面艺术，广告，影视后期制作等领域。

抠图问题提出至今已有几十年时间。早期的抠图技术以光学方法为主，直到最近十几年随着计算机和多媒体技术的飞速发展，基于数字的抠图方法才被学者广泛研究。传统的抠图方法往往被限定在解决某个特定物理环境下的抠图问题，这极大限制了抠图技术的使用范围。针对自然图像的抠图方法没有抠图环境限制，能对任意自然图像进行抠图处理，具有很广的适用范围，但该类方法往往计算复杂度高且抠图准确性难以保证，因此还处于研究和发展阶段。

本文针对自然图像抠图技术存在的上述问题，进行了下面的研究工作：

1. 系统研究了蓝屏抠图、自然图像抠图、环境抠图，阴影抠图及视频抠图的基本原理和相关方法。实验了鲁棒抠图，贝叶斯抠图和闭合场抠图三种经典的抠图算法。

2. 研究了基于样本的前景移除方法，并将其运用到抠图算法中。设计出一种基于背景还原先验假设的抠图算法，该算法能够在不影响抠图准确性的同时，提高抠图效率。该方法还能得到图像在指定前景对象移除后的背景还原图。实验表明，基于背景还原先验的抠图算法，能够准确地提取前景对象，具有良好的抠图效率。

3. 研究了样本颜色采样与图像中前景对象边缘梯度间的关系，设计出基于图像纹理方向采样的抠图算法，并结合颜色线模型对前/背景颜色样本进行采集和优化。实验表明，这种基于纹理方向先验信息的抠图方法，具有良好的抠图准确性。

关键词：图像处理，自然图像抠图，前景移除，前景提取演示系统

ABSTRACT

Digital image matting is an image processing technology which can precisely extract any specified foreground object in natural scene image. For example, it enables directors to transport an actor into a completely new location. Nowadays matting is widely used in magazines, 2D arts, advertisement and video post production. It is a crucial and frequently used operation in visual effects production.

Matting has been developed several decades ago, but it hasn't been extensively studied until recent years the computer and multimedia technique rapidly developing. Traditional matting methods are designed to solve a restricted version of the general problem in studio. Natural image matting can be used to process images with any natural background, but such methods have high computational complexity and can not guarantee the accuracy.

This article is for the aforementioned problems of matting technology, carried out the following work:

1. The current digital matting approaches, such as natural image matting, environment matting, shadow matting and video matting, are systematically studied. Three typical matting approaches, such as Robust matting, Bayesian matting and Closed-Form matting are experimented.

2. Exemplar-Based Inpainting technique is studied and added to the matting algorithm. This background recovery prior based matting method can improve the mask accuracy without raising the computational complexity. This method can get the foreground object removed digital image as the by product of getting the matting mask.

3. The relationship between data sampling and image gradient is studied and an image patch gradient direction prior based sampling strategy is designed. Combined this new sampling strategy with the color line model, a new matting method is designed. Experimental results show that this method has good matting mask accuracy.

Keywords: Image Processing, Natural Image Matting, Object Removal, Foreground Extraction System

目 录

| | |
|---------------------------|----|
| 第一章 绪论 | 1 |
| 1.1 研究背景 | 1 |
| 1.2 国内外研究现状和发展趋势 | 2 |
| 1.2.1 国内外研究现状 | 2 |
| 1.2.2 发展趋势 | 7 |
| 1.3 本文主要研究内容 | 8 |
| 第二章 基于背景还原的抠图方法 | 9 |
| 2.1 引言 | 9 |
| 2.2 基于背景还原的抠图算法 | 9 |
| 2.2.1 基于样本的背景还原算法 | 10 |
| 2.2.2 最近邻采样与掩膜平滑性 | 14 |
| 2.3 实验结果 | 18 |
| 2.4 本章总结 | 24 |
| 第三章 基于纹理方向采样的抠图方法 | 25 |
| 3.1 引言 | 25 |
| 3.2 纹理方向采样 | 26 |
| 3.2.1 颜色线模型 | 28 |
| 3.2.2 基于纹理方向采样的抠图算法 | 29 |
| 3.3 实验结果 | 32 |
| 3.4 本章总结 | 37 |
| 第四章 前景提取演示系统 | 38 |
| 4.1 系统简介 | 38 |
| 4.2 系统使用实例 | 39 |
| 第五章 总结与展望 | 42 |
| 5.1 本文研究工作总结 | 42 |
| 5.2 进一步的研究工作 | 43 |
| 致 谢 | 44 |
| 参考文献 | 46 |
| 攻硕期间取得的研究成果 | 50 |

第一章 绪论

1.1 研究背景

如何从静态数字图像或者数字视频中抠取出特定的前景对象是近二十年来图像和视频编辑领域的研究重点之一。这种将场景中的前景对象精确的提取出来的技术称为数字抠图^[1] (Digital Matting)。

抠图技术起源于照片和电影，在电影制作中导演可以通过抠图与合成技术将不同时间，不同地点拍摄到的影像合成到同一场景中实现特殊的艺术效果。抠图与合成技术现今已成为视觉特效制作领域中的一项关键技术。《侏罗纪公园》，《指环王》，《黑客帝国》等电影的成功就与出色的运用抠图技术密不可分。根据 2004 年因特网电影数据库的统计，当年 10 部最成功的电影作品中有 6 部同时获得了奥斯卡最佳视觉效果奖^[2]。除了电影视觉效果制作外，数字抠图技术在杂志、二维图形艺术、电视、广告、视频后期制作等领域也有广泛地运用。

如图 1-1 所示，通过抠图和合成处理，图中前景对象玩具狮子，被准确地提取出来并无缝地融合到 4 幅不同的场景中。



图 1-1 抠图与合成效果示意图

抠图技术经历了光学^[3]和数字两个阶段。上世纪 80 年代以来,随着数字技术的进步,光学抠图的方法逐渐被计算机和专门的抠图软件所取代。早期的数字抠图方法简单地模仿光学抠图中的相关技术,直到 1984 年学者 Porter 和 Duff 在文献[1]中首次从数学上定义了抠图问题,文献[1]中指出任意一幅观测图像 I 可以分解成一幅前景图像 F 和一幅背景图像 B 在通道 α 下的线性组合:

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

式中 α 为实数,取值范围为 $[0,1]$,当 $\alpha=0$ 或 1 时,图像中 α 所对应的像素分别属于绝对背景或绝对前景。自然图像中绝大多数区域都属于绝对前景或者绝对背景,然而在图像边缘、毛发等区域,像素颜色往往出现混合,此时的 α 值介于 0 到 1 之间。合成公式的定义标志着数字抠图技术的正式诞生。

式(1-1)中的 4 个变量中,仅有 I 是已知的。对于彩色图像,合成公式是由 I, F, B 三个通道的合成公式组成的线性方程组,因此彩色图像的抠图问题是已知 3 个方程 7 个未知量的严重的欠约束问题^[4]。抠图问题的这种严重欠约束性,使得求解精确的 α 变得非常困难。现有的抠图方法往往需要用户引入适当的提示信息,还需要将图像的颜色、空间分布等信息进行统计,得出必要的先验假设才较准确的 α 值。求解出精确的 α 掩膜后,用户仅需简单地将背景图像 B 用新背景图像 B' 进行替换,就能得到无瑕疵的融合图像。

1.2 国内外研究现状和发展趋势

1.2.1 国内外研究现状

数字图像抠图技术按是否引入了额外物理条件可分为额外信息抠图和自然图像抠图两大类^[5]。其中额外信息抠图的应用已非常广泛,然而自然图像抠图的方法由于已知信息少且场景信息复杂多变等诸多原因当前作为研究热点还处于快速发展阶段。

额外信息抠图方法只能用于特定场景。例如蓝屏技术主要用在摄影棚里,录制时需要被摄影人员面对纯蓝色(或其它单一颜色)的屏幕,然后通过电脑制作将被摄影人员合成到特定的场景中;闪频抠图需要快速对场景进行二次曝光,然后利用前景物体和背景物体在快速二次曝光时强度不同的特性进行抠图,该方法不能用于金属、玻璃等在曝光时有高光点的对象;阵列镜头抠图需要用多个镜头同时对前景物体进行拍摄,然后利用各个镜头的视角差信息估计出前景,该方法

同样只能用于摄影棚，在露天环境中效果较差。

自然图像抠图方法按算法思想可以分为基于采样，基于传播以及采样与传播相结合三大类方法。基于采样的方法首先需要采集前景/背景颜色样本，然后通过聚类、拟合、置信度计算等方式从采集到的样本中估计出每一个像素点的前景颜色和背景颜色，进而根据合成公式(1-1)得到 α 掩膜；基于传播的思想认为前景和背景像素的同一类样本间具有相似性和平滑性，可以通过这种性质估计出每一个像素点属于前景和背景的概率，从而得到 α 掩膜，例如 α -matting 抠图利用图像梯度场与颜色场的相关性，从求解 α -matting 方程出发计算出 α 掩膜；采样与传播相结合的方法认为，通过采样可以估计每一像素点的前景和背景信息，然后利用像素间的相似性和平滑性可以对估计出的每个像素点的前景和背景信息进行最优化求解，得到抠图掩膜值。

以下介绍一些典型的抠图方法，其中（1）至（5）属于借助额外物理条件的抠图方法，其余方法属于自然图像抠图。

1.2.1.1 蓝屏抠图

蓝屏抠图又叫做蓝屏技术，它在影视制作领域的应用已非常成熟。早期的蓝屏抠图要求前景对象进行拍摄时，它的背景必须为某种固定的已知颜色，通常为纯蓝。Smith and Blinn^[6]通过求解抠图方程组，分析了蓝屏抠图各种精确解的情况。在抠图方程组中，假设前景颜色中红、绿、蓝分量的相对比例，可以推算出每一点的解为：

$$\alpha = 1 - \alpha_1(C_b - \alpha_2 C_g) \quad (1-2)$$

抠图时可以通过调节参数 α_1 和 α_2 直到获得满意的结果为止。这种方法简单高效，但是需要有经验的专家对参数进行调整。

蓝屏抠图有三个主要缺陷：第一，蓝屏抠图需要固定颜色的背景；第二，运用蓝屏抠图只有在摄影棚里才能获得较好的效果；第三，作为前景的被摄影人员或物体不能包含和固定背景相同的颜色。

1.2.1.2 闪频抠图

照片增强^[7]和红眼矫正^[8]技术都需要用到闪光/非闪光图片对。在文献[9]中，作者发现在闪光图片和非闪光图片里明显变化的部分通常是前景物体，而背景物体变化通常不明显，即：

$$C^f = \alpha F^f + (1 - \alpha) B^f Z \quad (1-3)$$

$B^f \approx B$ ，然后通过解最大相似度函数 $\arg \max_{\alpha, f, b, f'} L(\alpha, F, B, F' | I, I')$ 求得 α 的掩膜值。

1.2.1.3 组合抠图

文献[10]中提出了一种组合抠图算法，抠图和融合这一对互逆的过程在绝大多数的算法中都独立进行处理。以抠图为例，通常认为从场景图像中得到前景的 α 掩膜，越接近真实值，视为该算法的效果越好。而组合抠图算法则将抠图和融合两个过程统一起来，认为当融合背景与真实场景对应有相似之处时，可以将真实场景和融合背景共同作为输入信息。然而该方法也存在一定的局限性，只有真实场景图像和融合背景有很大一片相似区域时该方法才能取得不错的效果。

1.2.1.4 散焦抠图

文献[11]引入了一种多传感相机，通过传感器阵列获得场景的深度信息。其中前景传感器可以获得 0.5m 处物体的精确图像，从 0.5m 开始图像散焦越发明显；背景传感器可以获得 5m 处物体的精确图像，从 5m 开始图像散焦越发明显。

通过传感器阵列获得场景对象的深度值 Z ，进而通过对像素的 Z 值聚类的方法获得三值图。由 1.1 节得知，彩色图像抠图问题是一个 3 方程，7 个未知变量的欠约束问题。通过传感器阵列，我们可以得到每个像素点的 9 个已知量 $I_{P\{r,g,b\}}$ ， $B\{r,g,b\}$ ， $I_{B\{r,g,b\}}$ ，此时的抠图问题变成过约束问题。进而通过梯度下降算法求解最优化问题可以估计出真实前景和真实背景颜色值。

散焦抠图的主要贡献是引入传感器阵列获得了场景的深度信息将前景抠取出来。然而该方法也有不足之处。第一，传感器通过分光的方法获得不同深度的信息，每个传感器仅能获得 25% 的光能量，因此需要在强光照条件下才能获得满意的效果。第二，前景传感器的焦距和背景传感器的焦距需要有一定的距离，因此在前景和背景距离很近的情况下该方法无法得到正确的结果。第三，通过梯度下降算法求解像素的真实前景和真实背景值需要很大的计算量。

1.2.1.5 阵列镜头抠图

阵列镜头抠图^[12]是运用一组水平排列或者垂直排列的镜头阵列，利用前景物体相对背景物体在各个阵列镜头中相对变化较小的原理，计算出前景物体的真实掩膜。该方法非常高效，优于闪频抠图和散焦抠图，对于毛发，树枝，半透明物体乃至烟雾都能得到很好的抠图结果。

该方法也有许多有待改进的地方。首先，该方法假设镜头阵列中不同传感器得到的场景信息中前景的变化总是相对背景较小的，这在实际情况中往往不成立。

其次，该方法仅仅能在室内场景中取得好的效果，不能用于户外。

1.2.1.6 AutoKey 系统

Autokey 系统^[13]是由 Mitsunaga 等人设计的，它是一种可以用于视频抠图的 Rotoscoping 过程。这个系统不断跟踪视频关键帧的轮转曲线（Roto-curve）使曲线紧贴前景轮廓线，然后通过对每一帧进行手工调整，得到高质量的掩膜。不仅如此，他们还提出了一种自适应的羽化方法避免了在将轮廓转换为不透明剖面时使用特殊的羽化操作。同对合成公式(1-1)两边取偏导，并假设 F 和 B 均为平滑的(即常数)，则有：

$$\nabla \alpha = \frac{1}{F - B} \nabla C \quad (1-4)$$

这里 $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$ 是梯度算子。式(1-4)称之为 AutoKey 约束。

Autokey 约束表明，在前景和背景光滑的假设下掩膜的梯度场与图像的梯度场呈线性关系。该假设决定了这种方法对硬性边界有较好的效果，不适合用于透明的或发丝之类的物体。

1.2.1.7 Knockout 系统

Knockout 系统^[14]是由 Berman 等人设计的。用户首先通过指定内外边界将图像分成三区域，对于未知区域内的每个像素，通过在前景区和背景区域采样估计该点的前景颜色值和背景颜色值。采样方法如下，对于未知区域内的某个像素，将已知前景区域边界上的像素进行加权求和得到前景色估计值 F' 。距离未知像素最近的权重设为 1，权重随着前景像素距离未知像素的距离增加而逐渐变小。两倍于最短距离时，样本的权重设为 0，用同样的方法可以得到未知像素的背景色估计值 B' 。

1.2.1.8 Ruzon 和 Tomasi 的方法

Ruzon 和 Tomasi^[15]将概率统计运用到前景抠图算法中。首先，该方法将未知区域分割成一些子区域，每个子区域及该子区域邻近的前景和背景区域构成正方形窗口（Box Window）。假设包含在窗口中的前景像素和背景像素分别取自于 RGB 颜色空间中的概率分布 $P(F)$ 和 $P(B)$ 的样本，然后通过颜色聚类将包含在窗口中的前景像素划分成一些集群（cluster），并且用均值为 \bar{F} ，对角协方差阵为 \sum_F 的无向高斯分布（Unoriented Gaoussian）来拟合每个像素集群；最后使用高斯混合模型

GMM (Gaussian Mixture Model) 来拟合前景/背景像素的分布。

接下来将前景像素集群分别和背景像素集群对建立集群对, 并且根据各种“相交规则”(Intersection Conflict) 和“夹角规则”(Angle Conflict) 对这些集群对进行取舍。假设集群对网络中未知区域的像素满足一种介于前景和背景分布之间的均值为 \bar{C} , 协方差为 \sum_c 的高斯分布 $P(C)$ 。其中 \bar{C} 是位置由掩膜值 α 决定, 位于连接每个集群对的均值 \bar{F} 和 \bar{B} 的直线上。此时最优的 α 值使得未知像素 C 的中间分布具有最大概率。

1.2.1.9 Hillman 的方法

Hillman 等人^[16]对样本进行了主成分分析 (Principal Component Analysis), 并使用向量来表示颜色样本, 由于颜色集群在 RGB 颜色空间中的分布趋向于扇形, 所以使用向量是合理的。算法首先循序扫描图像找到下一个未知区域内未被处理过的像素点 C , 以 C 为中心, 半径 r 范围内的所有未被处理的像素点构成未知区域 U , 这些像素使用相同的前景模型和背景模型。通过将背景边界上距离 C 最近的像素作为集合 B , 并且 B 内任意像素为中心, 半径 r 范围内其它背景像素也归入 B , 建立背景模型。前景模型与背景模型建立方法相似。对前景和背景颜色样本使用主成分分析找到主方向 (Major Orientation), 接着将颜色样本投影到主坐标轴上, 获得 α 值。

1.2.1.10 Bayesian 抠图

经典的贝叶斯抠图算法^[17], 是一种基于概率统计的方法。与 Ruzon 和 Tomasi 的算法类似, 贝叶斯抠图也是根据给定区域来构造前景和背景的概率分布, 但是该方法使用一个滑动的窗口来定义领域, 该窗口采用由内而外, 层层推进的方式从前景和背景区域向未知区域内滑动, 然后在一个贝叶斯框架 (Bayesian Framework) 内针对掩膜值的问题进行描述, 并使用 MAP (Maximum a Posteriori) 算法进行求解。不同于 Tomasi 算法所使用的高斯混合模型 GMM, 贝叶斯算法假设未知区域像素仅对应一个 F/B 集群对。当存在多个 F/B 集群对时, 需分别为每一对 F/B 进行上述的优化过程, 然后选取具最优对。

1.2.1.11 Poisson 抠图

若假设前景和背景是平滑的, 由公式(1-4)可以看出输入图象的梯度场与掩模值的梯度场成比例。泊松抠图^[18]通过直接在二维数据图像中求解泊松方程来重建掩模值。算法根据输入的提示图象 (trimap), 计算掩模的梯度场, 然后通过求解

满足 Dirichlet 边界条件的泊松方程来重建未知区域内每像素的掩膜值。全局泊松抠图 (Global Poisson Matting) 对合成公式(1-1)两边求梯度, 并假设 F 和 B 为常数, 梯度值为零, 得到梯度场的掩膜近似值。全局泊松抠图所得到的结果会在未知区域内 F 和 B 不满足平滑的假设时产生错误, 这种错误可以通过对梯度场交互式手工局部修改操作, 得到满意的结果, 这种修改被称为局部泊松抠图 (Local Poisson Matting)。

1.2.1.12 Erosion 抠图

文献[19]中, 作者提出了一种自适应 trimap 的建立方法。该方法首先根据对象的初始轮廓线, 建立一个固定宽度 (5 pixel) 的未知区域带, 然后对轮廓线上的每一点 p , 采用高斯函数拟合的方式计算出拟合半径 r_p , p 与高斯函数 g 的拟合程度越高, r_p 越接近最大值 r_{\max} 。然后对轮廓线上的每一点做参数大小为 r_p 的双向腐蚀膨胀, 得到优化的 trimap。接下来定义了一种最小能量函数的迭代的方法来估计未知区域的 α 值, 能量函数 E 由颜色误差能量 E_1 和平滑度误差能量 E_2 相加构成。计算未知区域 Ω_u 的迭代过程可以分为三步, 首先初始化前景 F , 背景 B , 及未知区域 Ω_u , 接下来对能量函数求梯度最优化得到更新的 α 值, 最后设定 α 的收敛门限, 迭代直到 α 收敛, 得到抠图结果。

1.2.1.13 Robust 抠图

鲁棒抠图系统 (Robust Matting)^[20]本质上是一种基于采样的抠图算法, 总体上可以分为最优化颜色采样和优化掩膜值两个步骤。算法的第一步是进行最优化颜色采样, 通过分析了以往的基于采样的抠图算法不能得到较好的抠图结果的原因, 作者定义了用于判断前背景样本对可靠性标准的距离比例 R_d , 以及样本点的权值 $\omega(F^i)$, $\omega(B^j)$ 。在三值图中, 沿前背景边界稀疏的采集前背景样本, 然后综合计算距离比例 R_d 和权值 ω 得到每一对样本的信度 $f(F^i, B^j)$ 。第二步为优化掩膜值, 鲁棒抠图将掩膜值的优化转化为图标记问题, 由颜色距离确定图的边权重, 由信度值 f 确定每一点到虚节点的权重, 然后通过解最优化图标记问题的拉普拉斯矩阵, 求得抠图掩膜。

1.2.2 发展趋势

通过以上对该领域研究现状的分析不难看出, 数字抠图方法虽然多但总结起来以下三点不足之处制约了这项技术的运用:

第一，需要特定条件。蓝屏抠图算法需要固定颜色的背景环境，前景物体不能包含有与背景相同的颜色，而且蓝屏抠图在抠出的前景物体边缘时常出现颜色残留，影响了抠图效果。散焦抠图算法需要特殊的传感器对前景，背景的深度信息进行采集，因此当前景与背景距离较近时便失去效果。闪频抠图，阵列镜头抠图等方法也都需要特殊的环境条件才能获得好的效果。

第二，自然图像抠图需要手工交互。自然图像抠图由于没有额外信息的引入，因此需要通过交互式的操作，得到三值图或者其它提示信息，这使得自然图像抠图技术的运用具有很大的限制。在处理多幅图片时交互式操作的工作量也很大。此外算法的性能对手工标记操作具有依赖性，对不同场景的图像需要有经验的标记操作才能得到好的抠图结果。

第三，自然图像抠图的计算复杂度高。虽然目前有一些算法如文献^[33]采用 GPU 并行计算的方法提高了计算速度，使处理速度能够达到实时化。但是在普通 PC 上处理一幅一百万像素点左右的图片仍然需要几秒至数十分钟不等的时间。这使得自然图像抠图难以运用在视频处理，现场直播等领域。

随着自然图像抠图技术的发展，已产生了大量的基于不同原理的方法。如何找到一种更为简便、通用、快速、准确的抠图方法，仍是该领域有待突破的地方。

1.3 本文主要研究内容

本文在分析理解了近年来国内外十余种数字图像抠图方法的基础上，重点研究了自然图像抠图的相关方法，提出了自己的一些想法并进行了相关实验。

第一章主要介绍了数字图像抠图的概念和相关方法。

第二章研究了基于样本的前景移除方法，并加入到抠图算法中。设计出一种基于背景还原先验条件的抠图算法，能够在不影响抠图准确性的同时，提高抠图效率。作为抠图的副产品，该方法还能够得到图片在指定前景对象移除后的背景还原图。实验表明，基于背景还原先验的抠图算法，能够较为准确地提取前景对象。

第三章研究了样本采集与图像梯度间的关系，设计出基于图像纹理方向的采样方法，并结合颜色线模型对前/背景颜色样本对进行了优化。实验表明，这种基于纹理方向先验采样的抠图方法，具有良好的抠图准确性。

第四章利用 vc++ 实现了数字图像处理实验平台，该平台能够实现图像复原，滤波，几何变换等基本操作，本文设计的抠图方法也集成到平台系统中。

第二章 基于背景还原的抠图方法

2.1 引言

抠图问题在数学上是已知 3 个方程 7 个未知量的严重的欠约束问题，因此求解合成公式(1-1)，需要提供额外的条件或者做出某些假设。引入额外物理条件的方法，例如蓝屏技术，虽然可以高效地求解前景掩膜值，却由于必须使用固定颜色的背景，使用起来受到限制。

相对于引入额外物理条件的抠图方法，针对自然图像抠图的可以直接对任意给定的图像进行抠图，因此将会有更加广阔的应用前景。一般的自然图像抠图方法，往往先假设像素点的颜色和空间分布满足某种规律，进而在一系列特定假设的基础上，求解出优化的前景掩膜值。现有的自然图像抠图方法从假设基础上可以分为基于采样、基于传播和两者相结合三类。

考虑到抠图问题与背景移除问题很相似，且抠取前景与移除背景互为逆过程，因此，本文尝试通过将前景移除的方法用于解决抠图问题。图像修复（Image Inpainting）的相关算法中，有很多方法都可以用来做移除前景处理，这些方法大致可以分为以下三类：复制粘贴图像纹理^[21,22]，基于几何部分差分方程的方法^[23-25]（PDEs），和基于图像局部自相似性的方法^[26,27]。本文在文献[28]的方法上进行了改进，将前景移除与抠图问题进行了结合，提出了基于背景还原先验信息的抠图方法。

2.2 基于背景还原的抠图算法

基于背景还原的抠图算法分为基于样本的背景还原，最近邻颜色采样和掩膜平滑性修正三个步骤。

背景还原阶段的输入为原图像和三值提示图，输出为三值图中未知区域被填充为背景提示图像；前景颜色采样阶段的输入为原图像、三值提示图及背景还原图，输出为未知区域内每个像素点前景颜色经采样和样本对选择后得到的前景/背景颜色样本对；掩膜平滑阶段的输入为原图像、三值提示图及每个像素点的前景/背景颜色样本对，输出为最终的 α 掩膜值。

算法基本流程如图 2-1 所示：

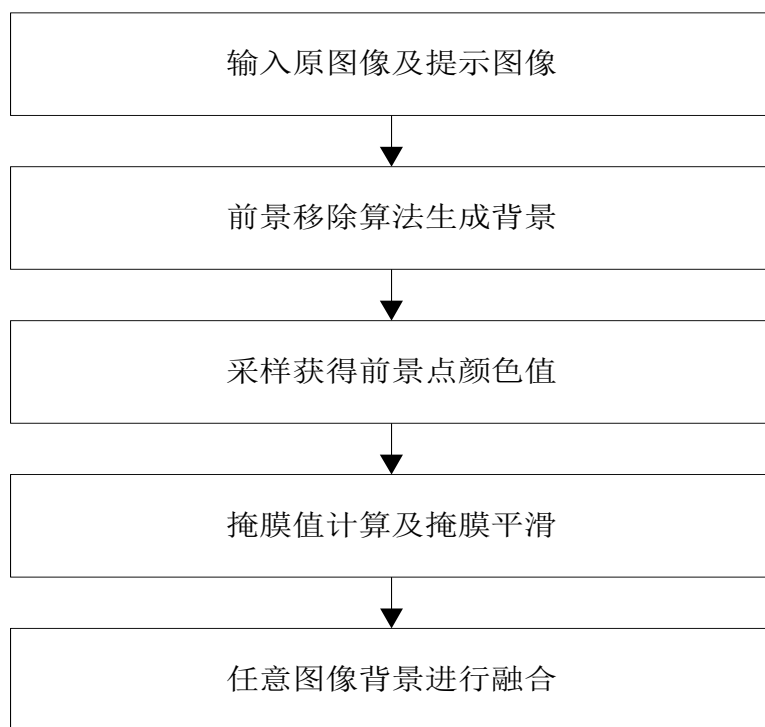


图 2-1 基于背景还原的抠图算法流程图

2.2.1 基于样本的背景还原算法

由图 2-2 演示背景还原过程：

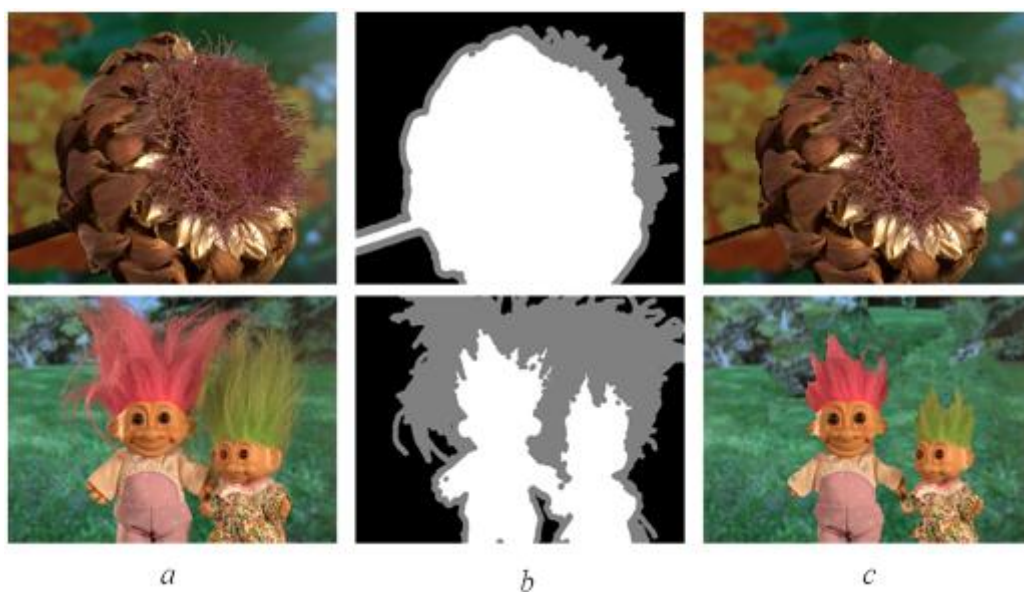


图 2-2 两张图片的背景还原结果

根据提示三值图 (Trimap) 将原图像分割成三区域, 区域 R_b (黑色) 叫做源区域, 区域 R_u (灰色) 叫做目标区域, 整个还原过程就是从源区域中选取最佳的样本填充目标区域。源区域与目标区域相接的轮廓线叫做填充边界, 每次在填充边界上选择优先级最高的点进行填充, 然后更新边界, 随着填充的进行, 填充边界逐步向目标区域缩小, 直到目标区域内的所有像素点均被填充时算法结束。图 2-2 中, a 图是原始输入图像, b 图是三值提示图 (Trimap), c 图是填充结果。由于三值提示图中, 纯黑色部分像素的掩膜值 $\alpha=0$, 纯白色部分的掩膜值 $\alpha=1$, 抠图算法需要估计的像素点在灰色区域内, 因此需要进行背景还原的像素点仅为灰色区域内的点, 这使得算法的计算量相比将整个前景对象全部移除大大减少。

与文献[28]类似, 本文的背景还原算法流程如图 2-3 所示:

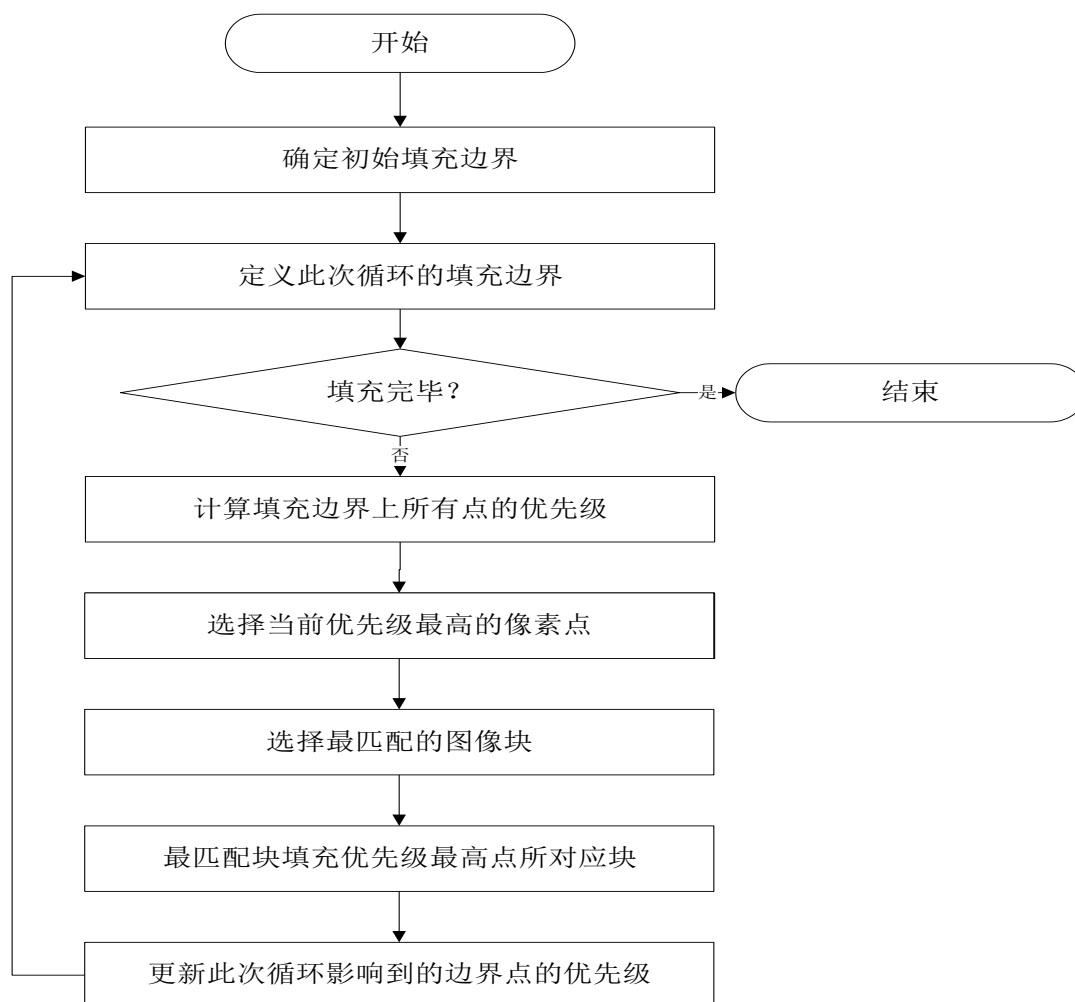


图 2-3 基于样本的背景还原算法流程图

(1) 计算填充边界点的优先级

填充顺序对于非参数的纹理综合方法至关重要，传统的方法倾向于采用‘洋葱剥皮’式由外到内的填充方式。考虑到纹理本身具有线性结构，定义新的填充次序，可以使结果优化。本文采用文献[28]中所定义的优先值计算方法如图 2-4 所示：

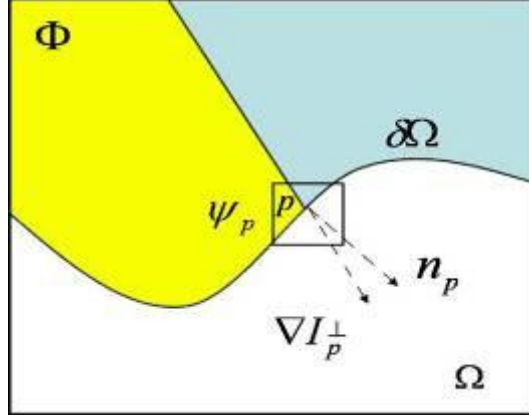


图 2-4 填充边界上的像素点优先级计算示意图

假设图像块 ψ_p 以属于填充边界 $\delta\Omega$ 上的像素点 p 为中心点（如图 2-4 所示），它的优先级 $P(p)$ 定义为下式中，两项的乘积：

$$P(p) = C(p)D(p) \quad (2-1)$$

其中 $C(p)$ 为信度项，定义如下：

$$C(p) = \frac{\sum_{q \in \Psi_p \cap \bar{\Omega}} C(q)}{|\Psi_p|} \quad (2-2)$$

$C(p)$ 的含义实际上是指的在图像块 ψ_p 内，不属于源区域 ϕ 的像素点的个数占图像块 ψ_p 内像素个数的百分比。式(2-2)使得优先级 $P(p)$ 倾向于取以 p 点为中心的图像块 ψ_p 中包含目标区域 Ω 像素少的点。

$D(p)$ 是数据项，定义如下：

$$D(p) = \frac{|\nabla I_p^\perp \cdot n_p|}{\alpha} \quad (2-3)$$

式中， α 是正规化常数，实验中 $\alpha = 255$ （灰度级为 0 到 255）， n_p 是垂直 p 点梯度方向的单位矢量， ∇I_p 为 p 点的梯度值，式(2-3)的含义是使得 $P(p)$ 倾向于取源区域的边缘垂直填充边界进入目标区域的 p 点。

(2) 区域填充

在计算出优先级最高的填充边界点（记为 \hat{p} ）后，需要在源区域 Φ 中找到与以 \hat{p} 为中心点的像素块距离最小的样本块，样本块的选择如下式所示：

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q) \quad (2-4)$$

式中距离 d 为直接在 RGB 空间中计算的均方和误差（SSD）。

当源区域 Φ 中找到了相匹配的块 $\Psi_{\hat{q}}$ 后，将 $\Psi_{\hat{q}}$ 相对应的像素块 Ψ_q 中相应的像素值逐一拷贝到对应位置。填充使得纹理和结构信息能够有效地传播到目标区域中，相对于传统的采用图像平滑处理进行填充的方法，这种方法不会使图像变得模糊。

(3) 更新优先级

将像素块 $\Psi_{\hat{q}}$ 填充到对应位置后，原来的填充边界 $\delta\Omega$ 发生了变化。进行第二次填充前，需要更新填充边界，并计算被前一次填充所影响的新的填充边界点上所有点的优先级 $P(p)$ ：

$$C(q) = C(\hat{p}) \quad \forall q \in \Psi_{\hat{p}} \cap \Omega \quad (2-5)$$

随着循环的进行，信度值 p 高的点逐渐减少，直至所有点均已被填充。

算法的伪代码如表 2-1 所示：

表 2-1 背景还原算法伪代码

| | |
|------|---|
| 初始化： | 提取三值图中初始的填充边界 $\delta\Omega^0$ |
| 循 环： | 循环至算法结束 |
| 1 | 检查填充边界长度 $\delta\Omega^t$ ，终止条件：若 $\delta\Omega^t=0$ ，程序结束 |
| 2 | 计算当前优先级 $P(p)$ ， $\forall p \in \delta\Omega^t$ |
| 3 | 找到当前具有最高优先级的块 $\Psi_{\hat{p}}$ |
| 4 | 找 $\Psi_{\hat{q}} \in \Phi$ ，使得距离 $d(\Psi_{\hat{p}}, \Psi_{\hat{q}})$ 最小 |
| 5 | 将图像数据 $\Psi_{\hat{q}}$ 填充到对应位置 $\Psi_{\hat{p}}$ |
| 6 | 更新填充边缘和优先级 $C(q) = C(\hat{p}) \quad \forall q \in \Psi_{\hat{p}} \cap \Omega$ |
| 输 出： | 目标区域内所有像素点均被填充的图像 |

图 2-5 所示为两幅经过背景还原算法处理后的图像：



图 2-5 经过背景还原算法处理后的图像

由图 2-5 可以看出，经过背景还原算法处理后，三值图中灰色区域被填充，原始图像的纹理与结构信息在灰色区域内被很好地保留下来。

2.2.2 最近邻采样与掩膜平滑性

前节得到背景还原后的图像，即合成公式(1-1)中 B 已知，因此只需要估计前景颜色值 F 便可以由合成公式推导出彩色图像掩膜值 α 计算公式：

$$\alpha^i = \frac{(I^i - B^i)(F^i - B^i)}{\|F^i - B^i\|^2} \quad (2-6)$$

由式(2-6)可以看出，要得到任意像素点 i 的掩膜值 α ，只需要精确估计出该点的前景、背景颜色样本对 (F^i, B^i) 。在前节通过背景还原已经估计出背景样本值 B 的基础上，后续还需要以下三个步骤：前景样本点颜色采样，颜色样本对修正，掩膜平滑修正。

(1) 前景样本点颜色采样

不同于已有的抠图方法，我们采用最邻近采样法获得点 P_u 的 F 值，最邻近采样法的思想简单直观，即三值图中未知区域 R_u 内每一像素点 P_u ，在绝对前景区域 R_f 内找到与 P_u 点空间距离最近的点，作为 P_u 的 F 值。

本文采取最近邻采样基于以下三点原因：

第一，以往基于采样的方法为了找到最接近真实情况的 F ，首先要采集相当数量的前景样本点 F 和背景样本点 B ，构成前景/背景样本点的集合，然后将集合中的前景样本 F 和背景样本 B 两两配对，形成颜色样本对 (F, B) ，最后还要通过计算某种相似度函数找到用于估计每个像素点 P_u 的最佳样本对 $P_u(F, B)$ 。这种方法虽然能够得到较好的抠图效果，但是计算复杂度往往较高。假设一幅分辨率为 1024×1024 的真彩色图像未知区域内有仅有 1 万个像素点（图像中未知区域所占比例不到 1%），每点采集 20 个 F ，20 个 B ，共 400 个样本对 (F, B) ，因此共需求解 1200 万次相似度函数，这样庞大的计算量使得这类抠图方法由于耗时过长，在实际中很难直接运用。

第二，最近邻采样法在本文的方法中并不会使抠图质量降低，这是因为后续步骤加入了样本对修正和通过求解能量函数平滑 α 掩膜两个过程。引入了样本对修正项 E_1 ，降低采样错误的样本对最终求取的 α 掩膜的影响；在求解能量函数过程中，引入了 α 掩膜局部平滑约束项 E_2 。 E_1 和 E_2 的引入使得由于在最邻近采样中由于采集到错误的 F ，或者背景还原中错误估计值 B ，导致的 α 偏离真实值的程度减小。该方法不仅降低了计算复杂度，而且减少了前述为估计真实的颜色样本对 $P(F, B)$ 而进行的不必要计算。

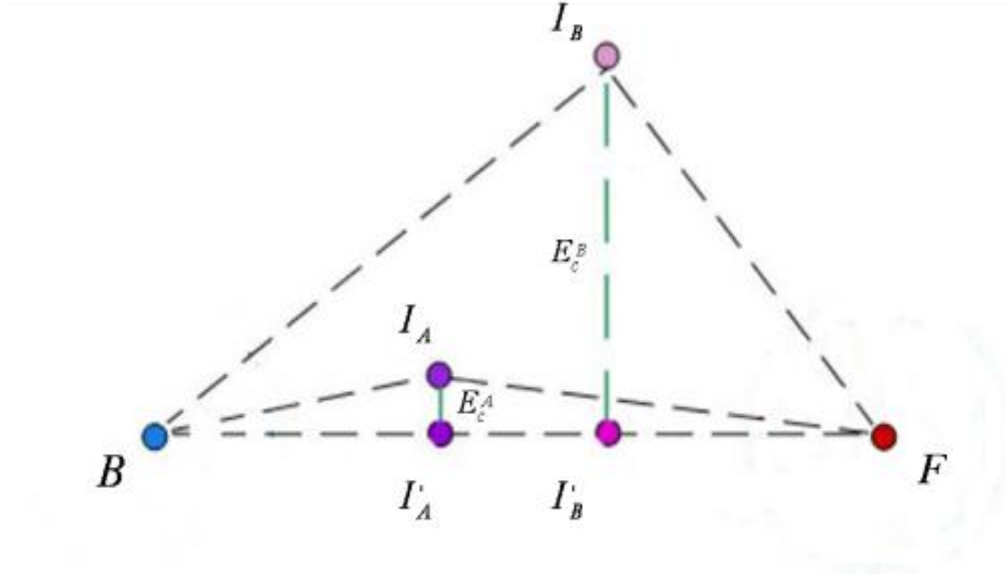
第三，最近邻假设认为物体的颜色在空间上具有某种联系，即像素的空间距离越接近，他们的颜色越容易相似。这个基本的统计规律是最近邻采样法的理论依据。

(2) 样本对修正

通过采样获得的前景颜色样本和通过背景还原估计出的背景颜色样本所形成的前景/背景颜色样本对往往偏离真实情况。用这些偏离真实情况的颜色样本来估计当前像素点的 α 掩膜值就会产生错误。因此在采样得到颜色样本后，还需要对其进行选择和修正。

文献[20]通过分析像素点、前景、背景三点颜色的空间分布，发现好的样本对与像素点的颜色在空间分布上越接近直线，估计出的 α 掩膜值的正确率越高。基于此定义了距离比率来区分颜色样本对的好坏。

本文在文献[20]的基础上定义了一种更加简单的方法来区分颜色样本对 (F, B) 的好坏如图 2-6 所示：


 图 2-6 任意颜色样本对 (F, B) 构成的颜色线

假设当前待估计的像素点为 I ，通过背景还原估计出点 I 的背景样本值 B ，通过前述最近邻采样估计出点 I 的前景颜色样本值为 F 。假设像素点 I 当前分别处在位置 A/B ，则 I_A/I_B 到直线 \overline{FB} 的投影点分别为 I'_A/I'_B ，到直线 \overline{FB} 的颜色距离分别为 E_c^A/E_c^B 。

由图 2-6 可以看出，点 A 到直线 \overline{FB} 的距离 E_c^A ，小于点 B 到直线 \overline{FB} 的距离 E_c^B ，根据颜色线理论^[29]我们认为颜色样本对 (F, B) ，对于点 A 是好的估计样本，对于点 B 是较差的估计样本。

式 2-7 及公式 2-8 给出了 E 的计算公式：

$$\hat{\alpha} = \frac{(I^i - B^i)(F^i - B^i)}{\|F^i - B^i\|^2} \quad (2-7)$$

$$E_1 = \|I^i - (\hat{\alpha} F^i + (1 - \hat{\alpha}) B^i)\| \quad (2-8)$$

I^i 表示像素点 P 颜色的观察值， $\hat{\alpha}$ 表示 P 点的初始不透明度， F^i 表示最邻近采样得到的样本值， B^i 为前节通过背景还原得到，本小节求得的样本对代价 E_1 称为数据误差，下一节掩膜值平滑误差 E_2 共同构成求解 α 掩膜的代价函数 E 。

(3) 掩膜平滑修正

与像素点 P 相邻的 8 个像素点，分别位于八个不同方向，其不透明度 α_p 至少在一个方向上满足平滑性，否则该像素点有很大概率会被感知为噪声点(个别需要用噪声做艺术处理的情形除外)，因此本文利用这一特性构造掩膜平滑性误差 E_2 用

于约束 α_p 。

首先定义转移概率 P_{pj} ， P_{pj} 反映像素点 p 与其相邻的像素点 j 之间的相似程度。若像素点 p 与其相邻的 8 个像素点间的颜色平滑度越高，则转移概率 P_{pj} 越大。因此 P_{pj} 应该是个比值，而且分母应该是 p 点到其所有 8 个邻域点的平滑度之和，分子应该为 p 点与 j 点的平滑度。用高斯函数来拟合平滑度，得到转移概率的定义如式 2-9 所示：

$$P_{pj} = e^{-\frac{\|C_p - C_j\|^2}{\delta^2}} / \sum_{j=1}^8 e^{-\frac{\|C_p - C_j\|^2}{\delta^2}} \quad (2-9)$$

式 2-9 中 C 为点的颜色值， δ 为参数，用于调整转移强度。

接下来需要用转移概率 P_{pj} 来约束 α ，并且定义平滑项约束函数 E_2 如公式 2-10 所示：

$$E_2 = \sum_{j=1}^8 (\alpha_p - \alpha_j)^2 P_{pj} \quad (2-10)$$

由公式 2-9 可以看出，当 p ， j 两点间转移概率高时， α_p 与 α_j 必须接近才能满足 E_2 最小，这正是我们所希望的。

平滑项约束计算的是像素点 p 的 3×3 邻域，在逐个计算 p_u 时，邻域间会有重叠，重叠部分在迭代过程中，使平滑性逐步迁移至整幅图像。

将整个代价函数定义为数据项误差 E_1 与平滑项误差 E_2 的线性组合：

$$E = E_1 + \lambda E_2 \quad (2-11)$$

采参数 λ 用来平衡数据项 E_1 与平滑项 E_2 的贡献，将公式 2-10，2-11 带入式 2-11，得到代价函数 E 的计算公式：

$$E = [I_p - \alpha_p \hat{F} - (1 - \alpha_p) \hat{B}]^2 + \lambda \sum_{j=1}^8 (\alpha_p - \alpha_j)^2 P_{pj} \quad (2-12)$$

注意到 E 是关于 α_p 的二次函数，容易求得 $\alpha_p (\min E_p)$ 如公式 2-13 所示：

$$\alpha_p (\min E_p) = \frac{\lambda \sum_{j=1}^8 \alpha_p P_{pj} - (I_p - \hat{B})(\hat{B} - \hat{F})}{(\hat{B} - \hat{F})^2 + 1} \quad (2-13)$$

式中 λ 为参数，用于调整数据约束 E_1 与平滑项约束 E_2 的约束力。 R_u 区域内 α_p 初始值设置为 0.5，代入以上公式计算出每一点的 α_p 值并遍历整个 R_u 区域，即完成一次迭代。用计算出的 α_p 作初始值进行第二次遍历，即完成第二次迭代。实验中经过三次迭代后的 α_p 值趋近于收敛，说明该算法的收敛速度快，收敛性也较好。

2.3 实验结果

本实验采用现有的测试数据库^[20,30-32]中评价最高的基于感知的在线测试平台对实验结果进行测试。文献[30]中， Rhemann 指出，一个好的测试环境对于技术的发展有着重要的贡献。好的测试环境包含以下三个要素：

- 第一，有挑战性的测试数据，使得现有的各种方法都会出现不同程度的错误。
- 第二，具有真实可信的理想结果作对比，误差测量采用基于感知的误差函数。
- 第三，所有数据均在线动态更新，便于将新的研究成果共享。

(1) 数据库的形成及其特点

平台中的数据包含 8 幅测试图和 27 幅训练图。8 幅训练图的真实掩膜结果被隐藏起来，这样可以防止某种方法通过设置极端的参数取得好的效果。27 幅训练图的真实掩膜值是公开的。如图 2-7 所示为数据库中的全部图片，其中前 8 幅为测试图片，后 27 幅为训练图片。

所有的 35 幅数据，是从大量的图像数据中选择出的最具代表性的场景。它们包含高纹理性的背景，不同级别的图像深度以及前景背景颜色异议性等极端条件。并且这些数据是将大量图像采用的文献[17,18,20,31]¹中的方法测试后，选择出结果最差的 35 幅。

为了防止结果倾向于某种类型的算法，所有用于测试的数据都是特殊摄影棚环境里获得的。首先固定传感器和被拍摄的前景对象，将背景替换成红，绿，黄，蓝四种不同颜色的纯色背景，然后运用蓝屏技术^[6]得到真实掩膜值；在谨慎的防止抖动和位置偏移的情况下，将背景替换为复杂场景；最后移走前景物体得到纯背景图像。

(2) 用户提示

现有的所有方法均需要输入原图像和提示图，这里采用目前被广泛接受的三值图的形式。数据平台中的三值图从提示信息由强到弱分为三个级别。

(3) 评价准则



图 2-7 感知数据库中的训练和测试图片

评价分为主观评价和客观评价两个方面，传统的客观评价标准采用的误差函数，均方和误差 MSE(Mean Squared Error)和绝对误差和 SAD(Sum of Absolute Difference)。为了使评价准则更加具有主观感知性，在建立平台时，参加主观感知测试的十几名裁判对不同掩膜及其与新场景融合后的效果进行打分。从打分的结果总结出以下两个评价准则，梯度域错误和掩膜的连续性，定义如下式：

$$\sum_i (\nabla \alpha_i - \nabla \alpha_i^*)^q \quad (2-14)$$

$$\sum_i (\varphi(\alpha_i, \Omega) - \varphi(\alpha_i^*, \Omega))^p \quad (2-15)$$

其中 $\nabla\alpha$ 和 $\nabla\alpha_i^*$ 分别为结果和理想结果的以 δ 为参数的一阶高斯滤波结果。 φ 为连通函数，定义详见文献[30]。

本实验的环境为 MATLAB，实验所有数据均来自感知测试平台^[30]，实验用测试数据共 8 幅图片，分别在 3 个层次的提示图条件下的测试结果如表 2-2 至表 2-9 所示：

表 2-2 样本 1-4 的 SAD 测试结果

| Image matting evaluation results Competition: Error type SAD | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Sum of Absolute Difference | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Our Method | 14.5 | 14.8 | 14.6 | 14.3 | 18 17 15 | 17 17 17 | 5 8 7 | 16 13 14 |
| Iterative BP | 14.8 | 14.1 | 15 | 15.4 | 16 16 17 | 15 16 14 | 10 15 18 | 13 12 11 |
| Easy matting | 15.5 | 15.5 | 15.3 | 15.5 | 13 13 14 | 14 13 15 | 18 18 17 | 18 19 18 |
| Bayesian | 16.6 | 16.6 | 17.5 | 15.6 | 17 18 18 | 18 18 18 | 15 16 13 | 14 16 10 |
| Poisson | 18.8 | 19 | 18.6 | 18.8 | 19 19 19 | 19 19 19 | 19 19 19 | 19 18 19 |

表 2-3 样本 5-8 的 SAD 测试结果

| Image matting evaluation results Competition: Error type SAD | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|--|--|---|
| Sum of Absolute Difference | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Our Method | 14.5 | 14.8 | 14.6 | 14.3 | 17 16 15 | 18 17 16 | 11 14 13 | 16 15 17 |
| Iterative BP | 14.8 | 14.1 | 15 | 15.4 | 14 13 14 | 15 16 18 | 16 16 15 | 14 16 16 |
| Easy matting | 15.5 | 15.5 | 15.3 | 15.5 | 15 15 16 | 16 15 17 | 15 15 14 | 15 14 14 |
| Bayesian | 16.6 | 16.6 | 17.5 | 15.6 | 18 18 18 | 17 18 15 | 17 17 19 | 17 19 15 |
| Poisson | 18.8 | 19 | 18.6 | 18.8 | 19 19 19 | 19 19 19 | 19 18 17 | 19 18 19 |

表 2-4 样本 1-4 的 MSE 测试结果

| Image matting evaluation results Competition: Error type MSE | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Mean Squard Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Our Method | 14.7 | 14.6 | 15 | 14.4 | 16 17 15 | 17 17 17 | 6 11 8 | 16 13 14 |
| Geodesic | 14.7 | 15 | 14.6 | 14.4 | 14 14 12 | 10 19 8 | 15 12 15 | 15 15 13 |
| Bayesian | 16.4 | 16.4 | 17.1 | 15.6 | 15 18 18 | 18 18 18 | 16 16 13 | 14 16 10 |
| Easy Matting | 16.7 | 16.6 | 16.6 | 16.8 | 17 15 16 | 16 16 16 | 18 18 18 | 18 19 18 |
| Poisson | 18.9 | 19 | 18.8 | 18.9 | 19 19 19 | 19 19 19 | 19 19 19 | 19 18 19 |

表 2-5 样本 5-8 的 MSE 测试结果

| Image matting evaluation results Competition: Error type MSE | | | | | | | | |
|---|---------------|----------------|----------------|---------------|--|---|---|--|
| Mean Squard Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Our Method | 14.7 | 14.6 | 15 | 14.4 | 17 16 15 | 18 17 16 | 12 14 13 | 15 15 17 |
| Geodesic | 14.7 | 15 | 14.5 | 14.4 | 16 17 18 | 14 14 14 | 18 19 17 | 18 17 17 |
| Bayesian | 16.4 | 16.4 | 17.1 | 15.6 | 18 18 17 | 16 16 15 | 17 17 19 | 17 18 15 |
| Easy matting | 16.7 | 16.6 | 16.6 | 16.8 | 15 15 16 | 17 18 18 | 16 16 16 | 16 16 16 |
| Poisson | 18.8 | 19 | 18.9 | 18.9 | 19 19 19 | 19 19 19 | 19 18 18 | 19 19 19 |

表 2-6 样本 1-4 的 Gradient 测试结果

| Image matting evaluation results Competition: Error type Grad. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|--|--|--|--|
| Gradient Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Our Method | 15 | 15.4 | 15 | 14.6 | 15 16 12 | 17 17 17 | 15 15 15 | 16 15 15 |
| Geodesic | 15.3 | 15.8 | 15 | 15.1 | 17 14 14 | 8 7 8 | 17 16 14 | 15 14 17 |
| Bayesian | 16.5 | 16 | 17 | 16.4 | 14 17 17 | 16 16 16 | 16 17 16 | 14 17 14 |
| Easy Matting | 16.6 | 16.6 | 16.6 | 16.6 | 16 15 16 | 15 15 15 | 18 19 18 | 18 18 18 |
| Poisson | 18.9 | 19 | 18.6 | 19 | 19 19 19 | 19 19 19 | 19 18 19 | 19 19 19 |

表 2-7 样本 5-8 的 Gradient 测试结果

| Image matting evaluation results Competition: Error type Grad. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|--|---|---|--|
| Gradient Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Our Method | 15 | 15.4 | 15 | 14.6 | 15 15 14 | 16 16 16 | 14 14 14 | 15 12 14 |
| Geodesic | 15.3 | 15.8 | 15 | 15.1 | 18 19 18 | 15 14 14 | 18 19 18 | 18 17 18 |
| Bayesian | 16.5 | 16 | 17 | 16.4 | 17 17 17 | 17 17 17 | 17 17 17 | 17 18 17 |
| Easy matting | 16.6 | 16.6 | 16.6 | 16.6 | 16 16 16 | 18 18 18 | 16 16 16 | 16 16 16 |
| Poisson | 18.9 | 19 | 18.6 | 19 | 19 18 19 | 19 19 19 | 19 18 19 | 19 19 19 |

表 2-8 样本 1-4 的 Connectivity 测试结果

| Image matting evaluation results Competition: Error type Conn. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|--|--|--|--|
| Connectivity Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Geodesic | 14 | 14.1 | 14.1 | 13.9 | 16 12 6 | 11 16 15 | 9 15 12 | 10 3 11 |
| Easy Matting | 14.7 | 14.3 | 14.1 | 15.8 | 11 11 11 | 16 14 16 | 3 4 14 | 19 18 19 |
| Our Method | 16 | 16.1 | 15.5 | 16.3 | 18 16 17 | 17 16 18 | 14 17 16 | 17 17 16 |
| Bayesian | 17.3 | 16.4 | 17.3 | 18.1 | 17 17 18 | 19 19 17 | 6 11 18 | 18 19 18 |
| Poisson | 17.3 | 17.5 | 17.1 | 17.4 | 19 19 19 | 18 17 19 | 19 19 19 | 14 11 14 |

表 2-9 样本 5-8 的 Connectivity 测试结果

| Image matting evaluation results Competition: Error type Conn. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|--|---|---|--|
| Connectivity Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Geodesic | 14 | 14.1 | 14.1 | 13.9 | 15 16 17 | 15 15 17 | 16 16 16 | 19 18 17 |
| Easy Matting | 14.7 | 14.3 | 14.1 | 15.8 | 16 17 15 | 16 18 19 | 17 16 16 | 16 15 16 |
| Our Method | 16 | 16.1 | 15.5 | 16.3 | 17 12 16 | 19 17 16 | 12 13 14 | 15 12 17 |
| Bayesian | 17.3 | 16.4 | 17.3 | 18.1 | 18 18 19 | 17 16 18 | 19 19 19 | 17 19 18 |
| Poisson | 17.3 | 17.5 | 17.1 | 17.4 | 19 19 18 | 18 19 14 | 15 17 17 | 18 16 19 |

表 2-2 至 2-9 分别为 8 幅测试用例与性能相近的 5 种算法的均方和误差(MSE), 绝对误差和 (SAD), 梯度域错误 (Grad.), 掩膜平滑性 (Conn.) 四种指标的比较结果。其中 Geodesic Matting 见文献[34], Easay Matting 见文献[35]。本章算法的平均排名为 14, 单一测试样本中的最低排名为 17, 最高排名为 5。

8 幅测试图片在三值图提示图等级为 Small 下利用本章方法提取出的 α 掩膜如图 2-8 和 2-9 所示:

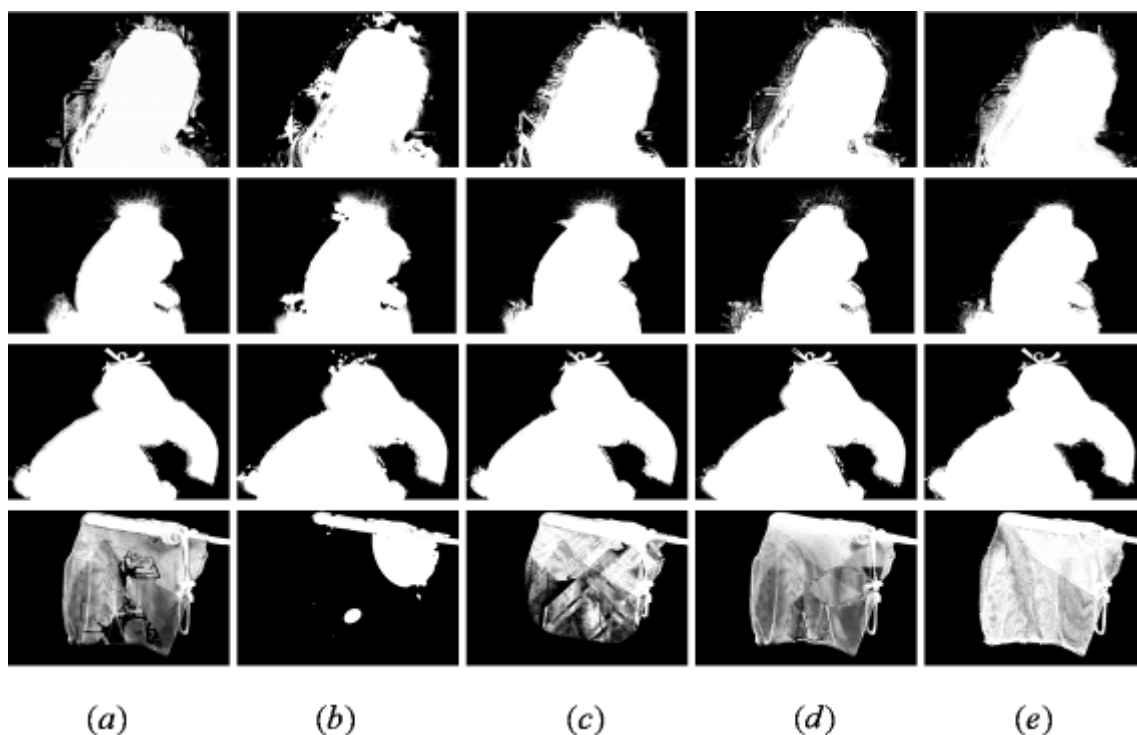
图 2-8 测试图片 1-4 的 α 掩膜

图 2-8 中, 第 a 至 e 列分别为本章方法, Poisson Matting, Bayesian Matting, Easy Matting, Iterative BP 的抠图结果。

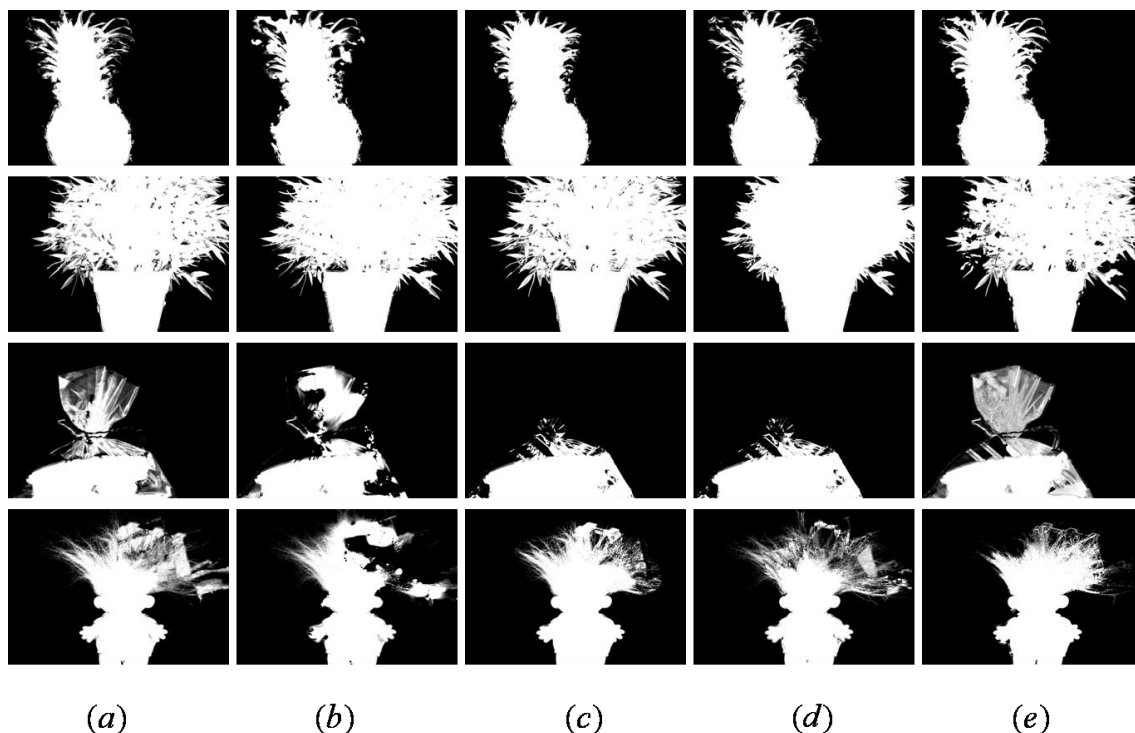
图 2-9 测试图片 5-8 的 α 掩膜

图 2-9 中，第 a 至 e 列分别为本章方法，Poisson Matting，Bayesian Matting，Easy Matting，Iterative BP 的抠图结果。以上 5 种抠图方法都出现了不同程度的估计错误。样本 net，plasticbag，troll 最为明显其中 net 和 plasticbag 的透明度等级为高，troll 的透明度等级为极高。

图 2-10 和图 2-11 显示了两个抠图和合成的结果示例，左图为原始图像，右图为经过本文算法提取出的前景图像。中间为将前景图像融合到新场景中的图像。



图 2-10 抠图效果示例 1



图 2-11 抠图效果示例 2

2.4 本章总结

本章详细介绍了背景还原技术中基于样本的背景还原方法，该方法能够从场景图片中将前景对象移除，移除前景对象后的背景还原图片能够较真实的反映场景信息，为后续抠图工作提供了先验信息。最近邻采样的方法原理简单，计算速度快，利用该采样方法估计像素点的前景颜色，结合已经还原的背景信息，得到了待估计像素点的颜色样本值，算法后续步骤中还利用邻近像素点间颜色和掩膜的相似性，加入了平滑因子，通过由颜色样本误差和平滑项误差构成的抠图代价函数，求解出较好的抠图掩膜。

基于背景还原的抠图算法在测试平台下的测试结果显示，抠图准确性优于现有方法如 Bayesian^[17]，Possion^[18]，但抠图质量在现有方法的整体排名不高，仅为 14。尽管抠图质量处于一般水平，但文中提出的用图像修复的方法得到背景信息去解决抠图问题的思想，将抠图问题和图像复原联系起来，相信该想法对后续研究工作会有积极的影响。

第三章 基于纹理方向采样的抠图方法

3.1 引言

鲁棒抠图^[20]方法取得成功之后,许多学者开始研究将采样和传播相结合的方法。其中最成功的方法有文献[36]中, Gastal 和 Oliveria 提出的共享采样抠图和文献[37]中, Kaiming He 提出了一种全局采样抠图。

文献[36]中, Gastal 和 Oliveria 提出了实时共享采样算法。该方法通过不停旋转采样方向,使得相邻像素总在不同方向上采集前/背景样本值,然后将局部区域内所有像素采集到的样本进行共享,从而增加了采集到正确样本的概率。文章中还定义了一系列的方法用于判定样本对的可靠程度,其中包括定义了反映样本对颜色可靠程度的颜色扭曲因子 $M_p(F, B)$; 反映样本对空间联系的因子 $E_p(s)$ 。后续还加入了一系列的样本修正和平滑策略。该算法虽然运算量大,但是其中每一步操作均相互独立,因此可以利用并行计算技术加速使得其最终执行速度非常快,甚至可以进行实时处理。

文献[37]中, Kaiming He 提出了一种全局采样思想,该方法通过将指定的前景样本集和背景样本集建立样本对代价谱,然后采用一种改进的随机搜索算法^[16]在建立好的样本对代价谱中检索到最优的颜色样本对。该随机搜索算法使得处理每一个像素点,只需要进行 10 次以内的查找就能找到最佳的样本对。作者后续还利用闭合场以及指导滤波器对结果进行了改进。

分析这两种方法,其取得成功的要素有两点:第一,对每一个待估计像素点 I ,直接或间接地选择了大量的样本对,如文献[20]中,采样每次旋转 18° ,即有 20 个不同的采样方向,因此,在文献[20]的采样策略中,相当于对每一个像素点分别采集了 20 个不同的前景,背景样本。共计可以构成 $20 \times 20 = 400$ 对不同的颜色样本对。第二,定义了准确的样本对判别标准,如文献[20]不仅考虑了像素点的颜色,空间分布,还考虑了相邻像素间的相似性。

本文通过观察发现图像中前景物体在边缘处往往表现出具有某种方向性,因此提出了一种基于纹理方向采样的抠图算法。

该方法可以分为纹理方向采样,样本对优化及掩膜平滑性优化三个步骤。纹理方向采样的作用是利用图像自身颜色的统计特性用尽量少的样本个数采集到尽

量准确的前景/背景样本颜色值；样本对优化过程至关重要，它保证了在采集到的大量的样本颜色对中，尽量选择正确的，排除错误的；掩膜平滑性优化从 α 掩膜的平滑性出发，使算法不会由于前面采样过程的错误导致结果的明显错误，该过程还可以提高掩膜的质量，使得融合后的图像更加真实。

3.2 纹理方向采样

自然图像抠图所关心的区域往往是图像中前景对象的边界区域。这一区域内的像素有以下三个特点：第一，边界区域内的像素点容易与背景中颜色相似的部分发生混合，难于分辨；第二，边界区域存在半透明像素如衣服，毛发及一些容易受光照产生颜色畸变的像素。这些像素点的掩膜值通常介于 0 到 1 之间，这使得精确求解正确的掩膜值变得复杂；第三，由于图像分辨率的原因，真实图像的颜色在低分辨率图像中往往发生混合形成亚像素点，求解这些亚像素点的掩膜值，使得抠图问题进一步变得复杂。

虽然单个像素点颜色复杂多变，但局部区域内的所有像素点形成的像素块往往满足某种统计规律，如图 3-1 所示：

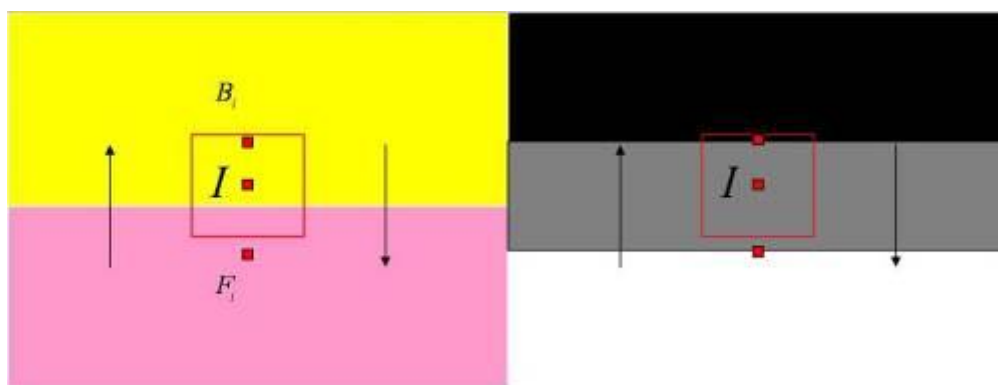


图 3-1 像素块 $P(I)$ 的梯度具有方向性

为了分析图像纹理与采样方向的关系，首先假设一个最简单的抠图问题。图 3-1 中不含任何半透明像素，前景和背景颜色也都为纯色。左图为假想的原始图像，其上半部分（黄色区域）属于背景物体，下半部分（粉色区域）属于背景物体。右图为左图对应的三值图，黑色部分为绝对背景 R_b ，白色部分为绝对前景 R_f ，灰色部分对应待求解区域 R_u 。

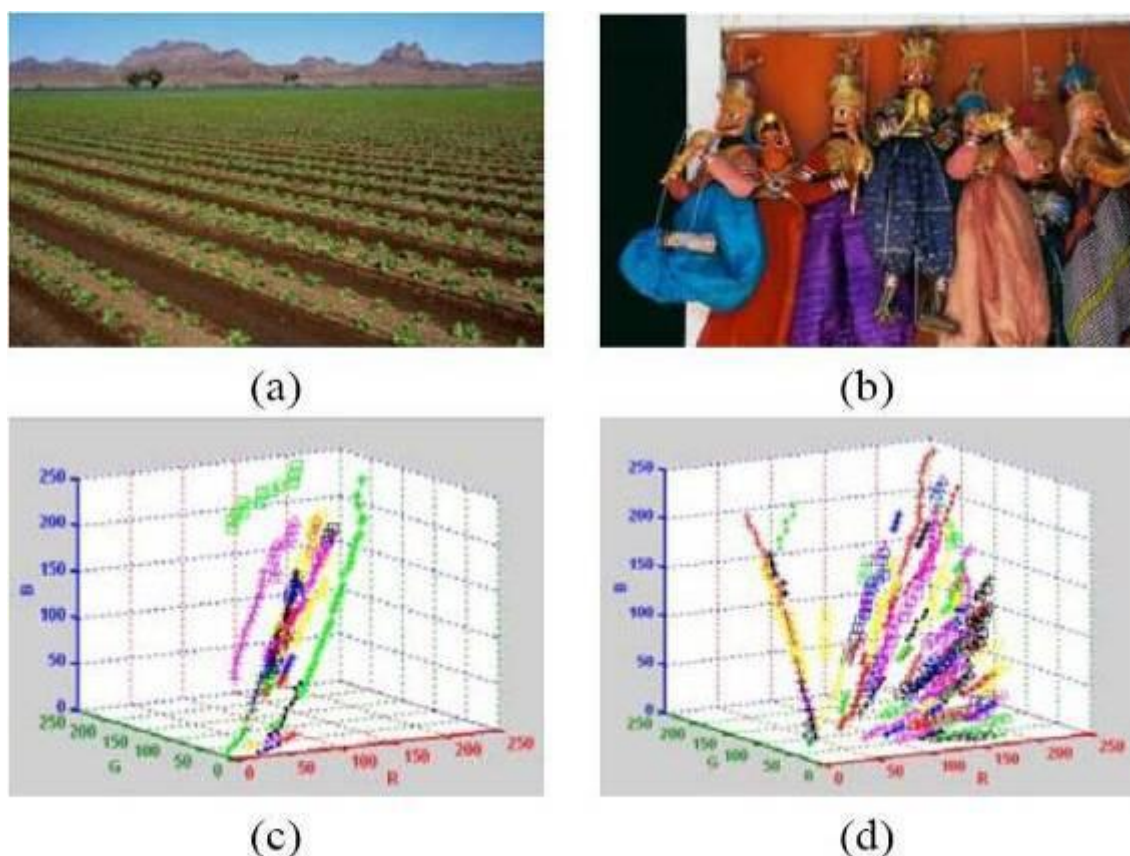


图 3-2 自然图像的颜色直方图

对于像素点 I ，其在水平方向 x 及垂直方向 y 上的梯度值均为 0：

$$\partial I / \partial x = 0, \quad \partial I / \partial y = 0 \quad (3-1)$$

而对于像素块 $P(I)$ ，其在水平方向 x 的梯度值为 0，在垂直方向 y 的梯度值不为 0：

$$\partial P(I) / \partial x = 0, \quad \partial P(I) / \partial y \neq 0 \quad (3-2)$$

图 3-1 左，前景区域(粉色)的边界方向恰好是 x 轴方向，正确的采样方向恰好是 y 轴方向。沿 y 轴方向（如图 3-1 中黑色箭头所示），在区域 R_u 与区域 R_b 、 R_f 的交界处分别采集到点 I 的背景颜色值 B_i ，前景颜色值 F_i ，进而通过公式 2-6 计算出点 I 的掩膜值 $\alpha_i = 0$ ，即 I 点属于绝对背景。

从假想图 3-1 所假设的抠图示例中可以得到以下两点假设。

假设一：在图像中，以物体边界处某一像素点 I 为中心像素点， I 邻域内所有像素点构成的像素块 $P(I)$ 的梯度方向与物体的边缘方向有相关性。

假设二：沿着与物体边缘垂直的方向采样得到的颜色样本值具有更高的可信程度。

3.2.1 颜色线模型

Omer 和 Werman^[29]首次提出了颜色线模型。图像中的两个像素点的颜色由于反射，光照及传感器在成像过程中出现扭曲，往往偏离其真实颜色。因此在图像中很难直接判断两个像素点是否属于同一物体。文献[29]中，作者通过对图像的颜色进行统计，发现颜色线对于颜色扭曲具有鲁棒性，于是利用颜色线对场景颜色进行重表达。

文献[29]统计了伯克利图像分割数据库[38]中 200 幅自然图像的 RGB 颜色直方图如图 3-2 所示：

发现自然图像像素点的分布在 RGB 颜色直方图中总体上呈近似直线。为了研究图像失真和颜色线的关系，在文献[29]中，作者进一步对场景引起的失真，传感器引起的失真以及其它失真这三种情况引起的图像颜色直方图进行了分析。如图 3-3 所示，为光照改变后物体颜色线的相应变化：

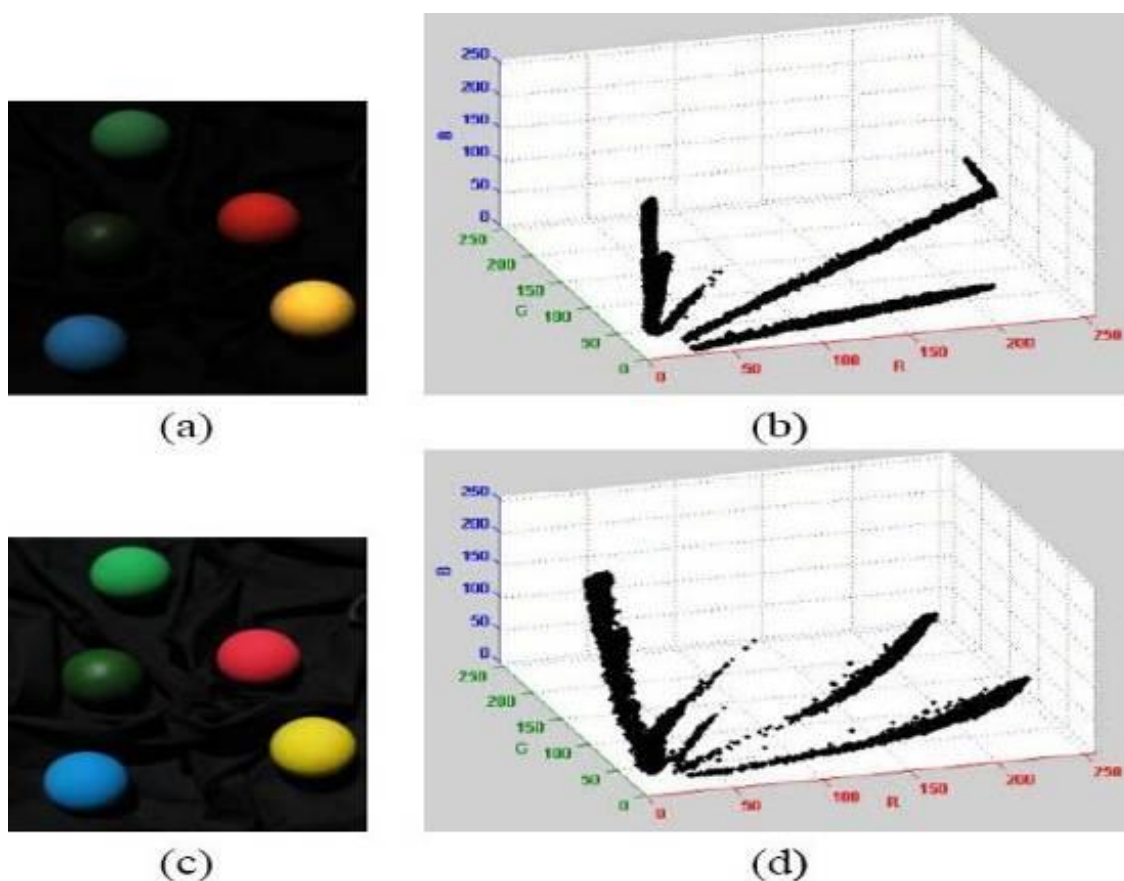


图 3-3 光照改变及其颜色线的变化

从图 3-3 中,可以看出,失真并不会影响图像像素在颜色直方图中呈线性分布。文献[31]中, Levin 基于 Omer 和 Werman 的统计,进一步分析,指出在图像的局部区域内像素点的颜色分布在的 RGB 空间中,总是满足颜色线模型,Levin 运用这种局部线性假设,推导出抠图问题的闭合场解。颜色线模型还解释了公式 2-8 中,代价函数 E_l 的合理性,可以作如下理解:假设当前像素点 I 在真实的自然环境中属于前景物体(颜色为 F),由于背景物体(颜色为 B)的影响而产生失真。则当失真后的像素颜色(即当前像素点 I 的颜色值)处于前景物体颜色 F 和背景物体颜色 B 所组成的颜色线上时,代价函数 E_l 最小。

3.2.2 基于纹理方向采样的抠图算法

基于前两节的分析和介绍,这一小节将详细介绍本文提出的抠图算法,算法具体流程如图 3-4 所示:

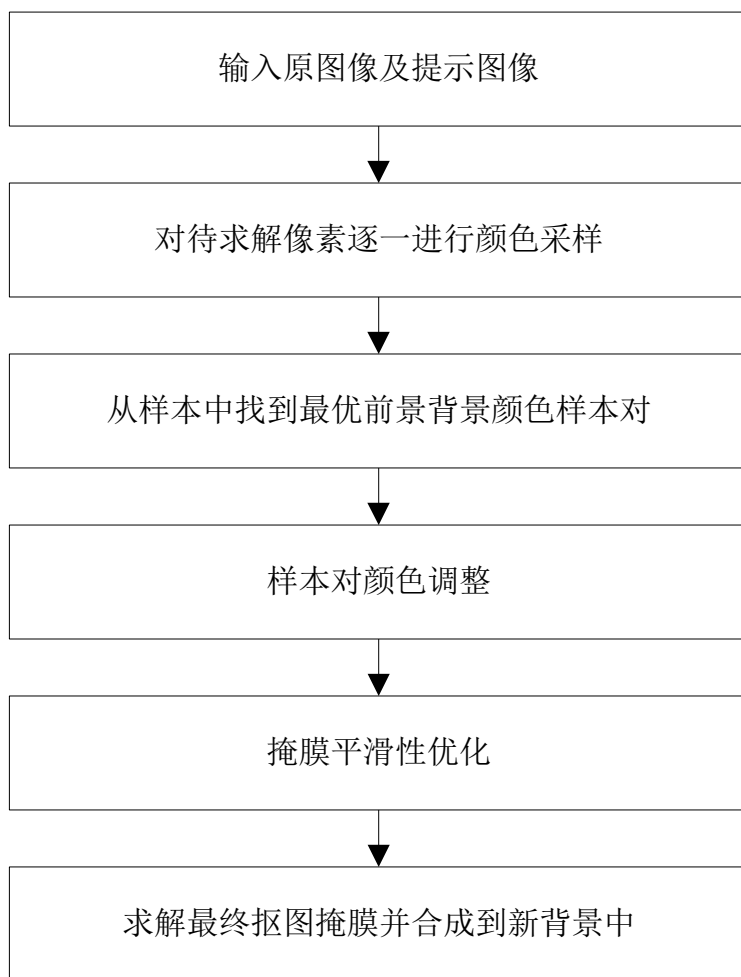


图 3-4 本节算法基本流程图

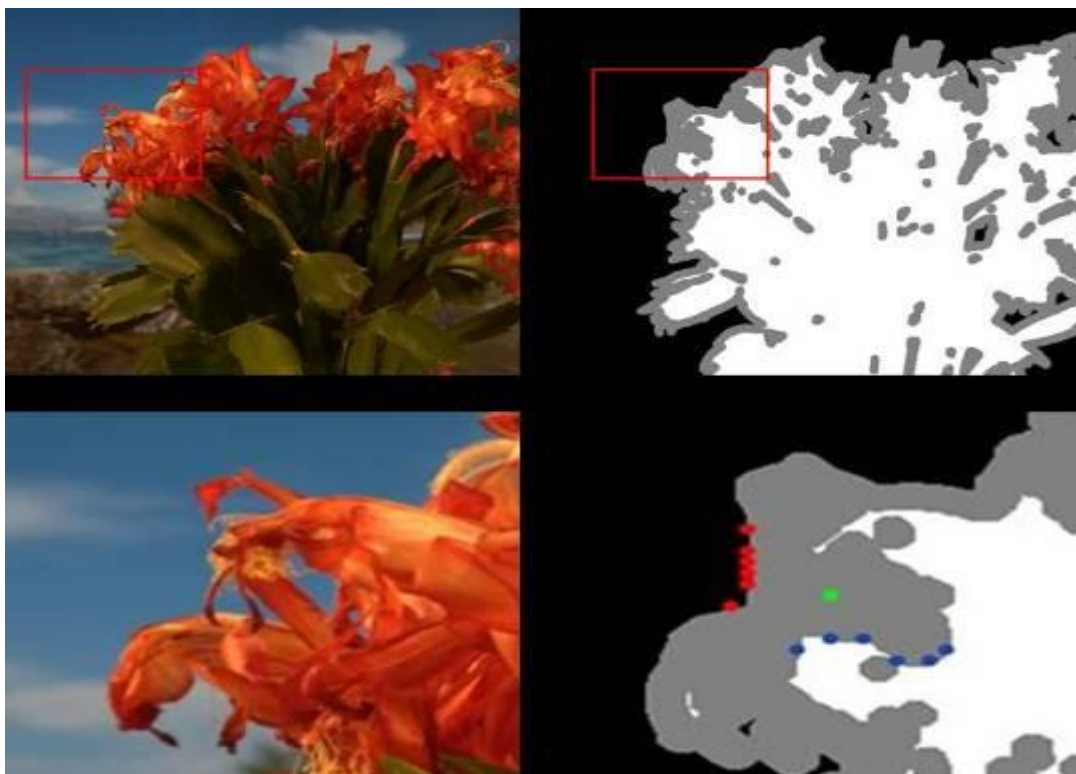


图 3-5 颜色对采样点示例图

以下分别具体介绍基本流程中的各部分内容：

(1) 输入图像及提示图

输入为任意自然背景图像及其三值提示图，三值图（Trimap）的定义与本文第二章中所采样的方法完全一样，如图 2-2 所示，其中黑色区域为绝对背景区域，区域内的掩膜值 $\alpha = 0$ ；白色区域为绝对前景区域，区域内的掩膜值 $\alpha = 1$ ；灰色区域为待估计区域，掩膜值需要通过后续步骤求解。

(2) 对三值图中未知掩膜值区域 R_u 中的像素点逐一进行颜色采样

假设当前正在处理的像素点为 i ，计算以当前待估计像素点 i 为中心点所包含的图像块 P_i ，分别在 0 、 $\pi/4$ 、 $\pi/2$ 、 $3\pi/4$ 、 π 、 $5\pi/4$ 、 $3\pi/2$ 、 $7\pi/4$ 这 8 个方向上的变化速度。其中， P_i 块的大小为 9×9 的矩形。沿着 P_i 变化速度最慢的方向的垂直方向发射射线，射线与前景轮廓线和背景轮廓线的交点为所采集到的样本点 f_1 ， b_1 ；找前景轮廓线和背景轮廓线上距离当前待估计像素点 i 最近的点为所采集到的样本点 f_2 ， b_2 ；在 f_1 、 b_1 、 f_2 、 b_2 两边分别稀疏的各取一个点，构成当

前待估计像素点 i 经过采样后的最终样本集合 $\{f_{1,...,6},\}, \{b_{1,...,6}\}$ 。 P_i 的变化速度通过求解下式得到：

$$D_n = \min_{k \in \{1,...,8\}} SSD(P_i - P_k) \quad (3-3)$$

其中，SSD 表示均方误差和(Sum of Squared Difference)， k 代表方向， n 代表 SSD 最小的方向，即 P_i 变化最小的方向由 n 的值表示。 f ， b 为三维向量，表示样本点颜色的 RGB 值。图 3-5 为对示例点 i 采样的示例图。图 3-5 中左上角为原图像，右上角为原图像对应的三值提示图。左下角为原图像在红色矩形内的局部图，右下为对区域 R_u 内绿色像素点 i 采样，得到前景样本点和背景样本点的示例图。图 3-5 中，绿色像素点通过最近邻采样和纹理采样的方法，共得到 6 个前景样本点，6 个背景样本点，这些点的颜色对组合有 36 种可能性。

(3)样本对颜色调整

首先，需要将(2)中采集到的样本集合 $\{f_{1,...,6},\}, \{b_{1,...,6}\}$ 进行两两组合，利用代价函数求解，得到代价最小的一对样本对，记为 $(f, b)_{best}^i$ 。代价函数如下式：

$$E(F^i, B^i) = \omega E_c(F^i, B^i) + E_s(F^i) + E_s(B^i) \quad (3-4)$$

其中 E_c 为颜色匹配误差， $E_s(F^i)$ 和 $E_s(B^i)$ 分别为前景样本和背景样本的空间距离误差， ω 为权值用来平衡颜色和距离误差对整个代价函数的影响。

E_c ， $\hat{\alpha}$ ， $E_s(F^i)$ ， $E_s(B^i)$ 的计算公式分别为：

$$E_c(F^i, B^i) = \|I - (\hat{\alpha} F^i + (1 - \hat{\alpha}) B^i)\| \quad (3-5)$$

$$\hat{\alpha} = \frac{(I - B^i)(F^i - B^i)}{\|F^i - B^i\|^2} \quad (3-6)$$

$$E_s(F^i) = \frac{Dis(F^i, I)}{\min Dis(F_{all}, I)} \quad (3-7)$$

$$E_s(B^i) = \frac{Dis(B^i, I)}{\min Dis(B_{all}, I)} \quad (3-8)$$

其中， $Dis(F^i, I)$ ， $Dis(B^i, I)$ 分别为样本点到当前处理点 i 的空间距离， $\min Dis(F_{all}, I)$ ， $\min Dis(B_{all}, I)$ 分别为当前处理点 i 到前景轮廓线和背景轮廓线的最短距离。

然后，利用文献[36]中的分享采样思想，对通过求解代价函数 E_c 找到的样本对 $(f, b)_{best}^i$ 进行颜色调整。调整方法如下，假设当前估计像素 i ，以 i 点为中心点的 5×5 共 25 个像素点构成的矩形区域内，共有 25 对样本对 $(f, b)_{best}^i$ ，将这 25 对样本利用公式 3-5 和 3-6 与 i 点做颜色距离测量，找到距离最小的前 3 对，将这最佳的 3 对进行颜色值的平均，记为 (F^i, B^i) ，颜色距离的测量同步骤三中 E_c 的计算。

(4)掩膜平滑性

由 3.2.2 节的颜色线模型得到启示，图像局部区域内的前景颜色值 F 和背景颜色 B ，均分布在直线上，则可以通过求解他们的均值，得到低频信息。将 F^i ， B^i 分别在 3×3 邻域内取颜色均值，得到 F_l^i ， B_l^i ，其中 l 代表低频，公式 3-9 求取低频掩膜值 α_l^i ：

$$\alpha_l^i = \frac{(I^i - B_l^i) F_l^i - B_l^i}{||F_l^i - B_l^i||} \quad (3-9)$$

(5)求解最终掩膜并融合到新场景中

最终的掩膜值由颜色误差 E_c 和低频掩膜 α_l^i 共同组成：

$$\alpha^i = \frac{\lambda \cdot \alpha_l^i + (I^i - B^i)(F^i - B^i)}{||F^i - B^i||^2 + \lambda} \quad (3-10)$$

逐一求解区域 R_i 内所有像素点的掩膜值 α ，得到图像最终的抠图掩膜后，即可将合成公式 1-1 中的背景 B 替换成任意新背景，实现抠图与融合。

3.3 实验结果

本章实验数据同第二章，采用文献[30]中提出的测试平台。数据库中 35 幅图片，其中测试图片 8 幅，训练数据 27 幅。测试图片为真实结果隐藏图，这样可以避免个别方法通过调节极端的参数取得好的实验效果，训练图片的真实结果对外公布，便于研究人员改进抠图算法。用本章所用的实验方法对 8 幅测试图片进行测试，结果与性能相近的 5 种方法比较如表 3-1 至表 3-8 所示：

表 3-1 样本 1-4 的 SAD 测试结果

| Image matting evaluation results Competition: Error type SAD | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Sum of Absolute Difference | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Our Method | 13.7 | 14.9 | 13.3 | 12.4 | 13 11 3 | 11 13 12 | 18 18 18 | 15 15 12 |
| Random Walk | 13.7 | 14.4 | 12.5 | 14.1 | 16 10 16 | 17 15 17 | 16 15 13 | 17 17 17 |
| Geodesic | 14.4 | 15 | 14.1 | 14 | 15 15 8 | 8 7 7 | 15 12 15 | 16 13 14 |
| IterativeBP | 15.1 | 14.4 | 15.3 | 15.6 | 17 17 17 | 16 17 15 | 9 14 17 | 13 12 11 |
| EasyMatting | 15.8 | 15.8 | 15.8 | 15.9 | 14 14 15 | 15 14 16 | 17 17 16 | 18 19 18 |

表 3-2 样本 5-8 的 SAD 测试结果

| Image matting evaluation results Competition: Error type SAD | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|--|--|---|
| Sum of Absolute Difference | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Our Method | 13.7 | 14.9 | 13.3 | 12.4 | 16 11 14 | 15 14 14 | 15 13 13 | 15 12 13 |
| Random Walk | 13.7 | 14.4 | 12.5 | 14.1 | 11 2 9 | 13 12 13 | 13 14 15 | 12 8 12 |
| Geodesic | 14.4 | 15 | 14.1 | 14 | 17 17 17 | 14 13 15 | 18 19 18 | 18 17 18 |
| IterativeBP | 15.1 | 14.4 | 15.3 | 15.6 | 14 13 15 | 16 17 18 | 16 16 15 | 14 16 17 |
| EasyMatting | 15.8 | 15.8 | 15.8 | 15.9 | 15 16 16 | 17 16 17 | 14 15 14 | 16 15 15 |

表 3-3 样本 1-4 的 MSE 测试结果

| Image matting evaluation results Competition: Error type MSE | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Mean Squard Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Geodesic | 14.7 | 15 | 14.5 | 14.6 | 15 15 13 | 10 9 8 | 14 12 14 | 15 14 14 |
| Our Method | 14.9 | 15.8 | 15.1 | 13.8 | 14 13 8 | 15 16 14 | 18 18 18 | 16 15 12 |
| Bayesian | 16.5 | 16.5 | 17.1 | 15.8 | 16 18 18 | 18 18 18 | 15 15 12 | 14 16 10 |
| Easy Matting | 16.8 | 16.8 | 16.9 | 16.9 | 17 16 16 | 17 17 17 | 17 17 17 | 18 19 18 |
| Poisson | 18.9 | 19 | 18.8 | 18.9 | 19 19 19 | 19 19 19 | 19 19 19 | 19 18 19 |

表 3-4 样本 5-8 的 MSE 测试结果

| Image matting evaluation results Competition: Error type MSE | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|--|--|---|
| Mean Squard Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Geodesic | 14.7 | 15 | 14.5 | 14.6 | 16 17 16 | 14 14 15 | 18 19 17 | 18 17 18 |
| Our Method | 14.9 | 15.8 | 15.1 | 13.8 | 17 15 15 | 16 15 14 | 15 14 14 | 15 15 15 |
| Bayesian | 16.5 | 16.5 | 17.1 | 15.8 | 18 18 17 | 17 17 16 | 17 17 19 | 17 18 16 |
| Easy Matting | 16.8 | 16.8 | 16.9 | 16.9 | 15 16 16 | 18 18 18 | 16 16 16 | 16 16 17 |
| Poisson | 18.9 | 19 | 18.8 | 18.9 | 19 19 19 | 19 19 19 | 19 18 18 | 19 19 19 |

表 3-5 样本 1-4 的 Gradient 测试结果

| Image matting evaluation results Competition: Error type Grad. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Gradient Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Geodesic | 15.2 | 15.6 | 15 | 15 | 17 15 14 | 8 7 8 | 16 15 14 | 15 14 16 |
| Our Method | 15.7 | 15.6 | 16.1 | 15.5 | 12 14 9 | 11 16 14 | 19 17 19 | 17 19 17 |
| Bayesian | 16.3 | 16 | 16.8 | 16.3 | 15 17 17 | 17 17 17 | 15 16 15 | 14 16 14 |
| Easy Matting | 16.3 | 16.3 | 16.5 | 16.3 | 16 16 15 | 16 15 16 | 17 19 17 | 18 17 18 |
| Poisson | 18.8 | 18.9 | 18.5 | 18.9 | 19 19 19 | 19 19 19 | 18 18 18 | 19 18 19 |

表 3-6 样本 5-8 的 Gradient 测试结果

| Image matting evaluation results Competition: Error type Grad. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|--|--|---|
| Gradient Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Geodesic | 15.2 | 15.6 | 15 | 15 | 18 19 18 | 15 14 14 | 18 19 18 | 18 17 18 |
| Our Method | 15.7 | 15.6 | 16.1 | 15.5 | 16 15 16 | 17 17 17 | 16 16 16 | 16 15 16 |
| Bayesian | 16.3 | 16 | 16.8 | 16.3 | 17 17 17 | 16 16 16 | 17 17 17 | 17 18 17 |
| Easy Matting | 16.3 | 16.3 | 16.5 | 16.3 | 15 16 15 | 18 18 18 | 15 15 15 | 15 16 15 |
| Poisson | 18.8 | 18.9 | 18.5 | 18.9 | 19 18 19 | 19 19 19 | 19 18 19 | 19 19 19 |

表 3-7 样本 1-4 的 Connectivity 测试结果

| Image matting evaluation results Competition: Error type Conn. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|---|---|---|
| Connectivity Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Plant (Little Transparent) Small user large | Pineapple (Little Transparent) Small user large | Plastic bag (Highly Transparent) Small user large | Net (Highly Transparent) Small user large |
| Our Method | 14.2 | 15 | 14 | 13.5 | 9 11 10 | 15 16 17 | 19 18 13 | 15 13 15 |
| Geodesic | 14.3 | 14.4 | 14.4 | 14 | 17 13 6 | 11 17 15 | 9 15 12 | 10 3 12 |
| Easy Matting | 15 | 14.8 | 14.3 | 16.1 | 12 12 12 | 17 14 16 | 3 4 15 | 19 18 19 |
| Poisson | 17.4 | 17.4 | 17.3 | 17.5 | 19 19 19 | 18 18 19 | 18 19 19 | 14 11 15 |
| Bayesian | 17.5 | 16.6 | 17.5 | 18.8 | 18 18 18 | 19 19 18 | 6 11 18 | 18 19 18 |

表 3-8 样本 5-8 的 Connectivity 测试结果

| Image matting evaluation results Competition: Error type Conn. | | | | | | | | |
|---|---------------|----------------|----------------|---------------|---|--|--|---|
| Connectivity Error | Over all Rank | Avg small Rank | Avg large Rank | Avg user Rank | Troll (strongly Transparent) Small user large | Doll (strongly Transparent) Small user large | Donkey (Medium Transparent) Small user large | Elephantt (Medium Transparent) Small user large |
| Our Method | 14.2 | 15 | 14 | 13.5 | 15 16 16 | 16 14 16 | 16 16 15 | 15 11 15 |
| Geodesic | 14.3 | 14.4 | 14.4 | 14 | 16 15 17 | 15 16 17 | 18 18 18 | 19 18 15 |
| Easy Matting | 15 | 14.8 | 14.3 | 16.1 | 17 17 15 | 17 18 19 | 17 15 16 | 16 15 17 |
| Poisson | 17.4 | 17.4 | 17.3 | 17.5 | 19 19 18 | 19 19 14 | 14 17 17 | 18 16 19 |
| Bayesian | 17.5 | 16.6 | 17.5 | 18.8 | 18 18 19 | 18 17 18 | 19 19 19 | 17 19 18 |

表 3-1 至表 3-8 分别列出了均方和误差 (MSE), 绝对误差和 (SAD), 梯度域错误 (Grad.), 掩膜平滑性 (Conn.) 四种指标的比较结果。左起第一列为方法名; 第二列为性能总体排名, 按三值图提示级别由 small, large, user 所得结果分别在第三至第五列; 剩余四列为各个样本在不同级别的提示图下表现的排名。其中本章算法的平均排名为 13, 单一测试样本中的最低排名为 17, 最高排名为 3。单一样本中表现最好的是 plant, 表现最差的是 plasticbag。参与排名的其它方法见参考文献[39-48]

8 幅测试图片在三值图提示图等级为 Small 下利用本章方法提取出的 α 掩膜如图 3-6 和 3-7 所示:

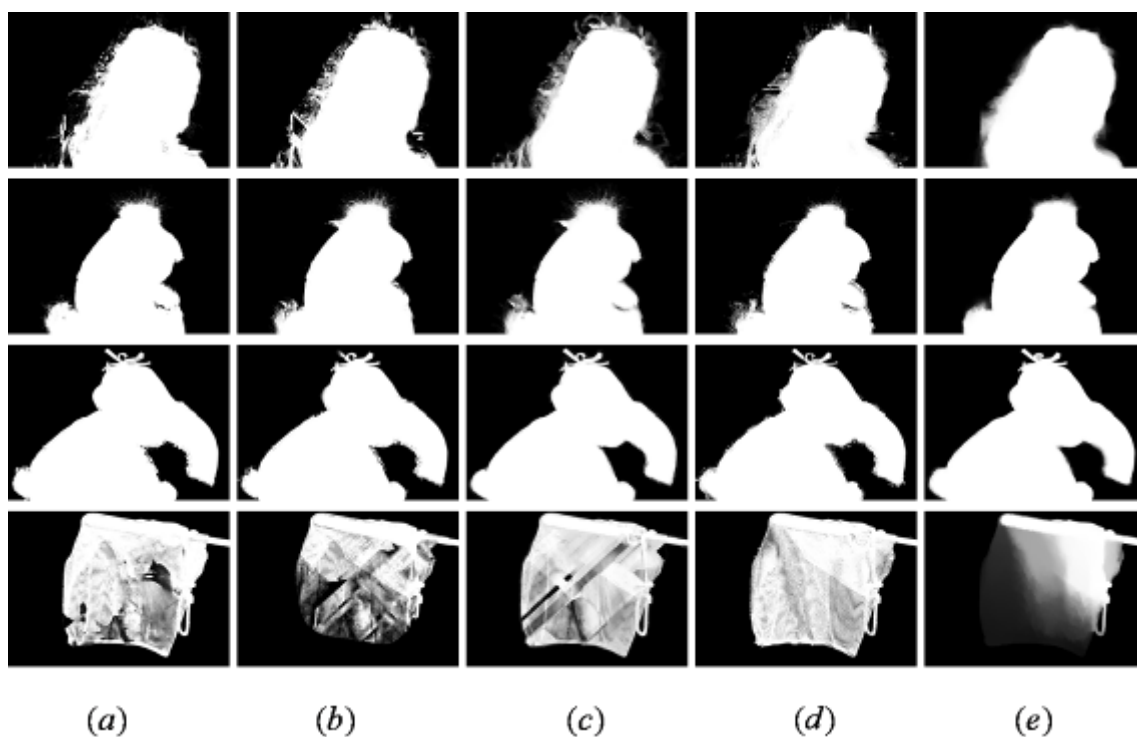


图 3-6 测试图片 1-4 的 α 掩膜

图 3-6 至 3-7 中, 第 a 至 e 列分别为本章方法、Easy Matting、Iterative BP、Geodesic Matting、Random Walk 的抠图结果。综合表 3-1 至 3-8 以及图 3-6 和 3-7 所示的数据分析, 本章所设计的抠图方法具有在如下三种情况下表现较差: 第一, 大面积的半透明像素 (plasticbag); 第二, 绝对背景区域含有大量与前景颜色相近的像素 (doll, troll); 第三, 提示图中位置区域宽度过大 (net)。另外, 从实验结果还可以看出, 本章方法的结果相对其它方法掩膜所含的半透明像素相对偏少。

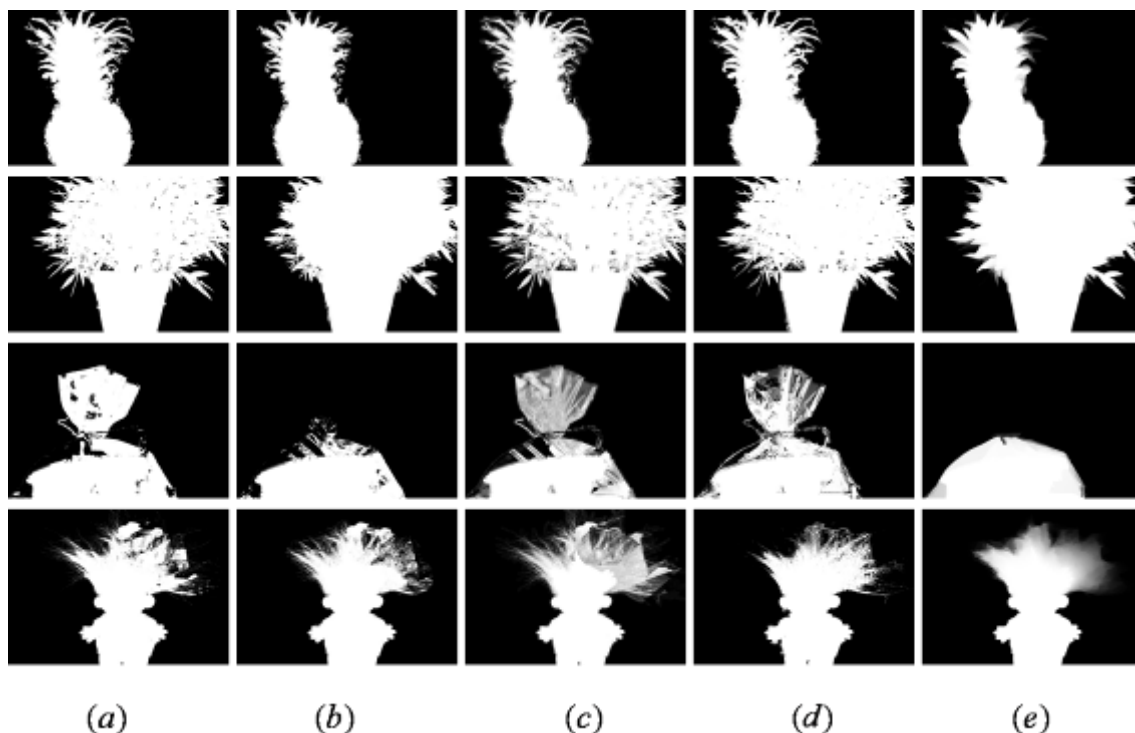


图 3-7 测试图片 5-8 的 α 掩膜

测试组的 8 幅图片是从大量的图像数据中挑选的具有高透明度，前景/背景二义性等极端条件的图像数据。从上图主观结果示例可以看出，不同方法对该组图片的处理结果均具有明显的错误区域。不同于测试组，训练组的测试结果更加接近实际使用时的性能效果。图 3-8 和图 3-9 显示了两个抠图和合成的结果的示例，左图为原始图像，右图为经过本文算法提取出的前景图像。中间位将前景图像融合到新场景中的图像。从融合结果可以看出，整个融合过程真实无缝，前景对象的边缘毛发区域也很好的保留。



图 3-8 抠图效果示例 1



图 3-9 抠图效果示例 2

3.4 本章总结

本章详细介绍了纹理采样的思想和方法，颜色线性模型的思想和方法，并把两者结合起来，提出了基于纹理采样的抠图算法。该方法在测试平台下的测试结果显示，实验结果虽然优于第二章中提出的方法，甚至在某些测试样本中取得了第 3 位的结果，但整体抠图质量在现有方法的整体排名为 13，与现有一流方法还有差距。文中提出的通过图像块的变化方向寻找采样方向以及利用颜色线模型优化抠图结果的思想，对后续研究工作有积极的影响。

第四章 前景提取演示系统

4.1 系统简介

本文实现了一个前景提取算法的演示系统。该系统实现了第三章中介绍的基于纹理采样的抠图算法，同时还集成了一些图像处理的基本算法包括几何变换、图像特效显示、图像增强等。演示系统既可以作为实用软件处理抠图问题，又可以作为学习 VC++ 下图像处理基本方法的平台。

该系统在 Windows 环境下使用 Microsoft VC++6.0 开发工具开发完成，使用 MFC 类库制作界面^[49]，还参考了教材数字图像处理^[50,51]中讲述的数字图像处理算法和相关程序。

系统的主界面如图 4-1 所示：

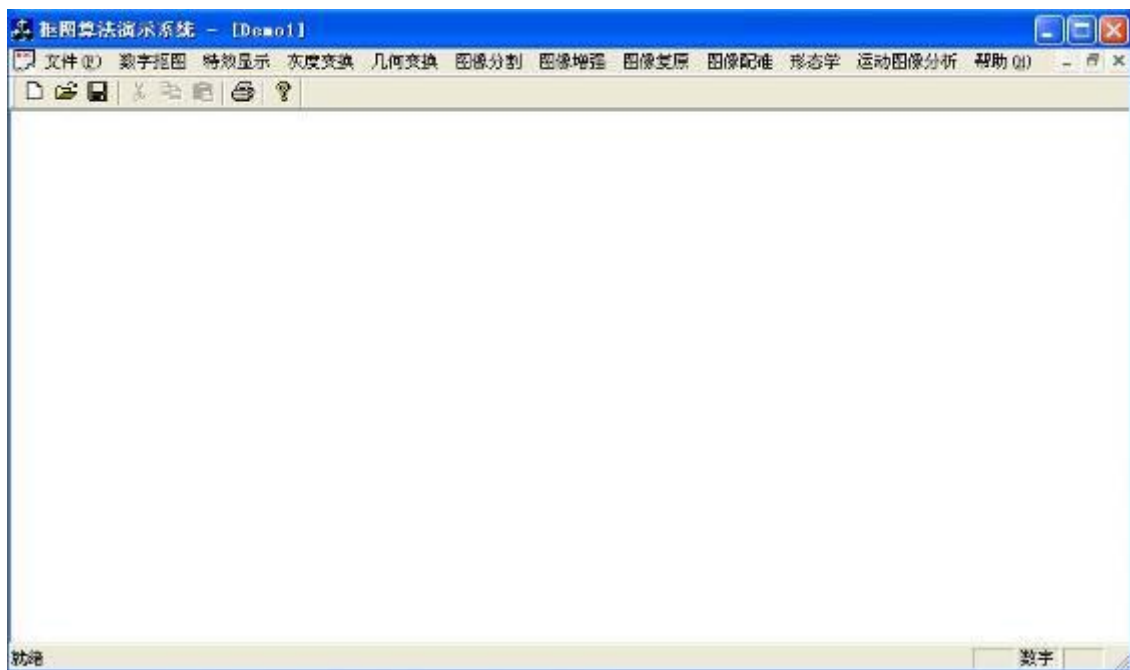


图 4-1 数字抠图演示系统主界面

图 4-1 中可以看出，整个界面简洁直观。上方的各个功能模块包含抠图，显示特效、灰度变换、几何变换、图像分割、图像增强、图像复原、图像配准、形态学和运动图像分析，功能十分丰富。其中抠图部分由本文作者设计，其余算法的

实现请参考文献[50]。

4.2 系统使用实例

本节先简要介绍系统的各部分功能，然后通过一个实例来演示数字抠图部分的使用方法和使用效果。

各个功能模块展开如图 4-2 所示:



图 4-2 数字抠图演示系统主界面

从图 4-2 中可以看到, 整个系统的算法涵盖数字图像处理教材中的大部分典型算法, 可以作为图像处理初学者在 VC++ 实验环境中的学习资源。

图 4-3 至 4-7 所示为数字抠图功能的使用实例,其中原始图像为任意大小的 24 位真彩色图像;提示图像为手工提前标定好的 3 值灰度图像,大小同原始图像;背景图可以为任意大小的 24 位真彩色图像,用户可以选定自己喜欢的任意背景。原始图像、背景图像,3 值提示图的格式目前仅支持 BMP 位图。

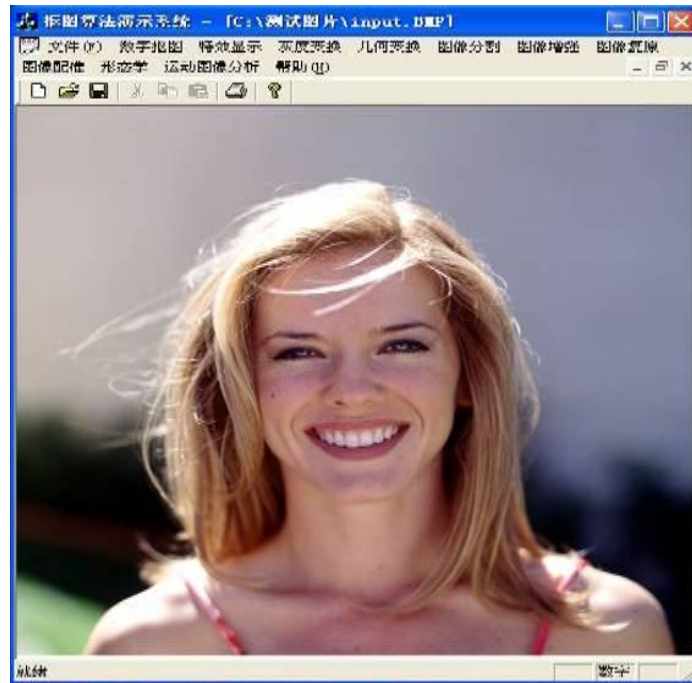


图 4-3 打开一幅原始图像

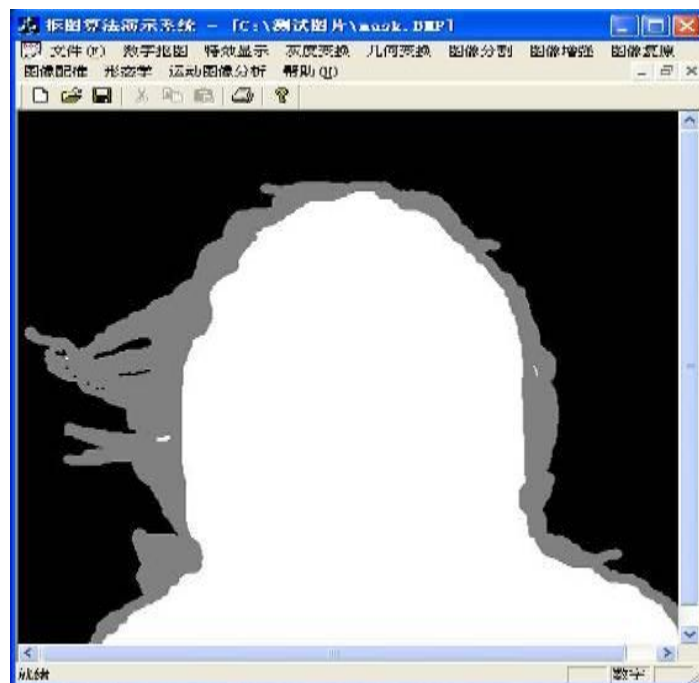


图 4-4 输入的提示图像



图 4-5 本文第三章方法计算出的抠图掩膜



图 4-6 抠取出前景物体融合到新场景中



图 4-7 原始图像与合成图的对比结果

图 4-3 为输入原始图像，图 4-4 为手工标记好的提示图像，图 4-5 为纹理方向采样算法估计出的前景掩膜值示意图，图 4-6 为将提取出的前景物体融合到新场景的效果图，图 4-7 展示了抠图前和融合后的效果对比。

此例中的前景物体边缘有明显的毛发，发丝部分被柔和地提取出来，通过具有半透明性质的掩膜合成到新场景。

第五章 总结与展望

5.1 本文研究工作总结

数字抠图技术已经运用到人们生活中的方方面面如电影制作，天气预报，数字影像编辑，数字娱乐等等。人们对抠图技术的性能要求越来越高，然而传统的抠图方法由于处理速度慢，条件限制多等缺陷已经无法满足日益广泛的应用需要。由于抠图问题本身的严重欠约束性，目前的算法在抠图质量，处理速度，简化人工交互这 3 个性能指标上无法同时取得好的效果，因此设计一种高效实用的抠图算法十分必要。本文在分析研究现有典型抠图方法的基础上，通过提出假设和实验验证，设计出基于背景还原和基于纹理方向采样两种抠图算法。

第二章主要介绍了基于背景还原的抠图方法。图像修复的相关技术已有诸如去字幕，对象移除等运用。本章提出的方法将图像修复与数字抠图结合起来，取得了较好的实验效果。

第三章主要介绍了基于纹理采样的抠图方法。通过观察图像中前景对象边缘的统计特性，简化了复杂的采样过程；结合颜色线模型中前/背景样本颜色的统计结论，提出了基于纹理采样的抠图方法。实验表明该方法在抠图质量上较第二章所提出的算法有明显提高。

第四章主要介绍了前景提取演示系统。该系统在 VC++ 环境下实现了第三章提出的抠图算法。在 MATLAB 实验环境下完成一幅百万像素级的图像的抠图过程需要耗时数分钟，用 VC++ 环境实现后，上述图像在一分钟内就能处理完成。该系统还集成了一些图像处理的基本算法，它可以作为学习在 VC++ 环境下做图像处理研究工作的资源。

5.2 进一步的研究工作

本文提出的相关方法虽然在某些测试用例上取得了抠图质量前三的效果，但整体抠图质量相比现有方法还有差距。其原因分析大致有三点：

第一，采样策略比较简单。文献[36]中，对每个像素点的估计采样数相当于本文的 10 倍却没有因此在计算效率上有丝毫损失。

第二，没有充分利用图像的邻域颜色信息。2012 年最新的基于学习的抠图方法^[48]（即将发表在 ICIP2012 上）将抠图问题和模式识别中学习问题结合起来。该方法的抠图质量在 SAD，MSE 等四个指标的综合测试结果排名第一。

第三，没有利用图像的全局特性。在闭合场抠图算法中，作者通过假定像素点前/背景颜色的局部邻域不变性，推导出全局拉普拉斯方程。在全局采样算法中，作者采用了全局采样策略，通过构造代价函数谱和快速搜素算法，解决了全局采样计算复杂度的问题。这两种方法都充分利用图像的整体特性。

在今后的研究工作中，应该注意横向比较其它抠图方法的优缺点，纵向发掘抠图与修复，图像分割等其它图像处理领域的联系，为设计更好的抠图方法找到思路。

致 谢

时光荏苒，日月如梭，三年的研究生生涯即将结束。它短暂而充实，紧张而不失活泼，有太多的事历历在目，宛如昨日。

在此论文完成之际，首先向尊敬的导师李宏亮教授表示衷心的感谢和诚挚的敬意。三年来，导师敏锐的思维、严谨的治学态度、渊博的学识、诚挚谦虚的品格和宽厚善良的处世方式，永远值得我学习和效仿。导师在我的学业上尤其是在论文的撰写过程中，倾注了大量的心血，给予了我许多教诲和指导，将使我终生受益。师恩如海！我将更加努力，不辜负恩师的期望。

李老师对本研究给予了大量的指导和鼓励。他勤奋的科研作风、与时俱进的科研精神和宽以待人的为人之道无不体现了一位学者的风采大师的风范。在此，致以诚挚的谢意。

感谢教研室老师和同学对我学习、生活上的帮助和关心。

特别感谢我的父母、姐姐和女友，每当我在学业上受到挫折时，总能从他们那儿得到精神上的安慰和鼓励，

再次衷心的感谢所有关心我、帮助我、支持我顺利完成学业的老师、朋友及家人们！

参考文献

- [1] T. Porter and T. Duff. Compositing Digital Images. In Proceedings of ACM SIGGRAPH, 1984, 253–259.
- [2] Jeremy A. Selan. Merging live video with synthetic imagery. Master's thesis, Cornell University, 2003.
- [3] Ron Brinkman. The Art and Science of Digital Compositing. Morgan Kaufman, 1999.
- [4] Peter Hillman, John Hannah, and David Renshaw. Alpha Channel Estimation in High Resolution Images and Image Sequences. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, 2001, 1063-1068.
- [5] 林生佑, 潘瑞芳, 杜辉, 等. 数字抠图技术综述. 计算机辅助设计与图形学学报, 2008, 473-479.
- [6] A. Smith and J. Blinn. Blue Screen Matting. In Proceedings of ACM SIGGRAPH, 1996, 259–268.
- [7] E. Eisemann and D. F. Flash Photography Enhancement via Intrinsic Relighting. In Proceedings of ACM SIGGRAPH, 2004, 673–678.
- [8] G. Petschnigg, M. Agrawala, H. Hoppe, et al. Digital Photography with Flash and No-flash Image Pairs. In Proceedings of ACM SIGGRAPH, 2004, 664–672.
- [9] J. Sun, Y. Li, S.B. Kang, et al. Flash Matting. In Proceedings of ACM SIGGRAPH, 2006.
- [10] J. Wang and M. Cohen. Simultaneous Matting and Compositing. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2007.
- [11] M. McGuire, W. Matusik, H. Pfister, et al. Defocus Video Matting. In Proceedings of ACM SIGGRAPH, 2005, 567–576.
- [12] N. Joshi, W. Matusik, and S. Avidan. Natural Video Matting Using Camera Arrays. In Proceedings of ACM SIGGRAPH, 2006, 779–786.
- [13] T. Mitsunaga, T. Yokoyama, and T. Totsuka. Autokey: Human Assisted Key Extraction. In Proceedings of ACM SIGGRAPH, 1995, 265-272.
- [14] A. Berman, A. Dadourian, and P. Vlahos. Method for Removing from All Image the Background Surrounding Selected Object. U. S. Patent 6.134.346, 2000.
- [15] M. A Ruzon and C. Tomasi. Alpha Estimation in Natural Images. In Proceedings of IEEE

- Computer Society Conference on Computer Vision and Pattern Recognition, 2000, 18-25.
- [16] Barnes, Connelly, Eli Shechtman, et al. The Generalized PatchMatch Correspondence Algorithm: 1-14.
- [17] YungYu Chuang, Brian Curless, David H. Salesin, et al. A Bayesian Approach to Digital Matting. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2011, 264-271.
- [18] Jian Sun, Jiaya Jia, ChiKeang Tang, et al. Poisson Matting. ACM Transactions on Graphics 23(3)2004. Special Issue in Proceedings of ACM SIGGRAPH, 2004.
- [19] Hongliang Li, King N. Ngan, and Qiang Liu. FaceSeg: Automatic Face Segmentation for Real-Time Video. IEEE Transaction Multimedia, 2009.
- [20] J. Wang and M. Cohen. Optimized Color Sampling for Robust Matting. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2007.
- [21] A. Efros and T. Leung. Texture Synthesis by Non-parametric Sampling. In Proceedings of IEEE International Conference on Computer Vision, 1999, 1033–1038.
- [22] A. Criminisi, P. Pérez, and K. Toyama. Region Filling and Object Removal by Exemplar-based Inpainting. IEEE Trans. Image Process. 2004. 1200–1212.
- [23] L. Demanet, B. Song, and T. Chan. Image Inpainting by Correspondence Maps: A Deterministic Approach. UCLA CAM R, Tech. Rep. 03-04, Aug. 2003.
- [24] M. Ashikhmin. Synthesizing Natural Textures. In Proceedings of ACM Symp. Interactive 3D Graphics, 2001, 217–226.
- [25] Y. Wexler, E. Shechtman, and M. Irani. Space-time Video Completion. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, 120–127.
- [26] S. Masnou. Disocclusion: A Variational Approach using Level Lines. IEEE Transaction on Image Processing. 2002, 68–76.
- [27] F. Bornemann and T. März. Fast Image Inpainting based on Coherence Transport. J. Math. Imag. Vis., 2007, 259–278, 2007.
- [28] A. Criminisi, P. Perez, and K. Toyama. Object Removal by Exemplar-Based Inpainting. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 2003, 721–728.
- [29] Omer, I. and M. Werman. Color Lines: Image Specific Color Representation. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, 946-953.

- [30] C. Rhemann, C. Rother, M. Gelautz, et al. A Perceptually Motivated Online Benchmark for Image Matting. 2009 IEEE Conference on Computer Vision and Pattern Recognition (June 2009): 1826-1833.
- [31] Levin Anat, Dani Lischinski, and Yair Weiss. A Closed Form Solution to Natural Image Matting. IEEE transactions on pattern analysis and machine intelligence 30, no. 2 (February 2008): 228-242.
- [32] C. Rhemann, C. Rother, A. Rav-Acha, et al. High resolution matting via interactive trimap segmentation. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2008.
- [33] Huang, Mengcheng, Fang Liu, et al. A GPU-based Matting Laplacian Solver for High Resolution Image Matting. The Visual Computer 26, (2010): 943-950.
- [34] Bai, Xue, and Guillermo Sapiro. Geodesic Matting: A Framework for Fast Interactive Image and Video Segmentation and Matting. International Journal of Computer Vision 82, no. 2 (November 2008): 113-132.
- [35] Y. Guan, W. Cheny, X. Liang, et al. Easy Matting: A Stroke Based Approach for Continuous Image Matting. Eurographics, 2006.
- [36] Gastal, Eduardo S. L., and Manuel M. Oliveira. Shared Sampling for Real-Time Alpha Matting. Computer Graphics Forum 29, (2010): 575-584.
- [37] Kaiming He, Christoph Rhemann, Carsten Rother, et al. A Global Sampling Method for Alpha Matting. Computer Vision and Pattern Recognition, IEEE Computer Society Conference. 2011
- [38] Berkeley segmentation dataset. <http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>.
- [39] Grady, Leo, Thomas Schiwietz, and Shmuel Aharon. Random Walks For Interactive Alpha Matting. In Proceedings of Visualization, Image and Image Processing, 2005.
- [40] Jue Wang, Michael Cohen. An Iterative Optimization Approach for Unified Image Segmentation and Matting. In Proceedings of IEEE International Conference on Computer Vision, 2005.
- [41] Sunyoung Cho, Hyeran Byun. Dynamic curve color model for image matting. Pattern Recognition Letters, Volume 33 Issue 7, May, 2012.
- [42] WuChih Hu, JiaJie Jhu, Cheng-Pin Lin. Unsupervised and reliable image matting based on modified spectral matting. Journal of Visual Communication and Image Representation, 2012.
- [43] C. Rhemann, C. Rother, M. Gelautz. Improving Color Modeling for Alpha Matting. In Proceedings of British Machine Vision Conference, 2008

- [44] Kaiming He, Jian Sun, and Xiaoou Tang. Fast Matting using Large Kernel Matting Laplacian Matrices. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.
- [45] Christoph Rhemann, Carsten Rother, Pushmeet Kohli, et al. A Spatially Varying PSF-based Prior for Alpha Matting. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.
- [46] Yuanjie Zheng, Chandra Kambhamettu, Yuanjie Zheng, et al. Learning Based Digital Matting. In Proceedings of IEEE International Conference on Computer Vision, 2009.
- [47] Shiming Xiang, Feiping Nie, Changshui Zhang. Semi-Supervised Classification via Local Spline Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010.
- [48] Zhanpeng Zhang, Qingsong Zhu, Yaoqin Xie. Learning Based Alpha Matting using Support Vector Regression. In Proceedings of IEEE International Conference on Image Processing, 2012.
- [49] 熊歆斌. Visual C++ 程序设计. 北京: 清华大学出版社, 2002.
- [50] 谢凤英. Visual C++ 数字图像处理. 北京: 电子工业出版社, 2008.
- [51] Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing. 北京: 电子工业出版社, 2008.

攻硕期间取得的科研成果

- [1] Bing-chaoFei, Hong-liangLi . A Background Recovery Based Matting Approach.
in IEEE International Conference on Computational Problem Solving(ICCP2011),
2011, 287-290.
- [2] 费炳超. 一种基于优化方向采样的抠图算法. 中国, 发明专利, 2012. 2, 已申请.