RGB图像中的文字去除和双向插值补全图像

**伍宇明 学号：51609020**

### 一、介绍

在实际图像应用场景中，很多时候由于实际需要，在图片或视频中嵌入文字，但由此也带来了在需要去除文字时的不便。如老式的电影视频由于技术限制，字幕被固化在视频中，成为视频图像的一部分，从而阻碍了不同语种间的交流。通过将文字从图像中去除，并补全修复被文字遮挡的部分是很有意义的。同时，在影视特技制作中，也需要大量的抠图技术，把多余物体剔除，去除文字的技术可以看作是这种抠图技术的一小部分。

本文通过利用锐化、滤波、二值化、膨胀处理等提取文字，通过横向和纵向两个方向的一维插值加权来对去除文字后图像的空缺部分进行补全。

### 二、问题的分析与解决

对RGB图像的去文字处理大体可分为两大部分：一、文字的提取；二、补全背景空缺部分。

1. 文字的提取

文字的提取涉及到边缘检测的知识。查阅相关的资料，我们得到了边缘检测的基本步骤，如图1所示。

平滑图像

锐化滤波

平滑滤波

边缘判定

边缘连接

二值图像

边缘图像

原始图像

锐化图像

图1 边缘检测

1. 平滑滤波：由于梯度计算易受噪声影响，因此第一步是用滤波去除噪声。但是，降低噪声的平滑能力越强，边界强度的损失越大。
2. 锐化滤波：为了检测边界，必须确定某点领域中灰度的变化。锐化操作加强了存在有意义的灰度局部变化位置的像素点。
3. 边缘判定：在图像中存在许多梯度不为零的点，但是对于特定应用，不是所有点都有意义。这就要求我们根据具体情况选择和去除处理点，具体方法包括二值化和过零检测等。
4. 边缘连接：将间断的边缘连接成为有意义的完整边缘，同时去掉假边缘。

本文在实际操作的过程当中，结合实际效果，对处理过程进行了修改，同时在应用中确定了具体方法，修改后的流程如下（已包含补全背景等后续处理）。

起始图像

锐化图像

平滑图像图像

二值图像

边缘图像

文字区域

填充图像

最终图像

尺寸缩放

用户截取

拉普拉斯

算子锐化

高斯滤波

二值化

膨胀处理

双向插值

加权平均

高斯滤波

回代原图

图2 图像处理流程

1. 尺寸缩放及用户截取：参考屏幕显示的分辨率，如果原图过大，先把图片先缩放一定尺寸，使图片能在屏幕上以完全尺寸显示，目的是为了在截取文字区域时所得的矩形尺寸是原尺寸，没有经过放大，保证截取区域的准确。然后用户需要框选截取文字区域，把文字区域图像保存为要处理的矩阵，记为text，以下均为对text的处理。
2. 拉普拉斯算子锐化：提取文字的重点在于边缘提取，因此在实际操作中，把锐化过程放在滤波前，为了获得更准确的边缘。此处用到了图像空间域增强的知识。拉普拉斯算子是基于二阶微分的。二维函数f(x, y)的二阶微分定义为



对于离散的二维图像f(i, j)，可以用下式作为对二阶偏微分的近似。



将上面两式相加就得到用于图像锐化的拉普拉斯算子。



对应的滤波模板如下。



1. 高斯滤波：拉普拉斯锐化对一些离散点有较强的响应，而相应的，高斯滤波对于去除离散点噪声有良好的效果。均值为0，方差为σ2的二维高斯函数为



离散化表示得到



经过调试本文选择sigma=0.75的3\*3高斯模板。

1. 二值化：经过锐化和去噪之后的图像边缘已经比较明显，此时直接二值化的效果已经足够好。阈值采用Matlab计算的自适应阈值。
2. 膨胀处理：图像膨胀是图像形态学处理的内容。经过二值化得到的二值图像与实际图像边缘有一定误差，主要是文字嵌入之后文字与图像之间的边缘存在过渡带。考虑到后面把文字去掉以后是用插值法把原来的文字区域的点取代，因此需要把文字部分完整抠掉，包括过渡带。所以采用膨胀处理，处理后得到的边缘一般会比实际文字略大。

至此，边缘检测部分的处理完成。

1. 补全背景空缺

得到的文字部分记录为一个逻辑矩阵，记为bw，其中的非零元素的位置即为要以插值法取代的目标点。

然后按行遍历，对每一行，以非文字部分的点作为有效点，以这些点为基础用一维插值法构建插值行函数，根据插值行函数得到背景目标点的RGB值。用同样的方法按列遍历，用插值列函数算得目标点的RGB值。把行函数得到的值与列函数得到的值加权平均。

由于一般文字都是横向排列，行函数可以的有效点比较分散，得到的插值函数可以更靠近实际，而纵向得到的列函数由于有效点集中在两端，中间是空缺的方块，所以插值函数对于图像来说不是很可靠，因而一般而言横向的权值更高。但对于纵向排列的文字或者图像有明显纵向取向的纹理，可以增大纵向的权值。本文只取行函数权值，插值选用效果最优的三次多项式插值法。

补全图片后，为了消除全空缺部分与其它部分的边缘感，再次应用了高斯滤波。最后把处理完成的text代回原来截取前的图像，图像处理部分完成。

1. 计算峰值信噪比

为了客观算法评估去掉文字的图片质量，需要寻找一个指标。本文采用的是广泛使用的PSNR(Peak Signal to Noise Ratio)峰值信噪比，方法是先选一幅源图像，用PhotoShop加入文字作为起始图像，经过处理后得到最终图像，把最终图像和源图像比较（只比较文字区域），计算峰值信噪比，PSNR越大则图像相似度越高。

PSNR的计算方法如下：

对于灰度图像，MSE表示当前图像X和参考图像Y的均方误差（Mean Square Error），H、W分别为图像的高度和宽度；



PSNR的单位是dB，数值越大表示失真越小。n为每像素的比特数，一般的灰度图像取8，即像素灰阶数为256.



对于RGB图像，计算PSNR有几种方法：

1：计算RGB三通道每个通道的PSNR值，再求平均

2：计算RGB三通道每个通道的MSE值，再平均，得到PSNR

3：将RGB转换成YUV颜色空间，仅仅计算y分量的PSNR

2和3比较常用，本文采用2，求MSE时对三个通道的所有像素点求方差和再除以图片尺寸再除以3，求PSNR时公式不变。

### 三、程序处理结果

在文字提取环节，本文的算法对于间隔较大，书写简单（尤其简单英文），颜色与背景差距显著的文字效果良好。

  

图3 几种文字提取的结果

在背景补全环节，本文算法对于背景颜色单调均匀，细节纹理较少的图片效果较好。根据实验，当PSNR大于3.5时图片修改痕迹不明显。



图4 图像处理结果，PSNR=4.57



图5 图像处理结果，PSNR=4.73



图6 图像处理结果，PSNR=3.88





图7 图像处理结果，PSNR=3.92

### 总结

本文结合了拉普拉斯算子锐化、高斯滤波、二值化取边缘、图像膨胀处理等多种图像处理方法，提取文字，利用双向插值求加权平均的办法补全文字去除后的背景，并引用PSNR峰值信噪比对样本图像评估图像处理质量。本文方法相对简单，在处理文字间隔较大，书写简单，颜色背景差异大，纹理细节较少的图片时效果明显。

同时方法也有局限性，在纹理细节较多时，会出现细节损失较大，图像处理痕迹明显，文字去除不干净等现象，PSNR会降低至2.5以下，可以考虑从插值法函数的方向以及引入微分法等方面优化。

### 参考文献

1. 李庆扬，王能超，易大义. 《数值分析》第五版. 清华大学出版社
2. 赵海滨等. MATLAB应用大全. 清华大学出版社
3. 张静，倪红霞，苑春苗，杨立红. 精通Matlab数字图像处理与识别. 人民邮电出版社
4. 季丽琴. 彩色图像内文字的自动提取与去除的研究. 苏州大学硕士论文2006
5. 陈柱. 视频图像字幕文字的检测与去除. 华中科技大学硕士论文2012
6. 两种常用的全参考图像质量评价指标--PSNR和SSIM.

CSDN博客http://blog.csdn.net/zjyruobing/article/details/49908979