# 第二章 复杂条件下的文本图像预处理

## 2.1 引言

相对与印刷体的光学字符识别，复杂条件下的文本图像有着截然不同的性质如分辨率低，大小不一，光照不均等。这些特性影响着从文本定位，分割到识别等各个过程。在将场景条件下的文本图像输入到各个模块前，对图像进行必要的预处理对于定位和识别的正确率提高有一定的帮助。对于初始的场景文本图像进行颜色空间变换使得图像更加适合于后续的聚类和显著性检测过程。得到文本词图像后，文本行可能不是以直线方式存在的，将词图像校正到水平位置有利于后续的词图像分割。而文本图像由于相机，场景的因素存在不同的噪声，去噪对于定位和识别的过程也相当重要。二值化则贯穿在文本定位和识别的过程，好的二值化能够较大提高识别结果的正确性。

## 2.2 复杂条件下的文本图像分析

复杂条件下的文本图像分析是进行后续研究的第一步，只有深刻认识了场景文本图像的一些特点，才能够提出有效的方法。我们从两个方面来对复杂条件下的文本图像进行分析。

1）场景分析

场景的复杂性是造成文本定位召回率和准确率低的一个重要原因。场景文本图像的获取一般是通过智能手机或者数码相机得到的，获取的途径存在多样性造成场景的多样性。这个多样性表现在场景可以是街道（室外），教室（室内），走廊，路边，玻璃窗户，书本封面等。场景多样性会导致场景干扰的多样性，场景的干扰如窗户，玻璃，草地，砖头，天空，栅栏等。有些干扰如果有和文本相同的边缘或者纹理，就会对文本定位的性能产生很大的影响。另一个场景的问题是场景存在着光照不均的情况，这个现象在图书封面场景中经常发生，因为拍照时背景光滑反射大部分光线。场景的多样性，干扰和光照不均的情况在图中可以看到

 

2）场景文本分析

在上面分析了场景的特性之后，这里分析场景图像文本的特性。场景文本是本课题研究的重点对象，场景文本不同于光学字符的方面在于文本的尺寸，颜色，光照和空间位置。

文本尺寸受拍摄远近和文本本身的影响，一般商店的名称，广告的尺度都比较大，可能占据的空间会超过整幅图像一半的大小，而一些物体的说明，一些告示，书本的封面通常比较小，占据整幅图像很小的部分。

文本颜色通常跟背景的差异都比较大，这样比较容易引起人们的注意。一个文本区域的颜色一般趋向于一致，但也存在不同颜色字符构成一个文本区域的情况，另外一个字符连通域内的颜色也可能存在多个。

光照不均的问题同样存在场景文本中，场景文本如果表面反射光线比较多的话，相应的区域会呈现比较亮，如果是在室外拍摄，太阳光照，遮挡会使文本呈现明亮不均的情况。

空间位置的变化，如果拍摄场景图像时，镜头对准了文本，文本一般会出现图像的中心。但是拍摄时一般没有考虑那么多因素，文本可能出现在图像的中心，边缘，上下左右任何一个位置。场景文本变换的例子在图2.2中可以看出。



## 2.3 颜色空间变换

最常采用的颜色空间为RGB，所有其他颜色空间都是从RGB空间转换而来的。

## 2.4 文本图像倾斜校正

文档图像的倾斜校正一般是在文本定位后进行。定位后得到词图像如果没有处于水平直线的位置，在后续采用投影图像进行分割的时候，容易造成切分出错。所以在把词图像放到分割模块前，先对词图像进行倾斜校正。

## 2.5 文本图像去噪

噪声存在场景文本图像的各个部分，其有可能是拍摄时引入，也有可能是在对图像进行处理的过程中引入的。这里的去噪一般是在词图像分割之前和对单个的字符进行处理。分割之前是为了不让噪声影响切分，而对单个字符的去噪是为了在特征提取阶段不生成无效的表征。

## 2.6 文本图像二值化

文本图像的二值化是一个精简信息的过程，在整个文本定位和识别系统占据重要地位。文本图像的二值化是将文本区域赋值为1，而将其余区域赋值为0。在分割时，不需要太多的文本区域内部的信息，只需要字符间的相对位置信息，所以一般分割的算法都是在二值化后的图像进行的。提取单个字符的全局特征时，字符内部局部的，细节的信息不是很关键，只需要保留有字符的整体，形状的特征即可，所以一般也是在二值化后进行提取。

二值化算法包括全局二值化和局部二值化, 全局二值化具有速度快但效果相对差的特点, 局部二值化算法具有速度慢效果好的特点。

## 2.7 本章小结

本章对红外弱小目标图像三个主要组成部分即背景、噪声和目标进行了特性分析，分析可知图像背景灰度比较平稳，处于图像频谱中的低频部分。图像受各种噪声的影响，因此需要对目标检测之前进行预处理。从分析中还知道目标的特性，一般目标没有形状，缺少纹理等信息。最后针对红外弱小目标检测提出了检测结果评价标准，分别是局部信噪比、局部对比度、检测概率、虚警率和平均虚警率，这些指标能将我们的算法与其他算法作对比以证明其有效性。上述对红外图像三个组成部分的介绍和性能评价指标的提出，为后续的研究奠定了基础。