1. **引言(P2-4)**
2. **牌照图象分割的理论及实现(P5-14)**

2.1 牌照图象分割的基本理论(P5-7)

2.2 牌照图象分割的实现(P8-14)

1. **字符串区域提取的理论及实现(P15-20)**

2.1 字符串区域提取的基本理论(P15-17)

2.2 字符串区域提取的实现(P18-20)

1. **字符串区域校正的理论及实现(P21-27)**

2.1 字符串区域校正的基本理论(P21-23)

2.2 字符串区域校正的实现(P24-27)

1. **字符区域分割的理论及实现(P28-38)**

2.1 字符区域分割的基本理论(P28-31)

2.2 字符区域分割的实现(P32-38)

1. **字符识别的理论及实现(P39-54)**

2.1 字符识别的基本理论(P39-47)

2.2 字符识别的实现(P48-54)

1. **车辆牌照字符识别系统中“字符识别”子系统的实现(55-58)**
2. **结论(59-60)**
3. **附录——牌照识别理论在“学生成绩单自动识别”系统中的应用(待定)**

**第1章．引 言**

车辆牌照是车辆的“身份证”号码，在交通系统中有着不可替代的作用。通过它可以检索车辆的许多重要信息,记录车辆的违章或进行收费等工作。若能用计算机实时地识别车牌就可以在无需为车辆加装其它特殊装置的情况下实现对车辆的自动监测，从而给交通系统的自动管理提供了极大的方便。因此车辆牌照自动识别系统的实现是推进交通管理计算机化的关键技术之一。

车辆牌照自动识别系统的实现过程是：在各交通路口、桥梁入口、主要街道上架设数字或模拟摄像机，以固定的角度拍摄来往行使车辆的动态或静态图象；识别系统的研究对象就是拍摄的数字或模拟录像带，如果待识别数据是模拟图象，则通过A/D转换器或其它数字化设备将其转换为数字图象，作为识别系统的输入。识别系统的任务是分析输入的数字图象，对其包含的车辆信息自动进行身份认证，即识别图象中车辆牌照的汉字、字母、数字字符，如图1－1所示，系统识别的结果应为“沪BC0255”。



（图1－1）

识别出车辆的牌照，就可唯一确定该部车辆。之后对此信息加以处理、分类，就能实现交通管理中违章处罚、流量管理、路口收费等各种任务的自动化管理。

关于车辆牌照自动识别的研究工作八十年代就已展开。中国、美国、德国、英国、日本以及意大利等国家都在进行有关的研究[1][2][3][4]，但是尚未见该类系统投入实用的报道。

本项目主要是对识别系统的算法进行研究和改进，希望能在原有理论的基础上实现更高的识别率和识别速度，推动该系统从试验到实用发展的进程。研究任务分为两部分，第一部分是车辆牌照的分割：将牌照从图象背景中分离出来，获得包含牌照字符串的小块图象区域，如图1－2所示；第二部分是牌照字符的识别：从前部分分离出的图象区域中确定牌照字符串的精确位置，对其校正识别，最后获得待识别车辆的牌照字符串信息。



（图1－2）

本人所研究的课题是“牌照识别系统中的字符识别”，实验样本采用照相机拍摄的照片，通过扫描仪数字化形成的256级灰度图象，为使研究样本具备普遍性，在随机的自然环境中，从各个角度、距离进行拍摄（后文提供的牌照图例均为实地拍摄照片经数字化后的灰度图象，以100％比例显示）。字符识别部分的研究对象是由前面牌照分割部分的输出图象区域（如图002所示），要求该区域为256级灰度图象，牌照字符串的所有图象信息包含在其中，所处的位置基本上在区域的中央，字符区域在图中所占比例不应过小，背景信息比较单一，不包含其它无效的字符。

牌照的字符识别主要包括五个环节：（一）牌照图象的二值化，通过该环节将灰度图象转变为二值图象，使字符串信息与背景分离。（二）字符串区域的提取，将牌照中的字符串从牌照的其它部分(例如底牌、牌照边框、固定装置及边框以外的其它景物)中分割出来。（三）字符串区域的校正，这一环节将消除牌照字符串图象的旋转、倾斜。(四)字符区域的分割，这一环节从校正后的牌照字符串中切割出独立的字符。（五）字符的识别，该环节对切割得到的独立字符进行识别，输出识别结果。

车辆牌照字符识别与普通的规则字符相比，存在着许多难点。造成这些难点的原因有自然因素、人为因素也有其自身因素。在自然因素中，如果光照较强，会使照片颜色偏亮，反之则偏暗；如果遇到阴雨或大雾天气，会使照片模糊不清。在人为因素中，如果车主使用不当，牌照会因碰撞或沾染污迹而发生变形、破损、字符模糊；再者，拍摄时因角度或距离的差异，造成照片中牌照区域有时较大，有时较小，有时发生水平或垂直倾斜，有时发出投影变形。在自身因素中，牌照底色与字符颜色存在多种组合，常见的有白底黑字、蓝底白字、黄底黑字等；还有，固定牌照的铆钉、牌照的边框也会给牌照代入噪声。上述这些因素互相影响，造成原本规则的牌照字符出现几何形变、投影形变、随机噪声及大小、色彩、字形在很大程度上的离散性，在照片图象中产生下列多种变化。

（图1－3） （图1－4）

图1－3显示的牌照区域较小；图1－4显示的区域较大。

（图1－5） （图1－6）

图1－5字符区域的底色为暗色调，字符为亮色调；图1－6字符区域的底色为亮色调，字符为暗色调。

（图1－7） （图1－8）

图1－7显示的字符区域整体颜色偏亮，且白色噪声很不规则，造成字迹模糊不清；图1－8的字符区域整体颜色偏暗。

（图1－9） （图1－10） （图1－11）

图1－9的牌照区域水平方向发生弯曲，上方的两个铆钉与字符区域溶和形成噪声；图1－10的字符区域发生水平倾斜；图1－11的字符区域发生垂直倾斜。

上面9幅图例仅是举例说明牌照各种因素干扰产生畸变的表现，在实际使用中还会遇到各种各样其它的干扰因素，这些干扰会对字符识别过程中的一个或几个环节造成影响，为了获得更理想的识别效果，就必须识别中解决这些干扰因素的影响。本文所做的工作主要是研究了牌照识别过程中各种基本理论及其应用，并对多种算法进行改进和完善，使其更加适合于高噪声环境中规则字符的提取与识别，从而实现较理想的识别率（实验室环境中系统综合识别率达到95％以上）。后文将按字符识别过程的先后次序详细阐述，第2－6章介绍牌照图象分割、字符串区域的提取及校正、字符区域的分割、字符的识别五个环节的相关理论及实现，每章分为两部分：前一部分主要研究相关理论；后一部分主要阐述实现该环节的算法及改进。第7章总体概括“字符识别”子系统实现的流程及相关算法。第8章在总结系统优缺点的基础上，提出今后对该系统加以改进的一些想法。第9章附录中介绍了牌照识别算法在“学生成绩单自动识别”中的具体应用。

**第2章 牌照图象的分割**

图象分割是图象处理的初级阶段，其目的是将有用信息（前景区域）与无用信息（背景及噪声区域）分离开，以便以后可以对前景区域进行深入的分析，其分割的结果将生成二值化图象。

**2.1 图象分割的基本理论**

图象的分割一般是依据其自身的图象特征来进行的。图象特征是指图象场中可用作标志的属性，其中有些是视觉直接感受到的自然特征，如区域的亮度、色彩、纹理或轮廓等，有些是需通过变换或测量才得到的人为特征，如各种变换频谱、直方图、矩等。常用的图象特征有幅值特征、直方图特征、变换系数特征、灰度边沿特征及纹理特征，相应的分割算法主要有灰度阈值法及边沿提取等。

2.1.1 幅值特征

幅值可在特定的象点上或某个区域内加以测量，我们不仅可以直接利用图象的灰度值或三刺激值，也可利用各种变换(线性的或非线性的，可逆的或不可逆的)形成的、位于变换域中的变换系数的幅值。

2.1.2 直方图特征

一幅数字图象可看作是二维随机过程的一个样本，能用联合概率分布加以描述。统计各象点的灰度值所得到的直方图。可用来估计图象的概率分布，从而形成一类特征。

定义图象f(x,y)灰度的一阶概率分布为P(b)=Prob(f(j,k)=b) , 0<=b<=L-1 ,式中b是量化所得的灰度值，其取值为0到L-1量化级别中的任一级。定义 P(b)≈N(b)/M 是一阶近似直方图，其中M是以(j,k)为中心的测量窗口中的象点总数，N(b)是该窗口内灰度值为b的象点数。通常假定图象是平稳的，取窗口大小为整幅图，则M为整幅图中的象点总数，N(b)是整幅图中灰度值为的b的象点数。

图象直方图的形状可给出图象特征的许多信息。例如，分布狭窄的直方图反映图象的对比度很低，双峰型直方图反映图中存在两个不同灰度区等。由一阶直方图表征的参数有：

1. 平均值 ：



(2)方差：

****

(3)能量：

****

(4)熵：

****

二阶直方图特征是以象点对的联合概率分布为基础的。设象点f(j,k)和f(m,n)分别位于坐标(j,k)和 (m,n)上，二者相距| j – m |,| k – n |,(或表示成极坐标r,θ),这时联合概率为：

P(a,b)=Prob{f(j,k)=a, f(m,n)=b}

式中a和b是图象灰度的量化值，由于数字图象中直线也离散表示，距离参数(r,θ)只能取一定的离散值。P(a,b)的估计值为：P(a,b)≈N(a,b)/M ,式中M是测量窗口中的象点总数，N(a,b)是窗口中满足距离参数的两点f(j,k)=a和 f(m,n)=b出现的总数。习惯上假定图象是平稳的，窗口取整幅图大。如果图象中象点对是与高度相关的，则P(a,b)的阵元将沿阵列的对角线聚集。下列参数常用来表征能量在对角线附近散布的程度：

1. 自相关系数：



1. 协方差：



(3)能量：



(4)熵：

****

在纹理分析中，二阶直方图测度非常有用。

2.1.3 变换系数特征

各种酋变换的变换系数是同图象f(x,y)一一对应的，但是，譬如Fourier变换的频谱 F（u,v）的平方值，即：



却与F（u,v）不是唯一地对应，这时即使f(x,y)的原点有了位移，M(u,v)的值仍保持不变。在某些应用中可利用这一特点。

2.1.4 采用灰度阈值法进行分割

这是最简单的分割方法,通过设置一门限(阈值)T，把象点按灰度级分为大于和小于T的两组，若水平的或垂直的相邻两点属于不同组别，则给予标记Le，否则给予标记Lb。于是把图象f(x,y)分割成二值化图g(x,y)，g(x,y)的取值规则为：

(1)若f(x,y) 和 f(x,y-1) 或f(x,y) 和f(x-1,y) 属于不同点组：g(x,y)=Le

(2)其它情况：g(x,y)=Lb

T是图象灰度级范围[zl,zh]中的某个值，当图象的直方图呈现若干块较明显的不同分布形状时，则T大致可选在两分块的边界处。显然，从直方图上妥善地选择T值，对正确划分出感兴趣的区域(或目标)和背景很有关系。

**2.2 牌照图象分割的实现**

牌照图象的分割是将灰度图象转变为二值化图象，使字符图象前景与背景区域得以分离。图2－1左侧为牌照灰度图象，图2－1右侧为分割后的二值化图象，以‘1’表示前景信息（黑色区域），以‘0’表示背景信息（白色区域）。

（图2.2－1）

这一环节的实现是通过牌照区域确定、区分图象前景与背景、图象二值化三部分来完成的。

2.2.1 牌照区域的确定

图象分割效果的优劣主要依赖于是否能确定图象特征的基本信息。对于牌照图象分割而言，其处理的对象是包含牌照字符串及其周围区域的灰度图象，牌照区域图象的幅值、纹理及边沿特征是相对比较稳定的，可以在某种程度上来预测；反之牌照周围的区域存在很多不确的干扰与噪声，因此要获得好的分割效果，就应先确定牌照区域的位置。



（图2.2.1－1）

“牌照区域的确定”只须大致估计其位置即可，至于精确定位则留给后面的处理来完成。从牌照纹理特征出发，注意到字符串灰度分布在水平方向变化最快，而其背景区域则不具备这一特征。由此可获得图象水平微分的投影图，如图2.2.1－1所示，牌照图象上方为水平微分的垂直投影图；牌照图象右侧为水平微分的水平投影图。水平微分的投影公式可表示为：

 (2.2.1-1)

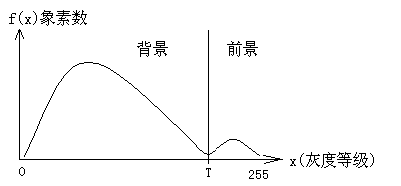
式2.2.1－1中*f(y,x)*为图象中点*(y,x)*的灰度值，*W*和*H*为图象的宽度与高度，*g(x)*为微分垂直投影值，*h(y)*为微分水平投影值，生成的结果如图2.2.1-1所示。以Amax[*g(x)*]为阈值（A为常数），从两侧开始搜索垂直投影*g(x)*，将第一次大于阈值的位置加以标注，可获得前后两点*x*=Pa与*x*=Pb，这两设定为牌照区域的左右边界点；用相同的算法也可确牌照区域的上下边界点*y*=Pc与*y*=Pd。四个边界点确定了牌照区域大致所处的位置，以便后文对此区域进行分析。

2.2.2区分牌照图象前景与背景

（图2.2.2－1）

由于牌照底色与字符颜色存在许多变化（如黄底黑字、蓝底白字等），造成牌照底色（背景色）与字符颜色（前景色）在灰度图象中既可以是低灰度级别（暗色调），也可以是高灰度级别（亮色调）——如图2.2.2－1所示，因此要将前景与背景区分开来必须首先确定前景为低灰度级别还是高灰度级别。



（图2.2.2－2）

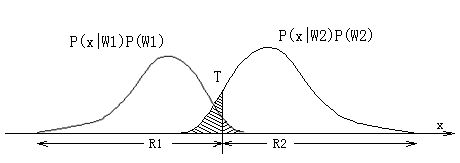
车辆牌照是按照统一的标准来制作的，因此其牌照字符的大小、笔画粗细在牌照区域中所占的比例固定不变。图2.2.2－2为牌照图象灰度直方图示意图，横坐标表示256级灰度等级，纵坐标表示牌照区域具有相同灰度值的象素点数目，*T*为二值化阈值。可以看到其灰度统计曲线呈双峰型，且背景象素点数目大于前景象素点数目，由此可得到确定前景灰度级别的判别公式：

 (2.2.2-1)

 (2.2.2-2)

上式中为牌照图象区域灰度级上的象素数目，*T*为二值化阈值，为大于0的常数（根据实际情况可调节）。如果统计结果满足式2.2.2－1，则前景颜色为高灰度级别；如果统计结果满足式2.2.2－2，则前景颜色为低灰度级别；如果两式皆不满足，则无法区分前景与背景亮度级别，因此只要确定*T*的位置，便可区分前景与背景的灰度级别。

从统计特性的角度看，如果前景与背景的概率分布函数确定，则根据Bytes最小风险判别准则，可以求得前景与背景的阈值。

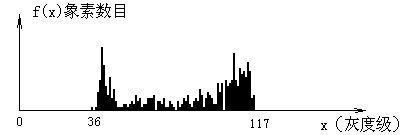


（图2.2.2－3）

图2.2.2－3所示为W1、W2两类的概率分布函数，T为区分两类的阈值，图中阴影面积之和为将两类错分的误差概率总和。改变T值，当误差概率达到最小值时，T值是进行二值化的最优阈值，参见式2.2.2－3。

 (2.2.2－3)

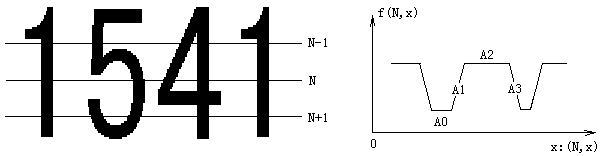
正规的车辆牌照都是按照统一标准制造的，所以其背景与前景在灰度图象中的概率分布函数也是唯一确定的，按式2.2.2－3可以方便的求得二值化阈值及前景与背景的灰度级别。然而牌照识别系统分割部分输出的图象区域中不仅包含牌照区域，也包含了牌照外的一些区域（见图2.2－1），这些区域的灰度信息给整个区域的灰度直方图引入了随机干扰，使得原本规则的概率分布函数曲线发生变形。

（图2.2.2－4）

图2.2.2－4中左侧为牌照图象区域，右侧为窗口区域（左图黑框内部分）对应的灰度直方图。在直方图中可看到其灰度概率分布的双峰特性仍然存在，但由于牌照外区域的干扰，使得双峰两侧，尤其是之间的部分产生较大的离散，如果依旧使用原有的分布曲线求取最小风险阈值T，将引入很大的误差，不仅使二值化效果变差，甚至可能造成背景与前景灰度级别判定的错误，给以后的识别工作带来致命的影响。

本文从牌照图象的纹理特征出发采用串线法有效地实现了背景与前景灰度级别的判定。



（图2.2.2－5）

图2.2.2－5中左侧为牌照字符区域的一部分，N-1, N, N+1为沿图象水平方向的第N-1, N, N+1条扫描线；右侧为第N条扫描现经过区域的灰度示意图，f(N,x)为N行x列象素点的灰度值，A0为低灰度级区域，A1为上升沿区域，A2为高灰度级区域，A3为下降沿区域。

以图2.2.2－4左侧牌照图象为例，牌照部分在图象区域中基本上处于中间位置，且所占的比例不会过小，因此选取一个矩形窗口（矩形窗口为图中黑线包围的区域，其宽度与高度分别为牌照区域宽度与高度的1/2，且位于牌照区域的中心，牌照区域为2.2.1节中求得的区域），以图象中心为起始点，采用螺旋扩展方式搜索域内方差值最大的窗口作为穿线扫描的区域。

 (2.2.2－4)

 (2.2.2－5)

 (2.2.2－6)

上式中为行列象素点，为该点的灰度值，为微分门限，A0,A1,A2,A3为图2.2.2－5中四个区域的标志。水平扫描从窗口区域的左上角开始，先列后行，进行逐行搜索。先计算同行中相邻象素间的微分值，如果微分值小于微分门限（即满足式2.2.2－4），且前一点属于下降沿区域A3或低灰度级区域A0，则点属于区域A0；如果微分值满足式2.2.2－4，且点属于上升沿区域A1或高灰度级区域A2，则点属于区域A2；如果微分值满足式2.2.2－5，则点属于上升沿区域A1；如果微分值满足式2.2.2－6，则点属于下降沿区域A3。上述算法实际是以提取灰度图象的边沿信息作为标定区域的依据，而附加在牌照图象上的随机噪声具备较多的边沿信息，因此算法对于噪声比较敏感，容易引起误差，解决这一问题的方法是先设置微分门限为较大数值（本文设其初始值为50），对窗口区域进行初步扫描，统计平稳区域A0,A2的数量，如果数量较小，可按一定步长逐次缩小微分门限，直到获得的平稳区域数量接近牌照字符的统计特性为止。如此，上文中区分前景与背景的判别式2.2.2－1及式2.2.2－2可改写为：

 (2.2.2－7)

 (2.2.2－8)

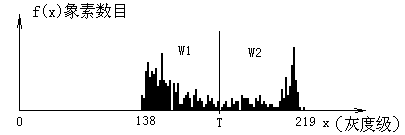
上式中为平稳区域A0,A1的长度（单位：象素），表示窗口区域，为大于0的常数。如果式2.2.2－7成立，则前景为低灰度级别；如果式2.2.2－8成立，则前景为高灰度级别。采用此算法，在所有实验样本中区分前景与背景，正确率达到100％。

2.2.3牌照图象的二值化

获得前景与背景的灰度级别后，需要对图象灰度进行规则化，然后再确定二值化阈值。规则化后的灰度图象中保证前景为高灰度级别，见式2.2.3－1：

 (2.2.3－1)

式中为图象规则化前点的灰度值，为规则化后点的灰度值，*height, width*为图象区域的高度及宽度（单位：象素）。图2.2.3－1左侧的图象为规则化后的灰度图，前景被设置为高灰度级别（白色），右侧为窗口区域对应的灰度直方图，与图2.2.2－4比较灰度值发生了颠倒。

（图2.2.3－1）

由于牌照外区域引入的噪声使得灰度直方图的分布很不规则，无法预先知道前景与背景的灰度分布函数，如果采用Bayes 判别函数不易获得最优阈值，所以本文主要使用聚类算法来求取最理想的分类二值化阈值。

 (2.2.3－2)

式2.2.3－2是描述聚类距离的多种公式之一，为两个不同类类中心间的距离，为类中各点到类中心的距离和。从上式可看出，聚类的效果越好，其比值就越大。本文聚类的对象是窗口区域的灰度直方图，归纳起来应是一维两类的聚类问题，其聚类边界就是二值化阈值*T*，因此式2.2.3－2可以表述为：

 (2.2.3－3)

其中：



 (2.2.3－4)



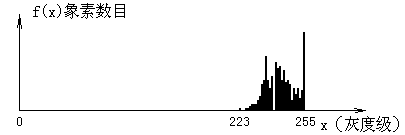
式011中比值为阈值*T*的函数，是背景区域灰度值的期望与方差，是前景区域灰度值的期望与方差（见图2.2.3－1左）。式2.2.3－4中*x*为灰度级别，为灰度级别对应的象素数目。

改变*T*值，将相应发生变化。当达到最大值时，对应的*T*值就是最优的聚类阈值，即二值化阈值：

 (2.2.3－5)

如式2.2.3－5所示，令的导数为零，便可得到二值化阈值。实际的算法将遍历所有灰度级别[0，255]对应的比值，求取其中的最大值以获得二值化阈值。

由于牌照图象窗口区域的灰度直方图基本上呈双峰分布，采用上述方法可以在前景与背景分布函数未知的情况下获得分类阈值，即使加入有限的噪声干扰，聚类方法也能得到较理想的结果（如图2.2－1所示）。在实际应用中大多数牌照的图象都能满足前面的要求，然而如受到强因素的干扰，可能会破坏牌照图象双峰分布的特性，此时使用聚类方法获得的阈值常会产生较大的偏差。

（图2.2.3－2）

图2.2.3－2左侧上方为牌照照片的灰度图象，右侧为窗口区域对应的灰度直方图。由于照片拍摄时的技术因素，造成图象中牌照区域的灰度值偏高（呈现白色），牌照底色与前景字符的灰度差别不大，在灰度直方图中可以看到原有的双峰特性已被破坏，其灰度分布杂乱无章，此种情况下逐次搜索[0,255]灰度范围内聚类比值的最大值易产生很大偏差。图2.2.3－3左侧为搜索并二值化后的图象，二值化的结果是失败的。为了避免这种情况，本文对聚类比值的搜索范围加以限制，以2.2.2节中确定的前景与背景候选区域的灰度平均值作为其自变量变化的上下边界。图2.2.3－2窗口区域的前景与背景灰度平均值分别为245和252，聚类后的二值化阈值为250，二值化结果见图2.2.3－3右侧图象，实现了较理想的前景背景区分效果。

（图2.2.3－3）

**第3章 字符串区域的提取**

字符串区域的提取实际上是在牌照图象中搜索字符区域集合的精确位置，即为字符串上下左右四个边界的确定。理想情况下，牌照中字符的分布是规则的，其所处的区域为一矩形，且与周围图象之间存在着间隙。而实际上，牌照受损变形，沾染了污迹或者因摄像过程中的环境与技术因素等都会造成字符串图象变形或与周围的区域、干扰粘连在一起，这些难点必须在区域提取中加以克服。

**3.1 字符串区域提取的相关理论**

字符串区域与背景区域区别的特征主要有边沿特征、纹理特征及结构特征，通常的区域分割算法也是从这此特征出发来实现的。

3.1.1 灰度边沿特征

图象灰度级的局部不连续性，被称之为局部边沿或边沿元。把诸边沿元沿其切线方向能聚集(联结)成大范围的线段，称为边沿或边界。这是一个有用的重要特征，反映了图象中物体或区域所占的物理限度。边沿检测技术，一直受到人们的重视和研究，寻找一种对噪声不敏感，定位精确，不漏检错检的检测方法，始终是人们努力的目标。

3.1.2 纹理特征

纹理区域是指这区域由紧密混杂在一起的某种单元(纹理基元)构成，这种单元或局部结构在比它更大的范围内大致作均匀重复排列。例如由种子或草地之类构成的图片，表现的是自然纹理。由织物或砖墙等构成的图片，表现的是人工纹理。在许多应用中，常常希望测量这些纹理的边界，以及纹理的粗糙度。后者反映局部结构的空间重复周期的大小，周期大的纹理粗。由于纹理是一种空间性质，涉及到象点的邻接特性，测量应限于相对均匀的区域。

分析纹理常有结构模型和统计模型两种，前者对于纹理基元的分布具有较多规律性的场合较适合，而统计模型则更适宜于分析自然纹理，因这时较少辨得出基元的形状。当然，这不是绝对的。

3.1.3 边沿提取

实践表明，图象的边界对人的视觉识别十分有用，人们常能从一粗糙轮廓图识别物体，而边界的表示又容易引入各种物体的识别算法中。

物体的边界在图象中是由灰度不连续性所反映的，人们希望直接从图中提取边界，但这在成象条件不好，边界形状复杂时是很难做到的。一般是先提取局部边沿(边沿元)，然后再组合这些边沿元成为精制的边界(通常采用松弛算法来实现)。

局部边沿在图象的总象点中只占不大的比例，那里局部灰度以一定方式迅速改变。可采用边沿算子来检测图中是否有局部边沿，它是小空间范围的数学算子。边沿算子的重要特性是检测到的边沿元方向应同最大的灰度变化方向垂直，而边沿元幅值应表示灰度变化的剧烈程度。边沿算子大致可分为三大类，即梯度算子、方向模板算子和拟合算子。梯度算子是数学上梯度算子的近似形式，而方向模板算子是对不同方向使用不同的检测模板。至于拟合算子则是将图象的局部灰度值同边沿的参数模型相拟合。在图象处理和计算机视觉文献中有着许许多多边沿算子可供借鉴。不同类别和种类的边沿算子各有特点，适用于不同的应用领域。

对牌照识别而言，由于牌照中的字符笔画通常与牌照底色间存在明显边界，因此依据边沿确定牌照中字符串的大致位置是一可取的办法。

3.1.4 已知形状的曲线检测( Hough 变换)

当通过边沿检测得到了一段段小的边沿元时，如何检测它们是否能连成某一已知形状的曲线(例如是否共直线，或共圆等)或者说怎样判断图中是否存在某种形状的曲线呢？Hough变换是一种有效的方法，它是把图象平面中的点的位置按待求曲线的函数关系映射到指定的参数空间，然后找出最大凝聚点完成变换。

例如，极坐标系中直线的方程为：r=xcosθ+ysinθ ，式中r是直线离原点的法线距离，θ是该法线对x轴的角度。可见直线的Hough变换在极坐标系(或叫参数空间)是一个点，而一点的Hough变换是一正弦型的曲线，共点的诸直线的Hough变换也是一正弦型曲线。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| r θ | 0 | 30 | 60 | … | … | … | 180 |
| r1 |  |  |  |  |  |  |  |
| r2 |  |  |  |  |  |  |  |
| r3 |  |  |  |  |  |  |  |
| … |  |  |  |  |  |  |  |
| -r3 |  |  |  |  |  |  |  |
| -r2 |  |  |  |  |  |  |  |
| -r1 |  |  |  |  |  |  |  |

(表3.1.4-1）

计算Hough变换的方法是将r-θ域量化成许多小格（见表3.1.4-1）。对于每一个(x0,y0)点代入θ的量化值，算出各个r，所得值(经量化)落在某个小格内，便使该小格的计数累加器加1，当全部(x,y)点变换后，对小格进行检验。有大的计数值的小格对应于共线点，其(r, θ)可用作直线拟合参数。有小的计数值的各小格一般反映非共线点，丢弃不用。

可以看出，若r和θ量化得过粗，则参数空间的凝聚效果差，找不出直线的准确r, θ值；反过来，若r, θ量化得过细，则计算量将增大，因此需兼顾这两方面，取合适的量化值。

对牌照识别而言，由于牌照中的字符通常呈水平排列，因此可采用Hough变换检测直线的办法找到字符串的上下边沿以实现分割。

3.1.5 区域生成

图象中各区域具有不同的特征值(如灰度值、颜色等)，从而可以彼此区分开。因此寻找互相连接在一起、并有相同特征的象点所形成的区域，也是实现图象分割的一种重要方法。对于某一区域应取什么样的特征、常与具体的应用有关。

因为区域是一两维连通区，故生成区域的方法大致有三类：

(1)局部法。即根据象点的性质或象点的邻近象点的性质决定象点分在什么区域。

(2)总体法。即根据遍及整幅图中大量象点的性质决定把一群象点划分到一区域中。

(3)分裂合并法。这利用状态空间技术分裂或合并区域，用图结构表示区域和边界，这时要用局部和总体的合并与分裂准则。

区域生成法的有效性大大地依赖应用领域和输入图片的内容。例如，当景物简单时（亮背景上一个黑团点），即使用简单的局部法也出奇地有效；而遇到象室外景物那样复杂情形，即使用了最复杂的方法也得不到满意的分割。

**3.2 牌照字符串提取的实现**

这部分所做的工作是从二值化后的图象中精确地确定字符串所在的区域。由于照片拍摄时的角度选取是任意的，所以照片图象中的牌照部分会产生旋转、水平垂直倾斜或投影变形，切割出的字符串也相应地发生变化，不再是矩形区域，而是一个四边形区域。图3.2－1左侧为牌照的灰度图象，中间为二值化后的图象（字符为黑色，背景为白色），右侧为字符串提取的结果，其中字符串上下边界以外的背景被设置为白色，两条垂直黑线为字符串的左右候选边界。

（图3.2－1）

字符串的切割主要包括上下边界的选取和左右边界的选取两部分。其中牌照字符串的左右边界精确来讲应该是字符串中第一个字符的左边界和最后一个字符的右边界，这两个边界的确定工作是在“字符提取”部分来完成的，所以目前字符串左右边界的确定只需要大致估计即可，本文以2.1.1节求得的牌照左右边界作为字符串左右边界的粗略估计。

3.2.1 字符串区域左右边界的确定

字符串上下边界的确定是以单个字符的上下顶点为基础的。由于牌照字符是规则的印刷体，在理想情况下其字符串字符的上下顶点是排列在同一条直线上的，实际情况下牌照因各种因素可能发生变形，且图象二值化过程中也会发生数值化误差，这些因素会造成字符顶点偏离原有的共线位置，形成一定的离散效应。虽然如此，这种离散性并不强烈，字符顶点仍然分布在其所属的直线两侧，所以以此为基础可确定出字符串的上下边界。

3.2.2 字符区域上下顶点的确定

牌照二值图象中的字符（除第一个字符汉字外）可看作是单一的连通区域（如图3.2－1所示），通过区域增长方法可实现连通区域的遍历，所以也可确定该区域的上下顶点。区域增长算法为：

1. 初始化种子P（P为连通区域E中的一点）及上下顶点Pup和Pdown。
2. 依次检测与P点相邻的八点。如果该点属于区域E，将该点与上下顶点进行比较，并更新Pup和Pdown的数值，然后将该点设置为种子迭代调用步骤2；如果该点不属于区域E，则结束本次迭代。

本文在2.1.1节确定的粗略牌照区域中选取水平线，沿水平线方向逐次检测每一象素，如果该象素属于前景，则将其设定为种子，使用上述区域增长方法求得该象素所属连通区域的上下顶点。当搜索完毕后对每个连通区域生成一对顶点。为了防止单一水平线无法穿越所有连通区域的现象，在牌照区域中等间距选择3－5条水平扫描线，使其尽可能覆盖所有的连通区域。算法的结果将产生顶点对集合其中为一个顶点对，它由上下顶点组成。

（图3.2.1－1）

图3.2.1－1左侧为区域增长后形成的图象，图中黑色＋号为连通区域的上下顶点。牌照图象中的连通区域不仅是单个字符，还有牌照外的背景、字符串“午H·3846”中的·号及牌照中的随机噪声，这些原因使得结果顶点对集合*P*产生离散性，所以在利用顶点对预测字符串的上下边界前，需要对顶点对集合的有效性进行校验。

 (3.2.1－1)

式3.2.1－1中为顶点对在垂直方向上的距离，为检测中心，是检测中心的偏心距，其值为常数，为*P*的子集，为选取中含顶点对最多的子集，所得的结果为顶点对集合*Q*。由于牌照字符串为规则字符，旋转、水平垂直变形都不会改变字符串上下边界平行的特性，即使牌照发生投影变形（字符串高度一侧较高，一侧较矮），其上下顶点垂直间距本上还是保持相同，而不会有较大误差；字符串以外的连通区域或大或小，上下顶点的垂直距离与字符串特性相差很远，式3.2.1－1可过滤出垂直距离离散度较小的顶点对集合，只要合理选择偏心距系数，可以有效的剔除噪声区域的顶点对。过滤的结果作为字符串顶点对的粗略估计，使后文上下边界的估测能获得更好的效果。

3.2.3 根据区域顶点拟合边界

3.2.2节获得的顶点对集合*Q*中的元素基本上是字符区域的上下顶点，可能也包含少量的噪声定，总体来这些顶点都分布在两条直线的附近，这两条直线就是字符串上下边界线。根据一些边沿点来拟合直线，可以采用3.1.4节介绍的hough 变换算法来实现。直线拟合的结果将产生两个直线方程：

 (3.2.3－1)

由于顶点对集合*Q*中含有噪声点，会造成拟合后的直线偏离原有的字符串边界线。要消除误差应尽量减少噪声点的干扰。

 (3.2.3－2)

上式中*a*为表征边沿点离散度的常数。采用hough 变换算法生成边界线后，使用式3.2.3－2进行校验，如果集合*Q*中存在不符合该式的边沿点，则将其删除，再重复进行直线拟合，直到集合*Q*中所有的元素点都符合上式为止，此时的直线确定为字符串区域的上下边界线。如图3.2.3－1中灰色直线所示。



（图3.2.3－1）

**第4章 字符串区域的校正**

原始的牌照图象是规则的矩形，字符串区域中的字符沿水平按一定间隔排列，字体垂直与底边界，且大小相等。由于拍摄是技术因素或牌照自身受损变形等原因，使得牌照图象中字符串区域出现水平倾斜、垂直倾斜及透视变形，其所处的矩形区域也相应变成不规则的四边形。这一环节的工作就是校正这些变形，使字符串区域恢复到初始的规则形态。

**4.1 区域串区域校正的基本理论**

图象区域的校正主要是通过几何校正来实现的。其相关的理论包括求取水平、垂直、投影变形的参数，空间坐标的变换，及重新确定在校正空间中各象点的取值。

图象的空间几何坐标变换是指按照一幅标准图象g(u,v)或一组基准点去校正另一幅几何失真图象f(x,y)。根据两幅图象中的一些已知对应点对(又称控制点对)，建立起函数关系式，将失真图象的坐标系(x,y)变换到标准图象坐标系(u,v)，从而实现失真图象按标准图象的几何位置校正，使f(x,y)中的每一象点都可在g(u,v)中找到对应象点。

4.1.1 二阶均值的方法

该方法针对单个字符的灰度图象进行校正。认为字符的旋转角度φ可由二阶均值求取即：



其中umn是从字符灰度图象中取得的二阶均值。根据φ可将图象旋转至合理位置。

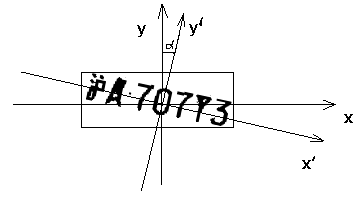
文献[5]中对邮包表面的字符逐个校正时就采用了该方法。

4.1.2 梯度法

该方法可在透视形变不严重的情况下用于对整个字符串图象的校正。它基于这样的观察结果即图象中各点的梯度方向将主要集中在与字符串走向相垂直的方向上。因此对图象f(x,y)各点取得p=∂f(x,y)/ ∂x, q=∂f(x,y)/ ∂x并令φ＝arctan(q/p)，对所取得的各个φ值按一定间隔量化，用一直方图统计量化后的φ值，最后将直方图中的最大值对应的φ作为字符串的旋转角度。按该角度旋转原图象后可实现字符图象的大致水平，采用类似方法可检测字符的倾斜角度(由透视形变引起)并用Shear变换较正字符的倾斜。(Shear变换将图象中的象点(i,j)变换为(x,y),其中y=j , x= i- (height-j)\*tan(θ) , θ是倾斜角度，height是图象高度)

4.1.3 投影法

该方法可在透视形变不严重的情况下用于对整个字符串图象的校正。首先将待校正的图象中的象点向正交坐标系x-y的x轴(检测字符竖向偏转角度)或y轴(检测字符横向偏转角度)投影，然后将该正交坐标系绕原点旋转一角度α得到一新的正交坐标系x’-y’再次作投影，如(图4.1.3-1)所示。重复上述操作直至找到象点投影结果方差最大时的坐标系，并认为此时对应的α为字符的竖向(采用x轴投影)或横向(采用y轴投影)偏转角度。根据α可将图象旋转至合理位置。



(图4.1.3-1)

文献[3]采用了类似的方法，但不是将象点直接投影，而是对原图象抽取了特征点后作投影。

4.1.4 象点取值的确定(内插)

图象经几何校正后，在校正空间中各象点的亮度值应等于被校正图象对应点的亮度值。但校正后的图象往往会出现某些象点挤压在一起，另一些则分散开了，使校正后的象点不落在坐标点上。为此常采用内插法来求得这些象点的灰度值。可采用的简单方法有最近邻点法和双线性内插法。

4.1.4.1 最近邻点法

该法取象点上下左右四个邻点中距离最近的邻点的灰度作为该点的灰度。这一方法计算简单，有一定精度。缺点是校正后图象的亮度存在明显的不连续性。

4.1.4.2双线性内插法

该方法利用上下左右四个邻点的灰度在两个方向上作线性内插。它的算法虽比最近邻点法复杂，计算量大但结果令人满意，没有灰度不连续的缺点。双线性内插的数学表达式是：

f(i+u , j+v)=(1-u)(1-v)f(i , j)+v(1-u)f(i , j+1)+u(1-v)f(i+1,j)+uvf( i+1,j+1)

双线性内插法具有低通滤波性质，使高频分量受损，图象轮廓模糊。

**4.2 字符串区域校正的实现**

前文提到由于照片拍摄时的角度不同，使得牌照图象会发生旋转、水平垂直变换。

（图4.2－1）

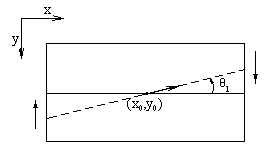
图4.2－1中左侧为发生旋转变化的牌照图象，产生的原因是拍摄时镜头方向与牌照平面垂直，而牌照在平面内以焦点为中心旋转了一定角度；图4.2－1右侧为发生透视变形的图象，产生原因是拍摄镜头方向与牌照平面法线之间有一定的夹角，在照片图象显示为近大远小的透视效果；图4.2－1中间为发生垂直倾斜的图象，它是旋转变化与透视变化合成的效果。字符串的校正工作就是先计算出牌照变化的角度，再通过几何变换恢复牌照字符串本来的规则字体（如图4.2－2所示）。



（图4.2－2）

4.2.1 水平校正

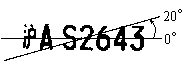
水平变换可用来校正牌照字符串的水平倾斜。上文已确定了字符串的上下边界线，从边界线的斜率可以求得其水平倾斜的角度。

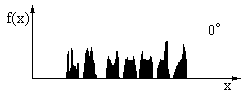
 

（图4.2.1－1）

图4.2.1－1左侧是牌照图象水平倾斜的示意图，虚线表示字符串相对于水平线逆时针旋转的角度为。右侧的水平校正公式可实现点到点的变换，其中为校正前图象中的点，为校正后图象中的点，原点位于图象的左上角，校正后虚线上的点沿垂直方向移动到水平线上，移动方向如图中虚线两侧的箭头所示，其旋转的基点为。实际上水平校正公式实现图象纵坐标调整，而横坐标保持不变。

4.2.2垂直校正

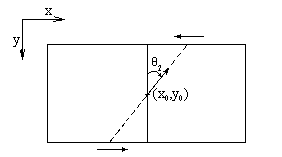


（图4.2.2－1）

图4.2.2－1上面的图象为牌照字符串二值图，其中的两条直线分别沿逆时针方向旋转了0°和20°；图4.2.2－1下方的两幅图为字符串向其上两条直线垂直投影的效果图，x方向为直线方向，f(x)为字符串中字符部分投影的黑色象素点数。由于字符串为规则字符，所以如果投影的角度与字符的垂直方向的倾角相同，则在投影图上黑色象素集中在字符宽度范围之内，字符间隙应基本不包括黑色象素点，即象素能量相对较集中；反之，黑色象素的分布将超出字符宽度范围，其象素点能量相对较分散，如图4.2.2－1下方投影图所示。由此，通过检测字符串沿不同角度的垂直投影，就可确定其垂直倾斜角度。

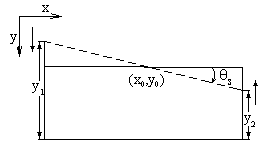
本文采用比较投影方差的方法来确定垂直倾角。因为牌照字符串一般垂直倾斜的角度不会很大，所以选取在-15°到15°范围内，等角度间隔搜索投影方差的最大值，将其对应的投影角度设定为字符串垂直倾角。

（图4.2.2－2）

图4.2.2－2左侧图象显示字符串以为基点在垂直方向顺时针倾斜的角度为。图4.2.2－2右侧公式可实现垂直方向倾斜的校正，校正后虚线上的点沿水平方向移动到水平线上，移动方向如图中虚线两侧的箭头所示，其旋转的基点为。垂直校正公式实现图象横坐标调整，而纵坐标保持不变。

4.2.3 透视校正

（图4.2.3－1）

图4.2.3－1中左侧为透视变形的示意图，其中虚线与实线包围的梯形区域为校正前字符串区域，假设其下边界与水平线平行，梯形右边界高度为，左边界高度为，图中完全由实线包围的区域是校正后的图形，可以看出透视变形使字符串上边界以为中心顺时针旋转的角度为。图4.2.3－1右侧的公式实现左右透视变形的校正，校正后的字符串上边界恢复水平状态；至于上下透视变形的校正，因字符串左右边界高度较小，相对于上下边界来讲透视变形不明显，可以不加校正。

4.2.4校正的合成

上述垂直、水平、投影校正三个公式合成在一起，可以实现字符串几何变形的校正。

 (4.2.4－1)

其中



 (4.2.4－2)



式4.2.4－1中，为水平校正矩阵，为垂直校正矩阵，为投影校正矩阵，这三个校正矩阵可以相互颠倒位置；为原图象的坐标矢量，为校正后图象的坐标矢量，其它的变量参见前面的说明。垂直旋转角度采用投影方差求得，垂直校正后字符串区域的左右边界基本上恢复到垂直状态；水平旋转角度选取字符串下边界的倾斜角，水平校正后字符串的下边界恢复水平状态，字符串区域呈梯形状；投影变换角度为字符串上下边界倾斜角的差值，投影校正后字符串基本恢复到规则字体状态。

原图象坐标矢量经三次校正矩阵的几何变换后，映射为坐标矢量，其中的矢量值不可避免地出现小数，不再对应于实际的象素点位置。因此必须对进行规则化，如果只是对矢量值简单地取整，会产生很大的量化误差。本文采用象素移交映射方法（具体说明在下一章）生成校正的图象（如图4.2.4－1所示）。



（图4.2.4－1）

**第5章 字符区域的分割**

字符区域分割是将字符从字符串区域中提取出来，形成7个独立的矩形区域，每个区域中应尽可能包含字符的全部信息，且不应携带噪声干扰。普通的OCR系统因其识别对象的图象质量较好，字符基本上是等间隔分布的，且很少发生字符粘连的现象。然而在摄取的牌照图象中，因采光、角度、天气、拍摄技术等多方面的因素常常会使本来分离的字符区域粘连在一起，效果交差的图象可能会使字符的大部分都溶和在一起。因此如何准确地分离字符区域在牌照识别中占有至关重要的地位。

**5.1 字符区域分割的基本理论**

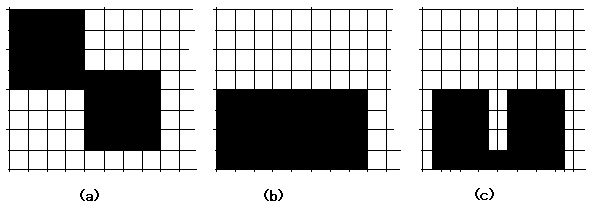
在文档分析领域中，字符分割方法可按照待分割图象是以灰度还是以二值化保存分为基于灰度图象的字符分割方法和基于二值化图象的字符分割方法。或者按照分割对象是手写还是机器打印的字符分为手写字符的分割方法及印刷体字符的分割方法。对于印刷体字符的分割方法又可分为单一字体的分割方法和多字体的分割方法。据此本文中牌照字符串的分割可认为是基于二值化图象的单一字体印刷体字符的分割问题。解决这类问题的具体方法可分为四类：(一)基于投影的方法，(二)基于拓扑的方法，(三)基于识别的方法，以及(四)基于背景细化的方法。

5.1.1 基于投影的字符分割方法



(图5.1.1-1)

这是一种最常用的方法。首先将单行文本图象中的象点作垂直投影取得如图(5.1.1-1)的结果。设定一阈值a，将投影图中小于a并满足其它附加条件的“波谷”作为两个字符间的间隙，从而将文本行分割成一系列字符。这一方法本身对字符图象的污迹干扰、字符的断裂、相邻字符的粘连没有辨别能力。文献[7]提出了一种新的投影方法，该方法将文本图象各象点与其相邻各点作AND运算，并将该运算结果作垂直投影。由于引入了粘连位置的部分信息从而减轻了字符粘连对分割的影响(见图5.1.1-2其中(a)表示两个斑点相粘连的图象，(b)垂直投影的结果，(c)新方法投影的结果)。然而这类方法不可避免地需采用后处理方法对取得的结果进行合并与分裂。



（图5.1.1-2）

(图5.1.1-2其中(a)表示两个斑点相粘连的图象，(b)垂直投影的结果，(c)新方法投影的结果)

5.1.2 合并与分裂

采用测定字符宽度和间距为基础的合并与分裂处理方法是较为常用的方法。该方法首先取得文本行中字符的间距、宽度及空隙，通常采用统计的方法来完成。当文本行中的字符间距固定时，该方法根据已测定的字符宽度和空隙把相邻的单独存在时宽度不足以成为字符的区域组合起来。若这些相邻区域较小，经组合后的整个宽度小于一个阈值，则设这些区域为合并候选区域予以合并。若某区域的宽度远大于正常的字符宽度则根据已测定的字符宽度和空隙分裂该区域。对整个文本行进行了合并与分裂处理以后，用一系列判据(例如各字符的高宽比，字符间的空隙与间距)检测处理后的结果是否合理，若结果不能被接受则需调整选用的阈值并重新对文本行的各部分进行合并与分裂。

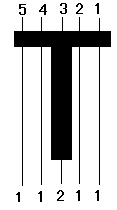
这一方法可在文本行的大部分字符不断裂不粘连的条件下进行，若文本行中的大部分字符出现断裂和粘连，则会使该方法完全失效。

5.1.3 基于拓扑的方法

该类方法利用了字符(字母与数字)的拓扑信息进行文本行的分割。具体方法有准拓扑分割算法和拓扑分割算法。

5.1.3.1 准拓扑分割算法

该算法的基本操作是求取字符的水平笔画计数。该计数值是通过计算一个给定位置处的垂直扫描所遇到的字符的不同笔画的次数取得的(如图5.1.3.1-1所示，字符顶部数字表示扫描序号，底部数字表示笔画计数)。每次扫描时笔画计数和扫描的次数均被统计和记录并存入一表格中。该表设有六项即当前的、上一次的和下一次的笔画计数和实际的扫描序号。将这一表格与存储的笔画计数参考模式匹配，通过寻找呈现的笔画计数模式和一组已存储的笔画计数参考模式之间最贴近的匹配作出分割决定。



(图5.1.3.1-1)

5.1.3.2 拓扑分割算法

该算法是以字符笔画前沿和后沿的检测及度量为基础的。在对二值化文本行图象从右至左的扫描过程中，对每一次扫描计算前沿比特数P、后沿比特数N和黑比特数L三个参数。一个前沿比特的定义是，在前一轮扫描时含有一个水平白比特近邻的一个黑比特。一个后沿比特的定义是，在前一轮扫描时含有一个水平黑比特近邻的一个白比特。在分割粘连、噪声字符时，这些参数代入下列方程组：



当Tn<=0时，则在第n次扫描时进行分割。需设计的问题是确定P、L和N恰当的定义以及选择U、V、W、X和Y系数恰当的值，从而实现在字符最为理想的区域内进行分割。各参数的选取方法见于文献[7]。

5.1.4 基于识别的方法

以识别为基础的分割过程将字符分割和字符识别合二为一。常用方法有(一)移动窗口算法和(二)闭环分割识别算法。

5.1.4.1 移动窗口算法

该算法首先在文本行中按正常字符的大小设定一窗口使其从左至右移动，每移动一步都将窗口中包含的图象块交由字符识别器识别，若该图象块可被识别则认为它是一字符，将其分割出来并将窗口移动至下一个字符可能出现的位置，若该图象块不可被识别则将窗口移动一步后再次交由字符识别器识别，重复上述过程直至得到满意的结果。为了防止将两个字符中的相邻部分错识为一个字符(例如将“O”的右半部和“C”的左半部的组合识别成“X”)需要保证文本行中的每一笔画都有合理的归属，因此当上述条件不能满足时需调整相关参数重新对整个文本行进行分割。

5.1.4.2 闭环分割识别算法

该算法将文本行图象作为输入，并将分类窗口的大小初始化为整个图象的大小。若该图象仅含有一个字符，则识别器可在第一阶段识出该字符。否则，从右侧缩小分类窗口，将其中图象交由字符识别器识别。这一过程不断重复直至分类窗口中的图象可被识别或分类窗口宽度已小于字符可能出现的宽度。对前一种情况可记录识别结果并将分类窗口左边缘移动至下一个字符可能出现的位置的左侧，同时恢复分类窗口的右边缘至原图象的最右端，重复上述过程。对后一种情况若发生在最初时刻则认为输入的文本行图象首部有错，算法将拒识整个图象，否则认为是先前的分类出现差错，算法将取消已完成的可疑分类，将分类窗口左边缘移动至可靠分类字符出现的位置的右侧，同时恢复分类窗口的右边缘至原图象的最右端，重复上述过程。

5.1.5 基于背景细化的方法

该方法采用细化算法对文本行中的背景作细化，并对细化所得的骨架端点作处理，以此达到分割字符的目的。多用于汉字文本的字符分割。

**5.2 字符区域分割的实现**

** **

（图5.2－1）

图5.2－1左侧为牌照图象，右侧为经过字符串校正后的二值化图象。这幅牌照图象由于质量较好，经字符串提取与校正后基本上已不存在噪声及变形，图5.2－1右侧图象黑色部分完全是字符区域，背景部分为均一的白色，字符之间没有粘连，图中显示出较清晰的七个字符“沪BC7280”。分析字符串的图象可以看出对图象的垂直投影进行直接分割就能准确地得到七个字符。

（图5.2－2）

实际上照片拍摄时因受到各方面的影响，使得结果图象对字符提取造成或多或少的困难。比较常见的是照片曝光过度（如图5.2－2所示），使照片图象中字符的笔画变粗，其影响是字符的轮廓模糊不清，字符之间粘连，甚至字符与牌照边框也有粘连现象。为了获得了较理想的结果，本文主要以字符串的垂直投影及轮廓投影为判别依据，对字符区域进行分割、合并、提取，实现了单个字符的有效分割。

牌照字符为规则的印刷体字，其字符串的长宽比、字符间隙、字符的长宽比及笔画宽度都是按照严格的标准制造的，在牌照照片数字化、提取和校正的过程中会代入一些干扰，但字符串参数之间的相互关系不会发生很大的变化。

 (5.2－1)

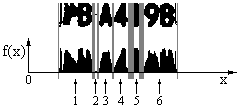
 (5.2－2)

 (5.2－3)

 (5.2－4)

式5.2－1中*L*为字符串的长度，*H*为字符的高度；式5.2－2中*Wc*为常规字符（不包括字符‘1’）的宽度；式5.2－3中*Wn*为字符‘1’的宽度；式5.2－4中*Dn*为字符‘1’两侧的字符间隙。上面四个公式为字符串参数关系式，其中的参数常量是实际图象的统计平均值，利用这些公式可使单一字符区域的估测更加准确。

5.2.1 粗分割



（图5.2.1－1）

图5.2.1－1中上部为字符串区域图象；下部为其垂直投影图，f(x)为字符串图象中第x列黑色象素点的数目。因为牌照字符串主要由字母和数字组成，字符区域的垂直投影f(x)>0，当f(x)=0时可以认为x列对应的区域是字符间的空隙部分，某些汉字（如‘沪’）的垂直投影f(x)可能出现为零的现象，此类情况将在2.4.3中加以解决。所以在粗分割中以f(x)＝A（A为常数，根据情况可变化）作为候选字符区域的切分标准。如f(x)>A，认为是字符区域；如f(x)<=A，则认为是字符间隙，投影图中f(x)连续大于A的区域设定为单个字符的候选区域。因为实际照片中牌照图象或大或小，遇到字符串较小时，二值化图象中字符的笔画可能仅有一个象素的宽度，因此垂直投影图中f(x)=1的区域也可能属于某个字符，为了使分割出的字符区域不丢失笔画，本文中参数A取值为零。图5.2.1－1中灰线之间的部分是粗分割的结果，共生成了六个候选字符区域（图中标注的1－6号区域）。

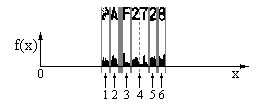
5.2.2 字符区域的确定

常规字符的确定：校正后的字符串上下边界间距离可认为是字符的高度H，根据式5.2－2计算出字符的字符宽度Wc的估计值，以此Wc为中心设置一定的浮动范围（本文中取字符宽度Wc的20％），搜索全部的候选字符区域，将满足条件的区域设定为字符区域，如果此类区域存在，计算它们宽度的统计平均值作为更高的精度值替换初始的字符宽度Wc。图5.2.1－1经过此步骤后区域3、4被确定为字符区。

字符‘1’的确定：根据前一步骤求得的字符宽度Wc可计算出字符‘1’的宽度Wn及其间隙Dn的估计值。以这些值为中心设置一定的浮动范围（选取其值的20％），搜索候选区域，将满足字符宽度及间隙条件的区域设定为数字‘1’区。经过此步骤图5.2.1－1中区域5被确定为字符‘1’区。

5.2.3 字符区域的分裂

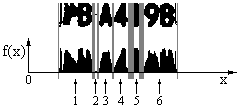
字符串图象经粗分割后，独立的字符区域已被提取出，剩余的候选区域可能是部分字符（区域宽度小于Wc），可能是由一个、多个字符及噪声组成的多成分区域（区域宽度大于Wc）。接下来的工作首先要对多成分区域进行分割，提取其中的字符区域。出现这种情况主要是由两个原因造成的：其一，字符之间虽然是分离的，但由于两个字符区域相距太近，间隙为零，使得垂直投影图中两区域连在一起组成了一个区域；其二，字符之间或字符与噪声（包括边界）相互粘连。

（图5.2.3－1）

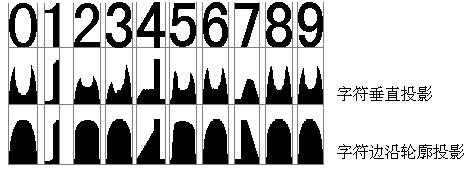
第一种情况比较好解决。虽然字符之间相距很近，但单个字符区域（不包括汉字）在二值图中仍保持连通区的特性，所以只要对多成分候选区域（区域宽度大于Wc）应用区域增长算法，如结果显示此区域为单一连通区，则此区域属于第二种情况；如此区域包含多个连通区，则以每个连通区域的左右边界作为新候选区域的边界拆分原来的区域，生成的新区域只包含一个连通区域，再按照2.4.3节的区域判定标准确定其是否为字符区域。图5.2.3－1中左侧为牌照图象，右侧为字符串区域的垂直投影。投影图显示区域4属于多成分区域的第一种情况，该区域为字符‘2’和‘7’的垂直投影，因为这两个字符相距过近，虽未粘连，但在投影图中不存在间隙，使用区域增长算法后区域4被切分为两个新候选区域（图中灰色虚线为新区域分界线），再按照区域判定标准确定两个新区域为字符区域。

对于第二种情况，需要对多成分候选区域垂直投影图进行分析。本文采用基于字符轮廓边沿特特征的字符分裂算法，实现粘连字符的有效拆分。牌照字符串中本身是存在固定的间隙的，只是因为干扰或技术方面的原因才造成字符间的粘连，两个字符不会完全混合在一起，一般情况下，字符之间仅有其边缘的小部分发生粘连，即使粘连现象比较严重，也不会使字符的两侧边沿完全混入其旁边的字符中，而且发生粘连处基本上是字符两侧边沿中比较突出的部分，在字符串的垂直投影图中可看出字符粘连处投影值很小，且均为该区域的局部极小值，从这两个特性出发，可确定字符间隙的位置。



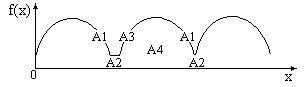
（图5.2.1－1）

图5.2.1－1中区域1及区域6是字符及牌照边框（噪声）粘连的情况，粘连处的投影值较小，但字符‘沪’与‘B’之间的粘连部分较多，其投影值也很大，已超过了字符笔画的宽度，当以特点的投影阈值来判定时，如阈值过高，可能会把字符的有效部分也误判为字符间隙；如阈值过低，则可能会丢失某些间隙位置，且投影阈值很难设定，这种算法是不可取的。当以投影值的局部极小值来判定时，此算法的优点是不会遗漏可能的字符间隙，缺点是搜索出的候选间隙中仍会包含很多字符的有效部分（如图5.2.1－1中字符‘B’、‘9’、‘8’的中间部分会被认为是候选间隙）。为了提高判别字符间隙的准确率，必须使候选区域中近可能少包含字符的有效部分，本文采用基于字符上下边沿轮廓的分裂算法，使得多成分区域候选间隙判定的有效性得到很大的提高。



（图5.2.3－2）

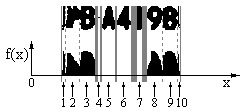
图5.2.3－2中第1行为字符0－9，第2行为字符的垂直投影图，第3行为字符的边沿轮廓垂直投影图，轮廓垂直投影值的定义是字符的上边沿与下边沿间的距离，其中字符上下边沿的定义是以象素点为单位沿字符某列自上向下搜索直到遇见字符区域为止，该点称为字符在此列的上边沿点；沿列自下向上搜索获得的点称为下边沿点。在字符区域内改变列值重复上述步骤得到的边沿点序列称为字符的上下边沿。从图中可看出，字符的垂直投影中大部分都存在两个尖峰及一个凹谷，采用局部极小值来拆分字符时很容易将单个字符分裂为两个区域，使得新生成的候选区域中有较多包含不完整字符的区域，这给分裂后字符区域的提取及后文不完整字符区域的合并带来了很大的困难。与垂直投影比较，字符轮廓垂直投影图具有较好的单峰特性，对其使用局部极小值分裂算法可以切割出完整的字符区域。



（图5.2.3－3）

图5.2.3－3为粘连字符轮廓垂直投影示意图，x为投影所处的列值，f(x)为轮廓投影值。当f(x)-f(x-1)<0时，边沿点属于下降沿区域A1；当f(x)-f(x-1)>0时，边沿点属于上升沿A3；当f(x)-f(x-1)=0且边沿点处于区域A1与区域A3之间（A1、A3次序不可颠倒）时，边沿点属于凹谷区域A2，区域A2的长度为凹谷区域长度L。搜索字符间隙时，沿多成分区域轮廓投影图的x方向自左向右搜索出所有的区域A2，将其设置为候选的字符间隙，以这些间隙为分界线生成新的候选字符区域A4。使用2.4.2节的判定标准确定新候选区域A4是否为字符区域，如区域A4符合判定标准且其两侧的区域A2长度L=1，则区域A4为字符区域；如A4符合判定标准且其两侧的区域A2长度L>1。则可适当以L长度为范围延伸A4，使其更符合判定标准，取得一个最优值设定为字符区域。

当然还存在少数字符，其轮廓投影具有多个尖峰（如汉字‘辽’及字母‘H’等），采用上述算法进行分裂仍然会产生不完整字符，但这种情况本身较少，即使出现，分裂后生成的不完整字符区域也很少，用2.4.4节的合并算法可以使不完整区域得到溶和。



（图5.2.3－4）

图5.2.3－4为多成分区域的轮廓投影图。其中区域4－7为2.4.1节中生成的区域，区域1－3、8－10为采用轮廓投影分裂算法对多成分区域进行分裂后新生成的候选区域（灰色虚线为分界线），使用2.4.1节标准对新候选区域判定，结果区域2、3、8、9被设定为字符区域。使用此算法实现了对粘连的字符及干扰的有效分裂，提取出的字符区域比较精确、完整。

5.2.4 区域的合并

这部分工作的任务是将粗分割及区域分裂生成的含有部分字符的区域溶和为完整区域。本文采用基于聚类的字符合并算法来重构字符区域。



（图5.2.4－1）

图5.2.4－1为字符候选区域的示意图。其中区域A1与A5为已确定的字符区域，区域A2、A3、A4为未确定的候选区域，这些区域可能是噪声区域也可能是包含部分字符的区域，这里定义为短区域，相邻的短区域所构成的集合定义为短区域集合，表示为{Ai | i= l…k}，图5.2.4－1中的短区域集合为{A2,A3,A4}。区域合并的工作是选取最可能是字符的几个相邻的区域，将其溶和生成新的字符区域。以短区域集合中的任意一个区域为聚类中心，结合其相邻的区域形成该集合的一个子集，称为字符集合（字符类）；字符集合在短区域集合中的补集，称为噪声集合（噪声类），因此区域合并可归结为两类的聚类问题。如以区域A2为中心，在短区域集合{A2,A3,A4}内生成字符集合{A2,A3}、{A2,A3,A4}，噪声集合{A4}、{}。

 (5.2.3－1)

式5.2.3－1为聚类效果判定式。其中为字符集合从区域An左边界到区域Am右边界的x方向距离，定义为类内距离；为字符集合从区域An左边界到区域Am右边界在x方向的中点位置，定义为类中心；*L*为预测的字符宽度，*C*为预测的字符位置。假设字符区域A1与A5为确定已知，根据牌照字符串印刷体的规律，可预测出这两个区域间存在几个字符及其宽度与位置。因为在2.4.1－2.4.3节中已完成大部分字符区域的确定工作，字符区域合并通常只出现在一两个字符中，所以上述的假设是可行的。根据聚类判定式，对每一个字符类评价其聚类的效果，选择出其中最符合式5.2.3－1的字符集合，将该集合中的区域合并为新的字符区域（以集合中An左边界为字符区域的左边界，Am右边界为字符区域的右边界），如果短区域集合中包含多个字符区域，则对应于式5.2.3－1将生成多个判别式，区域合并也演化为多类的聚类问题，其实现的方式与单字符区域合并算法相同。

（图5.2.4－2）

图5.2.4－2左侧为牌照图象；中间为经2.4.1-2.4.3节算法生成的候选区域，其中汉字‘沪’被分割为两个区域，其右侧还存在边框噪声区域；图中右侧为字符提取后的图象（灰色包围的区域为字符区域）。采用聚类算法使包含‘沪’字符信息的区域得到溶和形成完整的字符区域，而边框噪声区域被排除在外，在结果图象中字符串中所有的字符信息都完整地保存在提取的字符区域中，同时其它噪声与干扰得到了有效地抑制。

**第6章 字符的识别**

牌照字符的识别主要包括有限汉字（三十余个省、直辖市及自治区的名称缩写）识别，大写英文字母（A－Z）识别、阿拉伯数字（0－9）识别。字符识别的效果将直接影响到牌照识别系统输出的正确度，后文对其算法及应用加以详细阐述。

**6.1 字符识别的基本理论**

字符识别按照识别对象是印刷体字符还是手写字符可分为印刷体字符识别和手写字符识别两类，或按照识别对象的书写(印刷)时间和识别时间的关系分为实时(联机)识别与非实时 (脱机)识别两大类。牌照字符的识别可归为印刷体字符的实时识别，因此以下将着重介绍印刷体字符的实时识别技术。

目前常用的字符识别方法可归结为五类：（一）模板匹配方法，（二）基于字符图象统计特征的识别方法，（三）基于字符图象变换的识别方法，（四）对字符结构进行分析的识别方法，以及(五)应用人工神经元网络的识别方法。

6.1.1模板匹配方法

模板匹配方法是字符识别中最早被采用的一类方法。其基本思想是在将待识字符的图象f(x,y)的全部或局部与模板(基准字符图象)Ti(x,y),i=1,2,…,n(n为模板总数)的全部或局部按一定的规则进行匹配并根据匹配的结果确定待识字符的的类别。典型的模板匹配方法有简单模板匹配方法。此外穿线法、方向模板匹配法、特征点匹配法等方法也属于模板匹配方法。

6.1.1.1简单模板匹配方法

假设待识字符的图象为f(x,y) ，模板(基准字符图象)为Ti(x,y),i=1,2,…,n(n为模板总数)，它们有相同的大小和方向。

现在比较f(x,y)和Ti(x,y)的内容，若两者一致，则f(x,y)和Ti(x,y)之差为零，否则不为零，此时可以用下列两种测度之一来衡量f(x,y)和Ti(x,y)的相似程度：

 (6.1.1.1-1)

或

 (6.1.1.1-2)

其中D(i)是待识字符图象f(x,y)与第i个模板Ti(x,y)间的距离，M为字符或模板的宽，N为字符或模板的高。展开(6.1.1.1-1)有：



式中等号右边第一项与待识字符和模板的距离无关，第三项可以用对基准模板归一化为1的方法使其对距离大小判定不起作用，真正影响距离判定的是第二项，它是待识字符图象和模板的互相关，两者匹配时这项取值最大，因此可用下列相关函数作相似性测度：  (6.1.1.1-3)

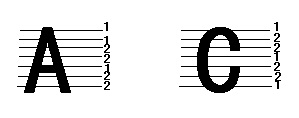
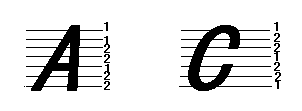
根据Cauchy\_Schwarz不等式知道(6.1.1.1-3)中0≤R(i) ≤1，并且仅在比值f (x,y)/Ti(x,y)=常数时R(i,j)取极大值（等于1）。由于图像中的平均分量常会使相关峰值同背景的差减少，并使相关峰变宽，给匹配程度的判别增加了困难，为此常先从图像和模板中各自减去均值再作相关处理。

当采用距离d(i)作相似性测度时取距离最小时对应的模板字符为识别结果，当采用相关函数R(i)作相似性测度时取最大相关时对应的模板字符为识别结果。

简单模板匹配方法采用的特征是图象中字符象点分布的整体情况，因而当待识字符的笔画位置、宽度与对应模板字符相差很小时该方法较为有效，此时简单模板匹配方法对存在缺损、污迹干扰和噪声干扰的字符图象仍有一定的识别能力，但是由于图象中字符的旋转、缩放和变形往往会较大范围地改变字符象点分布的整体情况，并且简单模板匹配方法匹配时并不考虑各象点在字符结构中的作用即没有使用字符结构方面的信息从而使该方法对旋转、缩放和变形字符的识别能力很差。

6.1.1.2穿线法

穿线法的基本思想用一组或几组与水平成0°、30°、45°、60°或90°的平行线贯穿字符，提取字符与每一平行线的交点次数作为识别依据。图6.2.1－1.2.1.2-1中显示了采用一组7条水平直线对大写字母“A”和“C”的穿线结果，其中“A”为(1,1,2,2,1,2,2)，“C”为(1,2,2,1,2,2,1)，根据这一结果可区分“A”与“C”。

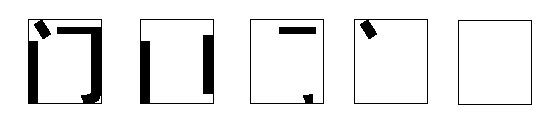
 

(图6.1.1.2-1) (图6.1.1.2-2)

穿线法本质上仍属于模板匹配方法，它采用的特征涉及了字符拓扑结构方面的信息，所以对字符的旋转、缩放、变形有相对于简单模板匹配较大的容忍度(如图6.1.1.2-2，虽然字符发生了倾斜但并不影响穿线结果)，但该方法放弃了对图象中字符象点分布情况的整体了解从而使得它对字符污损、噪声干扰的抵抗力不如简单模板匹配法。

6.1.1.3方向模板匹配法

该方法基于这样的观察：字符(尤其是印刷体汉字)主要由有水平、竖直、倾斜的笔画所构成，可提取这些笔画的信息用以识别。但是没有简单有效的方法能够准确地从字符图象中提取所有的笔画并找出它们之间的联系。因此采用折衷的方法，即沿着几个不同的笔画走向扫描字符图象，从中抽取出由单一走向的笔画构成的原图象的一系列子图象(如图6.1.1.3-1所示，其中左部第一个是汉字“门”的原图象，第二、三、四、五个图象分别是该汉字竖直方向笔画、水平方向笔画、左上到右下方向笔画和左下到右上方向笔画所构成的四个子图象)。通过对基准字符作上述处理取得的一系列子图象称为该字符的方向模板。识别时对待识字符图象也提取对应的子图象，以此作为识别的依据。



(图6.1.1.3-1)

笔画抽取之前要估计字符笔画的宽度。然后根据该宽度提取笔画。提取不同走向的笔画时扫描方向为走向的正交方向。例如为了提取竖直走向的笔画，需对原图象作水平扫描，若某一水平扫描线遭遇到一段宽度与估计所得的笔画宽度接近的线段则将该线段保留到竖直笔画子图象中，否则舍弃（具体方法可参见[14]）。

方向模板匹配法采用的特征较简单匹配法而言考虑了象点在字符结构中的作用，从而能够从字符图象中获取更多的信息，其结果是减少了匹配时将两个不同字符中不同部件包含的处于相同空间位置的点当作同一点的可能性，减小了误识率。

6.1.1.1.4特征点匹配法

特征点匹配法按照一定规则分别从待识字符图象和基准字符图象中提取特征点序列，并将它们配比，根据配比的结果按一定的判决准则来确定待识字符的类别。字符的特征点是字符图象中包含特定信息的象点。例如汉字笔画的端点、折点和交点等以及数字或字母笔画的局部最高点和最低点都可作为特征点。用特征点来表示字符图形可以大量减少字符描述中的冗余信息。这对提高识别速度是有利的，但是特征点的提取通常需花费大量时间，并且所获得的特征点易受字符变形、噪声等因素的干扰，降低了特征点的可靠性。

本文中对汉字的识别采用了特征点匹配方法，关于数字与字母的特征点匹配方法可参见[4]。

6.1.2基于字符图象统计特征的识别方法



模板匹配方法的最大弱点在于它对存在大范围旋转、位移、缩放的字符无识别能力。为此人们提出了基于字符图象自相关系数或各阶矩这类统计特征的识别方法。

6.1.2.1基于自相关系数的识别方法

字符图象f(x,y)的自相关系数R(a,b)可由下式取得：



其中M为图象的宽，N为图象的高，a和b为x和y方向上的偏移。基于自相关系数的识别方法在识别过程中需求取不同a与b时待识字符图象的自相关系数，组成一序列，并将该序列与预先求得的各基准字符的自相关系数序列比较，按一定的判决准则确定待识字符的类别。虽然将自相关系数序列作为识别依据可以部分或全部消除字符图象中由位移、倾斜与缩放带来的干扰，但实验表明[11]该类识别方法的识别效果不理想。

6.1.2.2基于各阶矩的识别方法

字符图象f(x,y)的各阶矩m(p,q)可由下式取得：



其中M为图象的宽，N为图象的高，p和q为x和y的阶数。基于各阶矩的识别方法在识别过程中需求取不同p与q时待识字符图象的矩，组成一序列，并将该序列与预先求得的各基准字符的各阶矩序列比较，按一定的判决准则确定待识字符的类别。实验表明[11]该类识别方法在较好印刷质量情况下可达到较理想的识别效果。

6.1.3基于字符图象变换的识别方法

该类方法通过对字符图象施行某种变换后提取的一系列特征值作为依据进行识别。常用的变换方法有K-L变换、Fourier变换和Hough变换。为了减小用于变换的数据量，通常不对原始图象直接施行变换而是对原图象进行某种预处理后，对预处理的结果进行变换。例如采用Fourier变换进行识别时，先将原图象向x,y轴投影后对投影结果施行变换，这样不仅减少了数据量而且由于投影的累积作用增强了识别方法的抗干扰能力。

6.1.3.1 基于K-L变换的识别方法

设随机图象F是N×N阶的，写成矢量f是维的即：



矢量中的各个分量都是一随机分量。通过K-L变换可以对这一多维随机过程的主要变动成分进行分析。根据分析的结果可将f变换到新的空间中，通过这一过程不仅可以将原数据的信息量压缩到适于处理的程度，而且可以将原数据中的差别更明显的表现出来。

基于K-L变换的识别方法就是将一系列矢量形式的字符图象f看作多维随机过程，通过K-L变化将其中的主要变动部分提取出来，一方面将矢量维数降低，另一方面增大不同字符矢量间的差别。可将变换后的矢量作为识别的依据。然而由于K-L变换涉及的矩阵尺寸依图象矢量f维数的增加而增加，使得具体应用中不可避免地涉及大型矩阵的运算，从而限制了该方法的适用范围。

6.1.3.2 基于Fourier变换的识别方法

该方法的具体实现方法主要有两种。

一种方法是先将原图象向x,y轴投影，将二维数据转变为一维数据，然后对投影结果施行一维Fourier变换，提取若干变换系数作为识别依据。采用投影结果作变换减少了数据量，简化了计算并且由于投影的累积作用增强了识别方法的抗干扰能力。当字符图象的象素总数较小时，这一方法的识别效果不佳。

另一种方法则采用了Fourier描述子。它首先提取字符的外边界，然后假设有一点在此边界上选定一起始位置开始沿着该边界移动直至回到起始位置，移动过程中产生一序列u(t)，u(t)=f(x,y)，其中x,y为移动点经过的坐标，设T为围绕边界一周的时间，则可按下式取得该字符的Fourier描述子序列：



n=0,1,2…，可证明Fourier描述子具有位移和缩放不变性，并且当移动点的起始位置发生了变动或字符旋转了一定角度后取得的Fourier描述子相对于原值仅在相位上发生相应的变动。因此采用若干Fourier描述子组成的序列作为识别依据的识别方法具有识别旋转、位移和缩放字符的能力。但是由于仅仅采用了字符的外边界，使得该方法对某些字符(例如“0”和“8”)的识别能力不高。此外当字符图象的质量不佳时，难以从中提取正常的边界，这也限制了该方法的使用。

6.1.3.3 基于Hough变换的识别方法

该方法主要用于识别汉字等主要由直线段组成的字符类别。通过Hough变换可以将原字符图象中的象点变换到探测直线的参数空间中去。直线参数空间中计数值大的参数项(r, θ)通常对应字符中一直线笔画，其位置与斜率由(r, θ)确定。由于通过Hough变换可以探测字符图象中笔画的位置与斜率，因此可将变换的结果作为识别依据。

6.1.4对字符结构进行分析的识别方法

事实上各种字符尤其是汉字都是由不同笔画按不同位置和次序组合而成的，这种组合形成了各字符特有的结构，这种结构是区分字符的本质特征。前述各种识别方法中采用了各种不同的特征来描述字符，然而这些特征并没有直接描述字符的结构，有一些是凭借人类经验获取的，缺乏理论依据，容易受到各种因素的干扰，因而建立在这种描述基础上的系统很难达到高的识别率。为了准确有效的描述字符结构，通常可采用PDL文法[9]和属性关系图方法[8]。

PDL(Picture Description Language)文法适用于简单形状的字符，它首先获得基准字符的PDL文法描述，根据该描述推断出一特定的文法，然后设计合适的自动机，识别时，通过一系列预处理获得待识字符的PDL文法描述，由自动机对描述进行句法归约以判定字符的类别。

属性关系图方法ARG(Attributed Relational Graph)可定义为：

G=(N,B,AN,AB,GN,GB)

其中：N为节点的有限集合，如汉字的部件或笔画；B为有向弧集合，用有序节点对表示；AN为节点的属性集合，如笔画类型和坐标；AB为弧的属性集合，如笔画的连接关系等；GN：N→AN为产生节点属性函数；GB：B→AB为产生弧属性函数。对于汉字的结构描述而言，因为汉字结构可分为笔划、部件和单字三个层次，所以可用两级ARG来描述。识别时需获得基准字符及待识字符的属性关系图，通过两者间的匹配决定待识字符的类别。

对字符结构进行分析的识别方法对字符的旋转、缩放、变形具有最好的容忍度，但实现困难，许多实现方法尚在探索之中。

6.1.5应用人工神经元网络的识别方法

采用人工神经元网络识别印刷体字符是实现字符识别的又一途径。通常采用的神经网络类型有：单层感知器，多层感知网络，Kohonen网络和Hopfield网络。其中单层感知器和多层感知网络在字符识别中有较好的表现[15]。

6.1.5.1单层感知器

单层感知器由一输入层和一输出层组成。每一层都可由若干节点组成，分属两层的节点之间由不同的权值连接，同层的节点间无任何连接(如图6.1.5.1-1所示)。输出节点将输入节点值的加权和作为函数f的输入，将f的输出作为节点的输出。f(x)可表示为f(x)=sgn(x-θ)，θ为设定的阈值。为了实现对不同类别样本的正确区分需采用已知样本集合对感知器进行训练，使得输入特定样本后，表示该样本类别的输出节点为1其余输出节点为-1。

若设输入样本由矢量X表示：



与第i个输出节点相连的权值由矢量Wi表示：



设第i个输出节点与第i类对应，其输出值为Oi且：



设训练次数为t，并设调整因子η且0<η≤1。则训练过程中每输入一X，可按照以下规则调整所有权值矢量：

(1)若X表示的样本属于第i类，且Oi=1则Wi(t+1)= Wi (t)；

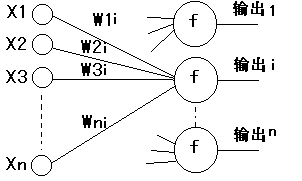
(2)若X表示的样本属于第i类，且Oi≠1则Wi (t+1)= Wi (t)+ηX；

(3)若X表示的样本不属于第i类，且Oi≠-1则Wi (t+1)= Wi (t)-ηX；

(4)若X表示的样本不属于第i类，且Oi=-1则Wi (t+1)= Wi (t)；

训练过程达到满意的程度以后可以将待识样本矢量作为输入，并根据输出判决样本类别。

由于对单层感知器训练的目的在于寻找区分不同类别的超平面，所以当类别数较多或者类别间的界面复杂时该方法无能为力。因此该方法使用于区分较少的类别种类。

在字符识别系统中可以用单层感知器来作粗分类或对粗分类后的结果在较小的类别空间中进行识别。字符识别的输入矢量X可有不同的取法，可以是字符图象矢量或图象分块后由每块子图象的均值组成的矢量。本文中即采用了将部分图象分块后由每块子图象的均值组成的矢量作为输入，用单层感知器区分“0”与“Q”取得了较好的结果。

(图6.1.5.1-1)

6.1.5.2多层感知网络

多层感知网络由一输入层，一输出层以及一个或一个以上的中间层(隐层)组成。每一层都由若干节点组成，分属两层的节点之间由不同的权值连接，同层的节点间无任何连接。采用多层感知网络区分样本类别前也需对网络进行训练。

多层感知网络的训练过程可分为一个前向的过程和一个反向的过程。在前向过程中，输入层将获得的训练样本矢量传递给最邻近的隐层，该隐层中的节点根据输入取得相应的输出传递给下一隐层(若包含一个以上隐层)并最终传递给输出层。反向过程取得输出层的实际输出与期望输出间的偏差后通过各层间的连接将该偏差反向传递到各层并最终到达输入层，各层根据分配到的偏差调整相关的权值以减小实际输出与期望输出间的偏差。正是因为误差反传的特点多层感知网络也称为误差反传网络。

与单层感知器训练的目的不同，多层感知网络的训练的目的在于寻找同一类别中不同样本的相似性。此外，即使是单隐层的多层感知网络也可以近似表示任何连续函数所以当类别数较多或者类别间的界面复杂时该方法仍可作出正确区分。

当训练过程达到满意的程度以后可以将待识样本矢量作为输入，并根据输出判决样本类别。

许多字符识别系统尤其是识别数字与字母的系统采用了多层感知网络。字符识别的输入矢量可有不同的取法，可以是字符图象矢量或图象分块后由每块子图象的均值组成的矢量。[15]中介绍了对字符图象提取简单特征后采用多层感知网络识别手写数字的实验，取得的结果明显优于直接将字符图象矢量作为输入时取得的结果。

**6.2 字符识别的实现**

** **

（图6.2－1）

这一环节的工作是识别字符区域图象中包含的字符信息。本环节的输入为“字符区域提取”环节输出的二值图象（如图6.2－1左侧所示灰色区域包围的七个矩形二值字符区域，）；其输出为识别的结果字符串（如图6.2－1右侧所示，每个字符区域生成一个结果字符合成为输出“沪A70773”）。 本文的字符识别是采用人工神经网络来实现的，网络输入的字符特征选用其灰度信息及纹理信息，使用BP神经网络进行初级字符分类，再用线性感知器进行次级分类。在实验室环境中获得较高的识别率，此方法比较适合于高噪声环境下规则字符的识别。



（图6.2－2）

牌照字符串的识别过程见图6.2－2，主要包括四个部分：首先将待识别字符区域规一化（汉字规一化为9×15象素灰度图，字母、数字规一化为5×9象素灰度图）；接着提取规一图的灰度特征及纹理特征；然后将结合的特征作为神经网络的输入，训练相应的权值矩阵，实现字符的初级分类；最后字母及字母数字神经网络的分类输出经过线性感知器的次级分类产生识别结果，而汉字及数字网络的分类输出直接作为识别结果。

标准的车辆牌照由七个字符组成（见图6.2－3，此类样本图例均为实际拍摄照片数字化后的灰度图象，以100％比例显示，后文提供的图例与此相同）。第一位字符是汉字（省名缩写），第二位字符是大写英文字母，第三位字符是大写英文字母或数字，其余字符是阿拉伯数字。为了提高综合识别率，根据其字符所处位置的差异构造了4个神经网络。其中汉字网络具有135个输入神经元及30个输出神经元；字母网络54个输入，26个输出；字母数字网络54个输入，36个输出；数字网络45个输入，10个输出。下面主要以字母数字网络为例，具体说明实现过程。

（图6.2－3）

* + 1. 规一化及特征抽取

对字符抽取的特征分为两部分，包括图象的灰度信息及纹理信息。



（图6.2.1－1）

第一部分选取图象的灰度信息：这些特征可使神经网络对字符的笔画缺损及随机噪声有较强的抗干扰性。分割后的字符串是7个二值化图象区域（图6.2.1－2左侧中背景为黑色的矩形区域）。首先对图象区域进行规一化，形成5×9象素点阵图。为使原图象有效灰度信息尽可能保留到规一图中，采用伪灰度及象素移交映射方法（见图6.2.1－1），即规一化的操作对象是二值图象而不是原灰度图象，先将二值化图象扩展为256级灰度图象，并映射到规一图上，再将其灰度值按差值比例分配到相邻的四个象素点：

 (6.2.1－1)

其中原图象点映射为规一图中点，和分别为相应的灰度值，而为与相邻的四点。

然后对规一图的灰度值规一化，形成数字字母网络的45个输入特征值：

 (6.2.1－2)

规一化后的图象如图6.2.1－2右侧所示。

（图6.2.1－2）

第二部分选取图象的纹理特征：这些特征可使神经网络不受字符变形的影响。处理的对象是规一化后的伪灰度图象。对图象水平扫描，记录字符的水平笔画数目，形成9个输入单元，例如数字0的笔画特征是（1,1,2,2,2,2,2,1,1）。最后结合两部分的特征，共计54个输入特征值。

* + 1. BP神经网络

（图6.2.2－1）



目前常用的神经网络主要有BP网络、Hopfield网络、Kohonen网络等，由于神经网络自身的复杂性，选用哪种类型网络并没有最优化的方式，主要是针对神经网络进行分类的样本类型、数量来决定。前文的特征选取仅考虑到图象的灰度及纹理特征，忽视了其边缘及轮廓特征，其优点是特征对于字符的噪声及变形不敏感，缺点是不易区分特征值相近的类（如0,D,Q；8,B，见图6.2－3及6.2.2－2，本人曾采用灰度图模板匹配的方式对前五类样本进行测试，结果显示这几类样本对模板的欧氏距离基本相同，无法有效区分）。所以选定的神经网络主要用来初级分类，将0,D,Q归为0类，将8,B归为8类；其它不同类之间的特征差距较大，不需要合并。鉴于此情况，选用BP网络（见图6.2.2－1）来实现字符初级分类，根据样本数及分类数确定用三层网络（实验显示单隐层网络在样本增加时，无法使输出达到收敛点，因此选用含两个隐层的三层神经网络），非线性输出函数采用，其中是系数。

（图6.2.2－2）

对于隐层神经元数的选择是一个十分复杂的问题，目前还没有很好的解析式来表示，可以说隐单元数与分类的要求，输入输出单元的多少都有直接的关系。普通的感知器需要较多的隐单元，由于BP网络的输入和输出之间是单调上升的非线性函数，它的输出是一个软函数，因此比感知器要求的隐单元数少，如何选取最佳的数目，主要参考下面几个公式：

 （6.2.2－1）

式（6.2.2－1）中为样本数，为隐单元数，为输入单元数。

 （6.2.2－2）

式（6.2.2－2）中为输出单元数，为输入单元数，为1－10之间的常数。

 （6.2.2－3）

式（6.2.2－3）中为输入单元数。

综合上述三个公式，确定字母数字网络的隐单元数为16，最后确定网络各层单元数分别为（N1＝54，N2＝16，N3＝16，N4＝36）。

BP网络的最大缺点是样本训练的收敛速度很慢，为了提高网络的收敛速度同时又要防止网络的振荡发散，主要采用下列二种方法：

1. 变步长算法

 （6.2.2－4）

 （6.2.2－5）

其中为权矩阵，为可变步长，为负梯度矢量，是动量因子，为相邻两次迭代误差能量的差值。当误差能量逐次降低时，可逐步加大步长；如误差能量提高，首先将权矩阵恢复为前次迭代值，减小步长，动量因子设为0，再重新计算。此算法可提高收敛的速度，但最大步长不可过大，一般应以初始步长10倍为限，否则易发生振荡。

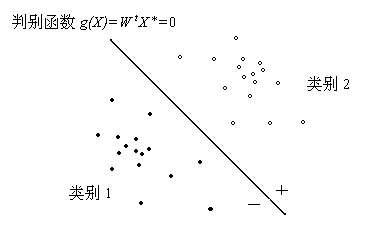
1. 改变输出函数因子

 （6.2.2－6）

当BP网络输出值接近0或1时，网络进入平坦区，常陷入局部极小值，此时可减小的值以使其跳出局部极小值或平坦区，加快神经网络的收敛。

以数字字母网络为例，选取了300多个样本进行训练，未使用上述方法网络经过18000余次迭代达到收敛点（收敛点为学习样本输出与教师差值小于0.1），采用了上述方法后网络经过4100余次迭代达到收敛，收敛速度明显加快，同时未使网络出现振荡现象。

* + 1. 线性感知器



（图6.2.3－1）

采用线性感知器对字符进行次级分类，主要是用来区分0,D,Q及8,B。注意到这几个字符在局部边缘特征上有较大的差异，例如0,D在左上角及右下角的边缘存在差异（见图6.2－3左侧）。因此，选取这几个字符的局部微分值作为线性感知器的分类特征，其判别函数采用线性分界线，仅用来进行两类区分（基于5×9规一图）。

Q与0,D区分的特征向量（右下角局部微分）：

 （6.2.3－1）

0与D区分的特征向量（左上下角局部微分）：

 （6.2.3－2）

8与B区分的特征向量（左侧中部局部微分）：

 （6.2.3－3）

感知器判别函数为：，其中为权向量，为特征向量。当时，特征属于类别1，当时，特征属于类别2。仍然使用上面训练BP网络的样本数据，根据其特征值训练权向量，最后生成的判别函数可实现两类区分。根据上述的特征向量，共构造了三个线性感知器：（Q/0,D）分类器、（0/D）分类器、（8/B）分类器，相互串联或并联区分（0,Q,D,8,B）五类字符。

6.2.4 结果与分析

测试数据选取随机摄取的50幅照片，各类字符共350个。对每类数据随机抽取50％作为训练样本，剩余的样本作为测试数据。考虑到训练样本数较少，对每个训练样本的规一化图象在水平及垂直四个方面进行半象素平移，使样本灰度值产生近10％的漂移，为样本提供人为的离散值。因实际条件限制，汉字仅采集到6类样本，分别为：沪、苏、浙、闽、辽、午。其它样本基本齐全，最后综合识别率见表1。

|  |  |  |  |
| --- | --- | --- | --- |
|  | 样本数 | 误识数 | 识别率 |
| 汉字网络 | 50 | 0 | 100％ |
| 字母网络＋感知器 | 50 | 0 | 100％ |
| 字母数字网络＋感知器 | 50 | 2 | 96％ |
| 数字网络 | 200 | 3 | 98.5％ |
| 总计 | 350 | 5 | 98.6％ |

（表6.2.4－1）

试验结果显示使用神经网络及线性感知器两级串联识别单个字符实现了较高的识别率，此方法对于被测样本的随机噪声、笔画缺损、有限变形具备很强的抗干扰性。图6.2.4－1中第1幅照片字符7上沾有污迹，第2幅照片分割后汉字‘沪’笔画缺损，第三幅照片字符具有明显的垂直向倾斜，最后的结果显示神经网络识别字符对多数干扰都不敏感，可实现准确识别。

原始图 提取的字符区域 识别结果

（图6.2.4－1）

原始图 提取的字符区域 识别结果

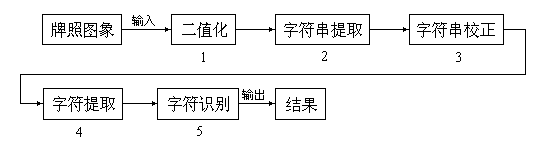
  

（图6.2.4－2）

图6.2.4－2为两幅识别错误的照片，第一幅照片第3,4,5位字符识别错误，第二幅照片字符1错识别为4，分析原因主要是因为技术原因，照片拍摄时曝光过度，带入的白色噪声将被测字符大部分有效信息掩盖，造成识别错误。在识别错误的5个字符中4个字符是在图6.2.4－2的两幅照片中，如果排除此类技术因素，总体误识率小于0.3％。在实际中只要加以有效限制，基本上不会出现此类噪声，如进一步扩大训练的样本数，相信该方法即使在实际使用中也能达到较理想的识别效果。

**第7章 牌照字符识别的系统实现**



（图7－1）

车辆牌照自动识别系统主要由“牌照分割”及“字符识别”两个子系统组成。其中符识别子系统的任务是识别输入图象中所包含的字符信息。系统的输入是前面“牌照分割”部分输出的包含牌照字符串的灰度图象区域（如图7－2左侧所示）；系统的输出是识别出的牌照字符（如图02右侧所示）。根据“字符识别”系统识别过程中各部分的不同功能，可划分为五个环节：（一）牌照图象的分割；（二）字符串区域的提取；（三）字符串区域的校正；（四）字符区域的分割；（五）牌照字符识别；

（图7－2）

牌照图象的基本参数：



（图7－3）

如图7－3所示，牌照图象水平自左向右定义为x方向；垂直自上向下定义为y方向；图象的坐标原点为图象的左上角；图象的宽度以参数W表示；图象的高度以参数H表示。

* 1. 牌照图象的分割

（图7.1－1）

这一环节的任务是将灰度图象转变为二值图象，使字符串信息与背景分离；其输入为256级灰度图象，图象中包含了牌照字符串的所有灰度信息，字符区域在图中占有较大的比例，背景色彩比较单一，不存在其它无效的字符；其输出是包含牌照字符串的二值化图象，字符区域设置为‘1’（以黑色表示），背景区域设置为‘0’（以白色表示）。首先采用水平灰度微分投影来粗略估计牌照区域在图象中所处的位置，获得牌照区域的上下边界Yup和Ydown，及左右边界Xleft和Xright，牌照字符串处于区域（Xleft<x<Xright；Yup<y<Ydown）之内。选取牌照区域中央的子区域，其宽度与高度为牌照区域宽度、高度的1/2，在此子区域中采用基于纹理特征的串线算法来确定前景（牌照字符）区域与背景区域分别所处的灰度级别，并对牌照图象灰度直方图进行规则化，使得字符区域在直方图中处于高灰度级别，背景区域处于低灰度级别。再采用一维灰度直方图聚类算法确定规则化后牌照图象二值化的全局阈值T，生成输出的二值化图象。图7.1－1左侧为本环节输入的牌照灰度图象，右侧为输出的二值化图象。

* 1. 字符串区域的提取

（图7.2－1）

这一环节的任务是确定字符串在牌照图象中的位置，即确定字符串区域的上下及左右边界线。其输入为“牌照图象二值化”环节输出的二值化图象（如图7.2－1左侧所示）；其输出为除去字符串上下边界线以外区域的二值化图象（如图7.2－1右侧所示，字符串上下边界以外的区域设置为背景色，两条垂直线标明字符串区域左右边界的粗略位置）。

首先对环节1中求得的牌照区域采用区域增长算法，求得该区域中所有连通区的上下顶点，获得顶点对序列（顶点对由上下顶点组成）；用聚类的方式删除顶点对序列中无效的元素；再采用霍夫曼变换求得对顶点对序列实现最佳拟合效果的上下两条直线，在求取直线的过程中按一定条件逐步删除离直线较远的顶点，反复迭代，最终获得的直线设定为字符串区域的上下边界线LINEup和LINEdown。对字符串区域的左右边界只进行粗略地估计，以环节1中求得的Xleft和Xright作为字符串的左右边界（精确的边界在字符区域提取中得到实现）。

* 1. 字符串区域的校正

（图7.3－1）

这一环节的工作是通过几何变换使字符串区域中的字符按水平排列，字符的大小基本保持一致，且将字体恢复到垂直状态。其输入是“字符串区域提取”环节输出的字符串区域的二值图象（如图7.3－1左侧所示）；其输出是校正后的字符串区域二值图象（如图7.3－1右侧所示，字符串从向上倾斜的变形状态恢复到水平排列的规则字体状态）。

首先进行水平校正，设定环节2求得的字符串下边界线LINEdown的倾角θ( LINEdown)为字符串区域的水平倾角，校正后字符串呈水平排列；接着进行垂直校正，采用投影方差算法求得字符垂直倾角θ( Vertical)，校正后的字符恢复到竖直状态；最后进行透视校正，水平透视角度为环节2求得的字符串上下边界线倾角之差θ＝ θ( LINEdown)－θ( LINEup)，校正后字符串中每个字符大小保持一致。在进行几何校正时采用象素移交映射算法，计算过程的中间图象象素灰度用浮点小数表示，而校正后的图象仍为二值图（字符区域为‘1’，背景区域为‘0’）。

* 1. 字符区域的分割

（图7.4－1）

这一环节的工作是在校正后的字符串区域中将单个字符区域精确地分离出来。其输入是“字符串区域校正”环节输出的校正后字符串二值化图象（如图7.4－1左侧所示，字符串区域为一水平放置的矩形），其输出是单个字符的二值化图象区域（如图7.4－1右侧所示，被灰色区域包围的七个矩形为分离出的单个字符二值区域）。

首先根据字符串区域的高度预测字符高度、宽度及间距等参数；依据垂直投影图进行粗分割，提取其中独立的字符区域；接着依据轮廓垂直投影图对粘连的区域进行分裂，提取其中完整的字符区域；然后采用聚类算法对不完整区域进行溶和，获得新的字符区域；最后对所有生成的字符区域利用垂直水平投影图检验其是否包含有效的字符。

形成的字符区域应为7个按一定间隔水平排列且大小基本相同的矩形区域。为了获得更精确的二值图象，依据每个字符区域对原灰度图象求取对应的局部阈值Ti(i=0…6)，七个字符区域按其局部阈值再次进行二值化，结果的二值图象还需进行去除噪声及修正边界的处理，修正后的字符区域将作为本环节的输出。

* 1. 字符识别

（图7.5－1）

这一环节的工作是识别字符区域图象中包含的字符信息。本环节的输入为“字符区域提取”环节输出的二值图象（如图7.5－1左侧所示灰色区域包围的七个矩形二值字符区域，）；其输出为识别的结果字符串（如图7.5－1右侧所示，每个字符区域生成一个结果字符合成为输出“沪A70773”）。

首先采用灰度移交算法将七个字符区域分别生成规一化图象（规一化包括区域规一及灰度规一），接着提取规一图象的灰度特征与纹理特征作为神经网络的输入，选取一定的样本训练其网络权值，用神经网络的识别结果进行初级分类（特征相近的字符集，如8,B；0,D,Q等在初级分类中作为同一类对待），最后再用线性感知器对初级类字符集进行次级分类，生成结果的字符串信息。

**第8章 总结**

“车辆牌照自动识别”与传统的OCR系统有许多相似之处。牌照字符串为印刷体字符，在图象中规则排列，其字符集中包括汉字、字母、数字，两者所处理的对象属于相同类型，因此这两种识别系统在识别算法上存在许多共同点。

与OCR系统比较，“车辆牌照自动识别”还具备一些优点，合理利用这些优点可使识别效果显著提高：

1. 每个牌照中的字符串只包含7个字符。
2. 牌照字符串为单一字体的印刷体字符。
3. 牌照的字符集有限：汉字（全国省、直辖市、自治区的简称）为30余个字符；字母（大写英文A－Z）为26个；数字（0－9）为10个。
4. 牌照字符串的高宽比、字符的高宽比、字符间隙符合一定的规则。

当然，“车辆牌照自动识别”因自然因素或人为因素也使其自身存在许多独特的难点。这些特征有些会影响系统中的“牌照分割”部分，有些会影响“字符识别”部分。从本文的“字符识别”出发，其分为五个环节：（一）牌照图象的分割；（二）字符串区域的提取；（三）字符串区域的校正；（四）字符区域的分割；（五）牌照字符识别。牌照图象不同的特征因素会直接或间接干扰不同的环节，从而对总体识别效果产生负面影响：

1. 牌照字符与背景在颜色上有很多变化（如白底黑字、蓝底白字、黄底黑字等），表现在灰度级别上，字符和背景都可能为高灰度级别或低灰度级别，给图象二值化时字符灰度级别的确定带来困难。
2. 牌照字符串上下边沿通常有起固定作用的铆钉，其颜色与字符可能会很相近，或者牌照上沾染的污渍、图象二值化带来的数值误差等都会使字符串的上下边沿的某些部分产生凸起或凹进的变形，给上下边界的确定带来较大困难。
3. 拍摄牌照图象时由于角度原因或牌照本身的扭曲，使得图象中字符串常出现幅度较大的水平垂直倾斜、透视变形或弯曲变形，给牌照的校正带来困难。
4. 由于拍摄照片时光线的影响，或技术因素，使得图象中字符区域的笔画过细，发生区域断裂现象；或字符区域的笔画过粗，发生区域溶和现象，严重的会使图象模糊、边沿不清晰，给字符区域提取带来较大困难。
5. 有些图象携带的噪声太多，或图象过于模糊使得字符区域中基本的信息被掩盖，对字符的识别也带来及其不利的影响。

上面举例说明了因“牌照自动识别”的独特性给“字符识别”部分带来的一些困难，其它如天气、环境、牌照破损等都会对“字符识别”部分的一个或多个环节造成不利影响。这些随机带来的负面因素使“牌照识别”系统较难实现，且无法获得如OCR系统一样的高识别率；同时，字符识别时OCR系统虽然会发生一些误识，但可利用人工进行校正，而“牌照识别”系统很难实现人工干预，一旦其发生误识的现象，会给交通的管理或处罚带来混乱，这些都是“牌照识别”系统无法进入实用领域的因素。

本文所做的工作是在原有识别理论的基础，对其各种算法进行分析并加以利用和形成一定的改进，使它们在“车辆牌照自动识别系统”的“字符识别”中获得更好的应用，提高系统的识别效果及适用的范围，对推动该系统的实用化作出一些贡献。

上文中提到的大部分难点，在本文的第2－6章中都获得较好的解决，不过其在实现中还存在一些缺点，表现为对部分难点处理的效果不理想，如在字符识别的几个环节的处理过程中不可避免地会产生数值化误差，体现在二值图象上会出现象素点偏移，当字符区域在图象所占的象素范围过小（如有些字符的宽度只有5个象素，当水平方向出现1个象素偏移或缺损时，会在结果中引入20％的误差）时，总体识别的效果不理想；还有如图象效果过于模糊，使得大部分字符的有效信息被掩盖，也会使识别的结果失效（在第6章识别结果分析中有详细的阐述）。

针对上述存在的问题，本文认为在今后的系统改进中样本采集时应作出一些改动，如缩短拍摄的距离，提高数值化的解析度，使得待处理的样本图象中字符宽度不小于10个象素；同时可用彩色图象来代替灰度图象，这样使图象中每个象素携带的信息增多，这样比较容易提取其中的字符信息，减少出现图象模糊的可能性；还有在研究算法时应提高“字符识别”环节以前4个环节的效果，这样会使最后的识别环节较易实现，最终能达到较高的识别率。