

利用模糊统计法识别货运列车车号的研究

王晓华 高庆吉 赵卫平

(东北电力学院自动控制系微机教研室, 吉林市 132012)

摘要 室外全天候货运列车车号识别系统, 由于受车厢新旧程度、环境光照变化等因素影响较大, 而且采集的车号图象二值化后, 车号数字常常变形或缺损, 因而用传统的模板匹配等方法已不能准确提取数字特征. 针对该问题提出了一种利用模糊统计法计算待识别数字各方向行程的方法来提取数字的内、外轮廓特征, 并引入模糊决策方法来克服特征提取过程的不确定性. 现场实验结果表明, 该方法对数字的变形或缺损有较强的适应能力, 已取得较高识别率. 该方法对其它的标准粗印刷体数字的识别也有意义.

关键词 字符识别 模糊决策 统计法 轮廓 投影

中图法分类号: TP391.4 文献标识码: B 文章编号: 1006-8961(2000)11-0963-05

To Recognize Bold Printed Numerical Characters with Shape Variation by Using Fuzzy Statistic Method

WANG Xiao-hua, GAO Qing-ji, ZHAO Wei-ping

(Automatic control Dep. North East China Institute of Electric power, Jilin 132012)

Abstract Affected by the wearing of frequent usage and variation of environmental brightness, binaried numerical character images of freight train become deformed and incompleted, which are sampled by freight train mark recognizing system under various weather conditions. As traditional methods, for example pattern plate matching, are not effective to extract the character features exactly, a new method called fuzzy-statistic is presented according to which the numerical character's figure features inside and outside them are extracted by computing their direction routes. By the mean time the fuzzy decision method is drawn on to overcome the indefiniteness during feature extracting. After utilize in some power plant, it appeared that the method is available for character's deformed and incompleted, also high recognized-rate was obtained. The method might be referred to recognize bold printed numerical characters and other courses.

Keywords Character recognition, Fuzzy decision-making, Statistic method, Figure, Projection

0 引言

大家知道, 火电厂入厂煤车检斤过程中需要记录各车厢车号, 以往这项工作由人工手抄完成, 但效率较低, 而本文开发的车号识别系统基本实现了这一过程的自动化. 一般货运列车车号是用一种标准粗印刷体数字显示的, 但由于受车厢新旧程度不同、环境光照变化较大等外界因素影响, 车号图象变化较大, 其二值化后的车号数字图象常常变形或缺损, 因此车号识别的难度在于目标对象变化较大, 因而要求采用一

种能够适应各类光照环境的, 并对有形变目标适应性较强的识别方法, 才能取得较好的识别结果.

经实验可知, 传统的模板匹配、不变矩等方法已不能适应如此大的对象变化, 故而提出一种基于结构特征的模糊统计识别方法, 其主要特点如下:

(1) 综合利用数字的结构特征和统计特征, 但结构特征的比重更大一些, 因为某数字区别于其它数字, 主要是依靠其结构特殊性. 譬如, 若发现某数字的中间具有较大的“圆洞”, 则该数字极有可能是“0”, 因为只有“0”的中间有较大的“圆洞”. 所以只要能准确识别出某数字区别于其它数字的特殊结构,

那么即使该数字的其它部分变形严重,甚至缺损,该数字也能得到正确识别.

(2) 在统计法的判决函数中引入模糊决策方法. 由于字符的统计特征和结构特征往往受成象噪声等因素的影响,在一定范围内波动,呈现一种不确定性,所以决策时,应当引入模糊决策方法进行不精确的推理.

(3) 采用人机结合的监督学习方式修改、扩充数字的模式库. 其中,特征匹配、数字待扩充模式特征矢量的计算和修改模式库等复杂的运算由计算机完成,而输入模式应归于哪一类别(教师信号),则由人来决定.

1 基于结构特征的模糊统计识别法的基本方案

统计法模式识别的大致过程如图1所示.

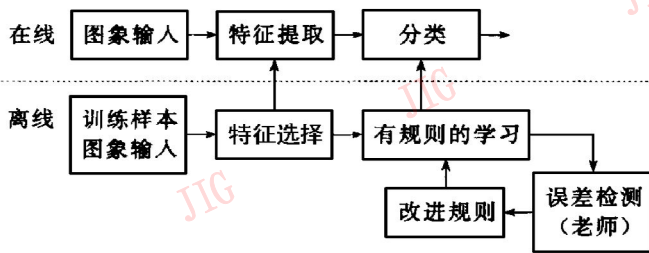


图1 统计学模式识别流程图

图中上半部分是识别部分,即对未知类别的图象进行分类;下半部分是分析部分,即根据已知类别的图象样本制定出判决函数及判决规则,并建立模式库,以便能够对未知类别的图象进行分类. 现对其中主要步骤叙述如下:

(1) 未知图象和训练样本图象均是包含有车号数字串的二值图象,并且经过字符分割后,已被分割成一些只包含单个数字的子图象.

(2) 模式识别时,必须引入各种特征,即与分类有关的各种因素. 在离线分析部分,这一步骤称为特征选择,即从数字的各种特征中选定有利于分类的部分特征;在线识别部分,这一步骤称为特征提取,即从待识别图象中提取离线分析部分所确定的特征.

(3) 框图的下半部分是监督学习部分,在该部分可依据训练样本制定出一些学习规则后,再输入新的样本,然后观察是否有误差,如果有误差,则进一步改进判决规则,直到获得比较满意的结果为止,本文采用人机结合的监督学习方法完成.

1.1 设定特征量

在详细分析了0~9这10个数字之后,本文提出一种采用2类投影图、4种行程图和6种剖图来全面刻划标准粗印刷体数字的方法(如图2(a)所示). 通过这12种特征图,每个数字都能被准确描述.

这里涉及3个基本概念:

(1) 投影 某一行(列)象素灰度值之和,就是这一象素行(列)的灰度累计投影值. 因二值图象中目标的象素值为1,背景的象素值为0,所以本文中投影值即某一行(列)目标象素个数. 由此可见,投影反映的是子图象中目标的统计特征.

(2) 行程 以数字的某一边界(左、右、上、下)为起始点,沿某一指定方向探测到目标为止所走过的距离,称为某方向的行程. 该行程描述了数字的外部轮廓特征,它反映的是数字的结构特征.

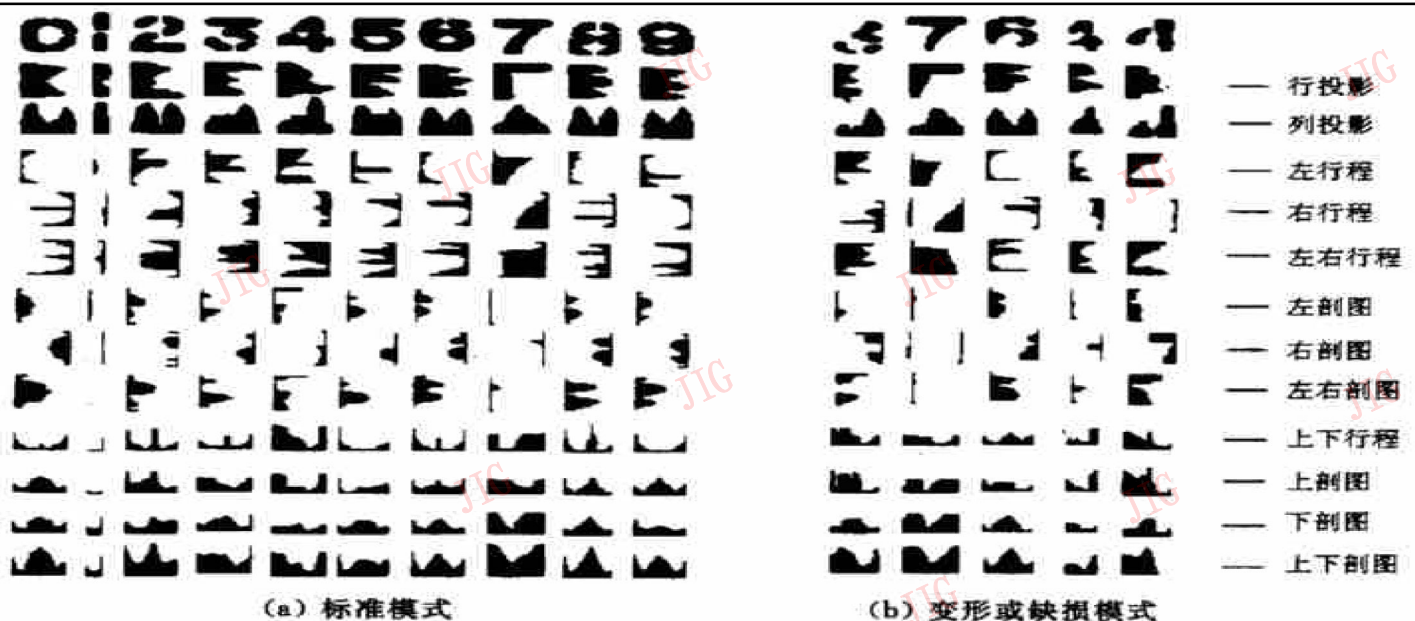


图2 数字特征图

(3) 剖图 将数字从中间(横向或竖向)剖开, 并以数字的中间位置为起始点, 沿某一指定方向(左、右、上、下)探测到目标为止, 其走过的距离称剖图. 该剖图则反映数字的内部轮廓特征.

数字的 2 种投影图分别为行投影图、列投影图; 4 种行程图为左行程图、右行程图、左右行程图、上下行程图; 6 种剖图为左剖图、右剖图、左右剖图、上剖图、下剖图、上下剖图, 各特征图具体含义如下:

行投影 子图象横向各点象素灰度值之和, 即横向灰度值为 1 的象素个数之和.

列投影 子图象竖向各点象素灰度值之和, 即纵向灰度值为 1 的象素个数之和.

左行程 从子图象左边界横向向右探, 遇到子图象中数字外边界为止, 所走过的距离.

右行程 从子图象右边界横向向左探, 遇到子图象中数字外边界为止, 所走过的距离.

左右行程 左行程与右行程之和.

左剖图 从子图象竖向中间位置开始, 横向向右探, 遇到数字内边界为止, 所走过的距离.

右剖图 从子图象竖向中间位置开始, 横向向左探, 遇到数字内边界为止, 所走过的距离.

左右剖图 左剖图与右剖图之和.

上下行程 从子图象上边界开始, 竖直下探, 遇到

数字外边界为止走过的距离, 与从子图象下边界开始竖直上探, 遇到数字外边界为止, 所走过的距离之和.

上剖图 从子图象横向中间位置开始, 竖直上探, 遇到数字内边界为止, 所走过的距离.

下剖图 从子图象横向中间位置开始, 竖直下探, 遇到数字内边界为止, 所走过的距离.

上下剖图 上剖图与下剖图之和.

1.2 特征选择

一类图象的特性往往是比较多的, 因此有必要对这些特性进行进一步选择, 尽量去掉一些误差, 而又能保留原来特性中的信息, 而选择时往往是利用原来的特性, 即通过一些方法, 找出某些综合性指标来, 这称为“特征选择”. 通过详细分析 0~9 这 10 个数字的 12 种特征图, 发现每个数字都有其独有的结构特征和统计特征, 并且某些特征在数字存在一定程度的变形或缺损时, 仍能可靠检测, 例如数字“4”的左右行程图中的顶部高峰(反映数字“4”左上角的空白), 见图 2(b). 依据这些特征, 就可对当前待识别图象进行正确分类. 本文对每个数字所选的特征如表 1 所示, 为清楚起见, 表中采用自然语言对数字特征作定性描述, 而在实际计算时, 每个特征峰是采用 6 个参数作定量描述. 表中数字“1”的特征选择与众不同, 这是因为对“1”的识别主要依据字符宽度.

表 1 数字特征表

行投影	列投影	左行程	右行程	左右行程	左剖图
0 上下两个峰					中间一宽峰
1	中间一宽高峰				没有一个峰
2 底部一高峰		中间一峰	中间一峰	中间一宽峰	偏上一峰
3 顶部一高峰	偏右一峰	偏上一宽峰		偏上一宽峰	偏下一峰
4 中下一高峰	偏右一高峰	顶部一宽峰、下部一窄高峰		顶部一宽峰、下部一窄高峰	
5 顶部一高峰	偏右一峰	偏下一窄高峰	偏上一窄高峰	上下各一窄高峰	
6 上中下三个峰			偏上一窄高峰	偏上一窄高峰	上下两个峰
7 顶部一高峰	偏右一高峰	偏上一峰	偏下一宽高峰	偏下一宽高峰	
8 上中下三个高峰					上下两个峰
9 上中下三个峰		偏下一窄高峰		偏下一窄峰	上部一峰
右剖图	左右剖图	上下行程	上剖图	下剖图	上下剖图
0 中间一宽峰	中间一宽峰		中间一宽峰	中间一宽峰	中间一宽峰
1 没有一个峰	没有一个峰				
2 偏下一峰	上下各一峰		偏左一峰	偏右一峰	
3 偏下一峰	偏下一峰		偏左一峰	偏左一峰	偏左一峰
4		左端部一宽高峰	两端各一峰	中间一宽峰	两端各一峰
5			偏右一矮宽峰	中间一宽峰	中间一宽峰
6 上下两个峰	上下两个峰		偏右一宽峰	中间一宽峰	中间一宽峰
7		左一窄峰、右一宽峰		左右两个宽峰	左右两个宽峰
8 上下两个峰	上下两个峰			中间一高峰	中间一高峰
9 上部一峰	上下两个峰			中间一矮峰	中间一高峰

选定特征之后,接着对其进行准确地定量描述,对某个数字特征的描述,也即对2种投影图、4种行程图和6种剖图中的特征峰进行描述,为每个特征峰选取6个参数,即,峰的位置、平均高度、宽度、最大值、最大值位置和相对高度(平均值以上部分的高度),这样,一个峰的形态就被准确地描述出来。

2 基于模糊决策的判决函数及判决规则

对于 m 类图象 Ω 的识别问题,判决函数是指一组定义在特征空间上的函数

$$g_j(\mathbf{x}) = g_j(x_1, x_2, \dots, x_n) \quad (j = 1, 2, \dots, m)$$

其判别规则如下:对于任意一个图象 \mathbf{x} ,若 $g_i(\mathbf{x}) > g_j(\mathbf{x}) (j \neq i, j = 1, 2, \dots, m)$,则待识别图象 $\mathbf{x} \in \Omega$, $g_i(\mathbf{x})$ 实质上就是图象与第 i 类目标的相似程度。

本文中待识别的是0~9这10个标准粗印刷体数字,也即属于10类图象的识别问题。所以 $m=10$,设 c_{ij} 为模式库中 i 类目标的 j 特征图中的特征峰数目,判决函数为

$$g_i(\mathbf{x}) = \sum_{j=1}^{12} \sum_{k=1}^{c_{ij}} u_{ijk} \times q_{ijk} \quad (i = 1, \dots, 10) \quad (1)$$

式中, q_{ijk} 为与 i 类目标的 j 特征图中第 k 个特征峰对应的权值; u_{ijk} 为判决系数,它决定是否要加上相应的权值 q_{ijk}

这里的关键问题是如何确定判决系数 u_{ijk} 。设 \mathbf{M}_{ijk} 为模式库中 i 类目标的 j 特征图中第 k 个特征峰的均值矢量, \mathbf{V}_{ijk} 为模式库中 i 类目标的 j 特征图中第 k 个特征峰的方差矢量。由于成像噪声等因素的影响,待识别图象 j 特征图中的第 l 个特征峰的特征参数矢量 \mathbf{X}_{jl} 不可能恒等于 \mathbf{M}_{ijk} ,而是以高斯分布的形式集中在 \mathbf{M}_{ijk} 的周围,呈现不确定性,所以应当引入模糊决策方法,利用隶属度概念进行不精确推理^[2]。

设待识别图象的特征矢量 $\mathbf{X}_{jl} = [x_{jl1}, \dots, x_{jl6}]$,定义 \mathbf{X}_{jl} 的第 p 个分量属于 \mathbf{M}_{ijk} 的第 p 个分量的隶属度 d_{ijkp} 为

$$d_{ijkp} = \begin{cases} \exp(-(x_{jl_p} - m_{ijk_p})^2 / 2v_{ijk_p}) & |x_{jl_p} - m_{ijk_p}| \leq 3v_{ijk_p} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

上述隶属度 d_{ijkp} 在0~1之间,若 \mathbf{M}_{jl} 属于 \mathbf{X}_{ijk} ,则 \mathbf{X}_{jl} 的每一分量均应在 \mathbf{M}_{ijk} 对应分量周围,即在逻辑上呈“与”的关系,设 \mathbf{X}_{jl} 属于 \mathbf{M}_{ijk} 的隶属度为 D_{ijk} ,则

辑上呈“与”的关系,设 \mathbf{X}_{jl} 属于 \mathbf{M}_{ijk} 的隶属度为 D_{ijk} ,则

$$D_{ijk} = \min\{d_{ijk_p}\} \quad (1 \leq p \leq 6) \quad (3)$$

令 $u_{ijk} = D_{ijk}$,这样不仅可以判别 \mathbf{X}_{jl} 是否属于 \mathbf{M}_{ijk} ,而且可判断出 \mathbf{X}_{jl} 与 \mathbf{M}_{ijk} 的接近程度,这样与标准模式更接近的特征峰就可以受到更多重视,并得到更大的权值。从而这一方面既能使具有一定程度变形或缺损的数字有可能被识别出来;另一方面,对结构相似的数字之间又不会造成很大干扰,因而避免了反复调整参数的麻烦。

最后,本方法的判决规则如下:

$$\text{令 } f = \max_{1 \leq i \leq 10} g_i, \text{ 若 } f = g_j, \text{ 则判决 } \mathbf{x} \in \Omega.$$

3 人机结合的监督学习

由于模式库中的标准模式是否具有广泛代表性和较高的准确性,对识别率有着极大的影响,因此,环境因素发生较大变化时,系统必须重新学习和更新模式库,才能保证较高的识别率。

学习过程如下:对于输入图象,当识别程序计算出一个识别结果,则调试人员即给出正确识别结果,然后通过计算机检测二者是否有误差,如果有误差,则对输入图象的12个特征图中的特征峰与模式库中的标准特征峰进行特征匹配,并将匹配结果作为待扩充模式样本暂存。全部训练样本输入完毕后,再利用各类目标各特征图中的各特征峰的待扩充模式样本,重新计算各特征峰的均值和方差矢量,然后修改模式库。

若识别程序计算的结果有误差,则说明输入图象的特征峰属于对应标准模式特征峰的隶属度较低,因此如何将输入图象的特征峰正确分类,即如何准确地将输入图象的特征峰同标准模式的特征峰进行匹配是修改、扩充数字模式的关键。

设输入的第 i 类图象的 j 特征图第 k 个特征峰的特征矢量为 \mathbf{X}_{ijk} ,其对应的标准模式的特征峰 \mathbf{H}_{ijk} 的特征矢量为 \mathbf{M}_{ijk} 、 \mathbf{V}_{ijk} , \mathbf{X}_{ijk} 的第 p 个分量 x_{ijk_p} 属于 \mathbf{M}_{ijk} 第 p 个分量 m_{ijk_p} 的隶属度为 d_{ijk_p} , \mathbf{X}_{ijk} 属于 \mathbf{M}_{ijk} 的隶属度为 D_{ijk} 。经过分析发现,若按式(3)计算 D_{ijk} ,只要 $\{d_{ijk_p} (1 \leq p \leq 6)\}$ 中有一个较小,则 \mathbf{X}_{ijk} 属于 \mathbf{H}_{ijk} 的隶属度 D_{ijk} 就较小,因此按式(3)计算的 D_{ijk} 无法将 \mathbf{X}_{ijk} 与 \mathbf{H}_{ijk} 准确地匹配,也就无法准确地计算出待扩充模式的特征均值和方差矢量。因此本

文认为在监督学习部分, D_{ijk} 应按式(4)计算

$$D_{ijk} = \sum_{p=1}^6 d_{ijk_p} / 6 \quad (4)$$

式(4)的依据是, $\{d_{ijk_p} (1 \leq p \leq 6)\}$ 中可能有一个或几个元素较小, 但不会所有元素都很小, 如果那样的话, 则 X_{ijk} 就不是与 H_{ijk} 匹配的特征峰了, 从而可通过求平均值来克服一定的干扰, 因此对于不太标准的数字来说, 按式(4)来计算 D_{ijk} 较为合理。

实验证明, 人机结合的监督学习方法进一步增强了系统的适应性。

4 实验结果

基于本文所述模式识别方法的车号识别系统, 已在东北某电厂通过厂方验收, 而且已投入生产实际运行。本文选取了晴天、阴天以及早晨、上午、下午、黄昏、夜间等各种光照条件下的 10 列国家标准货运列车, 共 326 节车厢的车号图象 ($326 \times 7 = 2\,282$ 个字符) 进行了实验, 实验结果如表 2 所示。

表 2 识别结果

待识别字符数	正确识别字符数	识别正确率
2 282	2 125	93.1%

5 结 论

本文以室外全天候货运列车车号识别为例, 提出了一种用于识别有变形或缺损粗印刷体数字的模糊统计识别法, 该方法利用一种计算待识别数字各方向不同起点行程的方法, 来提取数字的结构特征, 并引入模糊决策理论来克服特征提取过程的不确定性, 因而对数字的变形或缺损有较强的适应能力, 从而也取得了较高的识别率。该方法的缺陷在于对旋转较敏感, 这是由特征提取方法本身所决定的, 当待识别图象发生旋转时, 识别率将有所下降。

参 考 文 献

- 1 程民德, 沈燮昌. 图象识别导论. 上海: 上海科学技术出版社, 1983: 2~ 5.
- 2 文贡坚, 王润生. 基于模糊决策的快速识别多类目标的方法. 模式识别与人工智能, 1997, 10(2): 106~ 111.



王晓华 讲师, 1971 年生. 1993 年毕业于东北电力学院自动控制系获学士学位, 1999 年于东北电力学院自动控制系获工学硕士学位. 研究方向为图象处理、模式识别.



高庆吉 1966 年生, 1993 年毕业于东北电力学院, 工学硕士, 副教授. 现于东北电力学院自动控制系任教, 从事计算机图象处理与识别、计算机控制技术等方面的科研及教学任务.



赵卫平 东北电力学院自动控制系讲师, 1997 年于东北电力学院获工学硕士学位. 研究方向为图象处理、模式识别和介质导热系数.