

形态学运算在字符识别中的应用

杨智勇, 周定康, 周琪云

(江西师范大学 计算机应用研究所, 江西 南昌 330027)

摘要: 字符识别作为基本的图像识别技术, 在计算机输入、交通和安防等领域都获得了广泛的应用。在进行字符识别前, 需要先对图像进行灰度拉伸和阈值分割等一系列预处理以获得有效的目标字符特征。这些处理虽然可以增强目标字符的特征并滤除部分噪声, 但同时也经常会造成目标字符的笔画发生断裂。由于待识别字符的笔画为较细的线性结构, 采用常规的形态学开运算与闭运算对图像进行处理非但不能连接断裂的笔画, 字符笔画还有可能因为细小而被抹去, 使图像中的目标字符产生更严重的笔画断裂甚至缺失。讨论了数学形态学运算在字符识别技术中的应用, 利用形态学闭运算的变形对含目标字符的图像进行增强处理, 可在一定程度上有效消除字符的笔画断裂。实验数据表明, 该方法能有效消除笔画断裂, 进一步增强目标字符特征, 提高字符识别正确率。

关键词: 形态学; 闭运算; 图像处理; 字符识别

中图法分类号: TP391.43

文献标识码: A

文章编号: 1001-3695(2004)06-0099-02

Application of Mathematical Morphologic Process in Character Recognition

YANG Zhi-yong, ZHOU Ding-kang, ZHOU Qi-yun

(Research Institute of Computer Application, Jiangxi Normal University, Nanchang Jiangxi 330027, China)

Abstract: As a basic recognize technique, character recognition is widely used in computer input, traffic, security keeping and other fields. Before characters are recognized, we must do a series of preprocessings to the original image in order to extract their characteristics, such as Gray Stretch, Threshold Transform and so on. These preprocessings can strengthen the characteristics of target characters and get rid of noises. However, they often make the strokes of the target characters broken. Because the structure of the strokes is very thin, normal morphologic opening or closing process will break the strokes more seriously. By reforming the morphologic closing process, this paper proposes a process to eliminate the breaks of target characters. Experiment results show that the process is effective.

Key words: Morphology; Closing Process; Image Process; Character Recognition

1 引言

字符识别技术作为一种基本的图像识别技术, 在社会生产和科学研究中正获得越来越广泛的应用。例如在计算机输入设备方面, 字符识别被广泛应用于手写输入、OCR 扫描输入等; 在交通领域, 则被应用于汽车牌照识别、火车车皮编号识别、集装箱箱号识别等, 成为交通管理、监控的重要辅助手段。在前者及类似的办公应用环境中, 所采集的字符图像背景相对简单、光照均匀、干扰较小, 容易取得较为理想的结果; 而在后者及类似的工业生产应用环境下, 由于所获得的字符图像一般背景较为复杂、光照不均匀、环境干扰较大, 字符识别较难取得理想结果。

在字符识别过程中, 要对图像进行灰度拉伸和阈值分割等一系列预处理以获得有效的目标字符特征, 为后续的字符切分和模式识别打下基础。字符图像中目标字符的笔画均为线性结构、特征不很明显, 而在交通领域、工业生产等复杂的应用环境中, 由于背景相对复杂、光照不均匀、环境干扰大等因素, 字符识别较难取得理想结果。对含有目标字符的图像进行预处

理虽然可以增强目标字符的特征并有效滤除噪声, 但也往往会造成目标字符的笔画断裂。这种笔画断裂首先对字符的切分会产生不利的影 响, 导致字符切分的失败或错误, 更严重的是, 带有断裂笔画的字符在识别时可能会被误识, 导致识别错误的发生。

2 图像处理中的形态学运算

数学形态学是分析几何形状和结构的数学方法, 是建立在集合代数基础上, 用集合论定量描述几何结构的学科, 具有严格的数学理论基础。早期的数学形态学主要应用于微生物学和材料学领域。随着数学形态学逻辑基础的发展, 其应用开始向边缘学科和工业技术方面发展, 如工业控制、放射医学、运动场景分析等。目前, 数学形态学已成为一种应用于图像处理和模式识别领域的新方法。

数学形态学由一组形态学的代数运算子组成的。最基本的形态学运算子有: 腐蚀、膨胀、开和闭。在图像处理领域, 常用这些运算子及组合来进行图像分析及处理, 以描述特定区域的形状如边界曲线、骨架结构和凸形外壳等。

在图像处理的数学形态学运算中, 一幅图像称为一个集合。对二值图像, 一般取值为 1 的点对应景物, 取值为 0 的点

构成背景,这类图像的集合可直接表示。考虑所有取 1 值的点的集合 A,则 A 与图像一一对应。设有两幅图像 X 和 S:

对图像 X,位于其区域内的任意一点 a 称为 X 的元素,记为 $a \in X$;处于区域 X 以外的每一点 a 都不是 X 的元素,记为 $a \notin X$ 。

如果对 S 中的任意一点 b ($\forall b \in S$),都有 $b \in X$,那么称 S 包含于 X,记为 $S \subseteq X$ 。如果同时在 X 中至少存在一点 a ($\exists a \in X$) 而有 $a \in S$,则称 S 真包含于 X,记为 $S \subset X$ 。据定义可知,如果 $S \subset X$,那么必有 $S \subseteq X$ 。 $X \subseteq X$ 恒成立。

在 X 区域以外的所有点构成的集合称为 X 的补集,记为 X^c ,即 $X^c = \{a | a \notin X\}$ 。

X 和 S 的公共点组成的集合称为两个图像的交集,记为 $X \cap S$, $X \cap S = \{a | a \in X \text{ 且 } a \in S\}$ 。

X 和 S 的公共元素组成的集合称为两个图像的并集,记为 $X \cup S$,即 $X \cup S = \{a | a \in X \text{ 或 } a \in S\}$ 。

如果 $X \cap S \neq \emptyset$,那么称 S 击中 X,记为 $S \cap X \neq \emptyset$,其中 \emptyset 是空集的符号。否则如果 $X \cap S = \emptyset$,则称 S 击不中 X。

设 b 为一点,则 X 被 b 平移后的结果为: $X[b] = \{a+b | a \in X\}$ 。即取出 X 中的每个点 a 的坐标值,将其与 b 的坐标值相加,得到一个新点的坐标值 a + b。这些新点所构成的图像就是 X 被 b 平移的结果,记为 $X[b]$ 。

将 X 中的每个点取值相反数后所得的新图称为 X^v ,关于图像原点的反射,记为: $X^v = \{a | -a \in X\}$ 。

对给定目标图像 X 和结构元素 S,将 S 在图像上移动,满足 $S[x] \subseteq X$ 的点 x 的全体构成结构元素与图像的最大相关点集,我们称这个点集为 S 对 X 的腐蚀,记为 $X \ominus S$ 。也可以用集合的方式定义为: $X \ominus S = \{x | S[x] \subseteq X\}$ 。腐蚀在数学形态学运算中的作用是消除物体边界点。如果结构元素取 3 x 3 的黑点块,腐蚀将使物体的边界沿周边减少一个像素。

对图像 X 和结构元素 S,我们定义膨胀为: $X \oplus S = \{x | S[x] \cap X \neq \emptyset\}$,记为 $X \oplus S$ 。膨胀可以视为腐蚀的对偶运算。

对给定目标图像 X 和结构元素 S,如果 S 由两个不相交的部分 S_1 和 S_2 组成,即: $S = S_1 \cup S_2$ 且 $S_1 \cap S_2 = \emptyset$,定义 X 被 S “击中”的结果为:

$$X \oplus S = (X \ominus S_1) \cup (X \ominus S_2)$$

用 $X \oplus S$ 表示 X 对 S 的开运算,用 $X \ominus S$ 表示 X 对 S 的闭运算:

$$X \oplus S = (X \ominus S) \oplus S$$

$$X \ominus S = (X \oplus S) \ominus S$$

细化可视为是腐蚀的一种变形运算,它对图像进行腐蚀,但保证相连目标腐蚀之后仍然连接,可用来抽取字符骨架。目标图像 X 被结构元素 S 的细化用 $X \oslash S$ 表示,根据击中变换定义为:

$$X \oslash S = X - (X \oplus S)$$

3 形态学闭运算的变形及其在字符识别中的应用

在字符识别系统中,我们从所采集的图像中提取目标字符时,由于图像采集时硬件性能、采集方法、光照、污损等实际环境因素的影响,所采集到的图像一般均含有不同程度的噪声。在对目标字符进行切分、识别前,需要对图像进行滤波和灰度拉伸等一系列预处理,然后采用阈值分割的方法对目标字符特

征进行提取。预处理虽然可以增强目标字符的特征并滤除部分噪声,但往往也会造成目标字符的笔画断裂。从图 1 可以看出,经过预处理和阈值分割后,字符“5”产生了笔画断裂。这种发生了笔画断裂的字符切分后进行模式识别时将导致错误的产生。可见,要对字符特征进行有效提取,必须对字符图像进行进一步的处理,以尽可能去除笔画断裂对目标字符的切分和识别所造成的影响。

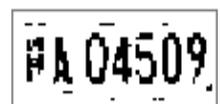
在图像处理过程中,可以利用开运算、闭运算来去除噪声或连接目标,从而达到恢复图像的目的。一般情况下,利用开运算可以消除散点和“毛刺”,即对图像进行平滑。而闭运算在选择适当的结构元素后则可以将两个邻近的分离目标进行连接。

对于含有目标字符的图像,由于目标字符一般为较细的线性结构,如果采用常规的开运算与闭运算对图像进行处理,目标有可能因为太细小而被抹去,使图像中的目标字符产生更严重的笔画断裂甚至缺失。如图 2 所示。S 定义如下(0 为黑色前景,1 为白色背景):

1, 0, 1
0, 0, 0
1, 0, 1

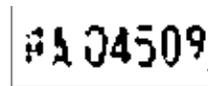


(a) 未经处理的灰度车牌图像

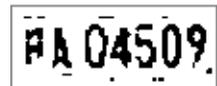


(b) 经过灰度拉伸、粗去噪等预处理后再进行阈值分割所得的结果

图 1



(a) 用 3x3 结构算子 S 对图 1



(b) 进行常规开运算、闭运算的处理结果

图 2

由图 1(a) 和图 2(a) 可知,开运算不但没能消除目标字符的笔画断裂,而且造成了更为严重的笔画断裂和缺失;而闭运算在由图 1(b) 和图 2(b) 中虽然没有产生开运算那样严重的后果,但并未有效消除目标字符的笔画断裂。鉴于此,我们考虑对常规开运算和闭运算进行变形,再用于对字符图像进行处理。

如前定义,设 $X \oslash S$ 表示细化。将开运算中的腐蚀用细化来替换,得到开运算的变形,记为 $X \oslash S$,有:

$$X \oslash S = (X \ominus S) \oplus S \tag{1}$$

同理,将闭运算中的腐蚀用细化来替换,可以得到闭运算的变形,记为 $X \oslash S$,则:

$$X \oslash S = (X \oplus S) \ominus S \tag{2}$$

根据细化的定义, $X \oslash S$ 与开运算相比不再能够有效消除散点或毛刺,细化将使得散点或毛刺至少会有一个骨架像素被保留下来。保留下来的至少一个像素的骨架经过后续的膨胀运算将重新成为散点和毛刺,虽然大小、形状上与原来的会有差别。也就是说,开运算经过变形后性能并未得到相应的提升。设 S 为如前定义的 3 x 3 结构元素。从图 3(a) 可见 $X \oslash S$ 运算既不能连接邻近目标(即消除笔画的断裂),也不能消除散点或毛刺。因此无论是常规的开运算还是如上定义的变形开运算 $X \oslash S$ 都不能达到消除笔画断裂的目的。

而闭运算的变形 $X \oslash S$ 与常规的闭运算相比在性能上则有很大提升。一方面其中的膨胀运算仍可以 (下转第 103 页)

差等信息。需要插入新表值时应进行如下操作:(以上面的ROADS表为例)

```
INSERT INTO SDO_GEOM_METADATA VALUES(
ROADS, GEOMETRY, MDSYS.SDO_DIM_ARRAY(
MDSYS.SDO_DIM_ELEMENT( X, 0, 100, 0.05),
MDSYS.SDO_DIM_ELEMENT( Y, 0, 100, 0.05) );
```

3 数据的上传和写入

MapInfo 公司是当今著名的 GIS 系列软件开发商之一,它开发出了从服务器、桌面系统到 ActiveX 控件等一系列的 GIS 产品,像 MapInfo Professional, MapX 等。由于该公司的产品简单易学、功能强大,而且具有比较强的二次开发能力,所以应用非常广泛。下面以 MapInfo Professional, MapX 为平台说明配电 GIS 数据的上传和写入。

3.1 上传 MapInfo 的地图文件

MapInfo 将地图存储为*.tab 格式的文件,如果要采用 Oracle 8i Spatial 实现空间数据和属性数据的一体化存储,就必须将地图文件存储到数据库服务器中,我们可以利用 MapInfo 公司的 Easyloader 工具进行上传,应用非常简单。

Easyloader 支持三种空间对象类型,MapInfo SpatialWare/Oracle 8i Spatial, XY 和 XY With MapInfo Key(MICODE)。如果服务器是 MapInfo SpatialWare/Oracle 8i Spatial, 则选择第一个,这样数据将作为空间数据存储,一般默认选择此选项。如果服务器不是 MapInfo SpatialWare/Oracle 8i Spatial, 则选择后两者,这时数据 XY 坐标系统的形式存储,并且在服务器上建立一张点集表。XY 和 XY With MapInfo Key 的不同是后者将提供一个 MapInfo 的主键作为空间索引,性能优于前者。

3.2 写入空间数据

动态的写入空间数据有很多方法,大大提高了系统的灵活性,方便数据进行实时更新,在 PowerBuilder 的编程环境的实现代码如下:

```
oleobject lyr, ftrfac, ftr
ftr = create OLEObject
```

```
ftr.ConnectToNewObject( MapX.Feature.4 )
ftrfac = Map1.object.FeatureFactory
ftr = ftrfac.CreateRegion( points, Map1.object.DefaultStyle)
lyr = Map1.object.layers.item( 1)
lyr.keyfield = "MI_PRINX"
ftr.keyvalue = lyr.AllFeatures.Count + 1
lyr.AddFeature( ftr) //添加图元
```

其主要思路是先创建一个空间实体,并赋予属性信息,找出需要添加的位置,然后进行添加,这样用户可以根据自己的需求添加任何电网图形。

4 结束语

Oracle 8i Spatial 可以实现属性数据和空间数据的一体化存储,有效地解决了以往配电 GIS 中可能出现的数据不一致的问题,并且可以很方便的对空间数据进行操作,应用非常方便。Oracle 8i 采用 OCI 这种底层接口,数据存取速度非常快,性能稳定,大大提高了配电 GIS 的效率和可靠性,相信这种用 Oracle 8i Spatial 的存储方式必将得到更广泛的应用。

参考文献:

- [1] 潘农菲. GIS 的空间数据在关系型数据库的实现理论及应用技术[J]. 计算机应用研究, 2002, 19(2): 92-94.
- [2] r Jayant Sharma, et al. Oracle 8i Spatial: 体验可扩展的数据库[M]. Oracle Corporation, 1999. 5.
- [3] racle Corporation. Oracle Spatial 用户手册和指南[M]. Oracle Corporation, 2000. 9.
- [4] 方正培训中心. Oracle 8.1.7 Spatial 培训[M]. 北京: 北大方正, 2001. 9.
- [5] apInfo Corporation. MapX 培训中心教材[M]. MapInfo Corporation, 2001. 8.

作者简介:

杜必强(1974-), 男, 讲师, 硕士, 主要研究方向为地理信息系统; 范孝良(1962-), 男, 副教授, 硕士, 主要研究方向为管理信息系统、地理信息系统; 许少伦(1978-), 男, 硕士, 主要研究方向为地理信息系统。

(上接第 100 页) 通过选取合适的结构元素实现邻近目标的连接, 另一方面其中的细化又保证不会像常规闭运算中的腐蚀一样使膨胀所得的连接处重新断裂, 而是可以保证连接处和原来的连续处一样保留下一个像素宽度的骨架结构。同样设 S 为如前定义的 3 × 3 结构元素。目标字符的断裂被有效去除、字符特征得到进一步的增强。

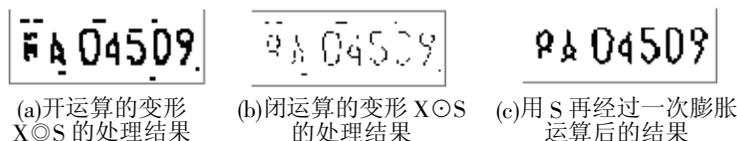


图 3

根据以上分析与实验, 通过采用变形闭运算 $X \odot S$ 对字符图像进行处理, 可有效消除预处理过程中产生的字符笔画断裂, 能在一定程度上提高字符识别的正确率。

4 实验及结论

采用如上定义的变形闭运算 $X \odot S$ 对实验数据进行增强处理, 样本为 1 639 例经过预处理后阈值分割了的二值车牌图像, 其中发生字符笔画断裂 453 例, 进行处理后除 72 例由于断裂严重无法采用上述方法恢复外, 对其余 381 例均取得了一定

的效果, 有效率达 84.1%。

在对图像中的字符进行处理时, 字符笔画经常由于噪声的干扰或对图像的预处理而发生断裂。本文针对这个问题, 提出将形态学闭运算变形后应用于字符笔画断裂的消除, 可以有效地增强目标特征、提高字符的识别正确率。该方法加以修改后也可应用于对其他类似的线性目标的处理。

参考文献:

- [1] 刘利军, 刘吉平, 周远英, 等. 基于数学形态学的地形图自动矢量化研究[J]. 计算机应用研究, 2003, 20(2): 83-85.
- [2] 郑玲, 叶大田. 灰度形态学方法在眼底图像处理中的应用[J]. 清华大学学报(自然科学版), 1997, 37(4): 76-80.
- [3] 阮秋琦. 数字图像处理学——数学形态学原理[M]. 北京: 电子工业出版社, 2001. 429-480.
- [4] 何斌, 马天予, 等. Visual C++ 数字图像处理——数字图像腐蚀、膨胀和细化算法[M]. 北京: 人民邮电出版社, 2001. 335-387.

作者简介:

杨智勇(1976-), 男, 硕士研究生, 主要研究方向为图像处理、模式识别、计算机网络与数据库; 周定康(1944-), 男, 教授, 硕士生导师, 主要研究方向为图像处理、模式识别、计算机网络与数据库; 周琪云(1964-), 男, 副教授, 硕士生导师, 主要研究方向为图像处理、模式识别、计算机网络与数据库。