

文章编号: 1000-5641(2017)03-0064-14

田径运动员号码牌图像的号码识别

赵丽科¹, 郑顺义^{1,2}, 马浩¹, 王晓南¹, 魏海涛¹

(1. 武汉大学 遥感信息工程学院, 武汉 430079;

2. 地球空间信息技术协同创新中心, 武汉 430079)

摘要: 田径运动项目中通常拍摄得到大量的图像, 如何快速获取特定运动员的图像成为普遍关注的问题. 为了快速检索包含特定运动员的图像, 本文提出了识别图像中运动员编号的方法, 依据运动员编号的识别达到快速检索的目的. 首先, 采用 DPM(Deformable Part Model)(可形变部件模型)进行人体检测, 缩小搜索范围, 接着按照运动员号码牌的先验知识, 采用两种方式进行运动员号码牌定位, 保障定位的可靠性; 然后对定位出的号码牌进行字符分割; 最后采用基于特征的 BP(Back Propagation)神经网络的方法进行号码牌识别. 实验结果表明, 在运动员号码牌几乎无遮挡的情况下, 使用本文提出的方法能有效地识别出完整号码牌; 在运动员号码牌存在部分遮挡时, 可以识别出未被遮挡部分的编号. 本文提出的运动员号码牌识别方法为检索特定运动员图像提供了思路, 大大减少了普遍采用的人工查找方式的工作量.

关键词: 田径运动; 字符分割; BP 神经网络; 字符识别

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3969/j.issn.1000-5641.2017.03.007

Research on the number recognition based on athlete number plate image

ZHAO Li-ke¹, ZHENG Shun-yi^{1,2}, MA Hao¹, WANG Xiao-nan¹, WEI Hai-tao¹

(1. School of Remote Sensing and Information Engineering, Wuhan University,
Wuhan 430079, China;

2. Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China)

Abstract: A lot of images are usually photographed in the sports of track and field, and manual operation is a general method to retrieve image containing certain athletes. In order to quickly retrieve images containing a particular player, a method based on the number to identify a player is proposed. It firstly applied DPM (Deformable Part Model) algorithm to narrow the search scope of number plate. Secondly, according to the prior knowledge, the position of number plate can be located by considering two different ways to ensure the reliability. Thirdly, the characters of number plate are segmented by connected component analysis method. Finally, feature-based BP (Back Propagation)

收稿日期: 2016-05-03

基金项目: 国家 863 计划项目(2013AA0630905); 中央高校基本科研业务费专项资金(2042016kf0012);
湖北省科技支撑计划项目(2015BCE080)

第一作者: 赵丽科, 女, 博士研究生, 研究方向为计算机视觉与数字摄影测量.

E-mail: zlk_lenci@163.com.

neural network is adopted to recognize the number plate. Experimental results show that the proposed method is efficient to identify the number plate of the players. The number plate recognition method provides a guideline of retrieving a specific player's images.

Key words: track and field sports; character segmentation; back propagation neural network; character recognition

0 引言

场景中字符的自动识别在图像检索、智能交通等领域有着广泛的应用,随着配备着高清摄像头的手机、数码相机等的大量使用,图像的获取更加方便快捷,图像场景中的字符提取与识别已引起了广泛的关注与研究^[1]. 体育赛场上,运动员的编号作为区分运动员的一个重要标志,广泛应用于体育比赛中运动员的身份识别^[2]. 运动员识别最常用的两种方式有人脸识别与运动员编号识别^[3]. 人脸识别需要众多样本参与,但是许多运动员仅有报名的图像,并且在运动中人的表情、姿态等会发生变化,人脸识别具有一定的难度. 与人脸识别运动员相比,运动员编号的号码牌上的字符通常由0—9组成(有时会有字母,若有字母,字母一般为运动员编号的第一位,用于表示运动员的分组,且每次比赛中包含的字母固定),运动员编号的识别就更为便捷^[4]. 不同于足球、篮球等运动中运动员的编号是直接印在球衣上,田径运动员的号码牌是贴在运动员衣服上的. 田径运动员跑步时身体姿态变化不定,伴随着一定程度的扭曲,同时伴随着手臂的摆动,手臂会出现在号码牌周围,因此会造成号牌一定程度的遮挡,这就导致运动员号码牌的定位、分割和识别存在一定的难度.

目前自然场景中的文本定位方法主要有基于区域和基于纹理特征定位这两种方式^[5]. 文献[6]采用Lab色彩空间进行分割,然后对连通区域进行分析定位文本;文献[7]利用边缘信息生成连通区域,采用由粗到细的方式定位文本;文献[8]采用 k -均值聚类的方式提取文本的连通区域. 总体来说,基于区域的定位方法快速简单,在彩色文本、低分辨率和噪声图像中被广泛应用^[9],但是在图像退化、变形或模糊等情况下不能取得较好的效果^[10]. 用于定位的纹理特征主要包括文本的边缘分布特性、文本周围的灰度较低以及文本内部较大方差等. 文献[11]提出了采用小波变换的基于纹理的文本定位方法;文献[12]采用支持向量机(Support Vector Machine, SVM)的方法定位视频中的文本;文献[13]采用AdaBoost算法和联合概率的方式检测自然场景中的文本. 基于纹理的定位方式对复杂背景下的文本定位能取得较好的结果,但该类方法需要详尽的扫描图像,造成文本定位计算复杂度大、耗时长^[14]. 不断有学者提出字符定位的方法,但是由于不同应用背景下字符文本的特性差别较大,没有普适的方法对每种应用都能取得较好的结果,需要根据应用的不同确定字符文本的定位方法.

字符分割是字符识别的基础,运动员号码牌的识别需要首先完成字符的分割,字符分割的精度直接影响字符识别的精度^[15]. 目前字符分割的方法主要有投影法^[16]、连通区域分析^[17]、模板匹配^[18]等. 投影法在字符无倾斜、旋转的情况下能够有效地分割字符,但对噪声敏感. 连通区域分析法在文本中有噪声、污点的情况下,容易出现分割错误;而结合先验知识的连通区域分析可以有效地改善分割结果. 模板匹配的方法是根据字符的宽度和位置等先验知识建立模板分割字符,这类方法对边框敏感,抗干扰性较差. 此外,也有一些算法将分割和识别进行结合,采用整体识别^[19]的方法,在图像质量较差时也能取得较好的结果,但计算复杂度大.

字符识别是文本分析的最后一个环节, 字符识别算法主要有模板匹配法^[20]、特征匹配法^[21]、神经网络^[22]和支持向量机^[23]等. 模板匹配法通过比较模板与待识别字符的差别进行识别, 计算速度快, 但抗噪能力较差, 只有在字符大小固定、无倾斜、无旋转的情况下才能有效识别. 特征匹配法通过对特征的匹配获得较好的区分相似字符的能力, 但是特征匹配的准则不易把握, 易受噪声影响. 神经网络具有良好的学习、容错和抗干扰能力, 字符识别率高, 但随着网络的复杂程度增大, 计算复杂度也增大. 支持向量机具有适应性强、效率高的特点, 在小样本条件下识别字符具有较高的准确率, 但是该方法对输入的参数要求比较严格, 选择的参数对后续影响较大.

通常大型的田径运动项目如马拉松比赛中, 会拍摄大量比赛的图像. 在查找包含某位运动员的图像时, 现阶段普遍采用的方法是人工查找每一幅图像, 这种查找方式工作量大、耗时费力. 针对缺乏有效检索方法这一问题, 本文提出了一种识别图像中运动员编号的方法: 首先采用 DPM 定位图像中的人体; 接着根据比赛中运动员号码牌的固定颜色组合方式进行号码牌定位(每次比赛中运动员身上所贴的号码牌的背景和字体颜色是固定的, 田径运动中最为常见的号码牌为白色区域、黑色字符组成, 也有一些其他的颜色组合方式, 本文以白色区域黑色号码牌为例, 根据白色区域、黑色字符等先验知识完成号码牌定位, 其他颜色组合方式可以采用类似的思路进行定位); 然后采用结合先验知识的连通区域分析法分割字符, 通过先验知识克服连通区域分析存在对噪声敏感的缺点; 最后, 根据运动员号码牌上字符倾斜、变形等特性, 同时考虑复杂度的影响, 采用基于特征的 BP 神经网络的方法识别号码牌中的各字符, 完成号码牌识别. 整体流程图如图 1 所示.

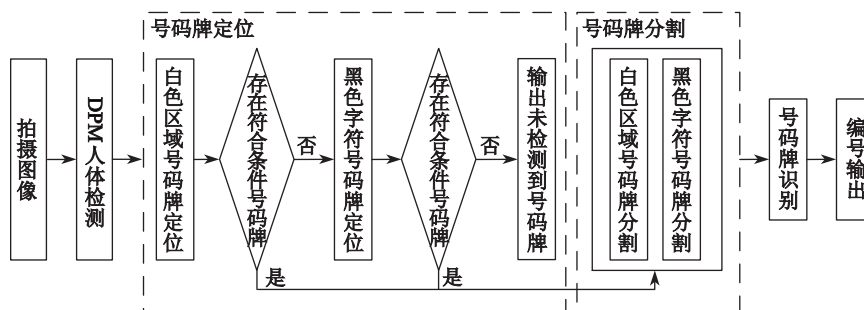


图 1 整体流程图

Fig. 1 The overall flow chart

1 人体检测

号码牌通常贴在运动员衣服上的胸部位置, 在人体中的位置相对固定, 如果能提取出人体在图像中的位置, 则可以避免在整幅图像中盲目搜索, 大大减小号码牌定位的搜索范围. 首先使用人体检测算法提取出图像中的各个个体, 然后针对各人体图像采用本文后续的方法进行运动员编号的识别. 运动场景中每一幅被定格的图像, 运动员的姿态近似为直立状态, 提取人体即提取图像中大体呈直立姿态的人体^[24]. 行人检测常用的方法有 HOG(Histograms of Oriented Gradients)(方向梯度直方图)、ACF(Aggregate Channel Features)(聚合通道特征), DPM 等, 其中 DPM 算法^[25]是一种基于部件的检测方法, 对目标的变形具有很强的鲁棒性, 对人体识别具有较高的精度. DPM 算法是一个混合模型, 由若干个组件模型构成, 其检测采用在不同分辨率上提取改进的 HOG 特征、SVM 分类器和滑动窗口检测方案, 通过求

解混合模型的响应得分, 即特征与待匹配模型的相似程度, 若得到超过了分类阈值, 则认为窗口中包含目标. 混合模型的综合得分计算方法为

$$\text{score}(x_0, y_0, l_0) = R_{0,l_0}(x_0, y_0) + \sum_{i=1}^n D_{i,l_0-\lambda}(2(x_0, y_0) + v_i) + b, \quad (1)$$

其中, $\text{score}(x_0, y_0, l_0)$ 表示锚点 (x_0, y_0) 处在尺度层 l_0 的检测分数, $R_{0,l_0}(x_0, y_0)$ 为主模型的响应分数, $D_{i,l_0-\lambda}(2(x_0, y_0) + v_i)$ 为第 i 个子模型的响应(根据子模型与主模型分辨率间一倍的关系, 子模型的坐标需要映射到更大的尺度层 $l_0 - \lambda$, 即 $(x_0, y_0) \rightarrow 2(x_0, y_0)$), v_i 为子模型 i 相对于 $2(x_0, y_0)$ 的位置偏移, b 为不同组件模型需要对齐所设置的偏移系数. 响应变换为

$$D_{i,l}(x, y) = \max_{dx, dy} (R_{i,l}(x + dx, y + dy) - d_i \cdot \varphi_d(dx, dy)), \quad (2)$$

其中, $D_{i,l}(x, y)$ 为子模型 i 在尺度层 l 的 (x, y) 位置的响应变换, $R_{i,l}(x + dx, y + dy)$ 为子模型在 $(x + dx, y + dy)$ 位置的响应得分. (dx, dy) 为相对 (x, y) 的偏移, $d_i \cdot \varphi_d(dx, dy)$ 为偏移 (dx, dy) 后所损失的得分, d_i 为偏移损失系数, $\varphi_d(dx, dy) = (dx, dy, dx^2, dy^2)$.

本文采用 DPM 算法进行人体检测, 人体检测结果如图 2 所示.



图2 人体检测结果

Fig. 2 Results of human detection

2 号码牌定位

人体检测获取每幅图像中的人体部分后, 需要对人体中的号码牌进行定位. 号码牌定位是后续处理的基础, 只有号码牌被准确定位才能提供可靠的分割、识别依据, 这对整体号码牌的识别至关重要. 由于田径运动员号码牌比较柔软, 容易发生变形, 所以不能依据边界线定位号码牌; 同时由于运动员所穿衣服的颜色可能与号码牌颜色一致, 较可靠的定位方法为分情况进行区域定位. 因此, 本文将号码牌定位分为两种情况: ①运动员衣服为非白色, 依据白色区域定位号码牌; ②运动员衣服为白色, 根据黑色字符定位号码牌.

2.1 白色区域号码牌定位

当运动员穿的衣服不是白色的情况下, 采用区域定位的方式, 直接锁定白色号码牌区域. 一些学者^[26-28]采用 HSV(Hue Saturation Value)、HIS(Hue Intensity Saturation) 等色彩空间进行定位, 但需要进行色彩空间转换, 计算量较大. 考虑黑白两色在该空间的范围, 本文采用的区域定位方式直接在 RGB(Red Green Blue) 空间进行, 同样能得到较好的效果且计算速度快. 理想情况下, R(Red)、G(Green)、B(Blue) 三个分量均为 255 表示纯白色; 然而由于

实际拍摄环境中受光照、角度、噪声等的影响,得到白色的 R、G、B 三个分量在一定的范围内波动,通常表现为 R、G、B 三分量值均较大且近似相等. 本文根据这一特点提取图像中的白色区域. 将 R、G、B 三个分量满足公式

$$(V_{\min} > 170) \vee (V_{\min} > 130, \Delta < 40) \quad (3)$$

的区域认为是白色区域.

公式(3)中,

$$\begin{cases} \Delta = V_{\max} - V_{\min}, \\ V_{\max} = \max(V_R, V_G, V_B), \\ V_{\min} = \min(V_R, V_G, V_B), \end{cases} \quad (4)$$

其中, V_{\max} 表示 R、G、B 三个分量中最大的值, V_{\min} 表示三个分量中最小的值. 当满足公式(3) 时, 判定该区域为白色区域, 其余部分为非白色区域. 按照这种方式对图像进行二值化处理, 图 3(b) 为人体图像二值化结果.

当运动员号码牌周围有其他白色因素的干扰时, 会对号码牌的定位产生影响. 根据数学形态学的闭运算具有一定的平滑功能, 能够检测图像中的边缘、漏洞和孤立点, 可以剔除比结构元素小的图像细节, 填充物体内部细小空洞等特点^[29], 首先对二值图像进行闭运算操作, 剔除图像中的细节, 将内部的小洞填充起来, 如图 3(c) 所示. 其次, 号码牌周围容易有手臂或者类似于白色的图标、字符等干扰元素, 若不消除, 会造成定位范围较大且包含白色字符, 对后续分割产生不利影响. 本文剔除号码牌周围的干扰轮廓的方法为: 定义一个矩形大小为 $w \times h$, 统计该矩形内黑色像素的个数, 若黑色像素个数小于阈值 Th , 则将该矩形内部的黑色像素予以剔除. w 、 h 、 Th 的选择依据为图像中的号码牌内部各个区域的黑色像素的比例较大而干扰位置黑色像素所占的比例较小. 去干扰后结果示例如图 3(d) 所示.

由于运动员身上的号码牌具有一些明显的特性: 号码牌的尺寸满足一定的范围; 标准号码牌为一个矩形, 宽高比固定; 号码牌贴在运动员的上半身; 号码牌中间位置为运动员的编号等. 针对号码牌的这些特性, 结合上述剔除干扰边缘后的二值图像, 获取黑色像素连通区域的最小外包围矩形, 依照下面几个准则剔除非号码牌区域.

(1) 剔除面积过大或者过小的连通区域, 由于号码牌尺寸的限制, 其面积需在一定的范围内.

(2) 号码牌在人体图像的胸部位置, 运动员的号码牌不会位于人体图像的顶部和底部, 剔除位于人体图像顶部或底部位置的连通区域.

(3) 连通区域的外包围矩形需要满足一定的宽高比限制, 虽然号码牌会发生变形或褶皱, 但其宽高比仍在一定的范围内波动, 剔除不满足宽高比要求的连通区域.

(4) 连通区域中黑色像素所占的比重在一定的范围内, 因为号码牌内存在数字, 并且数字所占整个号码牌的部分相对较小, 所以黑色像素所占的比例在一定范围内.

(5) 为了避免运动员衣服上存在类似号码牌颜色的区域, 判断连通区域内部的黑白像素跳变存在规律, 若连通区域无跳变或跳变过于频繁, 则予以剔除. 连通区域无跳变表明该区域中包含字符的可能性极小, 跳变过于频繁表示该区域中包含若干小字符或者图案等非运动员号码牌编号.

按照上述准则, 依次剔除不符合条件的连通区域, 号码牌定位结果如图 3(e) 所示.

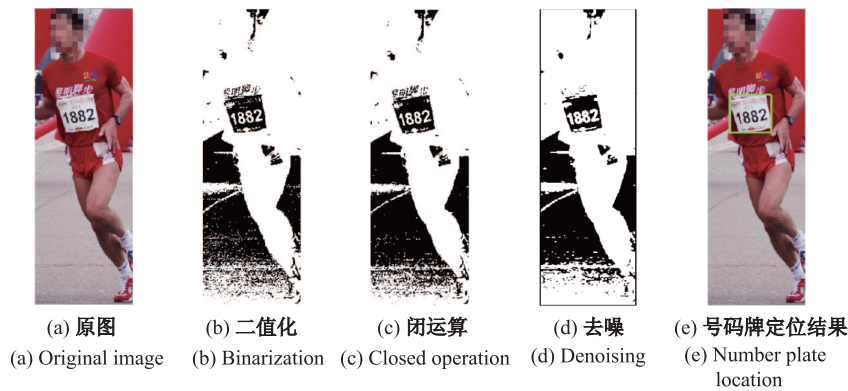


图3 白色区域定位

Fig. 3 Number plate location by white area

2.2 黑色字符号码牌定位

运动员身穿白色衣服时, 按照白色区域定位的原则显然不能得到号码牌的位置. 当采用上述区域定位的方式不能成功定位号码牌时, 采用如下检测黑色字符的方式进行定位.

同样在 RGB 空间直接进行黑色字符的检测, 理想的纯黑色 R、G、B 三个分量均为 0, 由于光照等原因的关系, 黑色表现为 R、G、B 三分量值均较小且近似相等, R、G、B 三分量满足公式

$$(V_{\max} < 80) \vee (V_{\max} < 130, \Delta < 40), \quad (5)$$

则认为该像素为黑色.

公式(5)中各个参数的意义与公式(3)中的各参数意义相同, 此处不再赘述.

限定了黑色像素的 R、G、B 分量之后, 对图像进行二值化处理, 图 4(b) 为二值化后的图像. 需要采用适合的方式将字符连接起来, 构成连通区域. 首先将不可能为字符的部分进行剔除, 剔除面积过大或过小白色连通区域, 图 4(c) 为剔除干扰的结果. 从图中可以看出, 一些明显不是号码牌上字符的干扰特征得以有效剔除, 为后续确定整体字符的连通区域提供了有利的条件. 根据数学形态学中闭运算的特点, 若选择合适的结构元素, 号码牌上数字可以被有效地合并在一起, 采用闭运算获得如图 4(d) 所示的结果. 在进行了上述操作之后, 图像中包含的连通区域为号码牌的候选区域. 同样根据号码牌在人体图像中的分布特性, 依照白色区域定位方式中的准则(1)—(3)进行候选区域的过滤筛选(准则(1)—(3)中的参数会做适当的调整), 图 4(e) 为号码牌的定位结果.

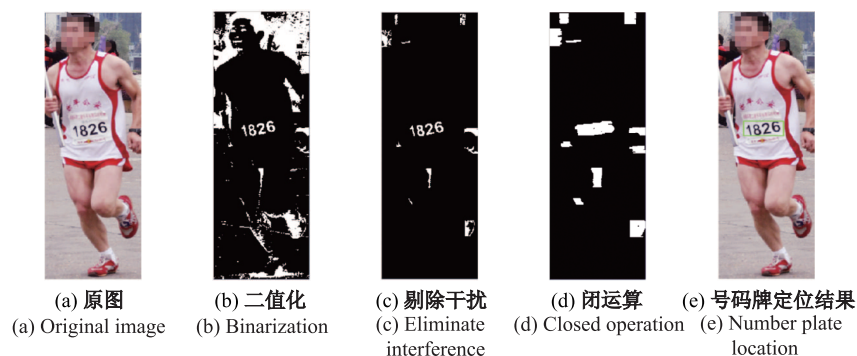


图4 黑色区域号码牌定位

Fig. 4 Number plate location by black area

3 号码牌数字分割

号码牌通常由多个数字组成, 识别的时候按照每个字符的特性进行判断, 因此需要首先完成号码牌中数字的分割. 由于定位得到的号码牌大小以及包含内容不同, 白色区域定位的号码牌包含了顶、底的广告信息, 黑色区域定位的号码牌通常只包含运动员编号, 所以需要根据号码牌的特性进行数字分割.

数字分割之前需要进行图像二值化, 二值化效果的好坏直接影响字符分割、识别的质量. 目前, 二值化方式大致可以分为两类: 全局阈值二值化和局部自适应阈值二值化^[30]. 全局阈值二值化对于目标和背景明显分离、光照分布均匀、噪声干扰较小的图像, 二值化效果较好; 局部自适应阈值二值化是由当前像素灰度值与该像素邻域内灰度特征确定阈值, 对光照不均匀、有突发噪声的情况能得到较好的结果. 以图像中的号码牌为例, 采用 OTSU 最大类间方差法(日本学者大津(OTSU)提出, 又叫大津法)和 Bernsen 二值化方法这两种方法二值化的结果如图 5 所示. OTSU 处理结果虽然能剔除号码牌中广告因素的干扰, 但同时造成了图像中的干扰与字符的黏连; Bernsen 二值化处理虽然保留了图像中的广告信息, 但是可以有效分离出字符周围的干扰, 为后续进一步确定字符分割提供了条件. 针对号码牌内部有时会出现手臂等干扰状况, 选择 Bernsen 局部自适应阈值的方法进行二值化.

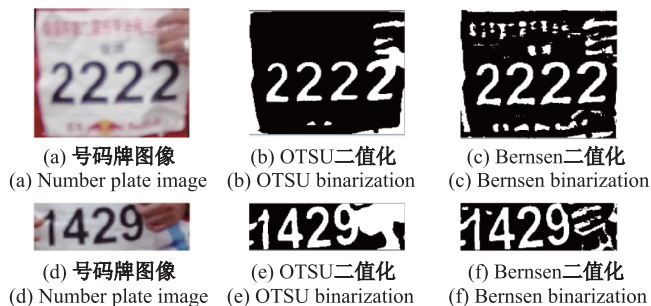


图 5 二值化结果

Fig. 5 Binarization

二值化后的图像需要采用合理的分割方法完成字符分割, 针对号码牌二值化后的特性, 辅以判断条件进行字符分割. 由于号码牌内部可能存在一些背景信息, 以及某些号码牌发生倾斜的情况, 本文提出基于先验知识的连通区域分析法进行字符分割. 如图 6、图 7 分别为白色区域定位号码牌、黑色字符定位号码牌的字符分割结果. 针对白色区域定位出的号码牌顶部、底部存在广告, 运动员编号位于号码牌的中间位置; 黑色字符定位得到的号码牌中通常只包含运动员编号的特性, 得到白色像素连通区域的最小外包围矩形, 外包围矩形需要满足如下要求.

(1) 通过白色区域定位得到的号码牌, 其外包围矩形不能位于号码牌的顶部和底部, 这是由号码牌顶、底的广告所决定的, 若某些比赛白色号码牌内部不存在广告信息以及通过黑色字符定位得到的号码牌, 即定位得到的号码牌内部只包含运动员的编号, 只需将阈值更改即可.

(2) 字符的大小限制, 号码牌大小相对固定, 外包围矩形过大或过小的连通区域均不可能为数字.

将满足这两个要求的白色像素连通区域保留, 其余部分剔除. 通常情况下, 若号码牌内噪声少, 且白色像素连通区域数目与运动员编号数目相同, 则满足这两个要求的白色区域即

为分割结果, 如图 6(a)、图 7(a) 所示. 若满足这两个要求, 但白色像素连通区域的数目与运动员编号数目不同, 则需要进一步处理. 本文提出如下方法进一步分割字符.



图6 白色区域定位的号码牌分割结果

Fig. 6 Segmentation of the number plate located by white area

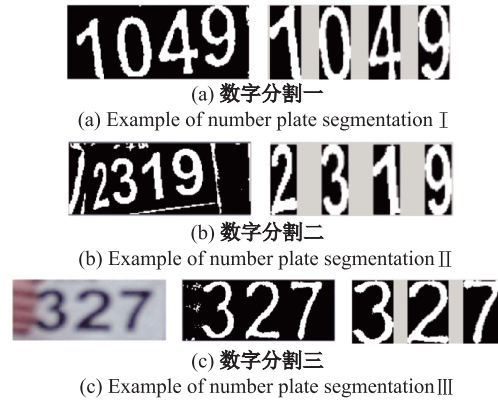


图7 黑色字符定位的号码牌分割结果

Fig. 7 Segmentation of the number plate located by black area

(1) 当白色像素连通区域的数目比运动员编号数目多, 则需要进一步处理. 通常情况下, 当得到的数目比运动员编号数目多(一般比运动员编号多一个数字, 多两个的基本没有)的情况下, 则认为错误发生在第一个以及最后一个的概率大过发生在中间的数字(边界容易有噪声干扰). 去除第一个与最后一个外包围矩形, 求出剩余外包围矩形起始行的平均坐标 $\overline{y_{start}}$ 、平均宽度 \overline{W} 和平均高度 \overline{H} . 求出所有外包围矩形的高度起始行 y_{start} 、宽度 W 和高度值 H 与平均值的差异程度, 其公式为

$$dis = \sqrt{(y_{start} - \overline{y_{start}})^2 + (W - \overline{W})^2 + (H - \overline{H})^2}. \quad (6)$$

将各外包围矩形得到的 dis 进行排序, 保留与运动员编号的字符数目一致的 dis 较小的外包围矩形, 得到最终的分割结果, 如图 6(b)、图 7(b) 所示.

(2) 当得到的白色连通区域的数目比运动员编号数目少时, 通常为号码牌发生了遮挡, 检测到的数字有缺失, 如图 6(c)、图 7(c) 所示. 针对此类情况, 虽然不能得到完整的运动员编号, 但根据号码牌识别的具体应用, 每个运动员具有唯一编号, 且每次赛事中所有运动员编号均可获得的特性, 将不完整的号码牌分割结果进行保留, 接着采用后续识别方法得到不完整的运动员编号, 与运动员编号进行对比, 可以得到该号码牌可能为某运动员.

4 号码牌数字识别

由于运动员号码牌分割得到的数字在一定程度上存在倾斜, 并且字符中可能伴有噪声, 如图 8 所示, 这些特性决定了数字识别需要具有较强的抗噪性、对畸变不敏感的识别方法. 神经网络方法具有较强的容错能力、自适应学习能力和准确性. BP 神经网络是最常用的神经网络方法之一, 通过对神经网络计算得到的输出和样本值进行误差分析, 不断反复修正神经网络中各个权值, 使网络的输出接近期望输出. 本文采用 BP 神经网络用于识别数字, BP 神经网络的结构有: 输入层、网络隐含层、输出层. Kolmogorov 定理表明, 在合理的结构和恰当的权值条件下, 三层网络可以逼近任意的连续函数, 因此本文构建一个三层的 BP 神经网络. 采用分割得到的二值图像进行字符识别, 若进行字符的细化, 则倾斜、变形以及噪声的影响对字符识别影响较大; 若直接采用输入整个二值图像作为输入神经元, 计算复杂度大, 影响字符识别的效率. 本文采用基于特征的 BP 神经网络进行字符识别, 在保证识别精度的同时降低计算复杂度. 输入层神经元的个数即为提取特征的数目, 本文提取 161 个特征, 具体特征提取方式见下文. 输出层的个数即为数字的个数 10. 目前隐含层神经元数目选取的方式有很多, 本文根据输入层、输出层神经元的数目, 结合实践经验, 采用隐含层神经元数目为 44. 每个数字有 200 个训练样本, 为了提高网络的泛化能力, 训练样本中包含规则、倾斜、变形的字符.



图 8 部分字符示意图

Fig. 8 Character sketches

具体特征的提取方式为: 首先将分割后的数字进行归一化(归一化大小为 42×24); 然后将归一化后的二值图像提取的特征作为输入神经元, 具体包括如下特征.

(1) 统计图像中所有白色像素的个数, 共 1 个特征.

(2) 图像中每一行白色像素的个数, 扫描每一行的每个像素, 若为白色像素则加 1, 直到到达图像的边界, 共有 42 个特征.

(3) 扫描图像的每一列, 统计该列白色像素的个数和, 共有 24 个特征.

(4) 将图像分为 $m \times n$ 的大小, 统计每一个区域内部白色像素的个数和, 如图 9(a) 所示, 本文选用 6×6 大小的区域进行统计, 共有 28 个特征.

(5) 统计每一行的分段数, 如图 9(b) 所示, 根据二值图像中每一行像素的值, 统计连续为白色像素段的个数, 共 42 个特征.

(6) 按照与 (5) 类似的方法统计每一列的分段数, 共 24 个特征.

BP 神经网络具体识别字符的过程如下.

第一步, 对字符图像进行大小归一化操作, 将字符的大小统一调整为 42×24 .

第二步, 对样本图像进行特征提取, 特征样本输入到 BP 神经网络学习, 建立识别模型.

第三步, 对字符图像进行特征提取, 采用建立好的识别模型进行数字识别.

第四步, 输出整个号码牌上各字符的识别结果.

按照上述方式采用 BP 神经网络进行字符识别, 获得的识别结果如图 10 所示.

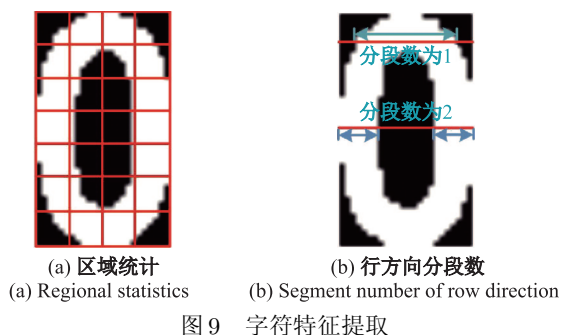


Fig. 9 Character feature extraction results

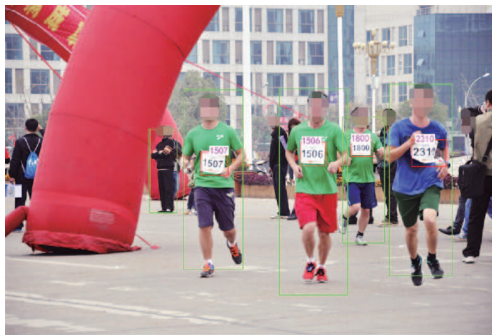


Fig. 10 Character recognition results

5 实 验

为了验证本文提出的运动员号码牌识别方法的可靠性, 在 CPU i7 八核 3.6 GHz、12 GB 内存、Windows 7 64 位操作系统的 PC 机上, 采用 Visual C++ 编程语言进行实验. 采用一组马拉松比赛现场拍摄的图像, 选用的马拉松比赛运动员号码牌为白色区域、黑色字符, 运动员编号为 4 位数字组成. 实验目的为输入包含马拉松运动员的图像, 输出运动员的编号. 这组马拉松运动员的图像库共有 1 042 张 ($4\,288 \times 2\,848$), 包含清晰可辨的人共 4 805 人, 其中包

含佩戴号码牌的运动员共1 924 个. 这组图像中号码牌矩形框的最大尺寸和最小尺寸分别为119×98、29×23, 其中几乎无遮挡的号码牌共有1 528 个, 严重遮挡的号码牌数目为396.

首先采用 DPM 算法对图像库中的照片进行人体检测, 共检测出人体 4 702 个, 正确检测出的人体的数目为 4 515, 检测有误的数目为 187, 未能检测出人体的数目为 240, 其中检测到佩戴号码牌的运动员数目为 1 827. 采用正确检测率、误检测率和漏检测率评价号码牌检测的结果. 正确检测率为正确检测人体数目与图像库中包含人体的实际数目的比值; 误检测率为检测错误的人体数目与图像库中包含人体的实际数目的比值; 漏检测率为未检测到的人体数目与图像库中包含人体的实际数目的比值. 人体检测结果如表 1 所示. DPM 是一种准确性非常高的检测器, 采用多部件模型, 对视角以及姿态变换鲁棒性较强, 能够有效地检测呈直立行走或奔跑的人体. 但由于拍摄距离的差异, 尺度过小对检测图像中的人体造成漏检, 马拉松现场背景拍摄中有时会出现人体结构的干扰会造成误检. 图 11 所示为人体检测的典型示例.

表 1 人体检测结果

Tab. 1 The efficiency of human detection

数据	正确检测率/%	误检测率/%	漏检测率/%
所有图像	93.96	3.89	4.99



图 11 人体检测结果示例

Fig. 11 Examples of human detection results

对检测出来的 1 827 个佩戴号码牌的运动员人体图像进行号码牌定位的实验, 其中检测出人体的号码牌共有 1 810 个, 正确定位号码牌的数目为 1 704, 定位有误的数目为 106, 未能定位的数目为 17. 号码牌的正确检测率、误检测率和漏检测率如表 2 所示. 采用本文提出的号码牌定位方法能够准确地定位出图像中的号码牌, 但当场景中存在类似于号码牌的形状时, 在定位过程中会发生误定位. 图 12 所示为号码牌定位的示例结果.

表 2 号码牌定位结果

Tab. 2 The efficiency of number plate location

数据	正确定位率/%	误定位率/%	漏定位率/%
检测到佩戴号码牌人体	93.27	5.80	0.93

将定位出来的号码牌进行分割处理, 只有当号码牌上有完整的数字时, 才有可能分割得到完整的编号. 由于受遮挡的影响一些号码牌上的数字并不完整, 则会造成无法分割得到完整的编号. 正确定位得到的号码牌中严重遮挡的有 309 个, 号码牌几乎无遮挡的有 1 395 个. 严重遮挡为号码牌中至少包含一个字符被遮挡, 且人眼也不可识别该字符的情况; 几乎无遮挡为号码牌上无

遮挡、遮挡的部位不是字符的部位, 或者遮挡较少的字符的部分, 且人眼可以识别该字符的情况, 如图13所示. 针对几乎无遮挡的号码牌, 只有当号码牌上的所有字符均被正确分割才认为是正确分割, 其中几乎无遮挡的号码牌正确分割数目为 1 297 个, 错误分割数目为 55 个, 漏分割 43 个号码牌. 严重遮挡的号码牌由于缺少字符或字符被严重遮挡难以有效分割, 严重遮挡的号码牌若能有效分割出未被遮挡的字符, 则认为该严重遮挡号码牌正确分割, 严重遮挡号码牌正确分割数目为 276 个, 分割错误数目为 19 个, 漏分割号码牌数目为 14 个, 号码牌的分割情况如表 3 所示. 大多数号码牌能够有效地完成分割, 但由于号码牌中存在干扰、字符遮挡造成的断裂、字符黏连严重等情况, 造成少量号码牌的误分割以及漏分割. 图 14 为图像分割的典型示例.



图 12 号码牌定位结果示例

Fig. 12 Examples of number plate location results

表 3 号码牌分割率

Tab. 3 The efficiency of number segmentation			
数据	正确分割率/%	误分割率/%	漏分割率/%
定位正确的号码牌	92.31	4.34	3.35
定位正确无遮挡号码牌	92.97	3.94	3.08
定位正确遮挡号码牌	89.32	6.15	4.53



(a) 严重遮挡

(a) Examples of severe occlusion



(b) 几乎无遮挡

(b) Examples of almost no occlusion

图 13 遮挡示例

Fig. 13 Occlusion sample



图 14 号码牌分割示例

Fig. 14 Examples of number plate segmentation results

将分割得到的数字采用训练的 BP 神经网络识别模型进行识别, 训练样本中包含了一些变形的数字. 几乎无遮挡情况下分割正确的号码牌中的所有字符均被完全识别的数目为 1 246, 其中有一个数字误识别则认为号码牌识别有误, 误识别号码牌数目为 51, 整个号码牌的识别正确率为 96.07%. 严重遮挡情况号码牌中分割正确的号码牌中的各字符均为正确识别的数目为 265, 误识别的数目为 11 个, 严重遮挡的号码牌识别的正确率为 96.01%. 号码牌识别结果如表 4 所

示. 号码牌中的运动员编号完全无误才认为识别正确; 运动员编号中可能存在某个数字变形过大, 以及分割过程中存在只包含字符的一部分, 该残缺的字符归一化大小后与其他字符类似造成识别错误, 运动员编号识别时某一数字识别失败则该号码牌被误识. 整体来看, 基于特征的 BP 神经网络方法能够高效地识别运动员编号. 图 15 为运动员编号识别的典型示例.

表 4 号码牌识别率

Tab. 4 The efficiency of number recognition

数据	正确识别率/%	误识别率/%
分割正确无遮挡号码牌	96.07	3.93
分割正确遮挡号码牌	96.01	3.99



图 15 号码牌识别示例

Fig. 15 Examples of number plate recognition results

目前识别运动员编号主要针对篮球、足球视频中运动员的识别, 而足球、篮球等运动方式是运动员所着球服的编号, 与田径运动员号码牌的检测差别较大, 缺乏有效的检测田径运动员号码牌的方法. 在字符识别过程中, 采用在车牌识别等领域广泛使用的基于特征提取的模板匹配^[31]方法与上述基于特征的 BP 神经网络方法进行比较, 其中模板匹配与 BP 神经网络采用相同的特征提取方式统计两种方法在图像库中运动员编号的正确检测率、误检测率以及漏检测率, 如表 5 所示. 由于人体检测、号码牌定位、号码牌分割以及识别各步骤积累的误差, 对整个图像库而言, 基于特征的 BP 神经网络方法的整体识别率为 78.35%. 由于基于特征的模板匹配方法对变形较大的字符识别鲁棒性低, 基于特征的 BP 神经网络方法进行运动员编号识别更为可靠.

表 5 图像库运动员编号检测结果

Tab. 5 Performance comparisons of number plate recognition

方法	正确检测率/%	误检测率/%	漏检测率/%
基于特征 BP 神经网络	78.53	3.22	18.24
基于特征的模板匹配	72.47	9.26	18.24

统计单幅图像的时间效率, 由于每幅图像中包含的运动员数目不同, 会造成图像耗时的差异. 统计现场拍摄的 1 042 张图像的处理时间效率, 其均值如表 6 所示.

表 6 各步骤时间效率

Tab. 6 Time efficiency of all steps

步骤	人体检测	号码牌定位	数字分割	数字识别	总体
时间/s	2.04	0.082	0.0002	0.067	2.19

单幅图像从开始预处理定位人体到最终检测出运动员号码牌的编号耗时 2.19 s, 各步骤均能快速有效地完成. 从实验结果可以看出, 号码牌几乎无遮挡情况下, 运动员号码牌识别能取得较高的精度, 严重遮挡的情况下能有效地识别出未遮挡字符的结果, 单幅图像耗时较短, 这些均为运动员检索图像中包含自身姿态良好的图像创造了条件.

6 结 论

针对缺乏快速检索田径运动中包含特定运动员图像的问题, 本文提出了一种识别运动员编

号的方法, 通过人体检测、号码牌定位、字符分割、数字识别得到运动员的编号. 人体检测的可靠性、号码牌定位的准确度、分割的效果以及识别的精度等通过马拉松现场拍摄的图像进行了实验验证: 当号码牌存在严重遮挡时, 能有效地识别出未被遮挡字符; 号码牌上各字符几乎无遮挡情况下, 运动员编号识别准确率较高. 以上这些为在大量图像中快速检索包含特定运动员图像提供了依据.

采用本文的方法在号码牌被严重遮挡情况下不能有效分割、识别字符, 下一步可以结合人脸识别进行运动员的识别, 改善运动员识别结果.

[参 考 文 献]

- [1] EZAKI N, BULACU M, SCHOMAKER L. Text detection from natural scene images: Towards a system for visually impaired persons[C]//Proceedings of the 17th International Conference on Pattern Recognition. IEEE, 2004(2): 683-686.
- [2] 杨靖. 运动员球衣号码研究[J]. 体育文化导刊, 2014, 9: 184-187.
- [3] MESSELODI S, MODENA C M. Scene text recognition and tracking to identify athletes in sport videos[J]. Multimedia Tools and Applications, 2013, 63(2): 521-545.
- [4] YE Q, HUANG Q, JIANG S, et al. Jersey number detection in sports video for athlete identification[C]//Proceedings of 2005 Visual Communications and Image Processing. International Society for Optics and Photonics. 2005: 1599-1606.
- [5] JUNG K, KIM K I, JAIN A K. Text information extraction in images and video: a survey[J]. Pattern Recognition, 2004, 37(5): 977-997.
- [6] MARIANO V Y, KASTURI R. Locating uniform-colored text in video frames[C]//Proceedings of the 15th International Conference on Pattern Recognition. IEEE, 2000(4): 539-542.
- [7] SONG Y, LIU A, PANG L, et al. A novel image text extraction method based on k-means clustering[C]//Proceedings of the 7th IEEE/ACIS International Conference on Computer and Information Science. IEEE. 2008: 185-190.
- [8] SHIVAKUMARA P, PHAN T Q, TAN C L. A Laplacian approach to multi-oriented text detection in video[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(2): 412-419.
- [9] KIM H K. Efficient automatic text location method and content-based indexing and structuring of video database[J]. Journal of Visual Communication and Image Representation, 1996, 7(4): 336-344.
- [10] HUANG Z, LENG J. Text extraction in natural scenes using region-based method[J]. Journal of Digital Information Management, 2014, 12(4): 246-254.
- [11] MAO W, CHUNG F, LAM K K M, et al. Hybrid Chinese/English text detection in images and video frames[C]//Proceedings of the 16th International Conference on Pattern Recognition. IEEE, 2002(3): 1015-1018.
- [12] LEE C W, JUNG K, KIM H J. Automatic text detection and removal in video sequences[J]. Pattern Recognition Letters, 2003, 24(15): 2607-2623.
- [13] CHEN X, YUILLE A L. Detecting and reading text in natural scenes[C]//Proceedings of 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2004(2): 366-373.
- [14] ZHAO X, LIN K H, FU Y, et al. Text from corners: a novel approach to detect text and caption in videos[J]. IEEE Transactions on Image Processing, 2011, 20(3): 790-799.
- [15] CHENG R, BAI Y. A novel approach for license plate slant correction, character segmentation and Chinese character recognition[J]. International Journal of Signal Processing Image Processing and Pattern Recognition, 2014, 7(1): 353-364.
- [16] CHEN Z, CHANG F, LIU C. Chinese license plate recognition based on human vision attention mechanism[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2013, 27(8): 488-196.
- [17] SULAIMAN N, JALANI S N H M, MUSTAFA M, et al. Development of automatic vehicle plate detection system[C]//Proceedings of the 3rd International Conference on System Engineering and Technology. IEEE, 2013: 130-135.
- [18] 王兴玲. 最大类间方差车牌字符分割的模板匹配算法[J]. 计算机工程, 2006, 32(19): 193-195.
- [19] 李文举, 梁德群, 王新年, 等. 质量退化的车牌字符分割方法[J]. 计算机辅助设计与图形学学报, 2004, 16(5): 697-700.
- [20] KO M, KIM Y M. License plate surveillance system using weighted template matching[C]//Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop. IEEE, 2003: 269-274.
- [21] 邓婷. 基于特征统计的车牌非汉字字符识别方法[J]. 广西师范学院学报(自然科学版), 2009, 26(4): 88-92.

(下转第86页)

- [9] 张英杰, 邵岁锋, Niyongabo Julius. 一种基于云模型的云变粒子群算法 [J]. 模式识别与人工智能, 2011, 24(1): 90-96.
- [10] 马颖, 田维坚, 樊养余. 基于云模型的自适应量子免疫克隆算法 [J]. 计算物理, 2013, 30(4): 627-632.
- [11] 李翠明, 龚俊, 牛万才, 等. 基于改进隶属云模型蚁群算法的喷涂机器人喷枪轨迹组合优化 [J]. 上海交通大学学报, 2015, 49(3): 387-391.
- [12] 孙儒泳. 动物生态学原理: 第三版 [M]. 北京: 北京师范大学出版社, 2001.
- [13] 崔志华, 曾建潮. 微粒群优化算法 [M]. 北京: 科学出版社, 2011.
- [14] 彭虎, 吴志健, 周新宇. 基于三角的骨架差分进化算法 [J]. 计算机研究与发展, 2015, 52(12): 2776-2788.
- [15] 李章维, 周晓根, 张贵军. 一种动态自适应差分进化算法 [J]. 计算机科学, 2015, 42(6): 52-56.
- [16] 孔祥勇, 高立群, 欧阳海滨, 等. 求解大规模可靠性问题的改进差分进化算法 [J]. 东北大学学报(自然科学版), 2014, 35(3): 328-332.
- [17] 胥小波, 郑康锋, 李丹, 等. 新的混沌粒子群优化算法 [J]. 通信学报, 2012, 33(1): 24-37.
- [18] 刘金梅, 屈强. 几类混沌序列的随机性测试 [J]. 计算机工程与应用, 2011, 47(5): 46-49.
- [19] POLI R, KENNEDY J, BLACKWELL T. Particle swarm optimization [J]. Swarm Intelligence, 2007(1): 33-57.

(责任编辑: 李 艺)

(上接第 77 页)

- [22] FOR W, LEMAN K, ENG H L, et al. A multi-camera collaboration framework for real-time vehicle detection and license plate recognition on highways[C]//Proceedings of 2008 IEEE Intelligent Vehicles Symposium. IEEE, 2008: 192-197.
- [23] ANAGNOSTOPOULOS C N E, ANAGNOSTOPOULOS I E, LOUMOS V, et al. A license plate-recognition algorithm for intelligent transportation system applications[J]. IEEE Transactions on Intelligent Transportation Systems, 2006, 7(3): 377-392.
- [24] 陈金辉. 静态图像行人检测算法研究[D]. 上海: 华东理工大学, 2015.
- [25] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008: 1-8.
- [26] 朱双东, 张懿, 陆晓峰. 三角形交通标志的智能检测方法[J]. 中国图象图形学报, 2006, 11(8): 1127-1131.
- [27] 王义兴, 黄凤岗, 韩金玉, 等. 基于颜色搭配与纹理特征的车牌定位方法[J]. 中国图象图形学报, 2009, 14(2): 303-308.
- [28] 常巧红, 高满屯. 基于 HSV 色彩空间与数学形态学的车牌定位研究[J]. 图学学报, 2013, 34(4): 159-162.
- [29] 向静波, 苏秀琴, 陆陶. 基于 Contourlet 变换和形态学的图像增强方法[J]. 光子学报, 2009, 38(1): 224-227.
- [30] MILYAEV S, BARINOVA O, NOVIKOVA T, et al. Fast and accurate scene text understanding with image binarization and off-the-shelf OCR[J]. International Journal on Document Analysis and Recognition, 2015, 18(2): 169-182.
- [31] 王建霞, 周万珍. 一种改进模板匹配的车牌字符识别方法[J]. 河北科技大学学报, 2010, 31(3): 236-239.

(责任编辑: 李 艺)