

文章编号:1001-7445(2009)03-0361-05

基于颜色和区域运动目标识别的研究

吴 晓,曹其新

(上海交通大学 机械与动力工程学院,上海 200240)

摘要:针对全自主足球机器人目标识别受光强变化的影响,实时性、准确性和鲁棒性存在的不足,提出了一种基于动态窗口 HSI(H 色度、S 饱和度、I 亮度)色彩空间模型的阈值向量位与及区域合并算法,并通过动态窗口减小搜索范围加快分割速度,实测目标识别的平均运行时间约 26 ms;另外,根据目标的矩不变及卡尔曼滤波的方法,提高了跟踪目标的准确性和鲁棒性,实验结果表明,目标跟踪准确率约 99.3%,使整套系统具有很高的实时性和很好的识别效果。

关键词:目标识别;图像处理;位与算法;动态窗口

中图分类号:TP242.62 **文献标识码:**A

Research of moving target recognition based on color and region

WU Xiao, CAO Qi-xin

(School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

Abstract: In view of drawbacks of autonomous robot soccer in aspects of real-time, precision and robustness due to influence of light intensity in the process of target recognition, a kind of algorithm based on regional merger and threshold vector bit-and by dynamic window HSI(Hue, Saturation, Intensity) color space model is presented. Dynamic window is used to narrow scope of search and accelerate speed of division with average operation time of approximate 26 ms for recognition of actual detected target. Furthermore, moment constant of target and Kalman Filtering Method is used to improve precision and robustness of traced target. The results show that precision in tracing of target is approximate 99.3% which endow whole set of system with high real-time performance and excellent effect of recognition.

Key words: object recognition; image processing; bit and algorithm; dynamic window

全自主足球机器人的重要目标是足球,快速、准确地跟踪足球并计算出位姿和运动趋势是视觉系统的主要任务,也是决策系统的基础^[1-2]。视觉处理主要集中于像素的分类和图像的连通区处理,本文根据足球机器人对目标识别和跟踪的特殊要求,提出了一种将实时性与抗干扰相结合,基于颜色和区域的目标识别算法,目前国内外全自主足球机器人目标识别的时间约在 20~90 ms 之间,识别准确率国外达到 99.6%^[2-3]。

收稿日期:2009-03-21;修订日期:2009-04-22

基金项目:国家自然科学基金资助项目(60443007);福建省教育厅基金资助项目(JB07154)

通讯联系人:吴 晓(1964-),男,江苏睢宁人,副教授,硕士;E-mail:wuxiao186@163.com。

1 目标识别方法

1.1 识别过程

如图 1 所示是全自主足球机器人识别并跟踪足球的过程,首先 directshow 将 USB2.0 的全向摄像机的视频流在内存中抓帧,获取一幅 RGB 图像;第二,先判断目标集 BLOB 是否为空(即上一帧是否被识别),如果不为空则根据帧间关联性进入动态窗口,缩小目标选区,然后将动态窗口内的像素通过颜色对照表映射到 HSI 颜色空间,如果 BLOB 为空,则把全局空间直接映射到 HSI 颜色空间;第三,离散设定球颜色的 H、S、I 向量域值,并和在线检测的图像向量与操作进行分割,将其结果进行游程编码,RL E 对原图像数据进行压缩,同时保留图像中目标的位置尺寸和颜色信息,并且计算复杂度和执行时间与图像尺寸成线性关系,有利于对压缩后的图像进行实时处理,识别出图像中目标的尺寸和位置,尤其机器人足球视觉系统只对图像中数量不多的小目标进行处理,经过 RL E 能够实现很好的压缩效果,使目标在识别阶段的数据量大大减少;第四,根据合并规则将编码后的图像进行连通区分析;第五,通过零阶矩、一阶矩不变的原则求出 BLOB ($m_{00}, m_{11}, x_p, y_p, T, B, L, R$) 识别目标的大小和方位,并跟踪目标^[3-5];采用 BLOB 结构表示视觉目标信息,它包含识别出物体的最小矩形框,其结构体中包含以下数据域: m_{00} (area)表示目标区域的零阶中心矩其值等于目标区域的面积,具有目标区域平移、旋转不变性; m_{11} 表示目标区域的二阶归一化中心矩,具有目标区域平移、旋转和缩放的不变性,也是目标跟踪的主要特征; x_p, y_p 表示目标区域的重心, T, B, L, R 分别表示目标区域最小矩形框的上(top)、下(bottom)、左(left)、右(right)边界^[6]。最后通过卡曼滤波器对 BLOB 参数进行修正,把结果通过 USB 通讯线发送给机器人运动控制系统。

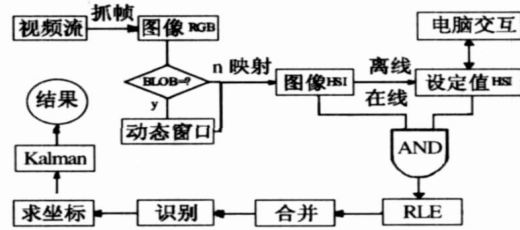
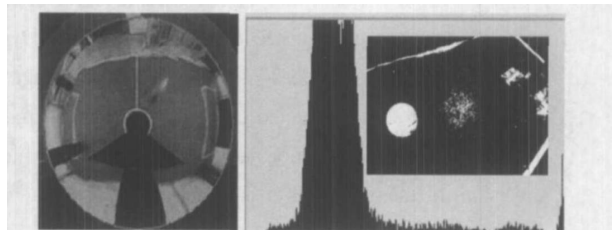


图 1 目标识别过程方框图
Fig.1 Target identification diagram

1.2 目标分析

在全景视觉图像中足球目标只是整个图像很小的部分(如图 2 所示),其红色通道直方图的占的像素很少(图 2b),且颜色并非真正的橙色。



(a) (b)
图 2 目标及其直方图
Fig.2 Target and histogram

另外由于图像畸变、光线变化等原因影响,有时球的图像并不是圆形而呈椭圆形(图 2a),甚至在某些场合还会出现分裂、孔洞、散斑等多目标现象(图 2b 上),致使 BLOB 数据不准确,机器人反应迟缓、摆动等鲁棒性较弱的现象。

针对出现的问题,本文采用颜色向量域值分割、区域合并规则、动态窗口和卡曼滤波等技术得以很好的解决。

2 基于颜色模型的区域分割

传统的彩色图像分割算法大致分为三类:(1) 阈值法;(2) 基于边缘的方法;(3) 基于区域的方法^[5]。

2.1 颜色模型

由于 RGB 色彩空间的三个分量之间相关度很高,改变亮度其他两个分量也会随之变化,不适合足球场上光照变化场合的识别。而 HSI 空间比 RGB 空间更贴近人眼观察彩色的方式,色度 HS 表征了目标的自身属性不随空间变化,亮度 I 表征场地属性随空间变化很大,通过转换可以有效地分离目标的本征和场地的属性^[6],因此采用 HSI 色彩空间^[7-8]。

RGB 空间到 HSI 空间的转换公式为:

$$H = \begin{cases} 2\pi - \arccos\left(\frac{B-G}{\sqrt{R^2 - RG - RB + GB}}\right) & B > G, \\ \arccos\left(\frac{B-G}{\sqrt{R^2 - RG - RB + GB}}\right) & \text{其他} \end{cases} \quad (1)$$

$$= \cos^{-1} \left[\frac{2R - G - B}{2\sqrt{R^2 - RG - RB + GB}} \right], \quad (2)$$

$$S = 1 - \frac{3 \min(R, G, B)}{(R + G + B)}, \quad (3)$$

$$I = \frac{1}{3} (R + G + B), \quad (4)$$

其中, RGB 值归一化为 [0;1] 范围内, 同样 $I = [0;1]$, $H = [0;2\pi]$, $S = [0;1]$ 。

2.2 区域分割

实时的 HSI 与离线的 HSI 学习得来的设定值进行比较, 如图 3 所示, 如果一个像素的 HSI 值落在圆柱体内, 那么该像素点就属于要找的颜色^[3], 而圆柱体空间是根据 HSI 三维的上、下阈值确定的。

由于 HSI 是三维的, 每维有最大、最小两个阈值, 要判断 1 个像素点的颜色是否被选中, 要进行 6 次判断, 可以想象, 对于一帧 640 × 480 的图像, 只一种颜色就要判断 1 843 200 次, 并且按照比赛规则最多同时要识别近 8 种颜色, 这样就使得计算量非常大, 实时性不能保证。

为了减小阈值分割方法计算量, 采用参考文献 [2] 位相与的方法, 一次判别多种颜色。首先定义三个一维数组分别为 Hclass [256]、Sclass [256]、Iclass [256] 数组长度均为 256 (按 RGB 为 8 bit 计算), 然后通过在不同光照条件下离线学习, 就可得出目标的设定域值范围, 比如实测目标球的 H 值 [12, 38]、S 值 [110, 178]、I 值 [45, 220]。写成一维数组的型式: Hclass [0, 0, ..., 1, 1, ..., 1, 0, 0, ...], 也就是第 12 位到第 38 位是 1 其余为 0, 其他同理。

假设某一像素的颜色为 H [23] (表示第 23 位值为 1, 其余为 0)、S [120]、I [80] 判断它是否为目标点, 位与的方法是: { H [12, 38] AND H [23] } AND { S [110, 178] AND S [120] } AND { I [45, 220] AND I [80] } = 1, 所以该点为目标点, 由于位与是一次操作可以得出结果, 且多像素可以同时位与操作, 所以计算速度提高了 6 × 8 = 48 倍。

2.3 区域归并

通过上述三维与的方法分割出目标的区域, 但在求连通域时目标会出现分裂、孔洞、散斑等多目标现象, 因此, 对满足颜色和几何条件的区域进行归并^[8]。

归并操作就是用区域的最小外接矩形替代被归并的区域, 并重新计算区域中像素的平均颜色。在实现时需要对分割出的区域进行反复的两两归并, 直到归并操作不再发生或满足一定的循环次数为止。利用 HSI 分量的均值和方差特征来进行区域合并, 设区域的第 i 个颜色特征矢量为:

$$V_i (\mu_H, \mu_S, \mu_I, \sigma_H, \sigma_S, \sigma_I), \quad (5)$$

其中, μ 是区域均值, σ 是区域方差, 其计算公式如下:

$$\mu = \frac{1}{N} \sum_{n=1}^N c(p_n), \quad (6)$$

$$= \frac{1}{N} \sum_{n=1}^N [c(p_n) - \mu]^2, \quad (7)$$

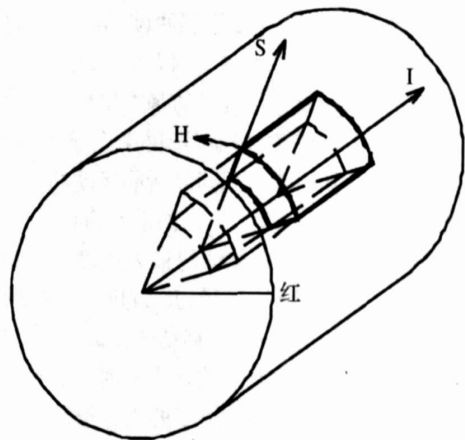


图 3 HSI 色彩空间

Fig. 3 HSI model

其中, $c(p_n)$ 代表像素 p_n 的颜色分量, $n = H, S, I$ 。

合并时, 首先找出尺寸最大的区域, 并计算此区域和其邻域之间的颜色特征的 Hamming 距离 d_j :

$$d_j = \sqrt{(\mu_j - \mu_i)^2 + (\sigma_j - \sigma_i)^2}, \quad (8)$$

其中 $i \sim j$, 然后把满足 Hamming 距离的小区域合并到大区域, 并重复以上步骤, 直到合并结束为止。

求出最大的目标区域后, 经游程编码可求出目标球的特征集 $BLOB(m_{00}, m_{11}, x_p, y_p, T, B, L, R)$, 以便进行判断所跟踪的目标, 并通过 USB 通信传送给机器人运动控制系统。

2.4 动态窗口

所谓的动态窗口是指目标足球的位置可能在的窗口, 它是根据帧间目标的关联性推断出的, 如图 4 所示。当上一帧已经定出目标的特征集 $BLOB(m_{00}, m_{11}, x_p, y_p, T, B, L, R)$ 后, 就不再进行全局扫描, 只进行动态窗口内局部扫描, 动态窗口远小于整幅图像, 使计算量大大减小。

当图像的特征识别周期足够短时, 由于小车运动的最高速度是有限的, 而且比赛场景的相邻两帧的时间间隔也是一定的, 所以每相邻两帧的图像中目标球的运行范围是有限的。因此, 可以采用动态窗口技术进一步缩小扫描区域。该技术的基本思想是在每次获取一帧图像之后, 以本次扫描得到的物体质心为中心, 以物体在前后两帧图像时间间隔所能运动的最大距离为扫描区域, 即形成一个 $L \times L$ 的动态窗口。

假设 V 为图像中心移动的最大速度, t 为采样平均间隔, S 为目标球的半径, R 为机器人两帧间移动的距离, 也即两帧图像中心移动的距离, r 为目标球可能移动的最大距离。则该物体下一时刻的搜索范围是以当前质心为中心, 边长为 L 的正方形区域。按照矢量图得:

$$\begin{aligned} d + R &= d + r & d &= d + r - R, \\ L &= d_{\max} + 2S = 2(R + r + S). \end{aligned}$$

在目标足球没有被踢中的情况下, 其运动速度是小于等于机器人小车的速度的, 所以这里假设目标的速度等于机器人即图像中心移动的最大速度 V , 则有:

$$R = r = V \times t, \quad L = 2(2 \times V \times t + S). \quad (9)$$

由于机器人的中心即为图像中心, 小车的最大移动速度为 2 m/s , 按照球场 $8 \text{ m} \times 6 \text{ m}$, 全像视觉分辨率为 640×480 , 解得小车的最大移动速度: $V = 2 \text{ m/s} = 2000 \text{ mm/s} \times 480 \text{ pixels} \div 8000 \text{ mm} = 120 \text{ pixels/s}$, 即每秒移动约 120 个像素, 按采样 30 fps 计算, 球的半径 $S = 80 \text{ mm}$, 则动态窗口 $L \times L$ 为:

$$L = 2(2 \times 120 / 30 + 80 \times 480 / 8000) = 25.6 \text{ pixels}$$

取 $L = 30 \text{ pixels}$ 则动态窗口为 30×30 邻域。

对于在上一帧图像中已经锁定的目标, 在当前扫描过程中就不必在一幅图像中从头到尾扫描整个图像, 只需要锁定局部区域(边长为 30 个像素的矩形区域)即可, 进一步减少了扫描时间。

2.5 目标跟踪

针对目标球的特征集 $BLOB(m_{00}, m_{11}, x_p, y_p, T, B, L, R)$, 可通过如下公式计算得到:

$$\begin{aligned} m_{00} &= \sum_x \sum_y f(x, y), & m_{10} &= \sum_x \sum_y xf(x, y), \\ m_{01} &= \sum_x \sum_y yf(x, y), & m_{11} &= \sum_x \sum_y xyf(x, y), \\ m_{11} &= \frac{m_{11} - y_0 m_{10}}{m_{00}^2}, & x_p &= \frac{m_{10}}{m_{00}}, & y_p &= \frac{m_{01}}{m_{00}}, \end{aligned}$$

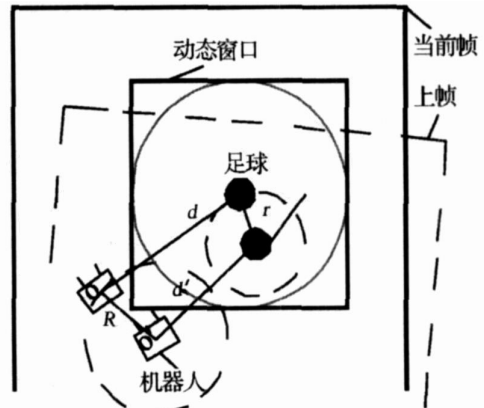


图 4 动态窗口示意图
Fig. 4 Dynamic window diagram

另外 (L, T) 和 (R, B) 分别指目标矩形选区的最左上角坐标和最右下角坐标, 根据目标球投影的椭圆形状和对称性, 得以下经验公式:

$$L + R + 2x_p, T + P = 2y_p, (L - R)(T - P) = 1.2m_{00} \quad (10)$$

如果直接使用普通矩 m_{00} 或是中心矩等不能保证图像同时具有平移、旋转和缩放不变性, 而利用它们的某些线性组合, 理论上能达到归一化中心矩具有平移、旋转和缩放不变性。我们采用彩色阈值分割、 RLC 连通域分析求出目标区域, 再用 I_1 来验证目标的正确性, 从而达到跟踪目标的目的。

为了更加准确估算出动态窗口的位置和大小, 就要对机器人的运动轨迹和足球的运动轨迹进行估算 (R, r) , 我们采用卡尔曼滤波的方法进行处理, 具体方法见参考文献[1]。



图5 实验装置

Fig.5 Test equipment

表1 视觉系统测试运行结果

Tab.1 Test results of vision system

扫描区域大小/像素	目标对象数目	连续扫描的图像帧数	算法平均运行时间/ms	识别目标丢失率/%
640 × 480	1	10	26.13	0
640 × 480	1	20	26.47	0
640 × 480	1	30	26.29	0
640 × 480	1	40	26.67	1
640 × 480	1	50	26.18	1
640 × 480	1	60	26.87	2

3 试验数据分析

在 VC++ 6.0 环境下编程以实现视觉系统的功能, 主要的功能模块有: 视频采集模块、动态窗口模块、彩色图像转换模块、彩色向量位与二值化操作模块、区域分割及目标识别模块、卡尔曼滤波模块, 最终计算出被识别目标的位置及姿态角等 BLOB 目标集信息。系统测试运行结果如表 1 所示, 运行环境为: 双核 3 G 的 CPU, 内存为 3 G。实验装置如图 5 所示, 利用彩色向量位与分割、动态窗口技术相结合的方法, 可以快速、准确地找到目标物体, 这种方式很好地满足了图像扫描的实时性要求, 可以节省更多的时间作一些其他处理, 如提高系统识别精度等。

从表 1 可以看出, 连续扫描目标的帧数与计算运行程序时间差别不是很大, 平均运行时间约 26 ms, 而帧数增加到一定程度会有一定概率的目标丢失。

4 结 语

本文在机器人识别与跟踪过程中, 采用一种改进的阈值向量位与的颜色判断方法, 并利用动态窗口限定扫描范围, 有效地提高了系统的运行速度, 经测试平均运行时间约 26 ms, 结合归一化矩不变、卡尔曼滤波的方法, 使目标识别更加准确且鲁棒性更强, 通过实测目标跟踪准确率达 99.3%, 与国外足球机器人目标跟踪准确率 99.6% 很接近^[2], 完全能够满足全自主足球机器人的比赛要求, 也可以推广到其他领域应用。

参考文献:

- [1] 吴晓. 基于新规则自主足球机器人的研究[J]. 湖南文理学院学报, 2008, 20(2): 61-66.
- [2] MAIK S K, MURTHY C A. Standardization of edge magnitude in color images [J]. IEEE Transactions on Image Processing, 2006, 15(9): 2 588-2 595.
- [3] 赵增荣. 基于 Robocup 中型组足球机器人的图像识别算法[J]. 山东理工大学学报, 2007, 21(1): 88-91.
- [4] 徐大宏. 足球机器人视觉系统中彩色目标分割方法[J]. 中南工业大学学报, 2002, 33(4): 428-430.
- [5] 黄晶. 基于快速彩色空间变换的足球机器人目标搜索[J]. 哈尔滨工业大学学报, 2003, 35(9): 1 036-1 038.
- [6] 王磊, 陈卫东. 面向自主机器人的实时彩色视觉系统[J]. 上海交通大学学报, 2007, 11(11): 1 645-1 648.
- [7] 王卫. 足球机器人彩色视觉图像的分割与识别[J]. 计算机工程与应用, 2005, 17(7): 68-71.
- [8] 艾海舟. 基于肤色和模板的人脸检测[J]. 软件学报, 2001, 12(12): 1 784-1 792.

(责任编辑 梁 健)