

类人足球机器人场上目标识别算法^①

李丽丽¹, 娄小平¹, 吕乃光¹, 林义闽²

¹(北京信息科技大学 仪器科学与光电工程学院, 北京 100192)

²(北京邮电大学 光通信与光电子学院, 北京 100876)

摘要: 在比赛过程中类人足球机器人的视觉系统需要对足球、球门以及对阵双方机器人进行识别. 考虑到算法的快速性及有效性, 采用基于颜色信息的算法对球及球门进行识别, 通过球及球门的颜色阈值提取图片中球与球门可能的位置, 再由球与球门的背景色或面积信息确定球与球门的正确位置. 对双方机器人的识别, 首先提取机器人的特征, 然后通过在线实时的监督学习方法训练一组级联分类器, 通过训练好的分类器对双方机器人进行检测. 实验表明算法能够快速有效地识别场上目标, 且算法具有较好的鲁棒性.

关键词: 类人足球机器人; 目标识别; 级联分类器

Objects Recognition Algorithm for Humanoid Robot Soccer on Playing Filed

LI Li-Li¹, LOU Xiao-Ping¹, LV Nai-Guang¹, LIN Yi-Min²

¹(School of Instrumentation Science and Photoelectric Engineering, Beijing Information Science and Technology University, Beijing 100192, China)

²(School of Optical Communication & Optoelectronics, Beijing University of Posts & Telecommunications, Beijing 100876, China)

Abstract: During the race, the football, goal and two sides of the robot need to be recognized by the vision system of humanoid robot soccer. Considering the rapidity and effectiveness of the algorithm, the ball and the goal are identified by the algorithm based on color information. The possible positions of the ball and the goal are extracted by the color thresholds of the ball and the goal in the picture, then the correct position is determined by the background color or area information. To recognize the two sides of the robot, first the features of the robot should be extracted and then through online real-time supervision learning method a cascade classifier is trained, which is used to identify robots. Result shows that the field objects can be recognized quickly and effectively and the algorithm has better robustness.

Keywords: humanoid robot soccer; objects recognition; cascade classifier

RoboCup 日趋得到人们的重视, 机器人若想取得比赛的胜利, 视觉系统对场上目标的识别工作是十分重要的, 为了使机器人更有效快速地识别目标, 类人组机器人比赛场地, 足球以及球门的颜色都已规定好, 因此, 很多参赛球队都采用基于色块的识别方法^[1-3]. 由于比赛现场环境较复杂, 光照的变化以及比赛场地各个位置反射光的差异, 使摄像头拍摄到的足球的颜色存在很大的差异, 因此颜色空间的选取很重要, 如果颜色空间选取不合适就很难提取各种颜色的属性值, 或者提取的值不准确, 这就会影响识别效果^[4].

另外, 由于场外观众及工作人员较多, 其服装的颜色可能对机器人识别足球产生干扰, 因此, 仅通过颜色分割识别足球及球门会产生无关目标, 如果加一些限定条件, 如足球场上的背景应该是绿色的球场, 图像面积阈值等, 则不满足限定条件的无关目标将被滤除.

球场上对阵双方机器人的形态在不断发生变化, 属于非刚体的识别, 霍夫森林[5]的方法得到广泛的应用, 由 Tim Cootes 提出的基于主动形状模型 (Active Shape Model, ASM) [6]的检测方法可以适应发生形变的目标检测, 本文采用在线训练分类器的方法,

① 基金项目: 北京市属高等学校人才强教计划(PXM2011_014224_113533); 北京信息科技大学研究生科技创新项目(5028223205)

收稿时间:2013-10-19;收到修改稿时间:2013-11-08

对机器人的特征模型进行不断地在线更新与学习, 然后通过分类器对对阵双方机器人进行识别。

1 颜色模型的选取

机器人的比赛现场通常用白色光源照明, 由于颜色在光照的影响下会发生很大变化, 且同一个场地的不同位置光照不同, 所以同一颜色的色标会在场地上发生很大范围的变化。因此, 颜色模型如果选择不合理, 会很难找到合适的颜色阈值分割图像, 识别目标也会变得很困难。

视觉系统的摄像机采集的图像一般是 RGB 颜色模型, 而 RGB 模型的三个颜色分量是关联变化的, 在光照条件发生变化, 照度和物体反射特性等不同条件下, R,G,B 三个分量的测量值分布分散, 很难确定 R,G,B 的颜色阈值范围, 造成误包含非指定的颜色目标或漏掉该识别的颜色目标。

RGB 颜色空间转变为 HSV 颜色空间, HSV 由色度 H (Hue), 饱和度 S (Saturation), 亮度 V (Value) 三个分量表示, 色度 H 表示不同的颜色, 饱和度 S 表示颜色的深浅, 亮度 V 表示颜色的明暗程度, 各分量的属性各不相关, H 能较准确的反应颜色种类, 对外界光照条件变化敏感度低。因此本文采用 HSV 颜色空间对图像做处理。

本文利用 HSV 空间的 H, S, V 分量进行足球的检测, 从 RGB 到 HSV 空间的变换很容易计算。将 R, G, B 归一化 $R, G, B \in [0, 1]$, 且将变换后的 H, S, V 归一化 $H, S, V \in [0, 1]$, 变换公式(1)-(4)如下:

$$V = \max(R, G, B) \quad (1)$$

$$\lambda = V - \min(R, G, B) \quad (2)$$

$$S = \begin{cases} 0 & V = 0 \\ \frac{V - \min(R, G, B)}{V} & V > 0 \end{cases} \quad (3)$$

$$\text{令 } r = \frac{V-R}{\lambda}, g = \frac{V-G}{\lambda}, b = \frac{V-B}{\lambda} \quad (4)$$

$$6H = \begin{cases} 1-g & V = R, \min(R, G, B) = B \\ 5+b & V = R, \min(R, G, B) = G \\ 3-b & V = G, \min(R, G, B) = R \\ 1+r & V = G, \min(R, G, B) = B \\ 5-r & V = B, \min(R, G, B) = G \\ 3+g & V = B, \min(R, G, B) = R \end{cases} \quad (5)$$

HSV 很容易计算, 而且是可逆的, 并且 HSV 空间能最大程度的满足颜色空间的均匀性, 自然性和完整性。

2 阈值分割图像

转换颜色模型后, 分析足球及球门的 H, S 和 V 各分量的直方图, 确定各分量的阈值, 通过阈值分割对图像进行二值化。

根据橙色球与黄色球门的 HSV 颜色模型中 H, S, V 的直方图如图 1, 找到各个分量的阈值, 并转换为 Opencv 环境下, 橙色足球三个分量的阈值范围为 (14,121,161) ~ (21,255,255), 黄色球门三个分量的阈值范围为 (25,100,150) ~ (35,255,255), 图像中各个像素与阈值进行比较, 在阈值范围内的像素作为前景, 其他像素为背景。

3 形态学处理

经阈值分割后的图像含有斑点噪声, 或者目标的最大连通域被噪声、阴影等分割成多个部分, 形态学处理中的腐蚀操作可以将斑点噪声腐蚀掉, 并且能够保证图像的较大连通区域依然存在, 而要找到的最大连通域则通常使用膨胀操作^[7]。膨胀是将图像与核进行卷积, 即计算核覆盖的图像区域的像素点的最大值, 并把这个最大值赋值给核的参考点处的像素, 这样就会使图像中的高亮区域逐渐增长, 这样可以找到目标的最大连通区域。腐蚀的操作和膨胀刚好相反, 腐蚀操作时, 核与图像卷积, 计算被核覆盖的图像区域的最小像素值, 并赋值给核的参考点, 这样一些斑点噪声就被腐蚀掉了。

4 足球和球门的识别

4.1 足球的识别

首先将摄像机得到的 RGB 图像转换到 HSV 空间, 根据橙色足球的 HSV 空间阈值对图像进行分割, 分割出的图像包括非目标区域和噪点, 对图像进行形态学处理, 腐蚀图像消除噪点, 膨胀图像得到足球的最大连通区域,



Software Technique • Algorithm 软件技术 • 算法 87

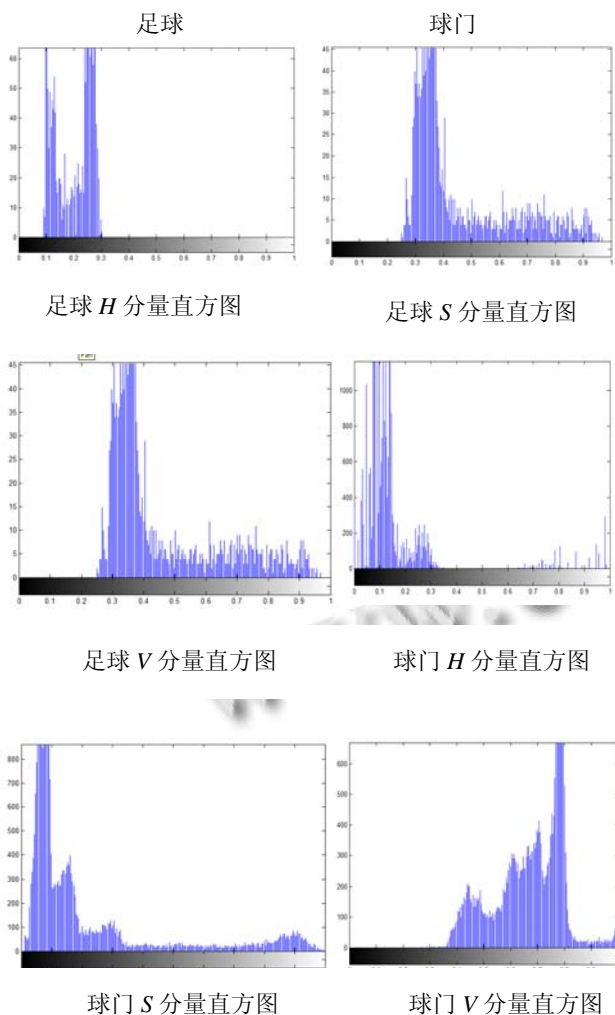


图 1 足球与球门颜色分量直方图

分析各连通区域的背景是否为绿色球场滤除非目标区域，如果在绿色球场的阈值范围内，则计算背景中一定区域内绿色区域所占的百分比，百分比超过 50%，则可判定该连通区域即为要识别的足球，并将其重心（绿色的“+”）计算出来，如图 3 所示。

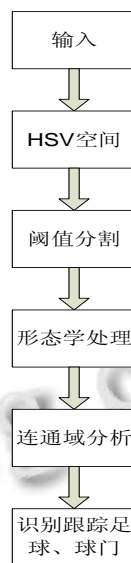
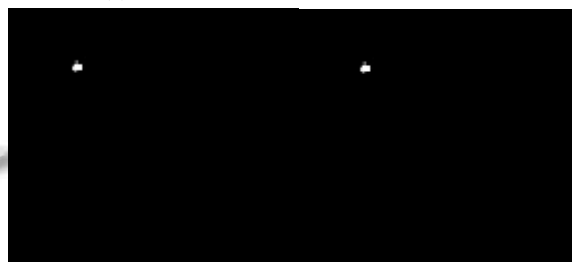


图 2 球与球门识别过程



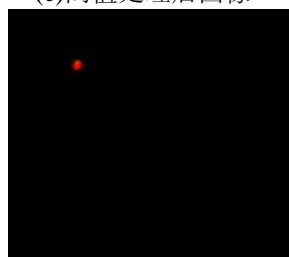
(a)原图

(b) HSV 空间图像



(c)阈值处理后图像

(d)形态学处理后的图像



(e)足球识别结果

图 3 足球的识别

4.2 球门的识别

首先将 RGB 图像转换到 HSV 图像空间, 对图像进行阈值分割, 依据要提取的球门是蓝色还是黄色 (以下以黄色球门为例) 设定不同的阈值, 图像被二值化. 二值化的图像中含有球门及在球门颜色阈值范围内的非球门物体, 接下来对二值化图像进行形态学处理, 先腐蚀, 将噪声滤除, 然后膨胀, 使球门区域形成连通区域. 对图像中的连通域进行分析, 对于球场上的球门及在球门颜色阈值范围内的非球门物体, 设定球门的连通域面积阈值, 在此阈值范围内的连通区域即为球门. 处理过程如图 4 所示.

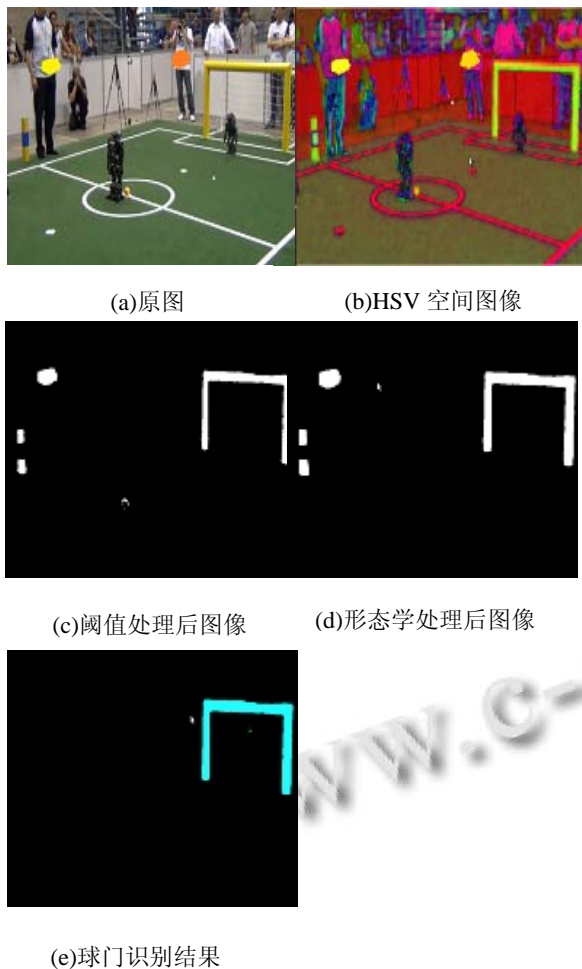


图 4 球门识别

图 3 和图 4 的结果表明足球和球门能够被准确地识别出来且实验过程中对算法的识别速度进行了测试, 每秒钟可以处理 15—20 帧图片.

5 对阵双方类人机器人识别

类机器人的姿态不断发生变化, 且时刻在球场移动, 背景在发生变化, 对于变化背景下的非刚体识别本文采用在线学习训练分类器的方法对机器人进行识别. 识别流程图如下:

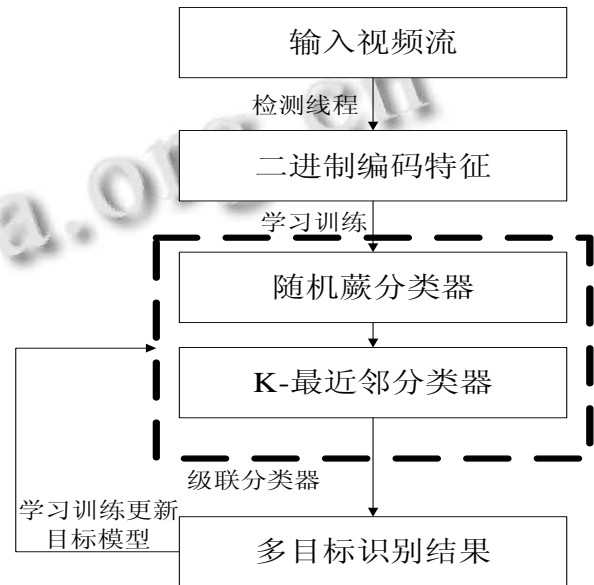


图 5 识别类人机器人过程框图

5.1 二进制编码特征提取

首先, 检测过程中需要尺度可变的扫描网格对待检测图片进行全面覆盖扫描, 以确定目标是否在图片中. 通过计算已经存储的目标特征与网格中提取出的特征的相似度来判断目标存在与否. 因此, 特征的提取至关重要. 本文首先将采集的图像转换到 HSV 空间, 然后提取出 H、S、V 通道的图像, 接着在每个扫描网格中随机提取 8 对点进行 H、S、V 值的比较, 比较结果用二进制码 0 和 1 表示, 这样每个网格提取出的特征就是 24 位二进制码. 特征描述如下:

$$f_1(X) = \begin{cases} 1 & H_x(i) > H_x(j), \\ 0 & H_x(i) \leq H_x(j) \end{cases} \quad (5)$$

$$f_2(X) = \begin{cases} 1 & S_x(i) > S_x(j), \\ 0 & S_x(i) \leq S_x(j) \end{cases} \quad (6)$$

$$f_3(X) = \begin{cases} 1 & V_x(i) > V_x(j), \\ 0 & V_x(i) \leq V_x(j) \end{cases} \quad (7)$$

i, j 代表网格中的点对, $f(X)$ 描述每对点的特征. 每一个网格提取出的图像特征为:

$$H \text{ 值特征 } F_1(X) = \{f_1^1, f_1^2, f_1^3 \dots f_1^8\}$$

$$S \text{ 值特征 } F_2(X) = \{f_2^1, f_2^2, f_2^3 \dots f_2^8\}$$

$$V \text{ 值特征 } F_3(X) = \{f_3^1, f_3^2, f_3^3 \dots f_3^8\}$$

不同尺度的 k_1, k_2, k_3 个扫描网格提取出的每个图像块的特征为

$$F(X) = \{F_1^1(X), \dots, F_1^{k_1}(X), F_2^1(X), \dots, F_2^{k_2}(X), F_3^1(X), \dots, F_3^{k_3}(X)\}$$

5.2 多分类器级联

分类器的设计采用基于随机森林的分类器^[8-9]与 K-最近邻分类器的级联.

随机森林由一系列的决策树组成, 每一棵决策树就是一个分类器. 待分类项的特征为 $F(x)$, 通过计算其后验概率值来判定其类别, 待分类项将被划为使其后验概率达到最大值的分类器 \hat{C} .

$$\hat{C}(F(x)) = \arg \max_k p(C_k | F(x)) \quad (8)$$

其中 $C_k, k=1, \dots, n$ 为各分类器的标记.

上式中 C 表示类的随机变量, 由贝叶斯公式可得

$$p(C_k | F(x)) = \frac{p(F(x) | C_k)p(C_k)}{p(F(x))} \quad (9)$$

先验概率 $p(C_k)$ 为均匀分布, (9) 式分母部分与类别无关, 则 (8) 式可简化为

$$\hat{C}(F(x)) = \arg \max_k p(F(x) | C_k) \quad (10)$$

由半朴素贝叶斯分类模型^[10], 将 $F(X) = \{F_1^1(X), \dots, F_1^{k_1}(X), F_2^1(X), \dots, F_2^{k_2}(X), F_3^1(X), \dots, F_3^{k_3}(X)\}$ 分成 M 组, 不同组之间的特征相互独立, 则 (10) 式的条件概率近似为

$$\hat{C}(F(x)) = \arg \max_k \prod_{M=1}^m p(F^M(x) | C_k) \quad (11)$$

由 (11) 即可得到待分类项的所属类别.

由训练好的随机森林分类器计算网格区域特征与目标模型的相似度, 即统计网格区域特征在随机森林分类器集合中的后验概率的和, 最终确定后验概率和较大的网格区域为检测目标的候选区域.

K-最近邻分类器, 对于待分类项根据欧氏距离找出它在训练集中的 K 个最近邻, 然后由这 K 个最近

邻中的大多数所属类别判定为待分类项的类别. 用训练好的最近邻分类器检测由随机森林输入的候选目标区域, 计算在线训练模型与候选区域的相似度, 找出相似度最大的候选区域, 作为最终被检测目标的位置.

$$D(p_i, p_j) = 0.5[NCC(p_i, p_j) + 1] \quad (12)$$

p_i, p_j 分别代表在线模型与候选区域, NCC 为归一化相关系数, 当 $D(p_i, p_j)$ 大于等于设定的相似度阈值 θ_{1MN} 时, 候选区域则被判定为目标区域.

5.3 学习训练分类器

在识别机器人的过程中, 它的姿态在不断发生变化, 需要对机器人特征进行在线学习不断更新分类器存储的目标模型, 保证准确识别机器人. 学习模块的关键是用已有的目标模型预测新的目标模型, 并通过检测结果的反馈修正被错误分类的目标模型, 以此不断更新目标模型^[11].

首先, 在第一帧图片中将要识别的对阵双方机器人的特征提取出来, 对目标区域进行旋转、平移、尺度变换, 并提取变换后的特征建立双方机器人各自的目标模型. 然后, 由目标模型对随机森林分类器和 K-最近邻分类器进行训练学习. 训练好的分类器在对目标进行检测过程中一方面根据检测结果去除目标模型中错误分类的样本, 一方面将新的样本加入目标模型, 不断重复此过程, 最终分类器的检测功能具有很高的鲁棒性.

5.4 实验结果及分析

实验过程中, 首先将算法应用到两段基准视频流, 如图(6)(a)所示, 实验将此算法与其他算法在同一段基准视频中的目标识别成功率进行对比, 算法包括 Fragments-based Tracker(FT)^[12], Online Boosting Tracker(OBT)^[13], Semi-Supervised Tracker(SST)^[14], Beyond Semi-Supervised Tracker(BSST)^[15], Tracking Learning Detection (TLD)^[16]. 如下图(6), 表(7)所示.

图 6 中(a)(b)为基准视频流, 红框内为感兴趣目标, (c)(d)中横坐标为视频帧数, 纵坐标为检测到的目标框与目标真实位置的重叠率, 当重叠率大于 50% 的时候则认为检测成功, 否则认为检测失败, (c)(d)图中不同颜色的曲线为不同算法在基准视频流中的检测结果, 由识别结果计算识别成功率, 见表 7.

目标定位的准确性由定位成功率来评价, 其定义为视频中目标位置与真实位置的面积覆盖率大于 50% 的

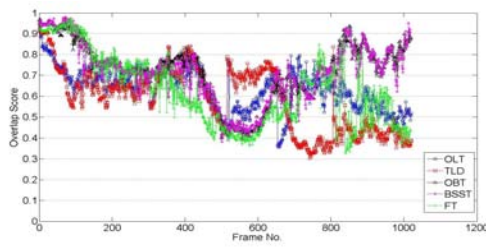
帧数除以总的视频帧数，其中 A 与 B 的面积覆盖率

如图 8 所示。

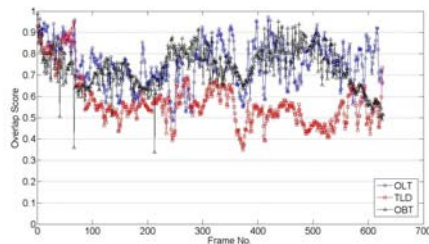


(a)基准视频流

(b)基准视频流



(c) a 视频流中不同算法的识别结果



(d) b 视频流中不同算法的识别结果

图 6 对比不同算法在 a 和 b 视频流中的识别结果

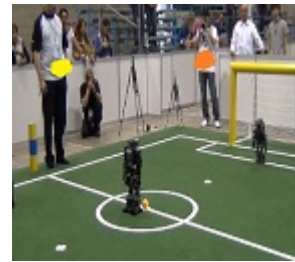
(Overlap Score, OS) 定义为

$$OS = \frac{A \cap B}{A \cup B} \quad (13)$$

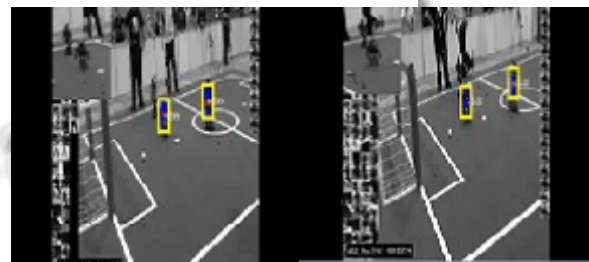
表 7 不同识别算法的成功率对比

方法	FT	BSST	SST	OBT	TLD	本文算法
a 视频	0.69	0.88	0.64	0.87	0.65	0.90
b 视频	0.15	0.40	0.47	1.0	0.76	1.0

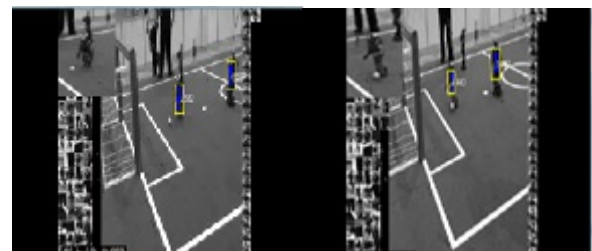
接着将算法应用到类人足球机器人在场上对对阵双方机器人的识别中，试验参数为 $k1 = k2 = k3 = 3$, $\theta_{KNN} = 0.6$, 识别结果如下图 8 所示，黄色框的位置代表算法检测到的目标位置，图中显示了从视频流中随机抽取四帧图片的识别结果，识别双方机器人的实验



(a) 足球场现场情况



(b)视频图像中的第 8 帧 (c)视频图像中的第 32 帧



(d)视频图像中的第 89 帧 (e)视频图像中的第 198 帧

图 8 双方机器人的识别

由以上实验结果可以看出本文的算法具备较高的准确率，在 a,b 基准视频流的实验中本文算法的识别成功率分别为 90%和 100%，明显高于其他算法，且本文算法在识别对阵双方机器人的实验中也得到了比较好的识别结果。究其原因，首先本文算法采用的级联分类器较单一的分类器检测效果更好，第一级分类器锁定目标的候选位置，第二级分类器从候选位置中确定目标的准确位置。其次本文算法中的分类器的训练方法采用在线学习训练的方法，保证分类器保存的目标模型能够随着机器人的移动不断更新，因此不会因为机器人的转身，四肢运动或摔倒而检测不到机器人的位置。综上所述，本文算法在目标识别过程中较其他算法具备较高的准确性。

6 结语

本文对类人机器人在比赛场上识别目标的算法进行了研究，球和球门的识别主要依赖颜色信息，识别

速度较快, 15-20 帧/秒, 且识别准确率较高 98.27%。双方机器人的识别算法较复杂, 达到了较高准确率 90%, 但是识别速度较慢 8-15 帧/秒, 在接下来的工作中需进一步提高算法的识别速度。

参考文献

- 1 卢丽,刘万春,刘俐.基于彩色图像分割的机器人足球目标识别.计算机工程与设计,2004,25(11): 2093-2095.
- 2 朱莹,洪炳镭,阮玉峰.全自主足球机器人快速目标识别与定位方法.哈尔滨工业大学学报,2003,35(9):1060-1063.
- 3 姜延,高庆吉,张利辉等.全自主型足球机器人目标识别与定位方法研究.东北电力学院学报,2002,22(3):12-16.
- 4 龚剑,方康玲.机器人足球视觉目标的快速识别.微计算机信息,2007,23(3-2):248-250.
- 5 Gall J, Lempitsky V. Class-specific hough forests for object detection. Decision Forests for Computer Vision and Medical Image Analysis. London. Springer, 2013: 143-157.
- 6 Cootes T, Baldock ER, Graham J. An introduction to active shape models. Image Processing and Analysis,2000:223-248.
- 7 布拉德斯基著,于仕琪译.学习 OpenCV.北京:清华大学出版社,2009:134-150.
- 8 Breiman L. Random forests. Machine Learning, 2001, 45(1):5-32.
- 9 Ozuysal M, Fua P, Lepetit V. Fast keypoint recognition in ten lines of code. Computer Vision and Pattern Recognition, 2007. CVPR'07 IEEE Conference. 2007. 1-8.
- 10 Zheng F, Webb GI. A comparative study of semi naive bayes methods in classification learning. Proc. of the Fourth Australasian Data Mining Conference, 2005. 141-156.
- 11 Vovk V, Gammerman A, Shafer G. Algorithmic Learning in a Random World. Springer, 2005:124-145.
- 12 Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference. 2006,1. 798-805.
- 13 Grabner H, Bischof H. On-line boosting and vision. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference. 2006,1. 260-267.
- 14 Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking. Computer Vision-ECCV2008, 2008: 234-247.
- 15 Stalder S, Grabner H, Van Gool L. Beyond semi-supervised tracking: Tracking should be as simple as detection, but not simpler than recognition. Computer Vision Workshops, 2009 IEEE 12th International Conference. 2009. 1.
- 16 Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409-1422.