

基于改进在线多示例学习算法的机器人目标跟踪

王丽佳^{1,2} 贾松敏¹ 李秀智¹ 王爽¹

摘要 提出基于改进的在线多示例学习算法 (Improved multiple instance learning, IMIL) 的机器人目标跟踪方法. 该方法利用射频识别系统 (Radio frequency identification, RFID) 粗定位 IMIL 算法的搜索区域, 然后应用 IMIL 算法实现目标跟踪. 该方法保证了机器人跟踪系统的连续性, 解决了目标突然转弯时的跟踪问题. IMIL 算法采用从低维空间提取的压缩特征描述包中示例, 以降低算法耗时. 通过最大化弱分类器与极大似然概率的内积, 选择判别能力强的弱分类器, 避免了弱分类器选择过程中多次计算包概率和示例概率, 进一步提高算法的实时处理能力. 计算包概率时该算法平等对待各示例, 保证概率高的示例对包概率的贡献度, 克服跟踪漂移问题. 跟踪过程中, 结合当前跟踪结果与目标模板间的相似性分数在线实时调整分类器, 提高了算法的自适应能力. 最后将本文方法在视频和移动机器人上进行实验. 实验结果表明, 该方法在目标运动突变及外观改变时具有较强的鲁棒性和准确性, 并满足系统的实时性要求.

关键词 改进的在线多示例学习, 目标跟踪, 射频识别系统, 压缩特征

引用格式 王丽佳, 贾松敏, 李秀智, 王爽. 基于改进在线多示例学习算法的机器人目标跟踪. 自动化学报, 2014, 40(12): 2916–2925

DOI 10.3724/SP.J.1004.2014.02916

Person Following for Mobile Robot Using Improved Multiple Instance Learning

WANG Li-Jia^{1,2} JIA Song-Min¹ LI Xiu-Zhi¹ WANG Shuang¹

Abstract An improved multiple instance learning (IMIL) algorithm is proposed for person following with a mobile robot. In the tracking process, radio frequency identification (RFID) provides a searching area for the IMIL algorithm, which then successfully detects the person of interest. In IMIL, compressed features are extracted to describe instances of bags from the low dimensional space so as to reduce the time complexity in the operation. Then, the most discriminative weak classifiers are selected from the weak classifier pool by maximizing the inner product between the weak classifier and the log-likelihood function. The scheme avoids computing the bag probability and instance probability many times, which further reduces the computational time. To deal with the drift problem, the bag probability equally depends on each instance. Furthermore, the classifiers are updated according to the similarity between the current tracking result and the target model, thus they can deal with appearance changes adaptively. The method is conducted on video sequences and a mobile robot. Experimental results demonstrate that the presented method can track the target accurately and robustly when there are abrupt motions and appearance changes, and satisfy the real-time requirement.

Key words Improved multiple instance learning (IMIL), person following, radio frequency identification (RFID), compressed features

Citation Wang Li-Jia, Jia Song-Min, Li Xiu-Zhi, Wang Shuang. Person following for a mobile robot using improved multiple instance learning. *Acta Automatica Sinica*, 2014, 40(12): 2916–2925

收稿日期 2013-08-19 录用日期 2014-07-16

Manuscript received August 19, 2013; accepted July 16, 2014
国家自然科学基金 (61175087, 61105033), 国家教育部留学回国人员
科研启动基金, 北京市自然科学基金重点项目 (KZ201110005004) 资
助

Supported by National Natural Science Foundation of China
(61175087, 61105033), Scientific Research Starting Foundation
for the Returned Overseas Chinese Scholars, Ministry of Educa-
tion of China, and the Key Program of Beijing Natural Science
Foundation (KZ201110005004)

本文责任编辑 孙富春

Recommended by Associate Editor SUN Fu-Chun

1. 北京工业大学电子信息与控制工程学院 北京 100124 2. 河北工
业职业技术学院信息工程与自动化系 石家庄 050091

1. College of Electronic Information and Control Engineering,
Beijing University of Technology, Beijing 100124 2. Depart-
ment of Information Engineering and Automation, Hebei College
of Industry and Technology, Shijiazhuang 050091

人跟踪技术是服务机器人领域备受关注的课题之一. 近年来, 许多国内外学者致力于机器人目标跟踪的研究, 并取得了一些引人瞩目的成果^[1–9].

视觉由于可以提供丰富的环境及任务信息而被广泛应用于机器人目标跟踪中^[1, 3–6]. Takashi 等^[1]设计了人跟踪机器人“ApriAttenda”, 该机器人装有立体视觉和超声波传感器, 其根据人体的区域信息、衣服的颜色和纹理信息识别目标. Yun 等^[3]利用装有彩色相机和激光传感器的移动机器人实现目标跟踪. 该机器人采用多层 Mean shift 算法跟踪目标, 然后利用激光传感器获得目标的距离信息. Bellotto 等^[4]结合从彩色相机获得的人脸、衣服颜色和身高等信息描述人体目标, 利用激光传感器检

测人体腿部以获得距离信息. Ozteta 等^[5] 设计装有立体视觉和激光传感器的移动机器人系统, 通过基于深度的直方图检测, 卡尔曼滤波和人体腿部检测等算法实现目标跟踪. 上述研究的前提是假设目标做平滑运动. 当存在运动突变时, 易发生目标跟丢现象. 此外, 视觉系统的视野较窄, 当目标突然转弯以致消失在相机视野外时, 易导致跟踪失败. 为了解决上述问题, Quadah 等^[6] 利用立体视觉和射频识别系统 (Radio frequency identification, RFID) 实现移动机器人的目标跟踪. RFID 系统由 8 根天线组成, 可以检测 360° 范围内携带标签的目标. 当目标运动突变时, RFID 仍能检测到目标, 引导机器人快速跟随目标.

此外, 目标的外观模型在目标跟踪过程中起着重要作用. 通常情况下, 目标模型可以分为两类: 静态外观模型和自适应外观模型^[10]. 静态外观模型在初始化阶段根据第一帧图像信息训练模型, 在跟踪过程中该模型保持不变^[11-13]. 当发生亮度改变、姿态变化、遮挡时, 采用此类模型的跟踪方法性能明显下降.

自适应外观模型方法将跟踪问题看成从背景中分离出目标的二值分类问题^[10]. 为了获得更好的跟踪性能, 该方法在跟踪过程中结合目标和背景信息实时更新描述目标模型的判别分类器^[14-16]. 目标模型更新中的关键问题是如何选取正、负样本. Grabner 等^[17] 提出采用在线 Boosting 算法学习判别分类器. 该方法将当前跟踪结果作为正样本, 在当前跟踪结果的邻域内采集负样本. 当跟踪结果是次最优时, 背景信息将以误差的形式被引入到目标模板更新过程中, 并造成跟踪漂移. 为解决这一问题, Zhang 等^[18] 提出实时压缩跟踪算法, 在当前跟踪结果的小邻域范围内采集多个正样本代替单个正样本用于更新目标模型. 然而, 该方法中多个样本无序组合, 使得判别分类器在目标模型更新过程中变得混乱, 从而影响跟踪性能. Grabner 等^[19] 提出半监督 Boosting 算法. 该方法在视频的初始帧标记样本, 跟踪过程中样本处于未标识状态. 该方法忽略了跟踪过程中的一些有用信息, 尤其是目标的运动信息. 同时, 该方法无法解决目标检测本身存在的歧义性问题.

近年来, 多示例学习算法被用于目标跟踪中, 并取得了一些引人瞩目的成果^[10, 20]. 在多示例学习中, 样本被表示为由多个示例组成的包^[21]. 每个包被标记为正或负. 若一个包中至少有一个示例为正, 则该包被标记为正包; 若包中所有示例均为负, 则该包标记为负包. 多示例学习的目的是通过对多个包组成的训练集进行学习得到判别分类器, 然后用该分类器尽可能准确地预测测试包. 在多示例学习算法中,

通过寻找包中最正确的示例确定跟踪目标, 克服了目标检测过程中存在的歧义性问题, 解决了由于目标外观改变造成的跟踪性能下降问题.

为了提高人体目标在运动和外观改变时的跟踪性能, 本文提出基于改进的在线多示例学习算法 (Improved multiple instance learning, IMIL) 的移动机器人目标跟踪方法. 其中, 移动机器人通过 RFID^[22] 和立体视觉传感器^[23] 获得目标位置信息. 本文主要工作如下:

1) 提出基于 IMIL 算法的移动机器人目标跟踪方法. 与传统的目标跟踪算法相比, 多示例学习算法在光照改变、姿态变化和遮挡等情况下的跟踪性能较好^[20]. 故本文利用改进多示例学习算法解决移动机器人目标跟踪系统中目标外观改变、遮挡等情况下的跟踪问题, 提高复杂环境下跟踪的稳定性和准确性.

2) 多示例学习算法存在实时性差, 分类器过更新等问题. 本文从自适应性和实时性两方面提高多示例学习算法的性能, 从而提高移动机器人跟踪目标的性能. 首先, 采用低维子空间提取的特征描述目标以减少计算耗时. 然后, 通过最大化弱分类器与极大似然概率内积的原则选择判别能力强的弱分类器, 避免多次计算示例概率和包概率, 进一步提高了算法的实时处理能力. 最后, 根据当前跟踪结果与目标模板间的相似性分数在线实时调整判别分类器, 适应跟踪过程中目标人体的外观改变.

3) 移动机器人跟踪目标时, 机器人和行人同时运动, 使得行人的帧间运动不规律. 而多示例学习算法假定行人的运动规律固定, 即在某一固定的候选目标搜索范围内搜索目标. 算法的跟踪性能受候选目标搜索范围影响较大. 搜索范围太小, 易导致目标跟丢; 搜索范围太大, 增加计算耗时. 为了解决这一问题, 本文利用 RFID 系统^[22] 获得目标的位置估计信息, 从而确定候选样本集搜索区域中心, 保证目标出现在候选区域. 然后在此候选区域应用 IMIL 算法实现目标的精确定位. 同时, RFID 的视野比立体视觉的视野宽. 当人体目标突然转弯以致消失在相机视野外时, 机器人将根据粗定位结果快速跟随人体目标.

4) 对本文提出的 IMIL 算法在经典视频中进行测试, 验证了该算法的鲁棒性和实时性. 然后对本文提出的移动机器人目标跟踪方法在 Pioneer3-DX 机器人上进行验证, 实验结果表明该方法对运动和外观改变时的人体目标跟踪有较强的鲁棒性.

本文内容安排如下: 第 1 节介绍 RFID 系统; 第 2 节介绍改进的在线多示例学习算法; 第 3 节简要说明基于 IMIL 的移动机器人目标跟踪方法的实现; 第 4 节对本文方法进行实验验证, 对比了 IMIL

算法与传统的多示例学习算法在不同视频上的跟踪性能. 并在 Pioneer3-DX 机器人上对本文提出方法进行了验证; 最后对全文工作进行总结.

1 RFID 系统

本文所采用的实验平台为 Pioneer3-DX 型移动机器人, 该机器人上装有立体视觉传感器和 RFID 系统, 如图 1 所示. RFID 系统由两根 8DBi 圆极化天线、读卡器和标签组成. 读卡器和天线安装在机器人上, 目标人体携带标签在机器人前方行走. 当携带标签的目标人体进入读卡器的工作范围时, 天线产生感应电流, 标签获得能量被激活并向读卡器发送自身编码等信息. RFID 系统依据读卡器和标签之间的信号到达时间延迟估计目标与机器人之间的距离, 根据每根天线获得的信号强弱判断目标方位, 最终完成目标的粗定位.

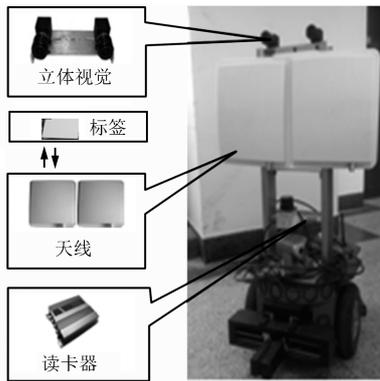


图 1 硬件平台

Fig. 1 The hardware platform

RFID 技术具有读写速度快、使用寿命长、可重复使用等特点, 可在非接触条件下实现移动目标识别^[22]. 本文将 RFID 技术用于移动机器人目标跟踪, 由其获得的目标粗定位结果辅助视觉系统完成目标的精确定位, 提高了目标运动突变时跟踪的稳定性, 并减少了后期图像处理过程中的计算耗时. 此外, 本文所采用的立体视觉传感器视野为 45° , RFID 天线的视野为 60° . 当目标快速转弯以致消失在相机视野外时, RFID 仍能检测到携带标签的目标, 为机器人提供导向信号, 使其快速跟随目标.

在 RFID 系统中, 读卡器在一帧内读取标签, 然后根据读取的观测值估计标签. 为了保证读卡器的读写速度, 读卡器在同一帧内只对标签读取一次. 在一般情况下一般的统计方法得到的估计值可靠性较差. 本文采用贝叶斯估计方法^[24] 实现标签的准确估计.

$$P_t(E|O) \propto P_t(O|E)P_{t-1}(E) \quad (1)$$

式中, $P_{t-1}(E)$ 为上一帧标签的估计值, $P_t(O|E)$ 为当前帧标签的观测值, $P_t(E|O)$ 为当前帧标签的估计值.

2 改进的在线多示例学习算法

2.1 在线多示例学习算法

在线多示例学习算法的思想是在 Online Boosting 框架下实现多示例学习算法^[10]. Online Boosting 算法是通过迭代弱分类器而产生最终的强分类器的算法. 该算法针对同一个训练集训练 M 个不同的弱分类器, 当存在一个测试样本时, 所有的弱分类器同时更新, 然后选择分类能力强的 K ($K \ll M$) 个弱分类器, 构成一个分类能力更强的强分类器. 各弱分类器在线调节自身参数, 分类正确率高的弱分类器具有较高的权重, 分类正确率低的弱分类器具有较低的权重, 进而适应时变的环境.

在多示例学习中, 训练集由带有标记的包组成, 即: $\{(\mathbf{X}_1, y_1), \dots, (\mathbf{X}_i, y_i), \dots, (\mathbf{X}_n, y_n)\}$, 其中 \mathbf{X}_i 为包, y_i 为包的标签. 一个包中含有多个示例, 即: $\mathbf{X}_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}\}$, 其中 x_{im} 为示例.

在线多示例学习算法^[10] 在当前帧目标位置的邻域内采样示例组成正包: $\mathbf{X}^r = \{x : \|l(x) - l_t^*\| < r\}$, 其中 l_t^* 为当前帧目标所在位置, r 为采集正示例的圆半径; 在以当前帧目标位置为圆心、内圆半径为 r , 外圆半径为 β 的圆环内采集示例组成负包: $\mathbf{X}^{r,\beta} = \{x : r < \|l(x) - l_t^*\| < \beta\}$.

根据正、负包中的示例学习弱分类器, 假设所有特征均服从高斯分布, 即 $p(v_k(x)|y=1) \sim N(\mu_k^1, \sigma_k^1)$, $p(v_k(x)|y=0) \sim N(\mu_k^0, \sigma_k^0)$, 且先验概率相等 $p(y=1) = p(y=0)$. 则弱分类器为

$$h_k(x) = \log \left(\frac{p(y=1|v_k(x))}{p(y=0|v_k(x))} \right) \quad (2)$$

式中, v_k 为示例的第 k 个特征, $(\mu_k^1, \sigma_k^1, \mu_k^0, \sigma_k^0)$ 为弱分类器的参数.

选择具有最大似然概率的弱分类器学习强分类器:

$$H = \sum_{k=1}^K h_k \quad (3)$$

式中, H 是由分类能力强的 K 个弱分类器构成的强分类器.

$$h_k = \arg \max_{h \in \Phi} \ell(H_{k-1} + h) \quad (4)$$

$\ell(H)$ 为所有包的极大似然概率:

$$\ell(H) = \sum_i \lg(p(y_i|\mathbf{X}_i)) \quad (5)$$

当 \mathbf{X}_i 为正包时, 包概率可由 Noisy-OR (NOR) 模型^[10] 表示:

$$p(y_i|\mathbf{X}_i) = 1 - \prod_{j=1}^N (1 - p(y_i|x_{ij})) \quad (6)$$

包中示例的概率为

$$p(y_i|x_{ij}) = \sigma(H(x_{ij})) \quad (7)$$

其中, $\sigma(x) = 1/(1 + e^{-x})$.

新一帧图像中, 在前一帧跟踪位置 l_{t-1}^* 的邻域内采集正样本构成候选样本集, 即: $\mathbf{X}^s = \{x : \|l(x) - l_{t-1}^*\| < s\}$. 强分类器在候选样本集中选择相似性分数最大的样本作为跟踪目标, 并在线更新弱分类器:

$$\begin{aligned} \mu_k^1 &= \lambda\mu_k^1 + (1 - \lambda)\mu^1 \\ \sigma_k^1 &= \sqrt{\lambda(\sigma_k^1)^2 + (1 - \lambda)(\sigma^1)^2 + \lambda(1 - \lambda)(\mu_k^1 - \mu^1)^2} \end{aligned} \quad (8)$$

跟踪过程中, 目标外观易受光照、姿态等因素影响而发生改变. 在线多示例学习算法能够实时更新分类器, 适应目标的外观改变, 改善跟踪性能. 同时, 在前一帧目标位置邻域内选择正示例组成正包, 当检测目标位置不准确, 或存在歧义性等问题时, 仍能准确跟踪目标. 然而, 该算法仍有一些缺点. 首先, 在搜索范围一定的情况下, 其实时性较差, 主要体现在两个方面: 1) 该算法采用维数较大的 Haar-like 特征描述目标, 增加了算法计算耗时; 2) 该算法通过对包概率和示例概率迭代学习 M 次得到弱分类器, 降低了算法的实时处理能力. 其次, 分类器更新过程中引入了更新程度参数 λ . 在实际应用中, 该参数需要预先给定, 其值过小易造成分类器“过学习”, 太大则无法适应目标外观的改变.

2.2 改进的在线多示例学习算法

本文在实时性和自适应性两方面对在线多示例学习算法^[10] 进行改进. 首先, 分别在目标描述和弱分类器选择两方面提高算法的实时处理能力. 然后, 在线实时更新分类器以提高算法的自适应能力.

2.2.1 目标描述

传统的多示例学习算法及其改进算法采用维数较大的 Haar-like 特征描述目标^[10, 20], 使算法的实时处理能力较差. 为了满足算法的实时性要求, 改进的多示例学习算法采用特征维数较小的压缩特征描述目标.

压缩感知理论是近年来诞生的一种新的信号处理理论^[18]. 该理论指出: 只要信号是可压缩的或在

某个变换域是稀疏的, 那么就可以用一个观测矩阵将其从高维空间投影到一个低维空间上. 自然图像具有可压缩的性质, 故 Zhang 等^[18] 采用一个满足有限等距性质 (Restricted isometry property, RIP) 的随机测量矩阵, 将原图像特征从一个高维空间投影到低维子空间. 该方法既保留了足够的信息描述目标, 又减少了计算时间.

随机测量矩阵 $R^{M \times Q}$ 中各元素 r_{ij} 为

$$r_{ij} = \sqrt{s} \begin{cases} 1, & \text{with probability } \frac{1}{2s} \\ 0, & \text{with probability } 1 - \frac{1}{s} \\ -1, & \text{with probability } \frac{1}{2s} \end{cases} \quad (9)$$

式中, s 以平均概率在 $2 \sim 4$ 中随机选取.

对于每个示例, 在图像平面提取其 Haar-like 特征 $\mathbf{Z} = (z_1, z_2, \dots, z_Q)$. 然后利用随机测量矩阵对其降维, 得到低维子空间的压缩特征 $\mathbf{V} = (v_1, v_2, \dots, v_M)$, $M \ll Q$, 如图 2 所示.

$$\mathbf{V} = \mathbf{RZ} \quad (10)$$

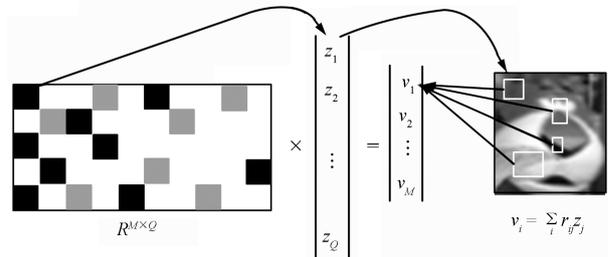


图 2 压缩特征的提取方法

Fig. 2 The method of extracting compressed features

由压缩感知原理可知, 一个满足 RIP 条件的稀疏随机测量矩阵不会把两个不同的稀疏信号映射到同一个采样几何中. 根据该方法提取的压缩特征可分性强, 适合在复杂环境下描述目标.

2.2.2 分类器学习

MIL 算法通过最大化所有包的极大似然概率学习分类器, 同时采用 Noisy-OR 模型计算包概率^[10]. 故在弱分类器的选择过程中, 需要计算 M 次示例概率和包概率, 增加了计算耗时. 为提高算法的实时处理能力, Zhang 等^[20] 提出加权多示例学习算法 (Weighted multiple instance learning, WMIL). 该算法采用最大化弱分类器与极大似然概率的内积原则, 在弱分类器池中选择具有最低错误率的 K 个弱分类器构造强分类器.

$$h_k = \arg \max_{h \in \Phi} \langle h, \nabla L(H) \rangle |_{H=H_{k-1}} \quad (11)$$

式中, $\langle h, \nabla L(H) \rangle = \frac{1}{N+L} \sum_{j=0}^{N+L-1} h \nabla L(H)(x_{ij})$ 为弱分类器与示例极大似然概率的内积.

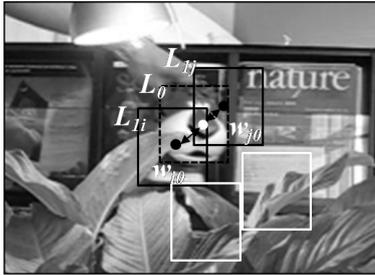
WMIL 算法在包概率的计算过程中, 引入权重以表示示例的重要性程度. 该权重根据当前示例位置与目标位置 L_i^* 之间的距离确定, 则包概率为

$$p(y|\mathbf{X}_i) = \sum_{j=0}^{N-1} w_{j0} p(y=1|x_{ij}) \quad (12)$$

$$w_{j0} = \frac{1}{c} e^{-|L(x_{ij})-L(x_{i0})|} \quad (13)$$

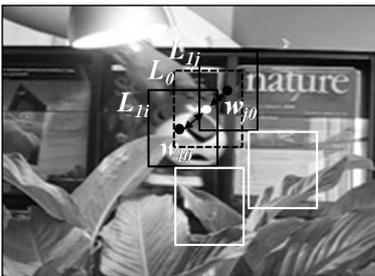
式中, $L(\cdot)$ 为示例的质心位置, c 为归一化常数.

该方法根据各示例的重要性计算包概率: 距离跟踪目标位置近的示例在包概率中所占权重大, 距离跟踪目标位置远的示例所占权重小. 如图 3(a) 所示, 位置 L_{1i} 处的示例与目标 (位置 L_0 处) 的距离大于位置 L_{1j} 处的示例与目标的距离, 故相应的权重 w_{i0} 小于 w_{j0} . 在准确跟踪目标的情况下, 示例概率大小与其权值成正比, 该方法提高了包概率计算的准确性. 然而, 当发生跟踪漂移时, 实际目标位置距离跟踪目标较远, 在 WMIL 算法中其在包概率中的权重较小. 如图 3(b) 所示, 跟踪目标的位置 L_0 偏离了目标的实际位置, 位置 L_{1i} 处的示例与实际目标相似度高于位置 L_{1j} 处的示例. 然而 L_{1i} 处的示例与跟踪目标的距离却比 L_{1j} 处的示例远, 故相应的权重 w_{i0} 小于 w_{j0} . 使得在包概率中 L_{1i} 处的示例重要性程度低于 L_{1j} 处的示例, 从而影响了跟踪性能.



(a) WMIL 构造原理

(a) The principle of the WMIL



(b) WMIL 的缺点

(b) The shortcoming of the WMIL

图 3 WMIL 示意图

Fig. 3 The illustration of the WMIL

为解决上述问题, 本文在计算包概率时平等对待各示例^[25].

$$p(y|\mathbf{X}_i) = \sum_{j=0}^{N-1} p(y=1|x_{ij}) \quad (14)$$

该方法中各示例概率的权重为 1, 包概率只与示例概率大小有关. 概率大的示例对包概率贡献大, 概率小的示例对包概率贡献较小. 当能准确跟踪目标时 (如图 3(a) 所示), 距离 L_0 较近的示例 L_{1j} 的概率大于示例 L_{1i} 的概率, 故示例 L_{1j} 对包概率的贡献较大. 当跟踪结果出现漂移时 (如图 3(b) 所示), 距离 L_0 较远的示例 L_{1i} 的概率较大, 其对包概率的贡献大于概率值小的示例 L_{1j} , 从而解决了 WMIL 算法中的跟踪漂移问题.

式 (11) 中极大似然概率 $\nabla L(H)(x_{ij})$ 为

$$\begin{aligned} \nabla L(H)(x_{ij}) &= \frac{\partial(H + \delta y_i)}{\partial \delta} \Big|_{\delta=0} = \\ &= \frac{\partial}{\partial \delta} \sum_{s=0}^1 \left(y_s \log \left(\sum_{m=0}^{N-1} \sigma(H(x_{1m}) + \delta y_i) \right) - (1 - y_s) \times \right. \\ &\quad \left. \log \left(\sum_{m=N}^{N+L-1} (1 - \sigma(H(x_{1m}) + \delta y_i)) \right) \right) \Big|_{\delta=0} = \\ &= y_i \frac{\sigma(H(x_{ij}))(1 - \sigma(H(x_{ij})))}{\sum_{m=0}^{N-1} \sigma(H(x_{im}))} - \\ &\quad (1 - y_i) \frac{\sigma(H(x_{ij}))(1 - \sigma(H(x_{ij})))}{\sum_{m=N}^{N+L-1} \sigma(1 - H(x_{im}))} \end{aligned} \quad (15)$$

根据分类器选择原则 (式 (11)), 在分类器池中选取判别能力强的 K 个弱分类器组成强分类器, 用于目标跟踪. 本文方法避免了在选择弱分类器前对示例概率和包概率计算 M 次, 节省了运算时间. 同时, 通过平等对待包中示例概率以得到包概率, 提高了发生漂移时的跟踪性能.

2.2.3 分类器更新

分类器检测到目标位置后, MIL 算法^[10] 按照式 (8) 以固定的参数 λ 更新分类器. 然而, λ 的选择限制了算法的应用. 当发生目标跟丢及跟踪漂移时, 如果 λ 太小, 易引起分类器过更新. 如果 λ 太大, 分类器更新过程无法适应光照、姿态等造成的目标外观改变.

本文根据当前跟踪结果与目标模板的相似性分数 H_{\max} 在线自适应调整分类器更新程度参数 λ . 设定高、低两个阈值, TH_s 为低阈值, TH_l 为高阈值. 1) 当光照或位姿改变等造成目标外观缓慢变化时, 当前跟踪结果与目标模型相似性分数大于高阈值.

此时, 分类器的更新主要取决于当前跟踪结果, 以克服外观改变对跟踪结果的影响. 2) 当目标被严重遮挡、发生漂移或误跟踪时, 当前跟踪结果与目标模型相似性分数小于低阈值. 此时, 分类器的更新主要依赖于目标模型, 以避免对目标模板的“过学习”. 3) 当前跟踪结果与目标模型的相似性分数在两个阈值之间时, 分类器同时依赖于当前跟踪结果和目标模型. 则式 (8) 的分类器更新规则中的参数 λ 按下式调整:

$$\lambda = \begin{cases} \lambda_s, & H_{\max} > TH_t \\ \lambda_m, & TH_s < H_{\max} < TH_t \\ \lambda_l, & H_{\max} < TH_s \end{cases} \quad (16)$$

2.2.4 算法实现

本文提出的 IMIL 算法用于视频目标跟踪, 该算法主要包括初始化、提取正负包、学习分类器、跟踪目标和更新分类器 5 部分. 算法实现的具体流程如下:

- 1) 初始化: 选取跟踪目标, 生成随机测量矩阵 R .
- 2) 在当前帧目标邻域采样正、负包, 利用式 (2) 学习包中示例得到 M 个弱分类器, 组成弱分类器池 $\Phi = \{h_1, \dots, h_M\}$.
- 3) 初始化强分类器 $H_0 = 0$, 令 $k = 1$.
- 4) 若 $k \leq K$, 由式 (15) 得 $\nabla L(H)(x_{ij})|_{H=H_{k-1}}$, 式中, $H_{k-1} = \sum_{m=1}^{k-1} h_m$ 为在弱分类器池 Φ 中选取的 $k-1$ 个最佳弱分类所构成强分类器. 否则转到步骤 11).
- 5) 初始化 $L^m = 0$, 其中 $m = 1, \dots, M$.
- 6) 令 $m = 1$.
- 7) 当 $m \leq M$ 时, 令 $i = 0$; 否则转到步骤 10).
- 8) 当 $i \leq 1$ 时, 令 $j = 0$; 否则令 $m = m + 1$, 并转到步骤 7).
- 9) 当 $j \leq N + L - 1$ 时, 循环计算 $j = j + 1$, $L^m \leftarrow L^m + h_m \nabla L(H)(x_{ij})|_{H=H_{k-1}}$; 否则令 $i = i + 1$, 并转到步骤 8).
- 10) 计算 $m^* = \arg \max_m L^m$, $h_k(x) \leftarrow h_{m^*}(x)$, $H_k = H_{k-1} + h_k(x)$. 令 $k = k + 1$, 并转到步骤 4).
- 11) 输出强分类器: $H(x) = \sum_k h_k(x)$.

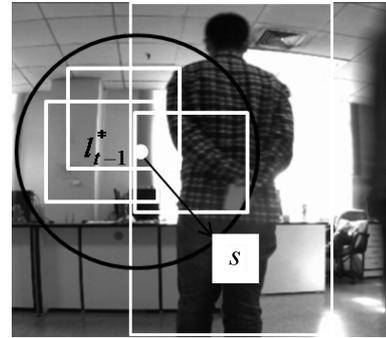
3 机器人目标跟踪

移动机器人目标跟踪系统中, 当目标运动突变或外观发生改变时跟踪性能显著下降. 为解决这一问题, 本文基于 RFID 和立体视觉传感器, 采用 IMIL 算法实现移动机器人目标跟踪. RFID 系统用于目标的粗定位, 立体视觉传感器提供包含场景三维信息的视差图和彩色图像^[26]. 在视差图中获得目

标的三维位置信息, 并发送给机器人控制器^[27] 控制机器人跟随目标.

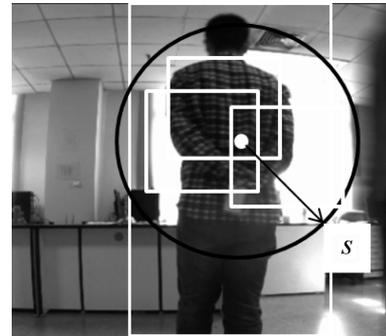
首先, 在机器人立体视觉传感器获得的彩色图像中手动框选目标, 完成 IMIL 算法的初始化和分类器的学习. 系统进入目标跟踪阶段, RFID 系统识别携带标签的目标完成粗定位. 在彩色图像中, 采用 IMIL 算法完成目标的精确定位. 最后, 在视差图中获得目标位置的三维坐标, 控制机器人跟随目标. 当目标消失在相机视野之外时, RFID 系统检测到携带标签的目标, 为机器人提供导向, 使其快速跟踪目标. 当目标被遮挡时, RFID 和立体视觉传感器均失效, 系统利用扩展卡尔曼滤波器^[25] 预测目标位置.

在 MIL 算法^[10] 的候选样本集的采集过程中, 目标运动突然变化时, 易发生候选样本集的搜索区域与目标区域交集过小, 从而造成跟踪漂移或目标跟丢, 如图 4(a) 所示. 为解决这一问题, 本文将 RFID 的粗定位结果作为 IMIL 的搜索区域采集候选样本, 以保证目标出现在候选样本集的搜索区域, 使机器人有效跟踪目标, 如图 4(b) 所示.



(a) MIL 中候选样本集的搜索区域

(a) The illustration of the searching area for sample set in MIL



(b) IMIL 中候选样本集的搜索区域

(b) The illustration of the searching area for sample set in IMIL

图 4 候选样本集搜索区域示意图

Fig. 4 The illustration of the searching area for sample set

4 实验结果及分析

4.1 视频目标跟踪结果

为验证本文提出的 IMIL 算法在视频目标跟踪

中的有效性,对 MIL^[10]、WMIL^[20]、IMIL 三种算法进行了比较. 其中 IMIL 算法利用第 2.2.1 节中的方法提取目标特征. 目标初始位置由手动给定. 各算法的参数设定保证其能达到最优跟踪性能. 在 MIL 中, $S = 35$, $r = 4$, $\beta = 50$, $M = 250$, $K = 50$, $\lambda = 0.85$; 在 WMIL 中, $S = 25$, $r = 4$, $\beta = 37.5$, $M = 150$, $K = 50$, $\lambda = 0.85$; 在 IMIL 中, $S = 25$, $r = 4$, $\beta = 37.5$, $M = 120$, $K = 30$, $\lambda_s = 0.05$, $\lambda_m = 0.25$, $\lambda_l = 0.85$. 本文从三个方面对跟踪结果进行了对比分析: 视觉对比、跟踪错误率、平均处理时间. 其中, 目标跟踪错误率指跟踪失败的帧数与总帧数的比率. 当跟踪到的目标区域与目标实际区域的重合度小于 50% 时, 该帧为跟踪失败帧. 其数学表达式为

$$F_{rate} = \frac{\text{Count}(\text{Overlap} < 50\%)}{\text{Frames}} \quad (17)$$

式中, $\text{Overlap} = \frac{\text{num}(R_S \cap R_T)}{\text{num}(R_T)}$ 是当前跟踪到的目标区域与实际目标区域的重合度. R_T, R_S 分别表示真实目标区域和实际跟踪到的目标区域. $\text{num}(\cdot)$ 为计算区域内的像素点个数. Count 为计算重合度小于 50% 的帧数. Frames 为总帧数.

实验 1. 在经典视频序列“Cliff bar”上进行目标跟踪. 该视频中目标尺寸不断变化, 背景与目标纹理相近, 特征区分度较低. 视频中存在目标突然运动, 在第 156 帧目标突然翻转, 在第 332 帧目标突然移动到另一位置. 实验结果如图 5 所示. MIL 算法的跟踪性能最差, 当目标发生旋转或突然快速移动时目标丢失; WMIL 算法在目标突然快速移动时跟踪效果不佳; 本文提出的 IMIL 算法在处理目标做机动时跟踪性能优于 MIL 和 WMIL 方法. 同时, 实验结果表明本文所采用的目标描述方法区分性强, 适合复杂背景下的目标描述.

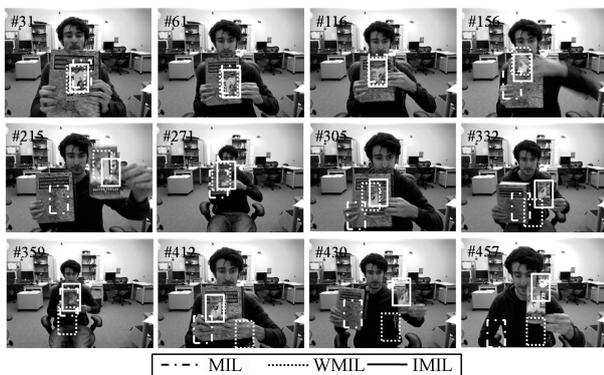


图 5 在视频序列“Cliff bar”上的目标跟踪结果

Fig. 5 The illustration of the tracking object location for the tested sequence “Cliff bar”

实验 2. 在经典视频序列“Tiger2”上的跟踪.

该视频中存在光照改变、姿态变化和突然运动. 实验结果如图 6 所示, MIL 算法跟踪性能最差, WMIL 算法在第 221 帧至 268 帧跟丢目标, 本文算法实现了目标的准确跟踪.

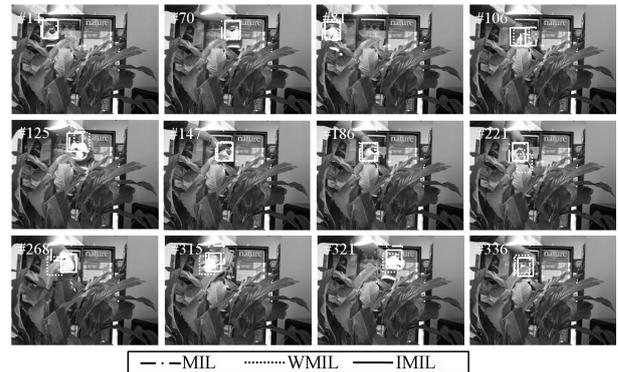


图 6 在视频序列“Tiger2”上的目标跟踪结果

Fig. 6 The illustration of the tracking object location for the tested sequence “Tiger2”

表 1 对比了本文算法与其他两种算法的跟踪错误率. 与其他两种算法相比, 本文算法具有较高的跟踪精度. 同时, 对三种方法的平均运算时间进行了对比, 如表 2 所示 (采用 Matlab R2009A 编程, 硬件环境为 Intel Core 2 Duo CPU, 2.33 GHz×2, 1.97 GB 内存). 结果表明本文算法降低了计算耗时. 综上所述, 本文算法具有较高的跟踪精度, 跟踪过程相对稳定, 且满足目标跟踪的实时性要求.

表 1 MIL、WMIL 和 IMIL 算法的跟踪错误率 (%)

Table 1 Failure rate (FR) (%) for MIL, WMIL, and IMIL

视频	MIL	WMIL	IMIL
Cliff bar	58.75	45.54	40.55
Tiger2	19.83	8.68	4.13

表 2 MIL、WMIL 和 IMIL 算法的运算时间 (ms)

Table 2 The computing time for MIL, WMIL, and IMIL (ms)

视频	MIL	WMIL	IMIL
Cliff bar	0.531	0.098	0.079
Tiger2	0.574	0.101	0.088

4.2 机器人目标跟踪

本文所提方法在 Pioneer3-DX 型移动机器人上验证, 用 VS2009 和 Openvc2.1 实现. 为了说明本文方法的有效性, 对 IMIL 算法与 MIL, WMIL 算法进行对比实验. 跟踪过程中, RFID 系统检测携带

标签的目标人体, 确定 IMIL (MIL, WMIL) 算法的候选区域中心. 然后, 分别利用 IMIL, MIL, WMIL 算法对目标进行跟踪. 最后根据立体视觉技术, 在视差图中计算所得目标位置的三维信息, 并发送给机器人控制系统^[26], 引导机器人跟随目标.

在移动机器人跟踪目标过程中, 行人和机器人的同时运动使得行人的帧间运动变化较大. 从而极易造成行人在当前帧图像中的位置超出上一帧图像中跟踪位置的搜索范围而使跟踪失败. 本文通过 RFID 系统对目标粗定位, 确定目标的候选位置. 然后, 在确定的搜索半径下, 利用 IMIL, MIL, WMIL 算法对目标精确定位. RFID 的引入, 保证了 IMIL, MIL, WMIL 算法的搜索范围与目标实际位置有交集, 从而提高目标跟踪的准确性.

由第 4.1 节实验可知, MIL 算法的实时处理能力很差, 而机器人目标跟踪过程中为了保证跟踪的连续性, 对算法的实时性要求很高. 跟踪开始, 由于 MIL 算法计算耗时, 机器人没有跟踪到目标时, 目标已经走出了 RFID 和立体视觉的视野, 使跟踪失败.

采用 WMIL 算法时, 示例概率在包概率中所占比重与其位置有关. 当发生偏移时, 正确示例距离中心位置较远使得其在包概率中所占比重减小, 从而造成偏移加剧. 虽然 RFID 的引入保证了候选区域内存在目标, 然而 RFID 对目标定位的可靠性不是很高, 会发生目标偏移候选区域中间位置的情况, 如图 7(b) 和 (c). 此时, 离中间近的示例概率在包中所占比重大. 而真正目标位置由于其离中间较远, 其概率在包中的比重被削弱, 从而造成跟踪偏移进而使跟踪失败. 同时, 在跟踪偏移和目标转弯的共同作用下, 目标极易走出 RFID 视野, 使跟踪失败, 如图 7(d).

在 IMIL 算法中, 示例概率在包中所占比重只与其自身概率有关, 概率大的示例其所占比重大, 概率小的示例其所占比重小, 解决了跟踪偏移问题. 另一方面 IMIL 算法采用自适应的模板更新策略, 以适应外观改变、遮挡等复杂情况. 采用 IMIL 算法的跟踪结果如图 8 和 9 所示. 图中第一行为机器人视角下的跟踪结果, 其中白色矩形框表示 RFID 系统检测结果, 深色矩形框为最终跟踪结果. 第二行为机器人目标跟踪的全局结果. 采用 IMIL 算法进行目标跟踪, 当目标人体被非目标人体完全遮挡时 (如图 8(c)), 系统无法检测到 ID 标签信号, 机器人根据扩展卡尔曼滤波器的预测结果跟踪目标. 遮挡结束后, 机器人重新跟踪到目标. 目标突然转弯的跟踪结果如图 9(c), 在跟踪过程中, 实际目标偏移 RFID 检测到的候选区域中间时, 采用 IMIL 算法仍能准确跟踪目标.

表 3 对比了在机器人目标跟踪系统中采用本文算法与 WMIL 算法的跟踪错误率和处理时间. 由实验结果可知, 采用本文方法进行机器人目标跟踪降低了跟踪错误率和处理时间.

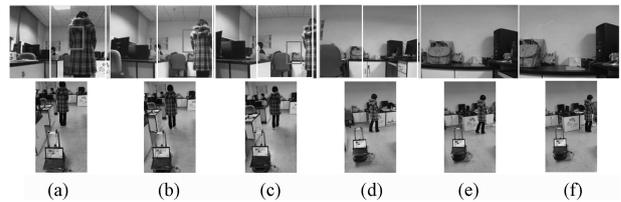


图 7 采用 WMIL 算法时, 移动机器人跟踪结果
Fig. 7 The illustration of person tracking for a robot using WMIL algorithm

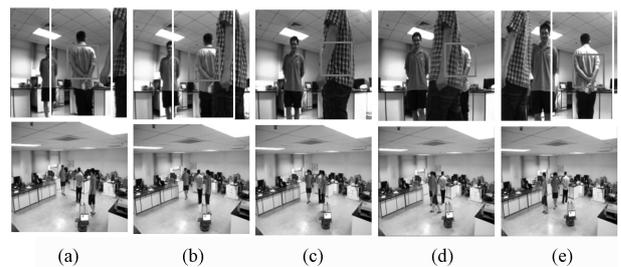


图 8 目标被遮挡时, 采用 IMIL 算法的跟踪结果
Fig. 8 The illustration of person tracking for a robot using IMIL algorithm when there is an occlusion

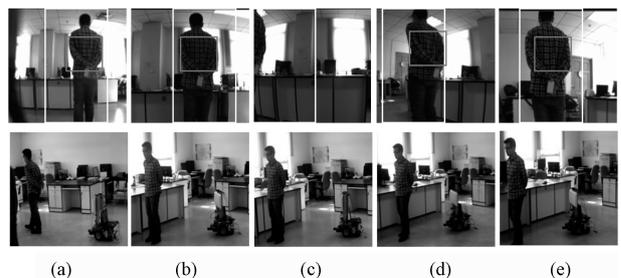


图 9 目标转弯时, 采用 IMIL 算法的跟踪结果
Fig. 9 The illustration of person tracking for a robot using IMIL algorithm when there is a sudden turn

表 3 移动机器人跟踪系统中采用 WMIL 和 IMIL 算法的跟踪错误率 (%) 和处理时间 (ms)

Table 3 Failure rate (FR) (%) and calculational cost (ms) for WMIL and IMIL using a mobile robot

	FR (%)	Calculational cost (ms)
WMIL	28.68	169
IMIL	3.23	144

5 结论

本文将多示例学习引入到机器人目标跟踪系统, 并结合 RFID 实现移动机器人的目标跟踪. 本方法具有以下几个特点:

1) 根据 RFID 粗定位结果确定 IMIL 算法的候选样本集搜索区域, 解决由于机器人和行人同时运动造成的目标运动不规律问题. 此外, 当目标消失在相机视野外时, 机器人根据 RFID 提供的定位信息快速跟随目标, 提高了跟踪性能.

2) 提出 IMIL 算法, 从目标特征表达和弱分类器选择两个方面改进 MIL 算法, 以提高算法实时处理能力. 首先, 利用随机测量矩阵对图像特征进行降维, 保留了足够的信息描述目标, 同时减少了运算时间. 然后, 通过最大化弱分类器与极大似然概率间的内积, 选择分类能力强的弱分类器以构成强分类器, 避免多次计算示例概率和包概率, 进一步提高算法的实时处理能力.

3) IMIL 算法在包概率计算时平等对待各示例, 解决了跟踪漂移问题.

4) 根据当前跟踪结果与目标模板间的相似度在线实时更新分类器, 解决了光照或位姿变化造成的外观改变.

实验结果表明本文方法提高了目标人体运动突变及外观改变时的跟踪鲁棒性和稳定性, 同时降低了计算耗时. 接下来的工作中, 将致力于解决长时间完全遮挡时的跟踪问题.

References

- 1 Takashi Y, Nishiyama M, Sonoura T, Nakamoto H, Tokura S, Sato H, Ozaki F, Matsuhira N, Mizoguchi H. Development of a person following robot with vision based target detection. In: Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. Beijing, China: IEEE, 2006. 5286–5291
- 2 Takemura H, Zentaro N, Mizoguchi H. Development of vision based person following module for mobile robots in/out door environment. In: Proceedings of the 2009 IEEE International Conference on Robotics and Biomimetics. Guilin, China: IEEE, 2009. 1675–1680
- 3 Yun W H, Kim D, Lee J. Person following with obstacle avoidance based on multi-layered mean shift and force field method. In: Proceedings of the 2010 IEEE International Conference on Systems. Istanbul: IEEE, 2010. 3813–3816
- 4 Bellotto N, Hu H. Multimodal perception and recognition of humans with a mobile service robot. In: Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics. Piscataway, USA: IEEE, 2008. 401–406
- 5 Martínez-Otzeta J M, Iburguren A, Ansuategi A, Tubío C, Aristondo J. People following behaviour in an industrial environment using laser and stereo camera. In: Proceedings of the 23rd International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems. Berlin, Heidelberg: Springer-Verlag, 2010. 508–517
- 6 Quadah N, Cadenat V, Lerasle F, Hamerlain M, Germa T, Boudjema F. Multi-sensor-based control strategy for initiating and maintaining interaction between a robot and a human. *Advanced Robotics*, 2011, **25**(9–10): 1249–1270
- 7 Andreas E, Bastian L, Konrad S, Luc van G. Robust multi-person tracking from a mobile platform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(10): 1831–1846
- 8 Choi W, Pantofaru C, Silvio S. A general framework for tracking multiple people from a moving camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(7): 1577–1591
- 9 Mekonnen A A, Lerasle F, Herbulot A. Cooperative passers-by tracking with a mobile robot and external cameras. *Computer Vision and Image Understanding*, 2013, **117**(10): 1229–1244
- 10 Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(8): 1619–1632
- 11 Wang Meng, Dai Ya-Ping, Wang Qing-Lin. A novel FAST-Snake object tracking approach. *Acta Automatica Sinica*, 2014, **40**(6): 1108–1115
(王蒙, 戴亚平, 王庆林. 一种新的 FAST-Snake 目标跟踪方法. 自动化学报, 2014, **40**(6): 1108–1115)
- 12 Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift. In: Proceedings of the 2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hilton Head Island, SC, USA: IEEE, 2000. 142–149
- 13 Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2006. 798–805
- 14 Liu X M, Yu T. Gradient feature selection for online boosting. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, USA: IEEE, 2007. 1–8
- 15 Avidan S. Ensemble tracking. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 494–501
- 16 Collins R T, Liu Y X, Leordeanu M. Online selection of discriminative tracking features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1631–1643

- 17 Grabner H, Grabner M, Bischof H. Real-time tracking via on-line boosting. In: Proceedings of the 17th British Machine Vision Conference. Guildford: BMVA Press, 2006. 47–56
- 18 Zhang K H, Zhang L, Yang M H. Real-time compressive tracking. In: Proceedings of the 2012 European Conference on Computer Vision, Berlin, Heidelberg: Springer-Verlag, 2012. 866–879
- 19 Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking. In: Proceedings of the 10th European Conference on Computer Vision. Berlin, Heidelberg: Springer-Verlag, 2008. 234–247
- 20 Zhang K H, Song H H. Real-time visual tracking via online weighted multiple instance learning. *Pattern Recognition*, 2013, **46**(1): 397–411
- 21 Ding Jian-Rui, Huang Jian-Hua, Liu Jia-Feng, Zhang Ying-Tao. Combining local features and multi-instance learning for ultrasound image classification. *Acta Automatica Sinica*, 2013, **39**(6): 861–867
(丁建睿, 黄碱化, 刘家锋, 张英涛. 局部特征与多示例学习结合的超声图像分类方法. 自动化学报, 2013, **39**(6): 861–867)
- 22 Liu Yu, Zhu Zhi-Yuan, Guan Qiang, Yang Yi-Ping. Research on experimental-design-based RFID application combinatorial testing optimization. *Acta Automatica Sinica* 2010, **36**(12): 1674–1680
(刘禹, 朱智源, 关强, 杨一平. 基于试验设计的 RFID 应用组合测试优化研究. 自动化学报, 2010, **36**(12): 1674–1680)
- 23 Su Lian-Cheng, Zhu Feng. Design of a novel omnidirectional stereo vision system. *Acta Automatica Sinica*, 2006, **32**(1): 67–72
(苏连成, 朱枫. 一种新的全向立体视觉系统的设计. 自动化学报, 2006, **32**(1): 62–72)
- 24 Jia S M, Sheng J B, Takase K. Obstacle recognition for a service mobile robot based on RFID with multi-antenna and stereo vision. In: Proceedings of the 2008 IEEE International Conference on Information and Automation. Changsha, China: IEEE, 2008. 125–130
- 25 Satake J, Miura J. Stereo-based multi-person tracking using overlapping silhouette templates. In: Proceedings of the 20th International Conference on Pattern Recognition. Istanbul: IEEE, 2010. 4304–4307
- 26 Zhang Juan, Pan Jian-Shou, Wu Ya-Peng, Liu Ji-Yan. Tracking and measurement of moving object in binocular stereo vision. *Computer Engineering and Application*, 2009, **45**(25): 191–194
(张娟, 潘建寿, 吴亚鹏, 刘继艳. 基于双目视觉的运动目标跟踪与测量. 计算机工程与应用, 2009, **45**(25): 191–194)

- 27 Jia S M, Wang S, Wang L J, Li X Z. Robust human detecting and tracking using varying scale template matching. In: Proceedings of the 2012 IEEE International Conference on Information and Automation. Shenyang, China: IEEE, 2012. 25–30



王丽佳 北京工业大学电子信息与控制工程学院博士研究生. 2008 年获得郑州大学硕士学位. 主要研究方向为机器视觉, 目标跟踪.

E-mail: wanglijia1981@hotmail.com

(**WANG Li-Jia** Ph.D. candidate at the College of Electronic Information and Control Engineering, Beijing University of Technology. She received her master degree from Zhengzhou University in 2008. Her research interest covers computer vision and target tracking.)



贾松敏 北京工业大学电子信息与控制工程学院教授. 2002 年获得日本国立电气通信大学博士学位. 主要研究方向为机器人分散控制, 机器视觉. 本文通信作者. E-mail: jsm@bjut.edu.cn

(**JIA Song-Min** Professor at the College of Electronic Information and Control Engineering, Beijing University of Technology. She received her Ph.D. degree from the University of Electro-Communications in 2002, Japan. Her research interest covers decentralized control for robot and computer vision. Corresponding author of this paper.)



李秀智 北京工业大学电子信息与控制工程学院讲师. 2008 年获得北京航空航天大学博士学位. 主要研究方向为机器人导航, 机器视觉.

E-mail: xiuzhi.lee@163.com

(**LI Xiu-Zhi** Lecturer at the College of Electronic Information and Control Engineering, Beijing University of Technology. He received his Ph.D. degree from Beihang University in 2008. His research interest covers intelligent robot navigation and computer vision.)



王爽 2014 年获得北京工业大学电子信息与控制工程学院硕士学位. 主要研究方向为机器视觉, 目标跟踪.

E-mail: wangshuang89@gmail.com

(**WANG Shuang** He received his master degree from the College of Electronic Information and Control Engineering, Beijing University of Technology. His research interest covers computer vision and target tracking.)