基于特征在线选择的目标压缩跟踪算法

李庆武 [1,2] 朱国庆 [1] 周 妍 [1,2] 霍冠英 [1,2]

摘 要 基于压缩感知理论的压缩跟踪算法能够有效地实现对目标的跟踪, 具有良好的实时性, 但该算法对目标特征没有进行在线选择导致跟踪鲁棒性不高. 本文提出一种基于特征在线选择的目标压缩跟踪算法. 首先, 在目标附近采样得到正负样本集合, 计算样本的多尺度矩形特征, 采用压缩感知中的随机投影矩阵对高维特征投影得到低维压缩域特征, 对压缩域特征进行在线选择提取最优特征, 剔除被污染的样本特征, 使用简单高效的朴素贝叶斯分类模型进行样本判断, 实现对目标的跟踪, 同时对跟踪中目标在摄像头中的尺度变化进行建模, 给出目标尺度变化的定量描述, 实现了适应目标尺度变化的多尺度跟踪. 实验结果表明本文算法具有更好的鲁棒性与更高的跟踪精度, 对目标跟踪中的遮挡、光线突变、尺度变化和非刚性形变等因素具有较好的抗干扰能力, 同时算法复杂度低, 可以满足实时性要求.

关键词 特征在线选择, 压缩感知, 尺度变化, 目标跟踪

引用格式 李庆武, 朱国庆, 周妍, 霍冠英. 基于特征在线选择的目标压缩跟踪算法. 自动化学报, 2015, **41**(11): 1961−1970

DOI 10.16383/j.aas.2015.c140809

Object Compressive Tracking via Online Feature Selection

LI Qing-Wu[1,2] ZHU Guo-Qing[1] ZHOU Yan[1,2] HUO Guan-Ying[1,2]

Abstract The compressive tracking algorithm based on compressive sensing theory can efficiently achieve real-time object tracking, but the algorithm does not select proper object features online, resulting in low tracking robustness. In order to solve this problem, an object compressive tracking algorithm with online feature selection is presented. Firstly, sets of positive and negative samples are obtained by sampling around the object, and the multi-scale rectangle features of the samples are calculated. Secondly, the compressive sensing random projection matrix is used to reduce the dimensionality of high dimensional features to obtain low-dimensional compressive domain features, and the compressive domain features are updated and selected online to extract the optimal feature to remove contaminated samples and update the classifier. Finally, a simple and efficient Bayesian classification model is utilized to achieve the object tracking. Moreover, changes of object scale in the camera are modeled and a quantitative description of changes in scale is given for multi-scale tracking which can adapt to change of the object scale. Experimental results show that the proposed algorithm can achieve a higher tracking accuracy and better robustness than several state-of-the-art algorithms and can well respond to the interferences such as block in the object tracking, light mutation, scale changes, non-rigid deformation and so on. Meanwhile, it has a low computational complexity and fully satisfies the real-time requirement.

Key words Online feature selection, compressive sensing, scale change, object tracking

Citation Li Qing-Wu, Zhu Guo-Qing, Zhou Yan, Huo Guan-Ying. Object compressive tracking via online feature selection. *Acta Automatica Sinica*, 2015, **41**(11): 1961−1970

基于视频的目标跟踪技术目前被广泛应用到智能安防、人机交互、智能交通、机器人导航等各个领域. 几十年来涌现了大量的目标跟踪算法, 但由于光照、形变、遮挡等因素影响, 目标跟踪仍是计算机视觉领域的热点与难点问题. 根据目标外观模型, 跟踪算法一般分为基于生成式和判别式模型两大类.

基于生成模型的跟踪算法主要通过学习得到的外观模型来表示目标并且在图像中寻找具有最小重建误差的图像块来表示所跟踪的目标. 2008 年, Ross 等[1] 提出 IVT (Incremental learning for robust visual tracking) 跟踪算法, 利用增量空间模型对子空间进行更新, 对目标外观变化与光线变化有较好的鲁棒性, 但对遮挡较为敏感. 近年来, 随着稀疏表示理论的发展, 基于稀疏表示理论的目标跟踪算法不断涌现. 2011 年, Mei 等[2] 提出一种基于

收稿日期 2014-12-05 录用日期 2015-05-27
Manuscript received December 5, 2014; accepted May 27, 2015
国家自然科学基金 (41306089), 江苏省自然科学基金 (BK20130240), 江苏省产学前瞻性研究项目 (BY2014041) 资助
Supported by National Natural Science Foundation of China (41306089), Natural Science Foundation of Jiangsu Province (BK20130240), Prospective Research Project on the Integration of Industry, Education and Research of Jiangsu Province (BY2014041)
本文责任编委 贾云得
Recommended by Associate Editor JIA Yun-De
1. 河海大学物联网工程学院 常州 213022 2. 常州市传感网与环境感知重点实验室 常州 213022
1. College of Internet of Things Engineering, Hohai University, Changzhou 213022 2. Changzhou Key Laboratory of Sensor Networks and Environmental Sensing, Changzhou 213022

L_1 范数的目标跟踪算法, 将目标作为一组目标模板与琐碎目标的稀疏表示, 可以解决部分遮挡等问题. 但是需求解多次的 L_1 范数最小化, 计算代价太高. 2013 年, 杨大为等[3] 将粒子滤波思想与稀疏表示结合提出一种基于粒子滤波与稀疏表达的目标跟踪方法, 该方法能有效地减小光照变化对跟踪效果的影响. 粒子数越多跟踪效果越好, 但是随着粒子数增加实时性变差, 难以实现实时跟踪. 姜明新等[4] 提出一种基于 ML (Maximum likelihood) 和 L_2 范数的视频目标跟踪算法, 该算法对目标发生遮挡、旋转、尺度变化、光照变化等异常变化时具有较强的鲁棒性, 同时在实时性方面也有了提升. 2014 年, 齐美彬等[5] 对基于 L_1 范数跟踪算法进行改进, 提出了一种基于 L_2 范数最小化的实时目标压缩跟踪算法, 实时性有了较大的提升. 此外, 随着仿生科学的发展许多学者将生物视觉特性引入到目标跟踪中, 2008 年王智灵等[6] 将蛙眼视觉特性引入到目标跟踪中, 利用蛙眼视觉认知的生理特性和外部特性提出了一种与之相应的模糊化区域理解的运动目标跟踪方法. 该方法对复杂背景下的目标跟踪问题进行了尝试, 取得了良好的效果. 李智勇等[7] 提出了基于蛙眼 R3 细胞感受野模型的运动滤波方法, 利用青蛙视觉中对目标清晰而背景模糊的特性提出了相应的数学模型, 对运动目标的滤波取得了良好的效果. 除了对动物的视觉特性进行借鉴外, 许多学者[8] 也将人眼视觉的注意机制引入到目标跟踪中, 使得基于生成模型的目标跟踪算法性能获得了较大的提高.

基于判别模型的跟踪算法提取正负样本特征, 将跟踪问题转化为一个二分类问题, 寻找一个最佳判别界限区分目标和背景. 2005 年 Collins 等[9] 提出一种特征在线选择的目标跟踪算法, 取得了良好的跟踪效果, 但该算法只用一个正样本 (目标所在区域) 和一些负样本 (背景区域) 来更新分类器, 当目标受到噪声干扰时容易产生跟踪漂移. 2011 年, Babenko 等[10] 提出了一种多示例学习 (Multiple instance learning, MIL) 跟踪算法, 其中训练样本通过 "包" 的形式给出, "正包" 中至少含有一个正样本实例, "负包" 中含有一些负样本实例, 但由于其特征选择计算复杂难以满足实时性要求. Zhang 等[11] 对 MIL 跟踪算法进行了改进, 提出一种高效的特征在线选择 (Online discriminative feature selection, ODFS) 算法, 大大提高了跟踪算法的实时性. 随着压缩感知理论的发展与完善, 基于该理论的各项技术已被广泛应用于各个研究领域, Zhang 等[12−13] 提出一种基于压缩感知理论的高效目标跟踪算法, 证明了满足压缩感知中有限等距性质 (Restricted isometry property, RIP) 的随机投影矩阵可以将多尺度图像高维特征投影到低维空间, 且低维特征空间能够保持高维特征空间足够的判别信息, 从而大大降低跟踪算法的复杂度. 但是由于该算法未对特征进行选择, 跟踪中容易产生漂移导致跟踪失败.

针对目前压缩跟踪算法中存在的鲁棒性问题, 提出一种基于特征在线选择的压缩跟踪算法. 定义边缘函数为正样本集合平均置信度与负样本集合平均置信度之差, 采用类梯度下降法选出使正样本平均置信度最大, 同时抑制负样本平均置信度的最优特征, 更新简单高效的朴素贝叶斯分类器参数, 在新的一帧图像中采集当前目标所在位置周围区域作为样本集, 并放入分类器计算置信度, 具有最大置信度的样本即为跟踪的目标. 同时对跟踪中目标在摄像头中的尺度变化进行建模, 给出目标尺度变化的定量描述, 实现了适应目标尺度变化的多尺度跟踪. 实验结果表明, 本文算法与目前优秀的跟踪算法相比, 具有较高的跟踪精度和鲁棒性, 同时能够达到每秒 25 帧的处理速度, 满足跟踪实时性要求.

1 压缩跟踪算法

随机投影与随机测量矩阵是压缩感知理论中两个重要的概念. 随机矩阵 $R \in \mathbf{R}^{n\times m}$ $(n \ll m)$ 可以将高维空间 $\boldsymbol{X} \in \mathbf{R}^m$ 投影到低维空间 $\boldsymbol{V} \in \mathbf{R}^n$.

$$\boldsymbol{V} = R\boldsymbol{X} \tag{1}$$

如果低维向量 $\boldsymbol{V}$ 可以完全地保留高维向量 $\boldsymbol{X}$ 的信息, 或者说保持原始空间中各样本 $\boldsymbol{X}$ 的距离关系则可以在低维空间实现对高维信号的分析. Johnson-Lindenstrauss[14] 推论表明, 一个高维向量空间, 将其投影到随机选取的适当维数的子空间, 其两点间的距离能够很好地保存, 即不改变高维向量空间的结构信息. Baraniuk 等[15] 证明了满足 Johnson-Lindenstrauss 推论的随机矩阵同样满足压缩感知理论中的有限等距性质条件. Zhang 等[12] 证明了如果随机矩阵 R 满足 Johnson-Lindenstrauss 推论, 那么如果 $\boldsymbol{X}$ 是可压缩的 (如图像、语音), 就可以通过最小化误差来从 $\boldsymbol{V}$ 中高概率恢复 $\boldsymbol{X}$, 即高维向量可以被投影得到的向量以最小误差重构出来.

通过选取适当的随机矩阵对 $\boldsymbol{X}$ 进行投影得到几乎全部保留 $\boldsymbol{X}$ 信息的 $\boldsymbol{V}$, 则可以通过低维空间来实现对高维信号的分析. 目前被广泛应用的满足 RIP 条件的典型测量矩阵为高斯随机矩阵 $R \in \mathbf{R}^{n\times m}$ $(n \ll m)$, 其元素 $r_{ij} \sim \mathrm{N}(0, 1)$. 但是由于高斯矩阵本身的稠密性, 当 m 取得较大值时, 其仍然需占用大量存储空间, 同时计算量也较大. 在压缩跟踪算法中选取一个非常稀疏的随机测量矩阵

$SR \in \mathbf{R}^{n\times m}$, 其元素服从如下分布:

$$sr \sim \begin{pmatrix} \sqrt{s} & 0 & -\sqrt{s} \\ \dfrac{1}{2s} & 1-\dfrac{1}{s} & \dfrac{1}{2s} \end{pmatrix} \tag{2}$$

Achlioptas[14] 证明了当 $s=2$ 或 3 时, 此矩阵满足 Johnson-Lindenstrauss 引理. 当 $s \gg 3$ 时, SR 服从非常稀疏投影. 由上式可以看出当 $s=3$ 时仅有 1/3 的原始数据被采样, 这样可以避免 2/3 的计算量.

在压缩跟踪算法中高维图像特征是由感兴趣区域与一系列不同位置和大小的矩形滤波器卷积而得到. 使用满足压缩感知原理中 RIP 条件的非常稀疏测量矩阵 (2) 对高维特征集合进行压缩, 得到压缩域特征. 稀疏测量矩阵只需离线计算一次, 在跟踪中始终保持不变, 避免了重复计算, 降低了计算复杂度能够满足实时跟踪的要求.

压缩跟踪算法是判别式在线跟踪算法, 将目标跟踪转化为二值分类问题, 即在线判别出目标与背景. Ng 等[16] 证明了目标特征大体上服从独立分布, 而朴素贝叶斯分类器对满足独立分布的随机变量分类表现出高精度和高效率, 具有最小的误分类率和耗时开销小的特点, 所以在跟踪中使用贝叶斯分类模型来对目标与背景进行分类, 实现对目标的跟踪. 样本 $\boldsymbol{X}$ 的特征向量 $f(\boldsymbol{x})=(f_1(\boldsymbol{x}),\cdots,f_k(\boldsymbol{x}))^{\mathrm{T}}$, 则朴素贝叶斯分类器 h_k 为

$$h_k(\boldsymbol{x}) = \log\left(\frac{\prod_{k=1}^{K} p(f_k(\boldsymbol{x}) \mid y=1)P(y=1)}{\prod_{k=1}^{K} p(f_k(\boldsymbol{x}) \mid y=0)P(y=0)}\right) = \sum_{k=1}^{K}\phi_k(\boldsymbol{x}) \tag{3}$$

$$h_k(\boldsymbol{x}) = \log\left(\frac{p(f_k(\boldsymbol{x}) \mid y=1)}{p(f_k(\boldsymbol{x}) \mid y=0)}\right) \tag{4}$$

Diaconis 等[17] 证明了高维随机向量的随机投影大体服从高斯分布, 所以可设条件概率 $p(f_k|y=1)$ 和 $p(f_k|y=0)$ 服从高斯分布, 即:

$$\begin{aligned} p(f_k|y=1) &\sim \mathrm{N}(\mu_k^+,\sigma_k^+) \\ p(f_k|y=0) &\sim \mathrm{N}(\mu_k^-,\sigma_k^-) \end{aligned} \tag{5}$$

式 (5) 中的参数 $(\mu_k^+,\sigma_k^+,\mu_k^-,\sigma_k^-)$ 的更新方式可以由最大似然估计方法推导得出:

$$\mu_k^+ \longleftarrow \eta\mu_k^+ + (1-\eta)\mu^+ \tag{6}$$

$$\sigma_k^+ \longleftarrow \sqrt{\eta(\sigma_k^+)^2 + (1-\eta)(\sigma^+)^2 + \eta(1-\eta)(\mu_k^+-\mu^+)^2} \tag{7}$$

其中, $\sigma^+=\sqrt{1/N\sum_{i=0|y=1}^{N-1}(f_k(x_i)-\mu^+)^2}$, $\mu^+ = (1/N)\sum_{i=0|y=1}^{N-1} f_k(x_i)$ 对于负样本参数 (μ_k^-,σ_k^-) 也采用同样的规则进行更新.

对每一帧图像, 在目标附近采样一些正样本, 在离目标较远的地方采样一些负样本, 使用正负样本来更新分类器参数, 为了预测目标在下一帧图像中的位置, 对下一帧图像中目标当前位置的附近区域进行采样, 并放入分类器, 具有最大响应的样本即为下一帧中目标所在的位置.

压缩跟踪算法简单高效, 满足实时跟踪要求. 但是对分类器进行更新时, 特征没有进行选取, 在跟踪中容易受到光照变化、遮挡等噪声干扰降低分类器性能, 跟踪结果会产生较大误差甚至出现跟丢这样灾难性后果. 针对这一问题本文提出一种高效的特征在线选择跟踪算法, 对压缩域特征进行在线选择, 选出最优特征对分类器进行更新, 提高算法的鲁棒性和精确性.

2 基于特征在线选择的压缩跟踪

2.1 特征在线选择

在目标跟踪中, Collins 等[9] 证明了在线选择目标与背景的判别特征能够极大地提升目标跟踪效果. 由于在跟踪中目标尺度不断变化, 目标提取窗口应适应尺度的变化, 才能够提取出好的正样本与负样本. 特征在线选择可以定义为从一个含有 $M\,(M>K)$ 个特征的特征池中选取 K 个特征使得样本 X^* 具有最大置信度即 $X^*=\arg\max_X(c(x))$. 可以使用排列组合方法选择 K 个特征使得置信度 $c(\cdot)$ 最大, 但其计算量为 C_M^K, 很难满足实时在线跟踪的要求. 本文提出一种高效的特征在线选择算法, 特征组合次数为 $MK \ll C_M^K$, 实验证明完全满足实时跟踪要求. 定义目标置信度为朴素贝叶斯分类器 $h_k(x)$ 的非线性函数, 如式 (8) 所示:

$$c(x) = p(y=1 \mid x) = \sigma(h_k(x)) \tag{8}$$

其中, 样本 $\boldsymbol{X}$ 通过其压缩域特征向量 $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}),\cdots,f_k(\boldsymbol{x}))^{\mathrm{T}}$ 表示, $y\in\{0,1\}$ 是分类标签, $\sigma(\cdot)$ 为 Sigmoid 函数, $\sigma(z)=(1/1+\mathrm{e}^{-z})$. 将样本空间分为两个部分 $R^+=\{X,y=1\}$ 和 $R^-=\{X,y=0\}$, 定义边缘函数为集合 R^+ 中的样

本平均置信度减去集合 R^- 中样本的平均置信度:

$$E_{\text{margin}} = \frac{1}{|R^+|}\int\limits_{x\in\mathbf{R}^+} c(x)\mathrm{d}x - \frac{1}{|R^+|}\int\limits_{x\in\mathbf{R}^-} c(x)\mathrm{d}x \tag{9}$$

在训练集中, 设正样本集 $R^+ = \{X_i\}_{i=0}^{N-1}$ (X_0 为当前帧的跟踪结果) 由 N 个样本组成, 负样本集 $R^- = \{X_i\}_{i=N}^{N+L-1}$ 由 L 个样本组成 ($L \approx N$). 由式 (3) 与式 (8), 式 (9) 可以写成:

$$E_{\text{margin}} \approx \frac{1}{N}\Big(\sum_{i=0}^{N-1}\sigma(\sum_{k=1}^{k}\phi_k(\boldsymbol{x}_i)) - \sum_{i=N}^{N+L-1}\sigma(\sum_{k=1}^{k}\phi_k(\boldsymbol{x}_i))\Big) \tag{10}$$

每个样本 X_i 都通过其特征向量 $f(\boldsymbol{x}_i) = (f_1(\boldsymbol{x}_i), \cdots, f_M(\boldsymbol{x}_i))^{\mathrm{T}}$ 表示, 则一个弱分类器池 $\Phi = \{\phi_m\}_{m=1}^{M}$ 可以由式 (4) 得到. 最优特征应使得类类之间的差别最大, 即从分类器池 Φ 中选出一个分类器子集 $\{\phi_k\}_{k=1}^{K}$ 使得集合 R^+ 中样本平均置信度最大, 同时抑制集合 R^- 中样本的平均置信度. 因此可以通过最大化边缘函数 E_{margin} 来选择弱分类器, 即:

$$\{\phi_1, \cdots, \phi_k\} = \arg\max_{\{\phi_1,\cdots,\phi_k\}}(\phi_1, \cdots, \phi_k) \tag{11}$$

使用贪心算法从 Φ 连续选择使 E_{margin} 最大的 k 个弱分类器:

$$\phi_k = \arg\max_{\phi\in\Phi} E_{\text{margin}}(\phi_1, \cdots, \phi_{k-1}, \phi) = \arg\max_{\phi\in\Phi}\Big(\sum_{i=0}^{N-1}\sigma(h_{k-1}(\boldsymbol{x}_i) + \phi(\boldsymbol{x}_i)) - \sum_{i=N}^{N+L-1}\sigma(h_{k-1}(\boldsymbol{x}_i) + \phi(\boldsymbol{x}_i))\Big) \tag{12}$$

在 $N+L$ 维数据空间中, 式 (12) 中 $g_{k-1}(x)$ 处最陡下降方向为

$$g_{k-1} = (g_{k-1}(x_0), \cdots, g_{k-1}(x_{N-1}), -g_{k-1}(x_N), \cdots, -g_{k-1}(x_{N+L-1})) \tag{13}$$

其中, $g_{k-1} = -\frac{\partial\sigma(h_{k-1}(x))}{\partial h_{k-1}} = -\sigma(h_{k-1}(x))(1-\sigma(h_{k-1}(x)))$, 因为 g_{k-1} 只在点 $(X_0, \cdots, X_{N+L-1})$ 处有定义, 其广义性能受到了限制. 本文选择与 g_{k-1} 最不相关的弱分类器 ϕ 来最大化目标函数 (12). 使用以下约束关系来选择 ϕ, 即:

$$\phi_k = \arg\max_{\phi\in\Phi}\{E_{SGSC}(\phi) = \|g_{k-1} - \phi\|_2^2\} = \arg\max_{\phi\in\Phi}(\sum_{i=0}^{N-1}(g_{k-1}(\boldsymbol{x}_i) - \phi(\boldsymbol{x}_i))^2 + \sum_{i=N}^{N+L-1}(-g_{k-1}(\boldsymbol{x}_i) - \phi(\boldsymbol{x}_i))^2) \tag{14}$$

但是通过分析可以发现单个梯度与弱分类器容易受到不良样本的影响, 从而影响到新的分类器的选择, 为了减少弱分类器对单个分类器的依赖引入中值滤波的思想, 对滤波器和梯度分别进行排序选择中值来减少不良样本对弱分类器选择的影响, 以提高特征选择能力的鲁棒性. 则 ϕ_k 可以表示为

$$\phi_k = \arg\max_{\phi\in\Phi}\{E_{MGMC}(\phi) = N(mid(g_{k-1}^+) - mid(\phi^+))^2 + L(-mid(g_{k-1}^-) - mid(\phi^-))^2\} \tag{15}$$

由上式可以看出通过从不同的正样本中选择出弱分类器能够有效地抑制由不良样本引入的噪声, 而且采用很多正样本的梯度来对特征进行选择有效地减轻了样本模糊的问题, 与传统判别模型采用单个正样本提取特征相比, 本文采用的方法具有很好的抗噪性能. 在每一时刻选取了 k 个最优特征 (弱分类器) 后, 采用传统的贝叶斯分类器更新方法将弱分类器线性组合后即可得到强分类器.

2.2 尺度更新

在视频序列图像中, 需要跟踪的目标距离摄像头的远近是会发生变化的, 图像中目标的大小尺寸也会跟着变化, 距离摄像头近则目标尺度变大; 反之则变小. 在二维图像中此现象体现在当目标尺度变小时, 其 y 轴坐标值变大; 反之亦然. 因此, 可以采用图像中目标位置相对变化值来衡量尺度的变化. 为了减少噪声对跟踪性能的影响, 增强目标尺度的稳定性, 取当前帧的前 n 帧目标位置相对变化的均值作为目标尺度. 由上述分析, 本文对摄像头进行建模, 依据目标距摄像头的距离与其尺度变化的关系给出下述尺度更新算法:

$$\begin{cases} \Delta l_t' = l_t - l_{t-1} \\ \bar{\Delta l_t} = \dfrac{1}{n}\sum\limits_{i=1}^{n}\Delta l_t' \\ s_{t+1} = (1-\lambda)s_t + \lambda \mathrm{e}^{-\overline{\Delta l_t}} \\ \sigma_{t+1} = s_{t+1}\sigma \end{cases} \tag{16}$$

式中, l_t 为目标位置的 y 轴坐标, l_{t-1} 为前一帧目标

位置坐标, $\Delta l_t'$ 为相邻两帧位置变化值, $\overline{\Delta l_t}$ 表示距离 t 帧前 n 帧平均位置变化量, s_{t+1} 表示第 $t+1$ 帧尺度大小, λ 为尺度影响因子用来调节尺度变化速度. σ 表示目标初始窗口大小, σ_{t+1} 为 $t+1$ 帧目标的大小.

3　算法步骤

本文算法流程框图如图 1 所示, 具体步骤如下: 令 $I_t(x) \in \mathbf{R}^2$ 表示在 t 帧图像中样本 $\boldsymbol{X}$ 所在的位置; $I_t(x^*)$ 表示在 t 帧图像中目标所对应的位置, 给定分类器 h_k, 则本文跟踪的步骤如下:

1) 在当前 t 帧图像中随机采集 $X^\alpha = \{x|\|I_t(x) - I_t(x^*)\| < \alpha\}$ 并标记作为正样本.

2) 在当前 t 帧图像中随机采集 $X^{\xi,\beta} = \{x|\xi < \|I_t(x) - I_t(x^*)\| < \beta\}$ 作为负样本, 其中 $\alpha < \xi < \beta$.

3) 计算正负样本的多尺度矩形特征, 并用非常稀疏矩阵对特征进行投影得到压缩域特征.

4) 用特征在线选择算法在特征池中选择最优特征.

5) 使用得到的最优特征对分类器 h_k 进行更新.

6) 在 $(t+1)$ 帧图像中, 对应 t 帧中目标的位置进行大范围的采样 $X^\gamma = \{x|\|I_{t+1}(x) - I_t(x^*)\| < \gamma\}$ (γ 一般取目标尺寸的 2 倍即可), 使用更新后的分类器对 $(t+1)$ 帧图像中采集的样本进行计算, 找出具有最大置信度即 $X^* = \arg\max_x(c(x))$ 的样本, 则目标在 $(t+1)$ 帧图像中的位置为 I_{t+1}^*.

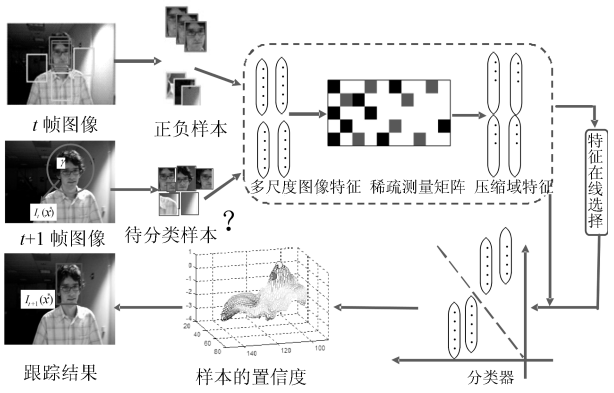


图 1　算法流程图

Fig. 1　Algorithm flow block diagram

4　实验与结果分析

为了验证本文跟踪算法的有效性, 选择了目前具有代表性的 CT[12] (Compressive tracking)、MIL[10]、ODFS[11] 跟踪算法, 对 4 个存在光照变化、尺度变化、遮挡、旋转、背景相似等干扰因素的图像序列作对比实验. 对于每一个待跟踪的目标都是在图像序列的第一帧手工选定. 实验中所有图像序列均来自 https://sites.google.com/site/trackerbenchmark/benchmarks/v10 中的标准测试序列, 分别对算法的跟踪精度进行定性与定量评价. 其中定量评价采用中心误差即目标实际中心与跟踪结果中心之间的距离来描述, 距离越小则误差越小, 跟踪精度越高.

本文算法的实现基于 Windows 操作系统, 采用 Matlab 作为软件平台, 计算机配置为 Intel Core 2 Duo CPU 2.10 GHz, 2.00 GB 内存, 可以达到 25 FPS, 满足实时跟踪要求. 对本文算法中主要参数的设置如下: 正负样本个数 N、L 一般都可以选择相似或相等, 本文均选择 50; 在特征选择上采用 $M = 150$、$K = 15$ 即可.

图 2 为对 4 种图像序列的跟踪结果. 在序列 Walking2 中, 存在着目标尺度变化、遮挡、低分辨率的影响, 给目标跟踪带来了很大的困难. 由图 2 (a) 中 ♯166 帧可以看到随着目标尺度的不断减小, ODFS 与 CT 算法都产生漂移现象, 跟踪出现较大误差, 而 MIL 与本文算法都能够对目标进行有效的跟踪. 当 ♯218 帧目标被遮挡时, MIL、CT、ODFS 跟踪发生丢失, 当 ♯299 帧时遮挡消失, 而 MIL、CT、ODFS 都已经将遮挡物作为目标, 而本文算法对目标尺度进行估计, 随着目标尺度的变小采样范围减小, 减少了不良样本的干扰, 同时特征进行在线选择, 提取出正负样本最佳判别特征, 始终能够对目标进行有效的跟踪.

在序列 Crossing 中存在着目标尺度变化、光照变化、背景与目标相似等干扰因素. 从图 2 (b) 中可以看出在 ♯12 帧由于受到背景干扰跟踪发生了漂移, 但还能够对目标进行跟踪. 当 ♯53 帧时, CT 跟踪丢失而 MIL 及 ODFS 算法也产生了漂移. 当 ♯120 帧时, MIL 与 ODFS 算法均产生了较大的漂移, CT 已经完全跟踪丢失, 而本文算法却始终对目标进行着准确的跟踪. 可以看出本文算法在目标尺寸改变时, 能够较好地做出定量的判断与估计, 满足对多尺度跟踪的要求.

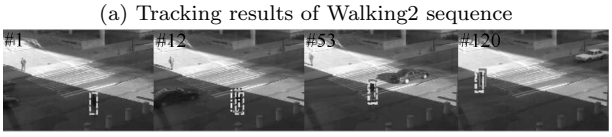
在序列 David3 中存在遮挡和旋转, 对此序列进行有效的跟踪, 要求跟踪算法必须能够适应目标的旋转、外观的改变. 由图 2 (c) 可以看出当目标被树木遮挡过后 (♯ 88、♯106 帧), MIL、CT 算法都发生了漂移, 且无法对目标进行有效地跟踪, 只有 ODFS 与本文算法采用特征选择机制仍然可以对目标进行有效的跟踪. 当 ♯139 帧时目标转身发生了形变, 本文算法与 ODFS 仍然能够对目标进行有效的跟踪.

在序列 Freeman 中存在左右旋转和前后旋转以及目标尺度的变化. 由图 2 (d) 可以看出, 当 ♯31 帧

时 ODFS 跟踪算法由于特征选择策略的原因发生了较大漂移, MIL 算法也出现了漂移. 在 ♯166 帧, ODFS 算法已经丢失了目标, 目标旋转后 MIL 与 CT 算法跟踪也出现了严重的漂移, 而本文算法却能够抵抗旋转产生的遮挡与非刚性形变的影响对目标进行有效的跟踪. 当 ♯190 帧时, ODFS、MIL、CT 算法均完全失去对目标的跟踪而本文算法能够适应目标的旋转与尺度的变化, 依然能够对目标进行有效的跟踪.

(a) Walking2 序列跟踪结果

(a) Tracking results of Walking2 sequence



(b) Crossing 序列跟踪结果

(b) Tracking results of Crossing sequence

(c) David3 序列跟踪结果

(c) Tracking results of David3 sequence

(d) Freeman 序列跟踪结果

(d) Tracking results of Freeman sequence

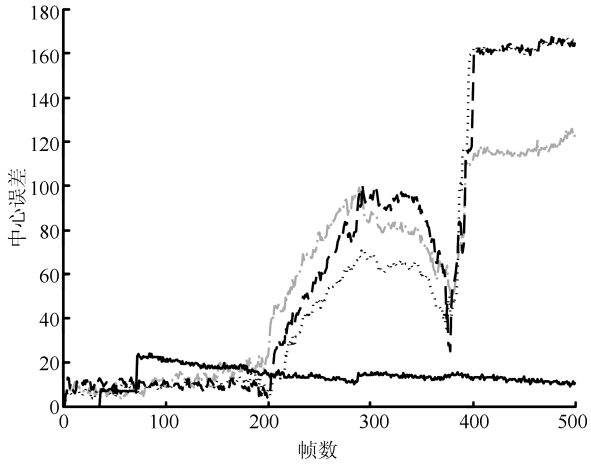
ODFS MIL CT 本文算法

图 2 图像序列跟踪结果

Fig. 2 Tracking results of the sequences

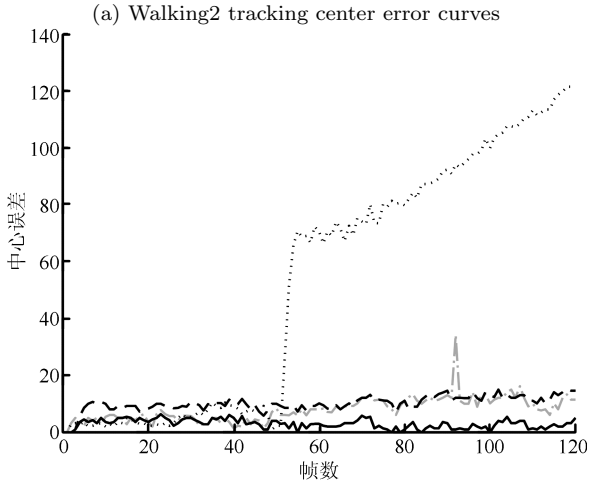
如图 3 所示的跟踪中心误差曲线定量评价表明了本文算法具有最高的精确性与稳定性. 通过上述具有代表性的 4 组标准测试序列实验后可以得出, 本文采用特征在线选择以及对摄像头进行建模适应目标尺度变化的方法, 对目标跟踪中的遮挡、尺度变化等难点问题具有良好的适应能力.

为了对本文算法进行全面客观的评价, 采用 Benchmark 中的算法性能评价方法, 分别从每帧跟踪精度 (One-pass evaluation, OPE)、时间鲁棒性 (Temporal robustness evaluation, TRE) 以及空间鲁棒性 (Spatial robustness evaluation, SRE) 对本文算法进行评价. 在测试序列中给定目标初始位置后, 使用跟踪算法对整个图像序列进行逐帧跟踪是目前最为常用的跟踪算法评价方法. 在实际应用中目标初始位置可能是不精确的, 存在一定的偏移, 同时在跟踪中可能会出现跳帧的情况, 这就需要跟踪算法具有较强的空间鲁棒性和时间鲁棒性. 为了衡量算法在实际应用中的表现, 对目标初始位置进行一定的偏移, 包含左、右、上、下 4 个方向以及对角



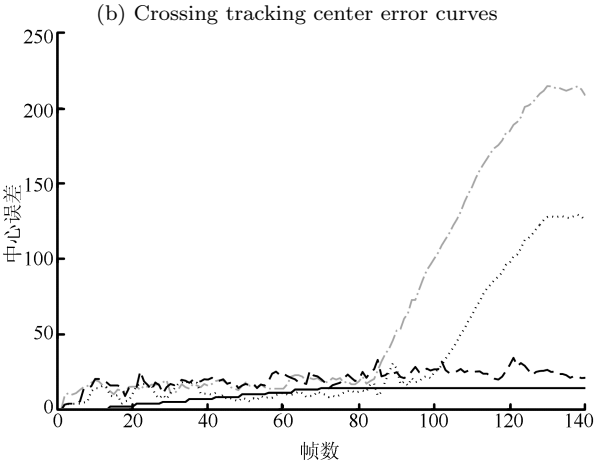
(a) Walking2 跟踪中心误差曲线

(a) Walking2 tracking center error curves



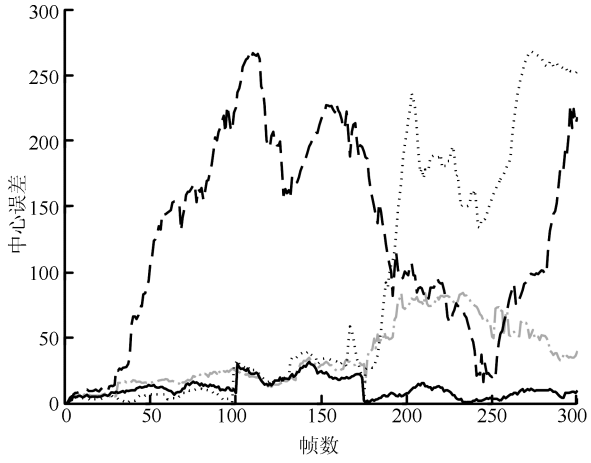
(b) Crossing 跟踪中心误差曲线

(b) Crossing tracking center error curves



(c) David3 跟踪中心误差曲线

(c) David3 tracking center error curves



(d) Freeman 跟踪中心误差曲线

(d) Freeman tracking center error curves

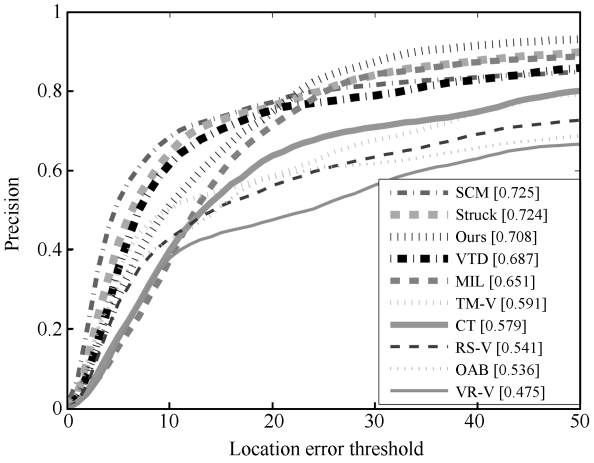
图 3　跟踪中心误差曲线

Fig. 3　Tracking center error curves

线 4 个方向, 共 8 种情况的偏移, 漂移量均为 10 %. 为了测试算法的跟踪鲁棒性, 本文选取 Benchmark 中具有代表性的 12 组测试序列, 序列中包含了背景相似、旋转、尺度变化、遮挡等影响跟踪鲁棒性的因素. 采用跟踪空间鲁棒性衡量标准, 从跟踪精度与跟踪成功率两个方面定量分析算法跟踪的空间鲁棒性. 跟踪精度曲线定义为一个图像序列中跟踪结果与目标中心误差小于给定阈值 d 的帧数占总帧数的比率. 本实验中令 $d = 20$ 来计算算法的平均跟踪精度. 跟踪结果的边界窗口为 r_t, 目标窗口为 r_a, 则重叠率定义为 $S = |\, r_t \cap r_a \,| / |\, r_t \cup r_a \,|$, 当重叠率 S 大于所给的阈值 t 时表示当前帧跟踪成功. 在实验中令 $t = 0.5$ 来计算算法的平均跟踪成功率, 本文采用文献 [18] 中定义的成功率曲线下的面积 (Area under curve, AUC) 来衡量算法的性能, 面积越大则说明算法的性能越好. 选取在 Benchmark 中具有代表性的 11 种跟踪算法进行对比实验, 12 种算法分别为: VR[9]、MIL[10]、CT[12]、RS[18]、OAB[19]、TM-V[20]、IVT[1]、VTD[21]、Struck[22]、SCM[23]、MS[24]、PD[25].

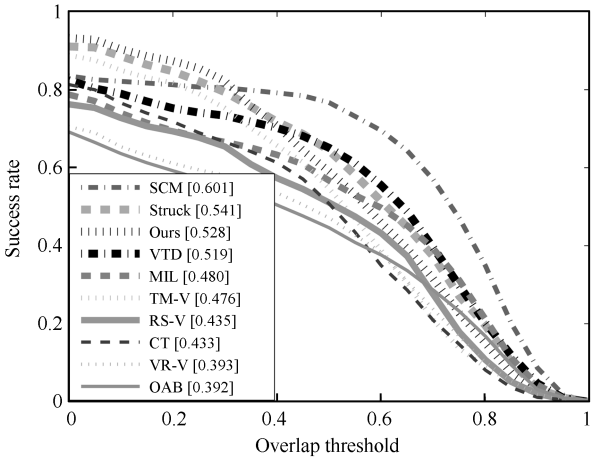
对图像序列按照在第一帧对目标位置初始化, 逐帧测试的结果如图 4 所示, 本文算法在 12 种典型的测试算法排名中性能达到了第三位, 这表明本文算法具有良好的目标跟踪性能. 在算法处理效率上本文算法能够达到 28 FPS 而 SCM 以及 Struck 在同样的硬件平台与实现语言下只能达到 0.01 FPS.

时间鲁棒性曲线如图 5 所示, 反映了跟踪算法在初始帧不是第一帧或跟踪图像序列中存在跳帧的



(a) 跟踪精度曲线

(a) Tracking precision curves

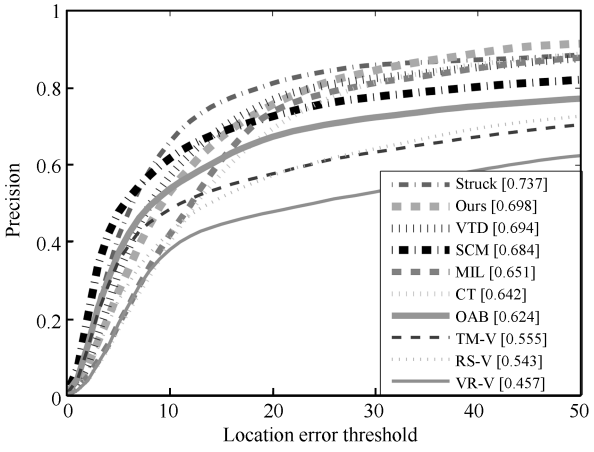


(b) 跟踪成功率曲线

(b) Tracking success rate curves

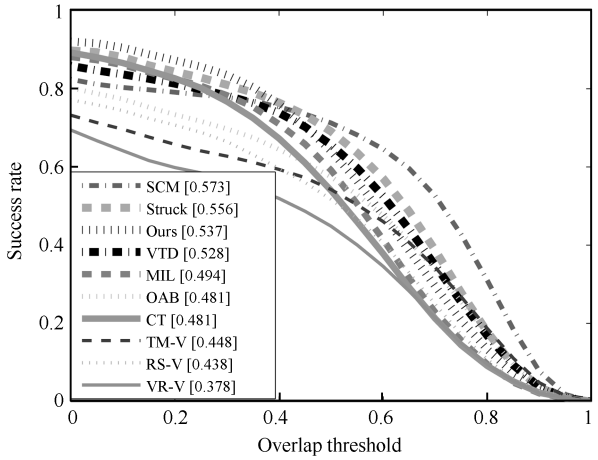
图 4　跟踪算法 OPE 曲线

Fig. 4　Tracking algorithm OPE curves



(a) 跟踪精度曲线

(a) Tracking precision curves

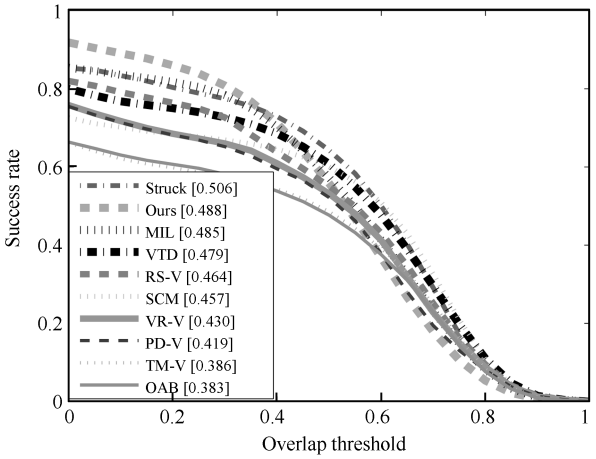


(b) 跟踪成功率曲线

(b) Tracking success rate curves

图 5 跟踪算法 TRE 曲线

Fig. 5 Tracking algorithm TRE curves



(b) 跟踪成功率曲线

(b) Tracking success rate curves

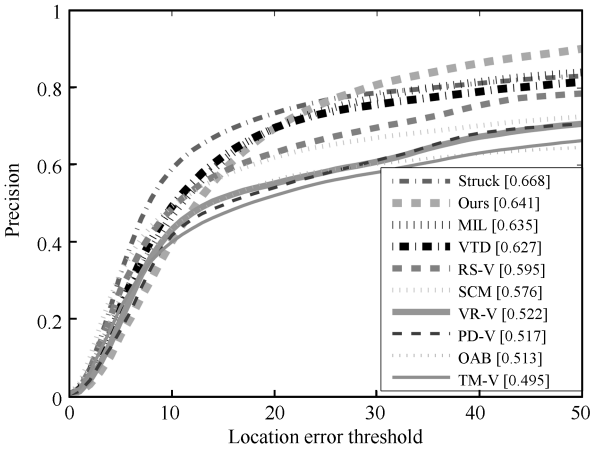
图 6 跟踪算法 SRE 曲线

Fig. 6 Tracking algorithm SRE curves

情况下跟踪算法的性能. 从图 5 可以看出本文算法无论是在重叠率还是误差精度上表现都是前三名, 而另外两种跟踪算法 Struck、SCM 都是花费巨大的时间消耗来获得较好的跟踪效果. 本文算法在保证了能够对目标实时跟踪的同时还具有很高的时间鲁棒性.

跟踪空间鲁棒性衡量曲线如图 6 所示. 在跟踪误差阈值小于 10、重叠率大于 0.4 时, 本文算法的跟踪精度比 MIL 稍微差些, 但是随着跟踪误差阈值取值的逐渐增大、重叠率的逐渐减小, 本文算法的跟踪精度提高很快, 这在实际应用中尤为重要.

综上所述, 本文算法在满足实时跟踪的同时具有良好的跟踪精度以及较强的鲁棒性. 跟踪成功率与精度前十的算法的综合性能如表 1 所示. 本文算法与在 Benchmark 中综合表现最好的跟踪算法 Struck、SCM 相比, 在精度损失不多且鲁棒性相差不大的情况下, 本文算法的跟踪速度能够达到实时, 而上述两种算法根本无法在实际应用中使用. 与同样采用特征在线选择的 MIL 以及 OAB 相比, 本文算法在提高算法性能的同时降低了时间复杂度. 与传统的 CT 算法相比, 本文算法明显提高了性能, 但运行时间没有优势.



(a) 跟踪精度曲线

(a) Tracking precision curves

表 1 算法综合性能表

Table 1 Algorithm integrated performance table

算法名称	平均跟踪成功率	平均跟踪精度误差	FPS
Struck	0.541	0.724	0.01
SCM	0.601	0.725	0.01
Ours	0.528	0.708	28
MIL	0.480	0.651	10.3
VTD	0.519	0.687	0.01
RS-V	0.435	0.541	10
VR-V	0.393	0.475	40
CT	0.433	0.579	33
OAB	0.392	0.536	8.5
TM-V	0.476	0.591	0.1

5 结论

本文采用非常稀疏投影矩阵对图像特征空间进行投影将其压缩到低维空间, 低维特征空间能够很好地保持高维信号的空间结构. 提出一种新的在线特征选择算法对低维压缩域特征进行在线选择, 提取最优特征对分类器进行更新. 使用分类器的非线性函数计算样本的置信度函数确定目标位置, 同时

对摄像头进行建模判断目标尺度变化使得算法能够多尺度对目标进行跟踪. 与原压缩跟踪算法相比, 极大提升算法鲁棒性与准确性, 本文算法因为特征在线选择使其复杂度稍有增加, 但不会对跟踪的实时性产生影响. 实验表明本文算法具有较高的抗形变、模糊、遮挡等能力, 具有较高的精度, 实时性较强, 同时算法具有较强的时空鲁棒性, 能够满足实际跟踪要求.

References

1 Ross D A, Lim J, Lin R S, Yang M H. Incremental learning for robust visual tracking. *International Journal of Computer Vision*, 2008, **77**(1−3): 125−141

2 Mei X, Ling H B. Robust visual tracking and vehicle classification via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(11): 2259−2272

3 Yang Da-Wei, Cong Yang, Tang Yan-Dong. Object tracking method based on particle filter and sparse representation. *Pattern Recognition and Artificial Intelligence*, 2013, **26**(7): 680−687
(杨大为, 丛杨, 唐延东. 基于粒子滤波与稀疏表达的目标跟踪方法. 模式识别与人工智能, 2013, **26**(7): 680−687)

4 Jiang Ming-Xin, Wang Hong-Yu, Wang Jie, Wang Biao. Visual object tracking algorithm based on ML and L_2-Norm. *Acta Electronica Sinica*, 2013, **41**(11): 2307−2313
(姜明新, 王洪玉, 王洁, 王彪. 基于 ML 和 L_2 范数的视频目标跟踪算法. 电子学报, 2013, **41**(11): 2307−2313)

5 Qi Mei-Bin, Yang Xun, Yang Yan-Fang, Lu Lei, Jiang Jian-Guo. Real-time object tracking based on L_2-norm minimization. *Journal of Image and Graphics*, 2014, **19**(1): 36−44
(齐美彬, 杨勋, 杨艳芳, 陆磊, 蒋建国. 基于 L_2 范数最小化的实时目标跟踪. 中国图象图形学报, 2014, **19**(1): 36−44)

6 Wang Zhi-Ling, Chen Zong-Hai, Xu Xiao-Xiao, Wu Liang. A fuzzy region understanding tactic for object tracking based on frog′s vision characteristic. *Acta Automatica Sinica*, 2009, **35**(8): 1048−1054
(王智灵, 陈宗海, 徐萧萧, 吴亮. 基于蛙眼视觉特性的运动目标模糊化区域理解跟踪方法. 自动化学报, 2009, 35(8): 1048−1054)

7 Li Zhi-Yong, He Shuang, Liu Jun-Min, Li Ren-Fa. Motion filtering by modelling R3 cell′s receptive field in frog eyes. *Acta Automatica Sinica*, 2015, **41**(5): 981−990
(李智勇, 何霜, 刘俊敏, 李仁发. 基于蛙眼 R3 细胞感受野模型的运动滤波方法. 自动化学报, 2015, **41**(5): 981−990)

8 Li Wan-Yi, Wang Peng, Qiao Hong. A survey of visual attention based methods for object tracking. *Acta Automatica Sinica*, 2014, **40**(40): 561−576
(黎万义, 王鹏, 乔红. 引入视觉注意机制的目标跟踪方法综述. 自动化学报, 2014, **40**(4): 561−576)

9 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

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 Zhang K H, Zhang L, Yang M H. Real-time object tracking via online discriminative feature selection. *IEEE Transactions on Image Processing*, 2013, **22**(12): 4664−4677

12 Zhang K H, Zhang L, Yang M H. Real-time compressive tracking. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: IEEE, 2012. 864−877

13 Zhang K H, Zhang L, Yang M H. Fast compressive tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(10): 2002−2015

14 Achlioptas D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *Journal of Computer and System Sciences*, 2003, **66**(4): 671−687

15 Baraniuk R, Davenport M, DeVore R, Wakin M. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 2008, **28**(3): 253−263

16 Ng A Y, Jordan M I. On discriminative vs. generative classifiers: a comparison of logistic regression and naive Bayes. In Advances in Neural Information Processing Systems. 2002, **14**: 841−848

17 Diaconis P, Freedman D. Asymptotics of graphical projection pursuit. *The Annals of Statistics*, 1984, **12**(3): 793−815

18 Kalal Z, Matas J, Mikolajczyk K. P-N learning: bootstrapping binary classifiers by structural constraints. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, California, USA: IEEE, 2010. 49−56

19 Grabner H, Grabner M, Bischof H. Real-time tracking via on-line boosting. In: Proceedings of the 17th British Machine Vision Association. Edinburgh, UK: IEEE, 2006. 47−56

20 Henriques J F, Caseiro R, Martins P, Batista J. Exploiting the circulant structure of tracking-by-detection with kernels. In: Proceedings of the 12th European Conference on Computer Vision. Florence, Italy: IEEE, 2012. 702−715

21 Kwon J, Lee K M. Visual tracking decomposition. In: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, California, USA: IEEE, 2010. 1269−1276

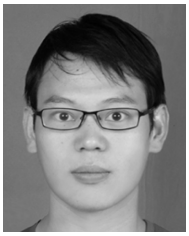
22 Hare S, Saffari A, Torr P H S. Struck: structured output tracking with kernels. In: Proceedings of the 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011. 263−270

23 Zhong W, Lu H C, Yang M H. Robust object tracking via sparsity-based collaborative model. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, Rhode Island, USA: IEEE, 2012. 1838−1845

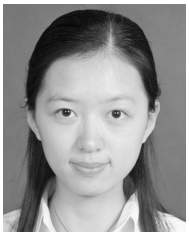
24 Collins R T. Mean-shift blob tracking through scale space. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, Wisconsin, USA: IEEE, 2003. II-234-40

25 Kwon J, Lee K M. Tracking of a non-rigid object via patch-based dynamic appearance modeling and adaptive basin hopping Monte Carlo sampling. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, Florida, USA: IEEE, 2009. 1208−1215

李庆武 河海大学物联网工程学院教授. 主要研究方向智能感知与图像处理, 目标检测与跟踪. 本文通信作者.
E-mail: Li_qingwu@163.com
(**LI Qing-Wu** Professor at the College of Internet of Things Engineering, Hohai University. His research interest covers intelligent perception and image processing, visual object detection and tracking. Corresponding author of this paper.)



朱国庆 河海大学物联网工程学院硕士研究生. 2012 年获得河海大学学士学位. 主要研究方向为数字图像处理, 目标跟踪. E-mail: zhu_gq0902@163.com
(**ZHU Guo-Qing** Master student at the College of Internet of Things Engineering, Hohai University. He received his bachelor degree from Hohai University in 2012. His research interest covers digital image processing and object tracking.)



周 妍 河海大学物联网工程学院讲师, 博士研究生. 分别于 2003 年和 2007 年在西安交通大学获得学士和硕士学位. 主要研究方向为数字图像处理和模式识别. E-mail: strangeryan@163.com
(**ZHOU Yan** Lecturer and Ph. D. candidate at the College of Internet of Things Engineering, Hohai University. She received her bachelor degree and the M. Sc. degree from Xi′an Jiaotong University, in 2003 and 2007, respectively. Her research interest covers digital image processing and pattern recognition.)

霍冠英 河海大学物联网工程学院副教授, 2012 年获得河海大学博士学位. 主要研究方向为声呐图像处理.
E-mail: huoguanying@hhu.edu.cn
(**HUO Guan-Ying** Associate professor at the College of Internet of Things Engineering, Hohai University. He received his Ph. D. degree from Hohai University in 2012. His main research interest is sonar image processing.)