

基于不确定性度量的多特征融合跟踪

顾鑫^{1,2} 王海涛¹ 汪凌峰² 王颖² 陈如冰¹ 潘春洪²

摘要 提出了一种新的基于特征不确定性度量的多特征融合跟踪算法. 首先, 针对粒子滤波跟踪算法中特征鉴别能力较弱且粒子分布相对分散时容易造成目标丢失的事实, 本文定义了一种新的特征不确定性度量方法, 该度量可以在线调整不同类型特征对跟踪结果的贡献. 同时, 针对乘性和加性特征融合跟踪算法方法中存在的缺陷, 提出了一种自适应的多特征融合方法, 融合的结果既突出了状态后验分布中目标真实状态对应的峰值, 又对噪声不敏感, 从而提高了目标跟踪的鲁棒性. 各种场景下的实验结果比较表明: 新的融合跟踪算法比单特征跟踪、乘性融合跟踪和加性融合跟踪有着更好的稳定性和鲁棒性.

关键词 目标跟踪, 不确定性度量, 粒子滤波, 多特征融合

DOI 10.3724/SP.J.1004.2011.00550

Fusing Multiple Features for Object Tracking Based on Uncertainty Measurement

GU Xin^{1,2} WANG Hai-Tao¹ WANG Ling-Feng² WANG Ying² CHEN Ru-Bing¹ PAN Chun-Hong²

Abstract This paper presents a novel tracking algorithm that fuses multiple features based on feature uncertainty measurement. It is based on the fact that tracking failure of particle filter often happens in the cases of low discriminative abilities of the observed features and disperse distributions of the sampled particles. To handle this failure, we first define a new feature uncertainty measurement method to adaptively adjust the relative contributions of different features. Then we introduce a self-adaptive feature fusion strategy to overcome the shortcomings of product and sum fusion ones. This strategy effectively sharpens the distribution of the fused posterior, and makes the tracking less sensitive to noises. Thereby, the tracking robustness is improved. An extensive number of comparative experiments show that the proposed algorithm is more stable and robust than the single feature, multiplicative fusion, and additive fusion tracking algorithms.

Key words Object tracking, uncertainty measurement, particle filter, multiple features fusion

目标跟踪在场景监控、运动捕捉、机器人自主导航、车辆自动驾驶系统、视频编辑和压缩、人机交互等众多领域有着广泛的应用^[1]. 在实际跟踪过程中, 各种类型的特征被用来表征目标, 如: 颜色^[2]、边缘^[3]、光流^[4]、纹理^[5]、运动^[6]、SIFT 特征点^[7]等. 然而, 由于背景的变化难以预测, 目标表面特性复杂, 目标特征会随着其自身运动而随时变化^[8]. 因此, 使用单一特征通常难以鲁棒地跟踪目标. 融合多种特征, 使特征之间取长补短, 是解决这一问题的有效手段.

近年来, 研究人员提出了大量基于多特征融合的跟踪算法, 其主要区别在于特征提取和融合策略的不同. 王永忠等^[9] 针对跟踪过程中的特征进行了

研究, 通过选择对当前场景区分性较高的特征模型进行目标的跟踪和模型更新. Birchfield 等^[10] 通过融合目标的颜色和梯度特征来跟踪人的头部运动, 它假设两特征在跟踪过程中对跟踪结果的贡献始终是固定的. 在此基础上, Li 等^[11] 提出了一种乘性融合策略, 在粒子滤波的框架下, 通过融合颜色、结构和边缘等多种特征, 实现对目标的跟踪, 用三种特征观测概率的乘积作为目标状态的观测似然概率. 由于采用了特征观测概率的乘积, 导致融合的结果对噪声非常敏感, 当场景中有相似目标或背景复杂时容易丢失目标. Wang 等^[12] 将该融合策略应用到粒子滤波跟踪算法中, 把颜色和边缘作为目标的特征描述, 用观测概率的加权和作为目标状态的观测似然概率. 这种加性融合的结果虽然一定程度上抑制了噪声, 但它并不能提高融合跟踪的可信度. 钟小品等^[13] 使用粒子滤波技术, 统计样本的二阶中心矩, 并求 Frobenius 范数以表征特征受噪声的污染程度, 最后适时切换乘性和加性两种融合策略.

针对加性和乘性融合中所出现的问题, 本文提出了一种新的多特征融合跟踪算法. 首先, 我们定义了一种特征不确定性度量, 该度量可以有效反映不同特征在跟踪过程中的不确定性; 然后, 为了克服加性和乘性融合策略的不足, 提出了一种新的融合策略, 该融合策略可以利用特征的不确定性自适应地调

收稿日期 2010-09-09 录用日期 2010-12-27
Manuscript received September 9, 2010; accepted December 27, 2010

国家自然科学基金 (60873161, 61005036, 61005013), 江苏省自然科学基金项目 (BK2010503), 苏州市科技局项目 (SYG201024) 资助

Supported by National Natural Science Foundation of China (60873161, 61005036, 61005013), Natural Science Foundation of Jiangsu Province of China (BK2010503), and Project of Suzhou Science Technology Bureau (SYG201024)

1. 南京航空航天大学自动化学院 南京 210016 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190

1. College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

节不同特征对跟踪结果的贡献; 最后, 将该融合策略应用于粒子滤波的跟踪框架中, 实现了对目标的跟踪. 对比乘性融合, 新的融合策略减少了对噪声的放大和对估计目标峰值的抑制. 因此能有效解决相似目标靠近时分辨困难的问题, 对目标相互遮挡或背景出现强干扰时有着较好的跟踪稳定性. 对比加性融合, 新的融合策略提高了其观测似然函数中对应于目标真实状态处的峰值的尖锐程度, 对后验概率分布的预测更加准确, 跟踪结果也因此更加可靠.

1 粒子滤波

为了定义贝叶斯跟踪问题, 考虑如下随时间变化的目标状态:

$$\mathbf{x}_t = f_t(\mathbf{x}_{t-1}, \boldsymbol{\xi}_t) \quad (1)$$

其中, \mathbf{x}_t 是目标在时刻 t 的状态向量, $\boldsymbol{\xi}_t$ 是过程噪声. 跟踪的目的是根据如下测量方程, 递归估计目标的状态:

$$\mathbf{z}_t = h_t(\mathbf{x}_t, \mathbf{e}_t) \quad (2)$$

其中, \mathbf{z}_t 是状态 \mathbf{x}_t 的观测值, \mathbf{e}_t 是观测噪声. 定义 $\{\boldsymbol{\xi}_{1:t}\}$, $\{\mathbf{e}_{1:t}\}$ 分别为过程噪声序列和测量噪声序列. 一般我们假设 $\{\boldsymbol{\xi}_{1:t}\}$ 是独立同分布的噪声序列, $\{\mathbf{e}_{1:t}\}$ 是独立同分布的噪声序列, 且两噪声序列和初始状态彼此相互独立.

从贝叶斯估计角度, 跟踪问题就是根据给定观察序列 $\{\mathbf{z}_{1:t}\}$ 递归估计目标状态 $\{\mathbf{x}_t\}$, 即构造目标状态的后验概率分布 $p(\mathbf{x}_t|\mathbf{z}_{1:t})$. 通常假设动态模型符合一阶马尔可夫状态转移, 并且观测关于状态是条件独立的. 求解过程可由以下两部分完成:

1) 预测过程

$$p(\mathbf{x}_t|\mathbf{z}_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})d\mathbf{x}_{t-1} \quad (3)$$

2) 更新过程

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \frac{p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{z}_{1:t-1})}{\int p(\mathbf{z}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{z}_{1:t-1})d\mathbf{x}_t} \quad (4)$$

对于非线性、非高斯的动态系统, 式 (3) 和式 (4) 的积分是无法解析计算的^[14]. 粒子滤波是基于序列蒙特卡洛和递归贝叶斯的估计方法^[15]. 它将积分运算转化为有限样本点的求和运算, 即后验概率分布 $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ 可近似表述为

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \sum_{j=1}^N w_{t,j} \delta(\mathbf{x}_t - \mathbf{x}_{t,j}) \quad (5)$$

N 是粒子的数目; $w_{t,j}$ 是时刻 t 第 j 个粒子的权值, 满足归一化条件, 即 $\sum_{j=1}^N w_j=1$; $\delta(\cdot)$ 是狄拉克函

数. 假如粒子从重要性分布 $q(\mathbf{x}_{t,j}|\mathbf{x}_{t-1,j}, \mathbf{z}_{1:t})$ 中采样得到, 则递归贝叶斯估计过程可看作如下权值更新过程:

$$w_{t,j} \propto w_{t-1,j} \frac{p(\mathbf{z}_t|\mathbf{x}_{t,j})p(\mathbf{x}_{t,j}|\mathbf{x}_{t-1,j})}{q(\mathbf{x}_{t,j}|\mathbf{x}_{t-1,j}, \mathbf{z}_{1:t})} \quad (6)$$

选取先验分布 $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ 作为重要性采样函数, 则权值更新过程 (6) 可以简化为

$$w_{t,j} \propto w_{t-1,j} p(\mathbf{z}_t|\mathbf{x}_{t,j}) \quad (7)$$

在算法的实际实现过程中, 选取椭圆描述目标状态 \mathbf{x} , $\mathbf{x} = \{c_x, c_y, l_x, l_y, \theta\}$. 其中, c_x, c_y, l_x, l_y 和 θ 分别是椭圆中心坐标、长轴、短轴以及偏转角. 在粒子滤波算法中, 式 (1) 的选择是很重要的. 由于难以获得任意目标的运动特性, 本文选取最简单也是最常用的一阶线性系统作为粒子滤波的状态转移模型:

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + W \quad (8)$$

其中, A 是状态转移矩阵, 在这里取其为单位矩阵. W 为均值为 0 的高斯噪声.

2 特征融合

本节通过分析粒子滤波中出现的四种情况, 定义了一种特征不确定性的度量. 同时, 借鉴乘性融合和加性融合模型, 提出了一种新的自适应特征融合跟踪算法.

2.1 特征不确定性

目标跟踪过程中, 针对不同的场景, 每种特征对目标位置估计的准确度是不一样的. 怎样度量各种特征在跟踪过程中的不确定性 (或确定性) 是一个棘手的问题.

在实际粒子滤波跟踪算法中, 当特征的鉴别能力较弱且粒子分布相对分散时容易造成目标丢失. 由此, 本文从特征的鉴别能力和粒子空间分布两个方面去分析特征在跟踪中的不确定性. 首先, 对特征的似然函数 (所有粒子在某种特征下的观测概率刻画了目标状态的一个似然函数) 而言, 不同的特征, 似然函数的尖锐程度是不同的. 似然函数的尖锐程度越高, 即观测概率值仅集中在少数几个粒子上, 说明该特征的鉴别能力越强, 此时不确定性越小. 反之不确定性越大. 其次, 对粒子的空间分布而言, 不同时刻, 粒子的聚散程度是不同的. 粒子滤波跟踪过程中, 目标的位置是由各粒子加权求和估计得到. 粒子空间分布越集中, 目标位置估计的方差越小, 不确定性越小, 反之亦然.

结合上述分析, 我们将单个特征作为一个虚拟的粒子滤波跟踪过程. 每种特征对应的粒子经过预

测后, 都重采样. 经过预测后, 观察概率值在粒子上的分布发生变化, 经过重采样后, 特征对应的粒子空间分布也发生改变. 各种特征在各时刻的分布总体表现为以下四种情况, 如图 1 所示:

1) 粒子空间分布很聚集, 则经加权求和后, 目标位置估计的方差很小; 观测概率值仅集中在少数几个优秀粒子上, 该特征鉴别能力较高; 经过重采样后, 优秀粒子被大量复制为下一步的跟踪提供好的条件, 该特征在跟踪中的不确定性最低.

2) 粒子空间分布很聚集, 目标位置估计的方差很小; 观测概率值在粒子上分布很均匀, 特征的鉴别能力不是很高; 该特征在跟踪中的不确定性高于 1).

3) 粒子空间分布很发散, 目标位置估计的方差很大; 概率值仅集中在少数几个粒子上, 该特征的鉴别能力较高; 该特征在跟踪中的不确定性也高于 1).

4) 粒子的空间位置很发散, 则经加权求和后, 目标位置估计的方差很大; 观测概率值在粒子上分布很均匀, 该特征的似然概率分布很平缓, 鉴别能力差; 该特征在跟踪中的不确定性最高.

根据上述分析, 定义了一种自适应的特征不确定性度量如下:

$$\beta_{t+1}^i = \sigma_t H(p_t^i) \quad (9)$$

其中, β_{t+1}^i 表示 $t+1$ 时刻第 i 种特征的不确定性. σ_t 是时刻 t 所有粒子空间位置的方差, 表示粒子空间分布的分散程度. $\sigma_t = \text{tr}(\Sigma)$, Σ 表示粒子两个位置坐标的协方差矩阵, σ 越大粒子的空间位置越分散. $H(p_t^i)$ 是时刻 t 第 i 种特征下所有粒子观测概率

值的熵, 表示观测概率值在粒子上的分散程度. 在信息论中, 熵是不确定性的量度:

$$H(p^i) = - \sum_{j=1}^N (p(\mathbf{z}^i | \mathbf{x}_j) \log_2 p(\mathbf{z}^i | \mathbf{x}_j)) \quad (10)$$

其中, $p(\mathbf{z}^i | \mathbf{x}_j)$ 表示第 i 种特征在第 j 个粒子状态下的观测概率值. $H(p^i)$ 越大, 则在此特征下观测概率值在粒子上的分布越均匀.

2.2 特征融合框架

在跟踪的过程中, 融合多种特征是跟踪鲁棒性的有效策略. 文献 [11] 提出了乘性融合的策略. 在目标状态给定的情况下, 假设目标的各种观测是条件独立的. 在乘性融合框架下, n 种特征融合后的观测概率如下:

$$p(\mathbf{z}^1 \cdots \mathbf{z}^n | \mathbf{x}) = \prod_{i=1}^n p(\mathbf{z}^i | \mathbf{x}) \quad (11)$$

文献 [12] 提出了加性融合的策略, 它与混合密度模型具有相同的形式. 在加性融合框架下, n 种特征融合后的观测如下:

$$p(\mathbf{z}^1 \cdots \mathbf{z}^n | \mathbf{x}) = \sum_{i=1}^n \phi^i p(\mathbf{z}^i | \mathbf{x}) \quad (12)$$

其中, ϕ^i 表示第 i 种特征所对应观测概率的权值, 满足: $\sum_{i=1}^n \phi^i = 1$.

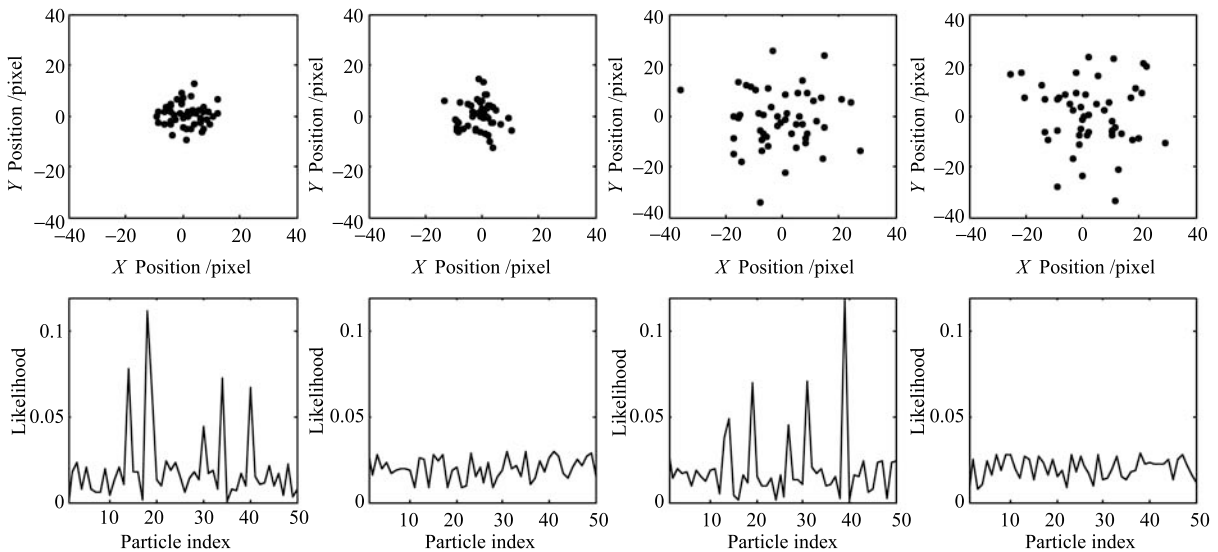


图 1 粒子滤波过程中的四种情况示意图 (上行: 粒子的空间分布; 下行: 概率值在粒子上的分布)

Fig. 1 The diagrammatic sketch of four cases in process of particle filter: the spatial distributions of particles (top row), and the distributions of probability values in particles (bottom row)

以上两种融合策略各有优缺点. 由图 2 可知: 对比两种单特征 (Feature 1, Feature 2) 的概率密度, 乘性融合 (Product rule) 使概率密度分布变得更加尖锐. 这说明融合后的概率密度鉴别能力更强了, 这一结论支持了多特征融合提高跟踪精度的观点. 然而它抑制了分布的多峰性, 放大了噪声. 当两个相似目标相互靠近甚至出现相互遮挡情况时, 或背景出现强干扰时, 乘性融合对目标位置的估计会发生很大的偏差, 甚至跟踪错目标, 这是跟踪过程中应尽量避免的. 对于加性融合 (Sum rule) 而言, 融合的结果虽然并没有提高概率密度的峰值, 但是它对噪声不敏感, 多峰也没有被抑制. 这说明对于复杂的密度估计问题, 加权求和是一种很有用的工具. 但是从概率的观点看, 加性融合模型并没有提高后验概率的置信度, 融合后概率密度的鉴别能力并没有增强, 并没有提高融合跟踪的可信度.

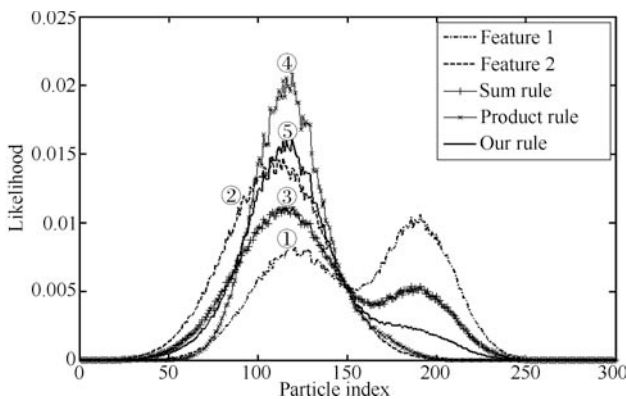


图 2 特征融合比较 (① 特征 1; ② 特征 2; ③ 加性融合结果; ④ 乘性融合结果; ⑤ 本文算法)

Fig. 2 The comparison of multiple feature fusion (① Feature 1; ② Feature 2; ③ Sum fusion result; ④ Product fusion result; ⑤ Our result.)

为了克服上述两种融合策略的缺点, 本文提出一种基于特征不确定性的自适应融合策略. 在假设 n 种观测特征是条件独立的情况下, 将每种特征的观测概率值加上一个正比于该特征不确定性的均匀分布, 做归一化处理, 再取乘积. 表达式如下:

$$p(\mathbf{z}^1 \cdots \mathbf{z}^n | \mathbf{x}) = \prod_{i=1}^n \left(\frac{p(\mathbf{z}^i | \mathbf{x}) + \beta^i U(\mathbf{x})}{1 + \beta^i} \right) \quad (13)$$

其中, β^i 表示特征 i 对应的不确定性, $U(\mathbf{x})$ 是一个离散的均匀分布. 当 $n = 2$ 时, 将式 (13) 展开得:

$$p(\mathbf{z}^1, \mathbf{z}^2 | \mathbf{x}) = \frac{1}{(1 + \beta^1)(1 + \beta^2)} \times$$

$$\left(p(\mathbf{z}^1 | \mathbf{x})p(\mathbf{z}^2 | \mathbf{x}) + \beta^1 U(\mathbf{x})p(\mathbf{z}^2 | \mathbf{x}) + \beta^2 U(\mathbf{x})p(\mathbf{z}^1 | \mathbf{x}) + \beta^1 \beta^2 (U(\mathbf{x}))^2 \right) \quad (14)$$

结合式 (14), 对于两种特征而言, 新的融合算法展开后为: 两特征的乘性融合结果, 加上两特征加性融合的结果, 再加上一个正比于两特征不确定性乘积的均匀分布. 本文的融合策略利用各种特征的不确定性调节其在观测概率中所占的比重, 其实质是将乘性融合、加性融合统一到一个自适应的框架下. 当 β^1 趋近于 0 时, 即特征 1 的不确定性达到最小, 式 (13) 的第 1 项趋近于 $p(\mathbf{z}^1 | \mathbf{x})$, 特征 1 在观测概率中所占的信息权重达到了其最大值. 当 β^1 趋近于 ∞ 时, 即特征 1 的不确定性达到最大, 式 (13) 的第 1 项趋近于 1, 即只有特征 2 在跟踪中起作用.

由图 2 可知: 对比乘性融合, 本文的融合策略 (Our rule) 对噪声放大较小, 没有抑制分布的多峰性, 能更好地处理相似目标靠近甚至相互遮挡情况. 对比加性融合, 本文的融合使观测概率分布变得更尖锐, 提高了后验概率密度的置信度, 增强了鉴别能力, 从而使跟踪结果更加可靠.

3 特征提取及算法实现

在实际应用中, 选取对目标描述差异较大且相互独立的颜色和边缘特征来表征目标. 本节首先给出两种特征的提取方法, 并由此给出目标候选区与模板相似度的计算和粒子的观测概率模型, 最后结合前文提出的多特征融合框架, 给出算法的具体实现.

3.1 特征提取

颜色是描述目标的一种主要特征, 前人对目标颜色特征的描述进行了大量研究. 文献 [1] 中提出了一种颜色直方图来描述目标的颜色特征, 它计算简单、处理快速, 对目标的部分遮挡、旋转等比较鲁棒. 其表达式为

$$\mathbf{h}^c(u) = \sum_{u=1}^{B_c} \delta(I(x, y) - u) \quad (15)$$

其中, $I(x, y)$ 是候选区像素点, B_c 表示颜色直方图的长度.

边缘是描述目标时另一种常用的特征. 本文选择在图像灰度信息下通过加权梯度方向直方图来描述目标的边缘特征^[16]. 首先, 对图像处理区域进行高斯平滑处理, 在椭圆内均匀采像素点 $I(x, y)$, 对于每个点求其边缘强度 G 和方向角 α :

$$G(x, y) = \sqrt{I_x^2 + I_y^2}, \quad \alpha(x, y) = \tan^{-1} \left(\frac{I_y}{I_x} \right) \quad (16)$$

以两个轴线为界将椭圆分为 4 个部分, 对于每个部分, 将边缘方向角量化成 B_e 级的方向直方图, 并将每部分的方向直方图对齐. 其次, 将每个点的边缘强度信息融进去, 得到各部分加权梯度方向直方图. 最后, 将四部分的直方图组合并归一化.

对于直方图描述, Bhattacharyya 距离是一种常用的相似度度量^[17]. \mathbf{h}_{tar} 和 \mathbf{h}_{mod} 分别表示当前归一化后的候选区直方图和模板直方图. 则两直方图的相似度为

$$\rho(\mathbf{h}_{\text{tar}}, \mathbf{h}_{\text{mod}}) = \sqrt{1 - \sum_{u=1}^B \sqrt{\mathbf{h}_{\text{tar}}(u)\mathbf{h}_{\text{mod}}(u)}} \quad (17)$$

在此基础上, 两种特征的观测概率定义为^[16]

$$p(\mathbf{z}^i | \mathbf{x}) \propto \exp(-\lambda_i \rho_i^2(\mathbf{h}_{\text{mod}}, \mathbf{h}_{\text{tar}})) \quad (18)$$

其中, $i \in \{1, 2\}$, 颜色是第 1 种特征, 边缘是第 2 种特征.

3.2 算法实现

跟踪算法总结如下:

步骤 1. 确定初始帧目标状态 \mathbf{x}_0 , 计算模板颜色、方向直方图 $\mathbf{h}_0^1, \mathbf{h}_0^2$, 令 $\{w_{0,j} = 1/N\}_{j=1}^N, \beta_0^1 = \beta_0^2 = 0.5$;

步骤 2. 预测: 由 $\mathbf{x}_t = A\mathbf{x}_{t-1} + W$, 根据上一帧的状态 \mathbf{x}_{t-1} 得到当前帧的预测状态 \mathbf{x}_t ;

步骤 3. 计算观测概率 $p(\mathbf{z}^1 | \mathbf{x}_{t,j}), p(\mathbf{z}^2 | \mathbf{x}_{t,j})$, 将两者分别做归一化处理, 根据式 (10) 求两特征观测概率的熵 $H(p_t^1), H(p_t^2)$;

步骤 4. 分别将颜色及边缘特征下对应的粒子重采样, 在给定处理框内计算两种特征下 N 个粒子的位置方差, 以防止个别粒子偏移较远, 引起方差过大;

步骤 5. 由式 (13) 计算两特征融合后的观测概率 $p(\mathbf{z}^1, \mathbf{z}^2 | \mathbf{x})$, 由式 (9) 计算两种特征的不确定性 β^1, β^2 , 并将两者做归一化处理;

步骤 6. 更新: 根据特征融合后求得的似然概率更新粒子的权值 $w_{t,j} = w_{t-1,j} p(\mathbf{z}_t | \mathbf{x}_{t,j})$, 计算目标的当前的估计状态: $\hat{\mathbf{x}}_t = \sum_{j=1}^N w_{t,j} \mathbf{x}_{t,j}$;

步骤 7. 根据粒子权值分布决定是否重采样, 若 $1/(\sum_{t=1}^N w_{t,j}^2) < N/2$, 则从 $\{\mathbf{x}_{t,j}\}_{j=1}^N$ 中重新采样 N 次得到 $\{\mathbf{x}_{t,j^l}\}_{l=1}^N$, 令粒子权值 $w_{t,l} = 1/N$, 否则不做处理;

步骤 8. 转步骤 2.

4 实验结果及分析

为了验证融合策略的有效性, 我们不仅将本文算法和单特征跟踪进行了比较, 还与乘性融合及加性融合跟踪进行了比较. 另外, 为了验证算法的普适性, 在数据源的选取上, 本文在公共视频集的基础上选择红外和光学两种不同的数据源. 因此, 实验包括不同数据源情况下本文算法与单特征跟踪 (实验 1 和实验 2) 的比较, 以及不同数据源情况下与乘性融合和加性融合的比较 (实验 3 和实验 4). 为叙述简洁, 定义单一以边缘为特征的跟踪为边缘跟踪, 类似地定义颜色跟踪、乘性跟踪和加性跟踪.

实验参数设置如下: 目标初始位置手动给定; 粒子数目 N 设为 100; 均匀分布 $U(\mathbf{x}) = 1/N$; 方向直方图取 $B_e = 18$; 光学视频下颜色直方图取 $B_c = 216$; 红外视频下灰度信息直方图取 $B_g = 64$; λ 取值如表 1 所示.

表 1 两特征的系数

Table 1 The coefficients of two features

对应视频	颜色特征 (λ_c)	边缘特征 (λ_e)
实验 1 视频序列	90	40
实验 2 视频序列	90	30
实验 3 视频序列	90	40
实验 4 视频序列	80	40

实验 1 是红外视频序列下的行人跟踪¹. 众所周知, 红外视频序列跟踪的难点是视频噪声较大, 导致目标易受背景干扰. 同时, 此视频还存在另一个非常棘手的问题: 行人之间的相互遮挡. 如图 3 所示, 当目标被行人部分遮挡时 (37 帧), 颜色跟踪跟丢目标, 而边缘跟踪和本文算法可以一直跟踪目标. 这是由于目标和其他行人的灰度信息十分相似, 所用的颜色特征不能准确地分辨出目标. 而边缘特征是基于目标梯度方向的信息, 虽然有部分遮挡, 但两者的边缘特征有着明显的区别. 因此, 这种情况下利用单一边缘特征可以跟踪目标, 而本文的融合策略在跟踪过程中可以自适应地提高颜色特征的不确定性, 从而减少它在跟踪过程中的信息权重, 可以一直跟踪目标.

实验 2 是光学视频序列下的车辆跟踪. 该视频序列受到诸多因素干扰. 首先, 车辆不断进出阴影区, 导致车辆表面颜色不断发生变化. 其次, 背景中存在树枝等边缘信息很强的干扰物. 这些干扰因素

¹ 视频选自 Otcbvs Benchmark Dataset Collection [Online], available: <http://www.cse.ohio-state.edu/otcbvs-bench/>, December 28, 2010

会严重影响边缘跟踪的结果. 如图 4 所示, 当车辆进入阴影区时, 其表面光照产生了很大变化, 虽然颜色跟踪在 121 帧没有跟丢目标, 但估计出的目标位置已经明显偏离了目标的真实位置. 而当车辆再次进入阴影区时 (346 帧), 颜色跟踪完全失效了. 对于边缘跟踪而言, 由于背景中树枝等强边缘信息的干扰, 从边缘获取的信息非常不可靠, 所以, 在第 246 帧, 边缘跟踪也完全跟丢了目标. 因此, 由于环境的复杂性, 使用单一特征很难获得鲁棒的跟踪结果. 而本文的融合策略在跟踪过程中可以根据外界环境的改变, 自适应地调整特征的不确定性, 进而调整其在观测概率中所占的权重. 所以, 在整个视频序列中, 本文的算法可以较鲁棒地跟踪目标.

为了进一步说明在跟踪过程中特征不确定性与

外界环境变化的关系. 我们给出了实验 2 中两种特征的不确定性 (见图 5). 第 1 行是不同阶段的视频样本, 第 2 行是两特征的不确定性. 明显可以看出, 特征的不确定性随着环境的变化时刻在发生着改变, 我们可以把整个视频序列大致分为五个阶段: 第一阶段, 目标刚进入视场区, 目标的颜色特征很明显, 此时, 其不确定性要明显小于边缘特征; 第二阶段, 目标长时间进入阴影区, 此时边缘特征不确定性较小; 第三阶段, 目标不断地进出阴影区, 颜色特征和边缘特征的不确定性交替变化; 第四阶段, 目标出阴影区, 边缘特征的不确定性要高于颜色特征; 第五阶段和第三阶段类似. 因此, 由于目标所处环境的复杂性, 利用单一特征跟踪通常难以获得满意的效果.

实验 3 选自红外视频序列下的行人跟踪 (视频

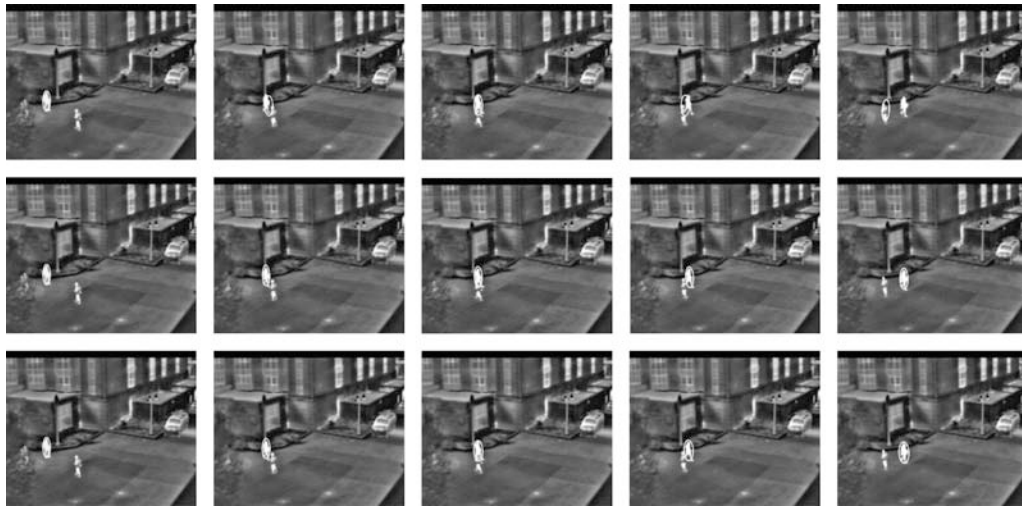


图 3 实验 1 跟踪结果 (第 1 行: 颜色跟踪; 第 2 行: 边缘跟踪; 第 3 行: 本文算法) (帧数: 1, 28, 37, 43, 61)

Fig. 3 Some results on experiment 1 by using single feature: color (the first row), edge (the second row), and proposed multi-feature fusion method (the third row) (Frames: 1, 28, 37, 43, 61)

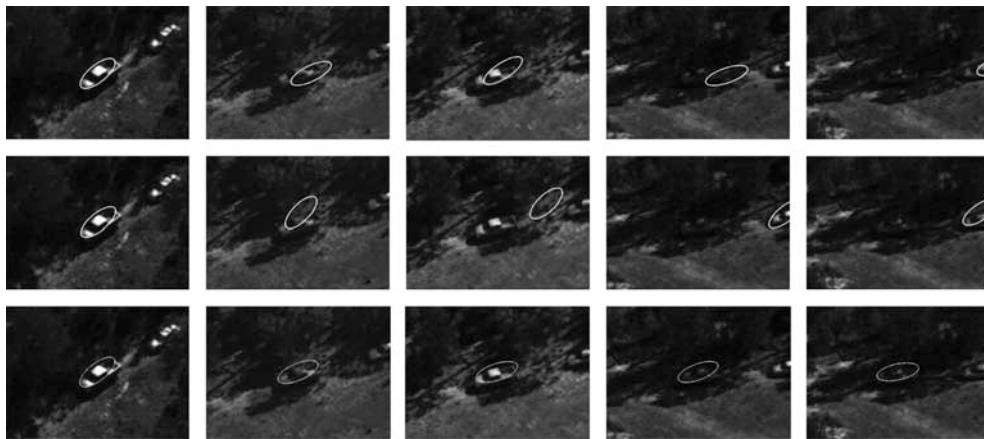


图 4 实验 2 跟踪结果 (第 1 行: 颜色跟踪; 第 2 行: 边缘跟踪; 第 3 行: 本文算法) (帧数: 6, 206, 246, 346, 376)

Fig. 4 Some results on experiment 2 by using single feature: color (the first row), edge (the second row), and proposed multi-feature fusion method (the third row) (Frames: 6, 206, 246, 346, 376)

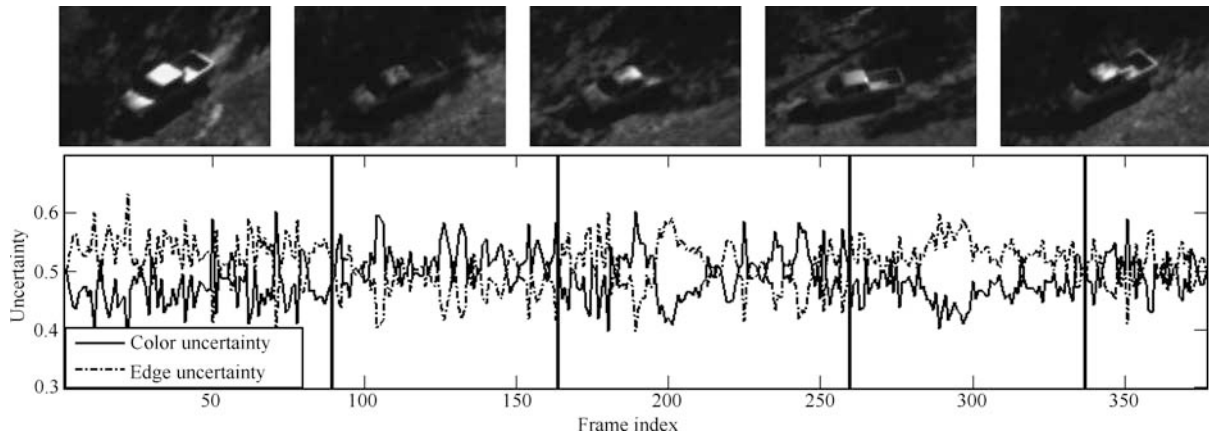


图 5 五个阶段中有代表的视频帧(上行), 两种特征的不确定性(下行)

Fig. 5 The representative video frames in the five stages (top row), and the uncertainty of two features (bottom row)

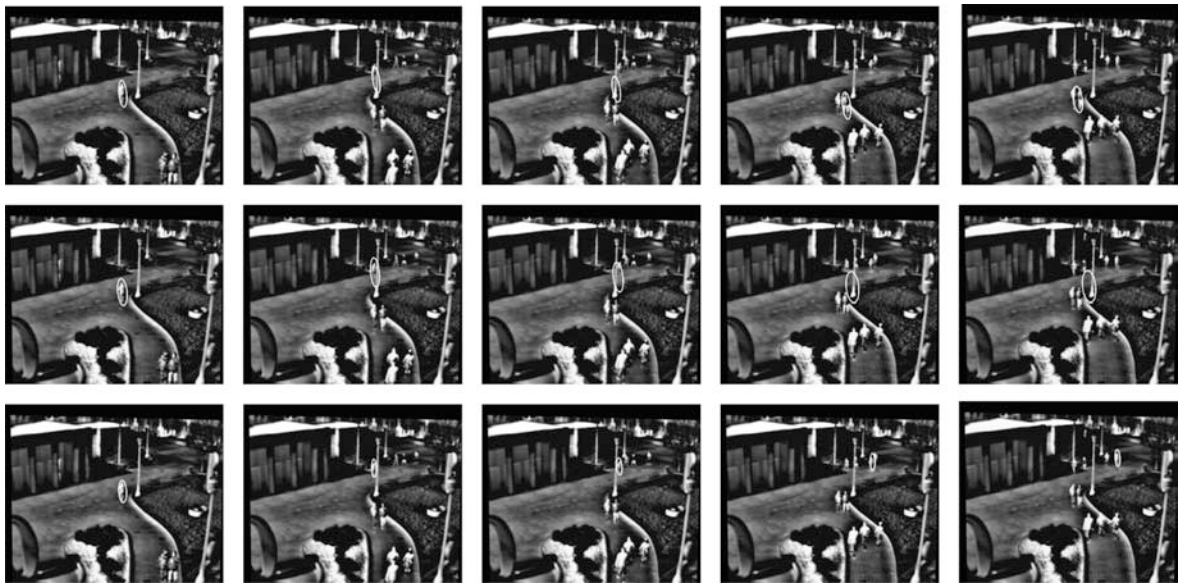


图 6 实验 3 部分跟踪结果(第 1 行: 乘性跟踪; 第 2 行: 加性跟踪; 第 3 行: 本文算法)(帧数: 251, 456, 486, 546, 386)

Fig. 6 Some results on experiment 3 by using product rule (the first row), sum rule (the second row), and proposed multi-feature fusion method (the third row) (Frames: 251, 456, 486, 546, 586)

出处同实验 1)。该跟踪主要难点来自其他相似行人的干扰和路灯的遮挡。如图 6 所示, 当目标经过路灯时(486 帧), 乘性跟踪和加性跟踪都跟丢目标。这是由于路灯对目标有部分遮挡, 从而使得路灯的干扰特征(边缘、灰度)和目标自身非常相似, 这种情况下, 乘性融合的结果放大了噪声, 抑制目标的真实分布; 而加性融合则降低了跟踪的可信度。因此, 两者都跟丢了目标。而本文的融合策略则有效地避免了这些问题, 当目标遇到和自身十分相似的行人或路灯干扰时, 两个特征的不确定性会自适应地进行调整, 融合的结果没有过分放大噪声, 可以将目标和行人或路灯有效地区分开来。另外, 这种融合策略还保持了确定性特征的优先权, 为下一步跟踪提供了很好的先验。

实验 4 选自光学视频序列下的车辆跟踪。该视频跟踪中遇到的难点主要是树木对车辆的遮挡和背景中边缘信息的干扰。如图 7 所示, 当目标经过对其有遮挡的树木时(116 帧), 加性跟踪跟丢目标。这是由于树木对车辆的遮挡导致车辆自身颜色特征发生突变, 而背景干扰又增加了边缘信息的不确定性, 这种情况下, 加性融合的结果并不能提高后验概率的置信度, 因此, 目标的丢失是必然的。目标刚出第二个遮挡区时(291 帧), 由于背景的颜色特征和目标自身很相似, 只能结合边缘特征对目标轮廓的捕捉跟踪目标。而乘性融合在相似特征干扰下对噪声太敏感, 最终跟丢目标。本文算法虽然由于外界干扰在个别帧对目标位置的估计偏差可能较大, 但可以马上恢复跟踪。其主要原因是: 本文的融合策略不仅

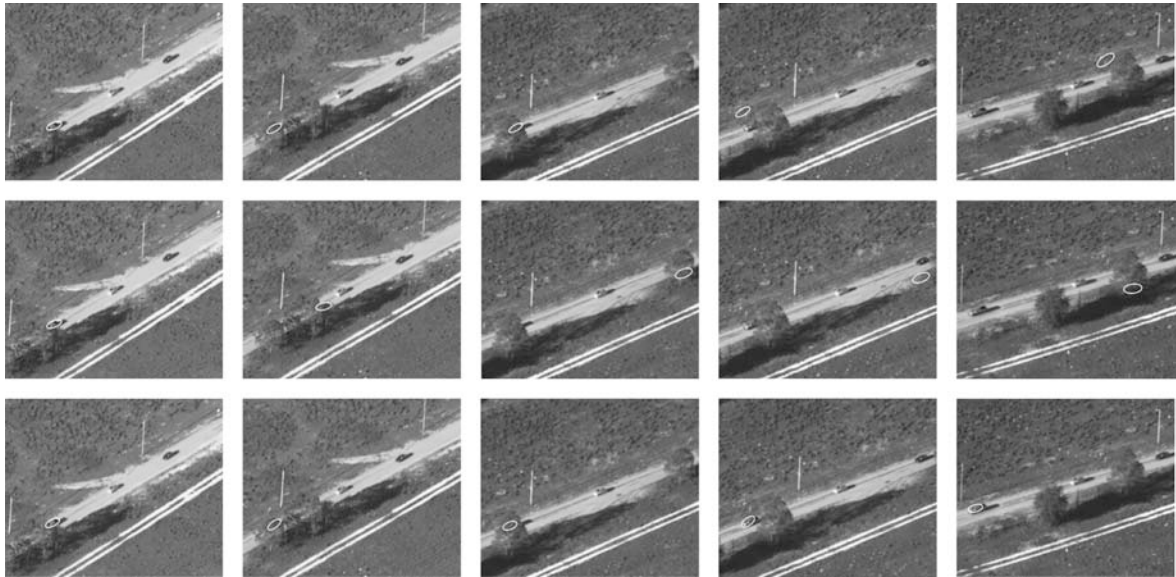


图 7 实验 4 部分跟踪结果 (第 1 行: 乘性跟踪; 第 2 行: 加性跟踪; 第 3 行: 本文算法) (帧数: 61, 116, 271, 291, 396)
 Fig.7 Some results on experiment 4 by using: product rule (the first row), sum rule (the second row), and proposed multi-feature fusion method (the third row) (Frames: 61, 116, 271, 291, 396)

表 2 实验结果对比 (“-” 表示未在此视频上做测试)
 Table 2 Comparison results (“-” interprets that we did not perform in this sequence)

实验	视频源	视频长度	颜色跟踪	边缘跟踪	乘性跟踪	加性跟踪	本文算法
实验 1	红外视频	共 150 帧	第 31 帧 (跟丢)	无丢失	-	-	无丢失
实验 2	光学视频	共 378 帧	第 346 帧 (跟丢)	第 246 帧 (跟丢)	-	-	无丢失
实验 3	红外视频	共 647 帧	-	-	第 456 帧 (跟丢)	第 486 帧 (跟丢)	无丢失
实验 4	光学视频	共 393 帧	-	-	第 291 帧 (跟丢)	第 116 帧 (跟丢)	无丢失

提高了后验概率的可信度, 而且可以在跟踪过程中自适应地调整颜色及边缘特征的不确定性, 即使在相似背景的干扰下, 仍然可以较鲁棒地跟踪目标.

表 2 给出了四组实验下颜色跟踪、边缘跟踪、乘性跟踪、加性跟踪和本文算法跟踪结果对比 (由于篇幅限制, 在这里并没有在每段视频上将所有算法都做测试). 实验结果表明: 无论是在光学或红外视频序列下, 由于光照、背景边缘等外界因素干扰, 单个特征 (如颜色、边缘) 的跟踪结果往往是不稳定的, 通常情况下难以确定哪种特征的跟踪效果更好, 通过融合多种特征, 使特征之间取长补短, 能够有效地解决这些难点. 同时, 对比乘性和加性两种传统的融合策略, 本文的算法既提高了后验概率密度的置信度, 又对干扰噪声不太敏感, 其跟踪鲁棒性更好.

本文中的所有算法均在 P4 3.4 GHz 1024 M 内存的计算机上用 Matlab 7.6 仿真实现. 需要指出的是, 本文算法的复杂度取决于粒子的数目、采样点

个数、所选特征个数、每种特征子模型的计算复杂度及目标的大小, 显然随着粒子数目的增加, 会带来整个算法计算量的增加, 从而对跟踪的实时性造成一定的影响. 表 3 给出了不同算法的跟踪速度 fps (帧/秒). 可以看出本文算法的复杂度和其他融合算法复杂度基本相近.

表 3 5 种算法的计算代价对比
 Table 3 The comparison of computation costs of the five methods

视频	颜色跟踪 (fps)	边缘跟踪	加性跟踪	乘性跟踪	本文算法
实验 1	25.8	24.7	11.5	11.4	10.9
实验 2	21.3	21.5	10.1	10.0	9.7
实验 3	24.9	24.1	10.9	11.0	10.3
实验 4	25.6	24.5	12.3	12.1	11.8

5 结论

本文通过分析粒子的概率及空间分布的特性,定义了一种特征不确定性的度量,并在此基础上提出一种新的自适应多特征融合跟踪算法.该算法将乘性和加性两种传统的融合方式统一到一个框架下,根据实际场景的变化,自适应地调节特征在跟踪中的不确定性,进而调整特征对跟踪结果的贡献.对常见不同源视频序列(红外和光学)的测试表明,本文算法对背景干扰、部分遮挡、光照变化等复杂环境均具有较好的鲁棒性.此外,本文跟踪算法亦可应用于其他跟踪环境,如基于多摄像头的融合跟踪、基于多源视频数据的融合跟踪等;同时,这一融合策略也可运用于计算机视觉的其他领域,如融合识别、检测等.

References

- Hou Zhi-Qiang, Han Chong-Zhao. A survey of visual tracking. *Acta Automatica Sinica*, 2006, **32**(4): 603–617
(侯志强, 韩崇昭. 视觉跟踪技术综述. 自动化学报, 2006, **32**(4): 603–617)
- Perez P, Hue C, Vermaak J, Gangnet M. Color-based probabilistic tracking. In: Proceedings of the 7th European Conference on Computer Vision. London, UK: Springer, 2002. 661–675
- Kim B G, Park D J. Unsupervised video object segmentation and tracking based on new edge features. *Pattern Recognition Letters*, 2004, **25**(15): 1731–1742
- Baker S, Matthews I. Lucas-Kanade 20 years on: a unifying framework. *International Journal of Computer Vision*, 2004, **56**(3): 221–255
- Bastos R, Dias J M S. Fully automated texture tracking based on natural features extraction and template matching. In: Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology. New York, USA: ACM, 2005. 180–183
- Du W, Piater J. A probabilistic approach to integrating multiple cues in visual tracking. In: Proceedings of the 10th European Conference on Computer Vision. Berlin, Germany: Springer, 2008. 225–238
- Lin Hai-Feng, Ma Yu-Feng, Song Tao. Research on object tracking algorithm based on SIFT. *Acta Automatica Sinica*, 2010, **36**(8): 1204–1208
(简海峰, 马宇峰, 宋涛. 基于 SIFT 特征目标跟踪算法研究. 自动化学报, 2010, **36**(8): 1204–1208)
- Yilmaz A, Javed O, Shah M. Object tracking: a survey. *ACM Computing Surveys*, 2006, **38**(4): 13–58
- Wang Yong-Zhong, Liang Yan, Zhao Chun-Hui, Pan Quan. Kernel-based tracking based on adaptive fusion of multiple cues. *Acta Automatica Sinica*, 2008, **34**(4): 393–399
(王永忠, 梁彦, 赵春晖, 潘泉. 基于多特征自适应融合的核跟踪方法. 自动化学报, 2008, **34**(4): 393–399)
- Birchfield S. Elliptical head tracking using intensity gradients and color histograms. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Santa Barbara, USA: IEEE, 1998. 232–237
- Li P H, Chaumette F. Image cues fusion for object tracking based on particle filter. In: Proceedings of the 3rd International Workshop on Articulated Motion and Deformable Objects. Palma de Mallorca, Spain: Springer, 2004. 99–107
- Wang X, Tang Z M. Modified particle filter-based infrared pedestrian tracking. *Infrared Physics and Technology*, 2010, **53**(4): 280–287
- Zhong Xiao-Pin, Xue Jian-Ru, Zheng Nan-Ning, Ping Lin-Jiang. An adaptive fusion strategy based multiple-cue tracking. *Journal of Electronics and Information Technology*, 2007, **29**(5): 1017–1021
(钟小品, 薛建儒, 郑南宁, 平林江. 基于融合策略自适应的多线索跟踪方法. 电子与信息学报, 2007, **29**(5): 1017–1021)
- Arulampalam M S, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 2002, **50**(2): 174–188
- Isard M, Blake A. Condensation—conditional density propagation for visual tracking. *International Journal of Computer Vision*, 1998, **29**(1): 5–28
- Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 886–893
- Perez P, Vermaak J, Blake A. Data fusion for visual tracking with particles. *Proceedings of the IEEE*, 2004, **92**(3): 495–513



顾鑫 南京航空航天大学自动化学院硕士研究生. 主要研究方向为目标跟踪和智能视频监控. 本文通信作者.

E-mail: xingu396@gmail.com

(**GU Xin** Master student at the College of Automation, Nanjing University

of Aeronautics and Astronautics. His research interest covers visual object tracking and intelligent video surveillance. Corresponding author of this paper.)



王海涛 南京航空航天大学自动化学院副教授. 主要研究方向为光电检测技术和计算机控制技术.

E-mail: htwang2002@126.com

(**WANG Hai-Tao** Associate professor at the College of Automation, Nan-

jing University of Aeronautics and Astronautics. His research interest covers photoelectric detecting technology and computer control technique.)



汪凌峰 中国科学院自动化研究所博士研究生. 主要研究方向为视频监控和医学图像处理.

E-mail: lfwang@nlpr.ia.ac.cn

(**WANG Ling-Feng** Ph.D. candidate at the Institute of Automation,

Chinese Academy of Sciences. His research interest cov-

ers video surveillance and medical image processing.)



王颖 中国科学院自动化研究所博士研究生. 主要研究方向为图像处理, 概率图模型和目标跟踪.

E-mail: ywang@nlpr.ia.ac.cn

(**WANG Ying** Ph.D. candidate at the Institute of Automation, Chinese

Academy of Sciences. His research interest covers image processing, probabilistic graphical models, and object tracking.)



陈如冰 苏州市计量测试研究所工程师. 主要研究方向为测试计量技术.

E-mail: chenrb@szjl.com.cn

(**CHEN Ru-Bing** Engineer at Suzhou Institute of Measurement and

Testing Technology. His research interest covers measuring and testing technologies.)



潘春洪 中国科学院自动化研究所研究员. 主要研究方向为图像处理和计算机视觉. E-mail: chpan@nlpr.ia.ac.cn

(**PAN Chun-Hong** Professor at the Institute of Automation, Chinese Academy of Sciences. His research in-

terest covers image processing and computer vision.)