

动态背景下基于光流场分析的运动目标检测算法

崔智高 王华 李艾华 王涛 李辉

Moving object detection based on optical flow field analysis in dynamic scenes

Cui Zhi-Gao Wang Hua Li Ai-Hua Wang Tao Li Hui

引用信息 Citation: *Acta Physica Sinica*, 66, 084203 (2017) DOI: 10.7498/aps.66.084203

在线阅读 View online: <http://dx.doi.org/10.7498/aps.66.084203>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2017/V66/I8>

您可能感兴趣的其他文章

Articles you may be interested in

Fresnel 波带片编码成像的高分辨重建

High-resolution reconstruction of Fresnel zone plate coded imaging

物理学报.2017, 66(5): 054203 <http://dx.doi.org/10.7498/aps.66.054203>

单像素探测频谱重构成像理论分析

Theoretical analysis of spectrum reconstruction imaging using single-pixel detection

物理学报.2017, 66(3): 034201 <http://dx.doi.org/10.7498/aps.66.034201>

贝叶斯迭代联合双边滤波的散焦图像快速盲复原

Fast Bayesian blind restoration for single defocus image with iterative joint bilateral filters

物理学报.2016, 65(23): 234202 <http://dx.doi.org/10.7498/aps.65.234202>

基于残差矩阵估计的稀疏表示目标跟踪算法

Visual tracking based on the estimation of representation residual matrix

物理学报.2016, 65(19): 194201 <http://dx.doi.org/10.7498/aps.65.194201>

基于激光散斑成像的零件表面粗糙度建模

Surface roughness modeling based on laser speckle imaging

物理学报.2015, 64(23): 234203 <http://dx.doi.org/10.7498/aps.64.234203>

动态背景下基于光流场分析的运动目标检测算法*

崔智高^{1)2)†} 王华¹⁾ 李艾华¹⁾ 王涛¹⁾ 李辉¹⁾

1)(火箭军工程大学, 西安 710025)

2)(清华大学自动化系, 北京 100084)

(2016年10月21日收到; 2017年1月24日收到修改稿)

针对现有动态背景下运动目标检测算法的不足, 提出一种基于光流场分析的运动目标检测算法. 首先根据前背景在光流梯度幅值和光流矢量方向上的差异确定目标的大致边界, 然后通过点在多边形内部原理获得边界内部的稀疏像素点, 最后以超像素为节点, 利用混合高斯模型拟合的表观信息和超像素的时空邻域关系构建马尔可夫随机场模型的能量函数, 并通过使目标函数能量最小化得到最终的运动目标检测结果. 该算法不需要任何先验假设, 能够同时处理动态背景和静态背景两种情况. 多组实验结果表明, 本文算法在检测的准确性和处理速度上均优于现有算法.

关键词: 动态背景, 运动目标检测, 光流场分析, 马尔可夫随机场模型

PACS: 42.30.Tz, 07.05.Pj, 02.50.-r

DOI: 10.7498/aps.66.084203

1 引言

运动目标检测是指从视频序列中提取出感兴趣的运动物体或区域, 是后期实现目标跟踪、行为分析的基础^[1-3]. 在实际应用中, 根据摄像机运动与否, 可分为静态背景下的运动目标检测两类: 应用于静态背景下的运动目标检测方法主要有帧差法和背景差分法^[4,5]; 而动态背景下由于摄像机的不规则运动会造成背景和前景目标的相对运动, 给目标的检测带来了非常严峻的挑战. 因而致力于研究摄像机运动下的目标检测方法具有非常重要的意义^[6].

动态背景下的运动目标检测方法主要分为基于背景补偿的方法、基于初始背景模型构造的方法和基于运动线索的方法三类^[7], 其中第三种是当前的主流方法和研究难点, 此类方法一般以视频序列获得的像素点运动轨迹作为运动线索的基本载体. 例如: Lee等^[8]首先根据像素点运动轨迹确定类似目标的候选关键区域, 然后计算候选关键区域的二

值化分割, 从而获得具有稳定外观和持续运动的假设组, 最后使用已被排序的假设组得到所有帧像素级的目标检测结果, 该算法的不足之处在于其准确率对物体的外观假设和位置先验依赖较大; Li等^[9]通过条件随机场模型将利用像素点运动轨迹获得的目标内部稀疏像素点进行有效集成, 算法不需要通过任何训练数据来获取先验知识和条件假设, 能够鲁棒处理前景目标形状和姿态的任意变化, 但是该方法在运动轨迹稀疏区域会出现大块的误检测; Zhang等^[10]首先根据运动目标的空间连续性和运动轨迹的局部平滑性建立目标样本集, 然后利用所有的目标样本集建立层状有向无环图, 图中最长的路径满足运动评分函数最大且代表了可能性最大的目标样本, 最后这些目标样本被用于建立目标和背景的混合高斯模型, 并利用最优化图割方法求解模型获取准确的像素级分割结果, 该算法要求每一环节中的参数设置都必须准确合理, 一般在实际场景检测中较难实现; Elqursh和Elgammal^[11]提出了一种基于轨迹聚类分析和颜色模型迭代学习的运动检测方法, 其中作者在使用运动轨迹时未用到

* 国家自然科学基金(批准号: 61501470)资助的课题.

† 通信作者. E-mail: cuizg10@tsinghua.edu.cn

未来信息, 即轨迹是在线延长的, 因此在视频序列的初始阶段, 由于运动轨迹缺乏足够的运动信息, 容易造成聚类错误并影响后续的颜色模型学习和前景背景分割; 高文等^[12]将目标检测问题视为一种更普遍的二分类问题, 并利用 1 bit BP 特征通过三个级联分类器实现运动目标检测, 该算法对摄像机微小晃动、背景模糊等复杂情况具有良好的检测效果, 但在背景剧烈变化时检测精度较低; 崔智高等^[13]首先利用多组单应约束对背景运动进行建模, 然后通过累积确认策略实现前背景轨迹的准确分离, 最后将轨迹分离信息和超像素的时空邻域关系统一建模在以超像素为节点的马尔可夫随机场模型中, 求解模型得到最终的前背景标记结果, 该算法的计算复杂度较高, 并且在前背景的边缘区域会出现较大的误检率.

众所周知, 像素点运动轨迹的提取基于帧间获取的光流场^[14,15], 即首先求得帧间光流场, 然后利用匹配方法^[16-18]获得像素点之间的匹配对应, 因此若直接以帧间光流场作为运动线索的基本载体, 则可以有效避免匹配过程中的误差累积和时间消耗. 基于上述思想, 本文提出一种基于光流场分析的运动目标检测方法. 算法首先利用光流的梯度幅值和矢量方向确定前景目标的大致边界, 然后根据点在多边形内部原理获得边界内部的稀疏像素点, 最后以超像素为节点构建马尔可夫随机场模型的目标能量函数, 利用混合高斯模型构建数据项, 利用超像素时空邻域关系构建平滑项, 并通过使目标函数能量最小化得到最终的运动目标检测结果. 本文所提算法不需要物体运动和场景估计的先验假设, 并且在静态场景和动态场景下均能实现准确、鲁棒的运动目标检测.

2 基于光流梯度幅值和光流矢量方向的目标边界检测

对于摄像机运动的场景, 背景所对应的光流场是由背景运动产生的, 而目标所对应的光流场则是上述运动与场景中目标运动叠加产生的, 二者的光流矢量存在着较大差异, 因而可通过对光流矢量的分析确定背景与目标的大致边界.

基于上述思想, 本文提出了一种基于光流梯度幅值和光流矢量方向的目标边界检测方法. 首先利用文献^[19]提出的算法计算视频序列的光流场, 若视频序列包括 N 帧图像, 则第 t 帧图像坐

标 (u, v) 处的光流场矢量可表示为 $f_t(u, v)$, 其中 $1 \leq t \leq N - 1$. 本文把获得的光流场分为两类: 由摄像机运动产生的背景光流场和由运动物体产生的目标光流场. 本节将通过对光流梯度幅值和光流矢量方向的分析, 获得背景光流场和目标光流场的大致边界.

2.1 光流梯度幅值确定边界

尽管目标运动和背景运动具有较大差异性, 但目标内部像素点的运动或者是背景内部像素点的运动则具有高度一致性, 具体表现在目标与背景边缘区域光流矢量梯度的幅值是较大的, 其余区域则接近 0. 因而可通过设置合适的阈值, 将梯度幅值超过阈值的像素点确定为边界点. 基于上述分析, 本文引入目标边界强度系数 $s_t^m(u, v) \in [0, 1]$,

$$s_t^m(u, v) = \frac{e^{\eta_m \cdot \|\nabla f_t(u, v)\|} - 1}{e^{\eta_m \cdot \|\nabla f_t(u, v)\|}}, \quad (1)$$

其中, $\|\nabla f_t(u, v)\|$ 表示像素点 $f_t(u, v)$ 的光流梯度幅值, η_m 表示将 $s_t^m(u, v)$ 控制在 $[0, 1]$ 范围内的参数.

2.2 光流矢量方向确定边界

背景内部像素点或目标内部像素点的光流矢量方向基本趋于一致, 而在背景与目标的边界区域, 光流矢量方向的差异则较为明显. 因此, 可将当前像素点的光流矢量方向与其 8 邻域像素点的光流矢量方向做比较, 获取最大的夹角值, 并将夹角超过阈值的像素点确定为边界点. 基于上述分析, 本文引入另一个目标边界强度系数 $s_t^a(u, v) \in [0, 1]$,

$$s_t^a(u, v) = \frac{e^{\eta_a \cdot \max \theta(f_t(u, v), \Phi_t(u, v))} - 1}{e^{\eta_a \cdot \max \theta(f_t(u, v), \Phi_t(u, v))}}, \quad (2)$$

其中, $\max \theta(f_t(u, v), \Phi_t(u, v))$ 表示像素点 $f_t(u, v)$ 与其 8 邻域像素点集合 $\Phi_t(u, v)$ 的最大夹角, η_a 表示将 $s_t^a(u, v)$ 控制在 $[0, 1]$ 范围内的参数.

一般情况下, 利用上述的其中一种强度系数即可实现目标边界的检测, 但在实际场景中往往存在各种噪声的干扰. 为了提高鲁棒性, 本文将光流梯度幅值确定的边界和光流矢量方向确定的边界进行融合处理, 并通过阈值判断得到目标边界的二值图, 如下式所示:

$$s_t(u, v) = \begin{cases} 1, & s_t^m(u, v) \cdot s_t^a(u, v) \geq \eta, \\ 0, & \text{其他}, \end{cases} \quad (3)$$

其中1代表边界点, 0代表非边界点, 阈值 η 取值范围为 $[0, 1]$.

3 基于点在多边形内部原理的目标像素点判断

在理想情况下, 通过上述步骤获得的目标边界应与目标实际轮廓相重合, 但由于图像噪声、光流估计误差、阈值判断等多种因素的影响, 二者的边界曲线往往存在较大误差, 并且经过上述步骤获得的目标边界通常不是闭合的. 图1给出了一个具体的例子, 其中图1(a)为People2视频序列^[20]中的第3帧图像, 图1(b)为其对应的目标实际轮廓, 图1(c)为利用上节提出的基于光流梯度幅值和光流矢量方向的目标边界检测方法得到的目标边界.



图1 目标边界检测结果示例

Fig. 1. An example of the object boundary detection.

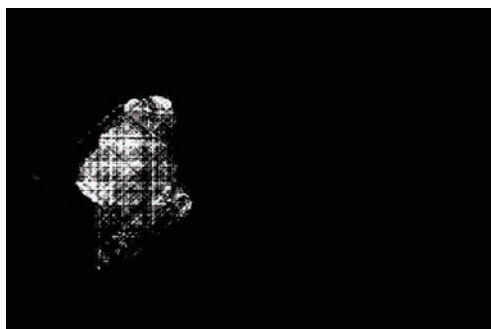


图2 目标内部像素点检测的实验结果

Fig. 2. Experimental result of internal points detection.

4 基于时空马尔可夫随机场模型的前背景像素标记

通过上述步骤只能获得稀疏的目标像素点. 为了进一步对每个像素信息进行前背景标记, 本文首

为了解决上述问题, 本文利用点在多边形内部原理确定目标内部的像素点. 其核心思想是从一点出发沿水平或垂直方向引出一条射线, 若该射线与多边形边的交点数目为奇数, 则判断该点在多边形内部, 否则判断该点在多边形外部. 基于上述点在多边形内部原理, 本文将每个像素点间隔 45° , 分别从8个方向引出射线, 若8个方向引出射线与目标边界交点为奇数的方向超过5个, 则认为该点在目标边界内部, 从而得到目标内部稀疏的像素点. 上述方法通过多个方向的综合判断得到像素点的位置, 可以有效避免部分边界不连续或者图像噪声造成的误判断, 增强算法的准确性和鲁棒性. 图2为图1所示图像对应的目标内部像素点, 其中目标内部像素点以白色星形显示.

先利用SLIC算法^[21,22]对视频序列进行过分割得到超像素集合, 然后以超像素为节点构建时空马尔可夫随机场模型的能量函数, 最后通过使能量函数最小化得到最终的前背景像素标记结果.

4.1 时空马尔可夫随机场模型能量函数设计

设第 t 帧图像对应的超像素集合为 $\mathfrak{R}_t$, 则 $\mathfrak{R}_t$ 中的每个索引为 i 的超像素 r_t^i 对应一个分类标签 $l_t^i \in \{0, 1\}$, 0表示背景, 1表示前景目标. 此时, 以超像素为节点可构建时空马尔可夫随机场模型的能量函数, 如下式所示:

$$E(L) = \sum_{t,i} \sum_{c=0}^1 A_t^c(r_t^i) + \sum_{t,i} \sum_{r_t^j \in \psi(r_t^i)} S_t^{i,j}(r_t^i, r_t^j) + \sum_{t,i} \sum_{r_{t+1}^j \in \zeta(r_t^i)} T_t^{i,j}(r_t^i, r_{t+1}^j), \quad (4)$$

其中, $L = \{l_t^i\}_{t,i}$ 表示视频序列所有超像素的标记结果; $\psi(r_t^i)$ 和 $\zeta(r_t^i)$ 代表超像素 r_t^i 的空域近邻集合和时域近邻集合; $A_t^c(\cdot)$, $S_t^{i,j}(\cdot)$ 和 $T_t^{i,j}(\cdot)$ 分别表示数据项、空域平滑项和时域平滑项势能函数。

4.2 数据项势能函数设计

数据项势能函数反映了超像素标记结果与第3节获得的目标内部像素点的符合程度。基于此, 本文首先计算每帧图像超像素 r_t^i 包含已获得的目标内部像素点的比例系数 p_t^i , 并将该比例与两个设定阈值进行比较, 从而将超像素初步分类为前景超像素和背景超像素, 如下式所示:

$$r_t^i = \begin{cases} \text{前景超像素, } p_t^i \geq T_1, \\ \text{背景超像素, } p_t^i \leq T_2. \end{cases} \quad (5)$$

然后本文利用包含超像素均值颜色和质心坐标的5维向量代表每个前景超像素, 并通过所有前景超像素为每帧图像构建前景混合高斯表观模型 g_t^1 。考虑到比例系数 p_t^i 越大, 其属于前景超像素的概率越高, 在混合高斯模型中的贡献也应更大, 为此本文在构建前景混合高斯表观模型时, 为每个前景超像素引入权重系数 λ_t^i ,

$$\lambda_t^i = \frac{p_t^i}{\sum_i p_t^i}. \quad (6)$$

背景混合高斯表观模型 g_t^0 的估计与上述过程类似, 只是权重系数 λ_t^i 表示为

$$\lambda_t^i = \frac{1 - p_t^i}{\sum_i (1 - p_t^i)}. \quad (7)$$

在每帧图像估计出前景和背景混合高斯表观模型后, 即可计算出该帧图像中每个超像素对应的数据项势能函数 $A_t^c(r_t^i)$,

$$A_t^c(r_t^i) = -\log(g_t^c(r_t^i)) \cdot \delta(l_t^i, c), \quad (8)$$

其中, $\delta(\cdot)$ 为 Kronecker delta 函数。(8) 式表明, 如果某个超像素被赋予更加符合其表观模型的标签, 那么它的数据项势能函数将更小, 从而使得整体能量函数最小。

4.3 平滑项势能函数设计

平滑项势能函数用于编码相邻超像素之间的标记连续性, 又可分为空域平滑势能函数和时域平滑势能函数两类。在空域平滑方面, 考虑到同帧图

像中各区域颜色是平滑渐变的, 因此相邻超像素应具有相同的分类标签。若设超像素 r_t^i 和其空域近邻超像素 $r_t^j \in \psi(r_t^i)$ 的质心坐标分别为 c_t^i 和 c_t^j , 均值颜色分别为 h_t^i 和 h_t^j , 那么空域平滑势能函数可定义为

$$S_t^{i,j}(r_t^i, r_t^j) = \left(1 - \delta(l_t^i, l_t^j)\right) \cdot \frac{1}{\|c_t^i - c_t^j\|} \times \frac{1}{e^{\|h_t^i - h_t^j\|}}. \quad (9)$$

在时域平滑方面, 考虑到视频序列的连续性, 时域近邻的超像素也应具有相同的分类标签。若设超像素 r_t^i 经过帧间光流补偿在下一帧图像的映射区域^[23,24]与 r_t^i 时域近邻超像素 $r_{t+1}^j \in \zeta(r_t^i)$ 的重合面积为 $Q(r_t^i, r_{t+1}^j)$, 超像素 r_{t+1}^j 的均值颜色为 h_{t+1}^j , 那么时域平滑势能函数可定义为

$$T_t^{i,j}(r_t^i, r_{t+1}^j) = \left(1 - \delta(l_t^i, l_{t+1}^j)\right) \cdot Q(r_t^i, r_{t+1}^j) \times \frac{1}{e^{\|h_t^i - h_{t+1}^j\|}}. \quad (10)$$

对每个超像素建立势能函数后, 本文利用图割算法^[25,26]求解(11)式能量函数最小化问题, 得到每个超像素的最优分类结果。

$$L^* = \arg \min_L E(L). \quad (11)$$

将上文所述的时空马尔可夫随机场模型应用到图2所示图像中, 可以得到图3所示结果。图中超像素之间的边界用黄色线段表示, 背景区域用暗红色表示, 运动目标区域则保持原有颜色。



图3 运动目标检测结果示例

Fig. 3. Experimental result of the final moving object detection.

综上所述, 本文提出的基于光流场分析的运动目标检测算法具体步骤如下。

算法1 动态背景下基于光流场分析的运动目标检测

输入: 视频序列, 图像帧数目 N

目标边界检测

1) 计算每帧图像的光流场 $f_t(u, v)$, $(1 \leq t \leq N-1)$

2) for $t = 1 : N - 1$ do

利用 (1) 式计算光流梯度幅值确定的目标边界强度系数 $s_t^m(u, v)$; 利用 (2) 式计算光流矢量方向确定的目标边界强度系数 $s_t^a(u, v)$;

利用 (3) 式确定得到目标边界的二值图 $s_t(u, v)$;

end for

目标内部像素点判断

3) for $t = 1 : N - 1$ do

利用点在多边形内部原理确定目标内部像素点

end for

前背景像素标记

4) 视频序列过分割得到超像素集合 $\sum_{t,i} r_t^i$;

5) for $t = 1 : N - 1$ do;

利用 (5) 式—(8) 式计算超像素 r_t^i 对应的数据项势能函数 $A_t^c(r_t^i)$;

利用 (9) 式和 (10) 式计算超像素 r_t^i 对应的空域平滑项势能函数 $S_t^{i,j}(r_t^i, r_t^j)$ 和时域平滑项势能函数 $T_t^{i,j}(r_t^i, r_{t+1}^j)$;

end for

6) 利用图割算法求解 (11) 式所示能量函数最小化问题

输出: 视频序列每帧图像的前背景二值标记

5 实验结果与分析

本文选择多个公开发布的视频序列进行实验测试. 实验数据分别来自标准视频库中的 Cat 和 Dog 序列, Hopkins 155 数据集^[27] 的 Cars1-Cars4 序列, Sand 和 Teller^[28] 提供的 VPerson 和 Vcar 序列, 以及 Changedetection.net 数据集中^[29] 的 Highway 和 Lab 序列, 选取视频包含多种复杂场景中的刚体和非刚体运动, 具有较好的代表性. 另外, 选取视频中 Highway 序列和 Lab 序列为静止相机拍摄的视频序列, 其余则为手持式相机拍摄的视频序列, 可验证本文方法在静态和动态两种背景下的有效性.

采用广泛使用的查准率 PR、查全率 RE 和综合评价 FM^[30] 对所提算法进行评价, 并与主模块算

法 (KS)^[8]、视觉和运动显著算法 (VMS)^[9] 和单应模型约束算法 (HC)^[13] 进行定性和定量对比, 结果如图 4 和表 1 所示. 实验中, 取 $\eta_m = 0.7$, $\eta_a = 0.4$, $\eta = 0.1$, 超像素初步分类参数 T_1 和 T_2 分别为 0.2 和 0.001.

如图 4 所示, 不同场景下各种算法都可以大致检测出感兴趣的目标区域, 但在检测的准确度上有所差异. 对比可以发现, 主模块算法 KS 虽然检测得到了前景目标的主体内容, 但在目标的完整性上误检区域较为明显, 比如 Cars1 场景中前轮胎区域和迎面驶来的小汽车的漏检, 以及在静态场景 Highway 中未检测出第二辆行驶的小汽车; 视觉和运动显著性算法 VMS 检测出的目标轮廓相对清晰, 但当运动目标旁边存在颜色相似的物体或者场景中含有视觉上较为突出的目标时, 也会出现明显的误检, 表现为在 Cat 场景中误检了场景中的两只碗, 同时猫的腿部白色区域出现了局部漏检, 在其他场景中也部分出现了上述问题; 单应性算法 HC 的检测结果相对完整, 但目标的过检测导致边界不清楚, 比如 Cars1 和 Highway 场景中将车身与阴影融为了一体, 除此之外, 该算法在操作过程中需要计算大量运动轨迹, 计算复杂度较高; 相比前三种算法, 本文方法在综合性能上更为优越, 算法采用光流梯度幅值和光流矢量方向两种方法来确定目标的边界, 在不同场景下得到的目标轮廓都较为清晰准确, 并且可以消除部分运动阴影的影响, 另外算法对前背景的表现信息建立混合高斯模型, 并结合超像素点的时空邻域连续性, 使检测的结果更加完整准确.

从表 1 数据可以看出, 不同场景下本文算法的查全率 PR 和查准率 RE 多数高于其他算法, 表明所提算法对前景目标的检测准确性明显提升, 综合评价指标值 FM 也稳居最高, 且基本达到了 90% 左右, 更充分说明了本文算法具有非常好的鲁棒性, 能够广泛适用于不同场景下的运动目标检测.

为进一步说明本文算法在处理速度上的优势, 同样在上述 10 组视频序列上进行对比实验, 得到 4 种算法在所有视频帧上的平均处理时间, 结果如表 2 所示. 值得注意的是, 动态背景下目标检测的输入通常是像素点运动轨迹或帧间光流场, 目前二者的计算度非常高, 尚不能满足实时要求, 这也是限制动态背景下目标检测运算速度的主要因素, 若想要提高处理速度, 可采用 GPU 加速的光流场或像素点运动轨迹^[31].

表1 四种算法的定量评估
Table 1. The quantitative result of four algorithms.

Video	KS			VMS			HC			Proposed		
	PR	RE	FM	PR	RE	FM	PR	RE	FM	PR	RE	FM
Cat	0.862	0.897	0.879	0.806	0.920	0.859	0.874	0.926	0.899	0.885	0.932	0.908
Dog	0.898	0.927	0.912	0.894	0.953	0.923	0.904	0.937	0.920	0.916	0.944	0.930
Cars1	0.816	0.685	0.745	0.899	0.788	0.840	0.843	0.957	0.896	0.923	0.906	0.914
Cars2	0.884	0.781	0.829	0.832	0.883	0.857	0.831	0.917	0.872	0.864	0.897	0.880
Cars3	0.902	0.803	0.850	0.824	0.902	0.861	0.852	0.950	0.898	0.892	0.925	0.908
Cars4	0.883	0.797	0.838	0.845	0.915	0.879	0.838	0.918	0.876	0.875	0.910	0.892
VPerson	0.870	0.715	0.785	0.827	0.842	0.834	0.836	0.855	0.845	0.902	0.884	0.893
Vcar	0.904	0.763	0.828	0.798	0.894	0.843	0.786	0.928	0.851	0.928	0.866	0.896
Highway	0.916	0.542	0.681	0.805	0.712	0.756	0.883	0.814	0.847	0.868	0.915	0.891
Lab	0.882	0.827	0.854	0.873	0.795	0.832	0.892	0.858	0.875	0.873	0.902	0.887

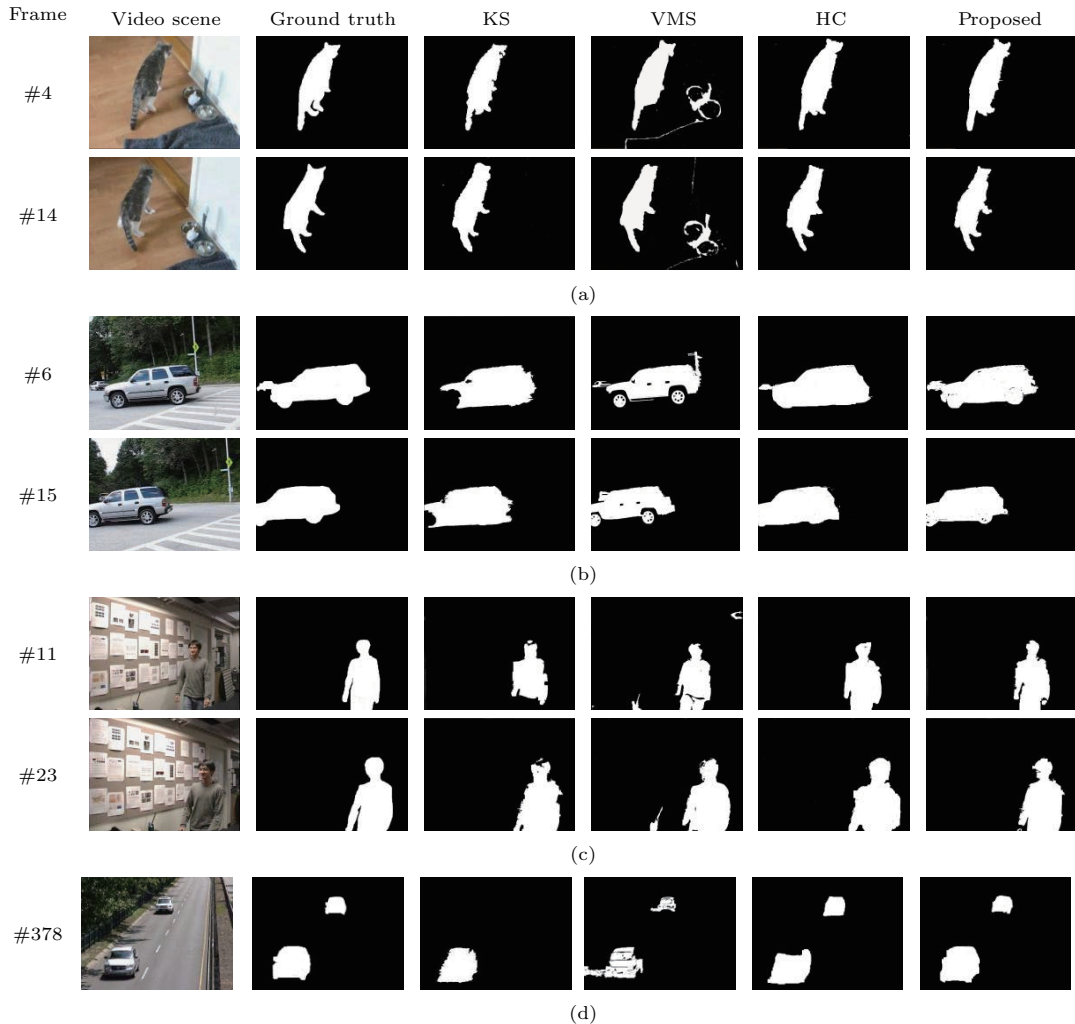


图4 (网刊彩色) 四种算法在不同场景下的前景检测结果 (a) Cat (dynamic scene); (b) Cars1 (dynamic scene); (c) VPerson (dynamic scene); (d) Highway (static scene)

Fig. 4. (color online) The experimental results of four algorithms in different scenes: (a) Cat (dynamic scene); (b) Cars1 (dynamic scene); (c) VPerson (dynamic scene); (d) Highway (static scene).

表2 四种算法的处理速度对比

Table 2. Comparison of processing time about four algorithms.

算法	KS	VMS	HC	本文算法
处理时间/s·frame ⁻¹	130	77	95	58

6 结 论

本文提出一种基于光流场分析的运动目标检测算法。算法以相邻视频帧的光流矢量为基础, 首先通过光流梯度幅值和光流矢量方向共同确定目标的边界, 得到相对清晰准确的前景轮廓; 然后利用点在多边形内部原理对像素点进行多方向判断, 获得前景目标内部较准确的稀疏像素点; 最后以超像素为节点, 将利用混合高斯模型构建的数据项和利用超像素时空邻域关系构建的平滑项统一纳入到马尔可夫随机场模型中, 并通过图割最优理论求解模型得到最终的前背景区域分割结果。本文算法具有非常广泛的适用性, 在静态背景和摄像机任意运动产生的动态背景均能实现准确、鲁棒的运动目标检测。

参考文献

- [1] Radke R, Andra S, Kofahi A, Roysam B 2005 *IEEE Trans. Image Process.* **14** 294
- [2] Ren Y, Chua C, Ho Y 2003 *Mach. Vision Appl.* **13** 332
- [3] Sheikh Y, Javed O, Kanade T 2009 *Conference. on Computer Vision and Pattern Recognition(CVPR)* Miami, USA, June 20–25, 2009 p1219
- [4] Chen L, Zhu S, Li X 2015 *International Symposium on Computers & Informatics* Beijing, China, January 17–18, 2015 p742
- [5] Bi G L, Xu Z J, Chen T, Wang J L, Zhang Y S 2015 *Acta Phys. Sin.* **64** 150701 (in Chinese) [毕国玲, 续志军, 陈涛, 王建立, 张延坤 2015 物理学报 **64** 150701]
- [6] Sun S W, Wang Y F, Huang F, Liao H Y 2013 *J. Visual. Commun. Image Represent* **24** 232
- [7] Li A H, Cui Z G 2016 *Moving Object Detection in Videos* (Beijing: Science Press) p15 (in Chinese) [李艾华, 崔智高 2016 视频序列运动目标检测技术 (北京: 科学出版社) 第15页]
- [8] Lee Y, Kim J, Grauman K 2011 *International Conference on Computer Vision(ICCV)* Barcelona, Spain, November 6–13, 2011 p1995
- [9] Li W T, Chang H S, Lien K C, Chang H T, Wang Y C 2011 *IEEE Trans. Image Proc.* **22** 2600
- [10] Zhang D, Javed O, Shah M 2013 *Conference on Computer Vision and Pattern Recognition(CVPR)* Oregon, Portland, June 25–27, 2013 p682
- [11] Elqursh A, Elgammal A 2012 *European Conference on Computer Vision (ECCV)* Florence, Italy, October 7–13, 2012 p228
- [12] Gao W, Tang Y, Zhu M 2014 *Acta Phys. Sin.* **63** 094204 (in Chinese) [高文, 汤洋, 朱明 2014 物理学报 **63** 094204]
- [13] Cui Z G, Li A H, Feng G Y 2015 *Journal of Computer-Aided Design & Computer Graphics* **27** 621 (in Chinese) [崔智高, 李艾华, 冯国彦 2015 计算机辅助设计与图形学学报 **27** 621]
- [14] Wang J, Adelson E 1994 *IEEE Trans. Image Process.* **3** 625
- [15] Cremers D, Soatto S 2004 *Int. J. Comput. Vision* **62** 249
- [16] Yoon S, Park S, Kang S 2005 *Pattern Recognit. Lett.* **26** 2221
- [17] Adhyapak S, Kehtarnavaz N, Nadin M 2007 *J. Electron. Imaging* **16** 13012
- [18] Di S, Mattoccia S, Tombari F 2005 *International Workshop on Computer Architecture for Machine Perception* Palermo, Italy, July 4–6, 2005 p193
- [19] Bouguet J 2001 *Intel Corporation* **5** 10
- [20] Brox T, Malik J 2010 *European Conference on Computer Vision (ECCV)* Crete, Greece, September 5–11, 2010 p282
- [21] Achanta R, Shaji A, Smith K 2012 *IEEE Trans. Pattern Anal. Mach. Intell.* **34** 2274
- [22] Achanta R, Shaji A 2010 *EPFL Technical Report* **1** 149
- [23] Vazquez A, Avidan S, Pfister H 2010 *European Conference on Computer Vision (ECCV)* Crete, Greece, September 5–11, 2010 p268
- [24] Fulkerson B, Vedaldi A, Soatto S 2009 *International Conference on Computer Vision (ICCV)* Kyoto, Japan, September 27–October 4, 2009 p670
- [25] Boykov Y, Veksler O, Zabih R 2001 *IEEE Trans. Pattern Anal. Mach. Intell.* **23** 1222
- [26] Boykov Y, Funka L 2006 *Int. J. Comput. Vision.* **70** 109
- [27] Tron R, Vidal R 2007 *Conference on Computer Vision and Pattern Recognition (CVPR)* Minneapolis, USA, June 18–23 2007 p1
- [28] Sand P, Teller S 2008 *Int. J. Comput. Vision.* **80** 72
- [29] Goyette N, Jodoin P, Porikil F 2012 *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* Providence, Rhode Island, June 16–21, 2012 p1
- [30] Cui X, Huang J, Zhang S, Metaxas D 2012 *European Conference on Computer Vision (ECCV)* Florence, Italy, October 7–13, 2012 p612
- [31] Sundaram N, Brox T, Keutzer K 2010 *European Conference on Computer Vision (ECCV)* Crete, Greece, September 5–11, 2010 p438

Moving object detection based on optical flow field analysis in dynamic scenes^{*}

Cui Zhi-Gao^{1)2)†} Wang Hua¹⁾ Li Ai-Hua¹⁾ Wang Tao¹⁾ Li Hui¹⁾

1) (*The Rocket Force of Engineering University, Xi'an 710025, China*)

2) (*Department of Automation, Tsinghua University, Beijing 100084, China*)

(Received 21 October 2016; revised manuscript received 24 January 2017)

Abstract

To overcome the limitation of existing algorithms for detecting moving objects from the dynamic scenes, a foreground detection algorithm based on optical flow field analysis is proposed. Firstly, the object boundary information is determined by detecting the differences in optical flow gradient magnitude and optical flow vector direction between foreground and background. Then, the pixels inside the objects are obtained based on the point-in-polygon problem from computational geometry. Finally, the superpixels per frame are acquired by over-segmenting method. And taking the superpixels as nodes, the Markov Random field model is built, in which the appearance information fitted by Gaussian Mixture Model is combined with spatiotemporal constraints of each superpixel. The final foreground detection result is obtained by finding the minimum value of the energy function. The proposed algorithm does not need any priori assumptions, and can effectively realize the moving object detection in dynamic and stationary background. The experimental results show that the proposed algorithm is superior to the existing state-of-the-art algorithms in the detection accuracy, robustness and time consuming.

Keywords: dynamic scene, moving object detection, optical flow field analysis, Markov random field model

PACS: 42.30.Tz, 07.05.Pj, 02.50.-r

DOI: [10.7498/aps.66.084203](https://doi.org/10.7498/aps.66.084203)

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 61501470).

[†] Corresponding author. E-mail: cui zg10@tsinghua.edu.cn