

基于双时间尺度运动分析的户外小目标跟踪

叶喜勇^{1,2}, 陶霖密¹, 萨如汗³

(1. 清华大学计算机科学与技术系, 北京 100084; 2. 海军大连舰艇学院航海系, 辽宁大连 116013;
3. 纽约州立大学布法罗分校计算机系, 纽约布法罗 14260)

摘要: 根据人类视觉发现并跟踪目标的原理, 提出了一种基于双时间尺度信号分析的户外小目标跟踪方法. 在短时间尺度信息分析中, 通过时空联合滤波, 滤除运动前景信息中的瞬时随机噪声, 提取高置信度候选运动小目标信息; 在长时间尺度信息分析中, 通过分析背景运动和目标运动的运动特性, 提出了小目标运动轨迹平滑性约束条件, 基于目标的运动信息和姿态信息构建了能量函数, 利用基于全局优化的 Viterbi 快速算法在长时间尺度上实时求解目标运动的时空三维轨迹. 不同场景条件下远程小目标跟踪实验表明了我们所提出方法的有效性和鲁棒性. 该方法的算法复杂性低, 适用于对小目标的实时检测与跟踪.

关键词: 目标跟踪; 小目标; 实时; 视频监控

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2015)07-1257-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2015.07.002

An Outdoor Small Target Tracking Method Based on Duo Temporal Scale Motion Analysis

YE Xi-yong^{1,2}, TAO Lin-mi¹, SA Ru-han³

(1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;
2. Department of Navigation, Dalian Naval Academy, Dalian, Liaoning 116013, China;
3. Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260, USA)

Abstract: In this paper, we propose a small target tracking method based on duo-temporal-scale signal analysis (DTS). The proposed tracking method exploits the natural difference between objects and noise in two phases: short-temporal-scale noise reduction phase (STS) and long-temporal-scale motion analysis phase (LTS), which simulates the small target tracking process in human vision system. In STS, we propose a new space-temporal noise filter to remove random noise or blocks that are easy to be distinguished in appearance, and then we get the candidate targets with high probability to initialize moving targets' trajectories. In LTS, we propose a new trajectory smoothness constraint function based on different motion patterns between small targets and noise. Finally we get the small targets' trajectories by optimizing smoothness constraint function in long temporal scale using Viterbi algorithm. Experiments on different kinds of outdoor scenarios show the efficiency and robustness of the proposed method.

Key words: target tracking; small targets; real-time; visual surveillance

1 引言

目标检测和跟踪在视频监控及相关领域中具有重要的应用, 近几十年来一直是计算机视觉领域研究的热点之一. 而远程目标在图像中成像面积小. 相对于大面积目标跟踪, 由于小目标缺乏纹理、形状等信息, 增加了目标跟踪的难度. 当前, 绝大多数小目标跟踪相关的工作主要是基于红外波段的^[1,2], 在可见光波段的小目标检测和跟踪的工作主要针对空中或者海面等背景相对简单的目标^[3], 近年来, 由于航拍视频的兴起, 还出现了

在航拍视频中对地面小目标跟踪的工作^[4]. 本文研究可见光波段视频中的室外远程小目标检测与跟踪, 其特点在于目标在图像中成像的面积小, 无可用的纹理特征, 难以用现有的方法检测与跟踪. 这一类的小目标主要具有以下特点:

(1) 目标在图像中的尺寸非常小, 目标/背景对比度低, 无稳定纹理特征. 很难在图像中提取到小目标的颜色, 边缘, 光流场等明显的特征信息. 尤其是在背景复杂, 或者雾天等低对比度的情况下, 人眼也很难通过仅仅几帧图像发现目标.

(2)室外场景中,场景本身通常含有运动信息.例如,树枝的摇摆及云和雾等的运动.这些背景运动与目标运动类似,因此很难通过短时间尺度分析将目标运动从背景运动分离出来.

(3)室外场景条件下,运动物体容易被树木等物体所遮挡.

因此,对室外场景中的小目标进行自动、稳定、实时地检测与跟踪,依然是计算机视觉研究领域富有挑战性的问题.

对运动目标跟踪的基本要求是要获取目标在三维时空当中的运动轨迹等与运动相关的参数,主要包含两个子问题:①目标检测及运动轨迹的初始化;②利用跟踪算法最终获取目标的运动轨迹.

目标轨迹的初始化用于初始化目标在图像中的位置,主要有基于目标特征和基于前背景分割方法.基于目标特征匹配的方法主要是随着跟踪与检测结合的跟踪方法发展而来,依赖于目标某些稳定的纹理特征.具体方法主要有基于特征点^[5,6]、基于颜色、梯度等的特征模版^[7]、颜色直方图模版^[8]、轮廓特征^[9,10]等.另一种是利用前景分割的结果对目标跟踪算法进行初始化.文献[11]利用混合高斯模型的分割结果对目标的位置进行初始化.在许多场景条件下,由于背景动态特性比较复杂,使得前背景模型分割结果中存在较强的噪声.为消除噪声的影响,文献[12,13]对前景分割的结果进行处理,提取前景区域内的颜色、立体视觉等信息,需要用到与目标纹理相关的信息对目标轨迹进行初始化.

给定目标的初始位置信息,利用各种跟踪算法对目标进行跟踪/匹配,可以进一步获取目标的相关运动参数信息,最终得到目标的运动轨迹,从而实现对目标连续跟踪.目标跟踪算法总体上可以分为两类,一类是利用历史信息对目标当前位置等运动参数进行估计,例如扩展卡尔曼滤波^[14]、Mean Shift跟踪^[15]、粒子滤波^[16~19],这一类方法主要是利用模型的马尔可夫特性,对目标的状态信息进行迭代求解,存在着跟踪轨迹偏离真实目标的潜在风险.另一种跟踪算法同时利用过去和将来的信息,在尽可能长的时间窗口上对目标轨迹进行优化求解.其中多假设跟踪(MHT, Multi-Hypothesis Tracking)^[20]和联合概率数据关联滤波器(JPDFAs, Multi-Hypothesis Data Association Filters)^[21]等方法需要对每个阶段的所有状态进行联合求解,运算复杂度随时间窗口长度呈组合式增长,因此在实际应用中只能取较小时间窗口求解目标的运动轨迹.基于动态规划的跟踪算法^[22,23]效率要比前两个要高,因此在同样条件下可以取相对较长的时间窗口,但是强噪声条件下利用动态规划进行求解时,仍然难以满足实时

性的要求.

上述初始化与跟踪方法,都依赖于目标的某些稳定的纹理信息,并通过特征检测与提取,实现目标的检测与跟踪.由于缺乏稳定特征,特征匹配等方法无法启动,而基于粒子滤波器等利用历史信息迭代求解的算法易于引起跟踪漂移问题,使跟踪轨迹偏离实际目标.类似地,由于目标纹理信息缺失,场景中存在强噪声等原因,传统的基于动态规划等全局优化跟踪算法也难以得到正确的目标运动轨迹.因此,上述方法适合于成像面积大、对比度高的场景.在远程或者大面积监控的条件下,目标成像面积小、对比度低、无稳定纹理特征,难以利用上述方法进行检测与跟踪.

尽管如此,小目标运动,依然存在一定的运动规律.人类视觉在对远程运动小目标进行观察时,首先将注意力聚焦于场景中的运动目标,注视一个方向一段时间以后,可以将运动噪声与运动小目标分离,从而发现并跟踪目标.因此结合人的视觉认知原理,发掘小目标的规律运动与背景随机运动之间的差别,机器视觉可以模拟人类视觉对户外远程小目标跟踪的过程,从而实现对小目标的检测与跟踪.

根据上述分析,针对低信噪比的户外远程小目标跟踪场景,本文通过对目标的运动与持续噪声的运动分析发现,小目标运动的时空轨迹是平滑的,通常是一阶或者二阶的运动;而噪声由于其随机性,其运动通常存在较大的高阶分量.因此,根据目标和噪声的运动特性,我们提出了一种基于双时间尺度信号分析的小目标跟踪方法.首先利用前背景分割技术,得到运动前景信息,然后在短时间内,根据噪声与目标在短时间时空域的不同特性,通过时空联合滤波处理,滤除前景中的瞬时随机噪声和边缘噪声等,从而得到可能性高的预选目标,最后通过分析目标与背景的运动特性,建立平滑性约束方程,在长时间尺度上进行目标检测和跟踪.

雨天、低对比度等复杂场景条件下远程小目标跟踪的实验表明,目标在图像中的面积低至6~8像素时,本文的方法仍能够实时、稳定地跟踪目标.

2 基于双时间尺度信号分析的小目标跟踪方法

在户外远程小目标跟踪场景条件下,由于图像信噪比低、成像面积小、对比度低、缺乏区分于背景的有效目标信息,尽管如此,人类视觉系统还是能够发现并跟踪运动目标的.通过对人类视觉系统发现和跟踪目标过程的分析,我们发现,人类视觉在对这类目标进行跟踪时,其过程可以简化如下:首先是人类视觉对运动目标有很高的灵敏度,视觉系统在短时间内提取场景中的运动信息,选择可能性较高的预选目标;其次,通

过注视预选目标,视觉系统根据运动目标与背景运动的不同运动模式,发现并跟踪运动目标。

本文主要根据上述人类视觉系统发现和跟踪小目标的过程,提出了双时间尺度的小目标跟踪方法(DTS, Duo-Temporal-Scale Tracking Method),将目标跟踪分成短时间尺度信号分析和长时间尺度目标轨迹分析两个阶段(图 1)。在短时间尺度信号分析阶段(STS, Short-Temporal-Scale Signal Analysis Phase),提取场景中运动信息,分析运动小目标与干扰噪声的短时间特性,利用时空联合滤波的方法进行去噪处理,滤除瞬时随机噪声及边缘噪声等,将保留的目标信息和持续噪声作为输入送到下一处理阶段。在长时间尺度轨迹分析阶段(LTS, Long-Temporal-Scale Trajectory Analysis Phase),通过长时间尺度的信号分析,提出一种运动平滑性能量约束条件,并利用基于 Viterbi 的快速求解方法,在长时间尺度上进行全局优化求解,获取目标运动轨迹。并将目标运动速度、形态等参数进行反馈,作为下一步长目标轨迹分析的初始化信息和轨迹连续性判决信息。

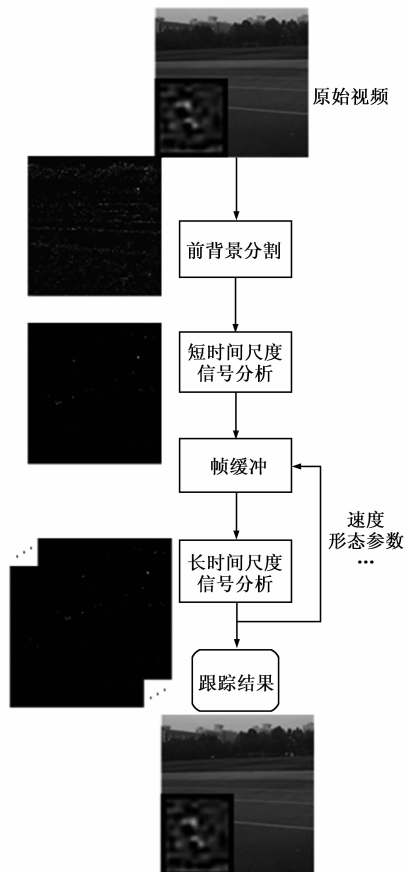


图1 小目标跟踪信息处理流程

2.1 前背景分割

通过前面分析我们得知,在户外远程小目标监控条件下,无法提取稳定的目标特征点信息或区分能力

强的模版、轮廓等信息。因此在视频的预处理阶段,我们采用前背景分割的方法来提取目标运动信息。混合高斯模型^[11]能够较好地动态背景的图像序列进行建模,他将每个像素点的色彩变化使用多个高斯的多项式分布表示:

$$p(\mathbf{x}) = \sum_{i=1}^M \omega_i N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (1)$$

式中, \mathbf{x} 为图像中某个像素点的颜色向量, $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ 表示均值为 $\boldsymbol{\mu}_i$, 方差为 $\boldsymbol{\Sigma}_i$ 的正态分布。 ω_i 为权重系数。 M 为高斯分布的个数,在实际应用中, M 通常取 3~5 的范围。其中, $\omega_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$ 根据当前帧像素的颜色进行动态更新,最后根据设定的门限 T ,将 ω_i 从大到小排列得到 $\{\omega'_1, \omega'_2, \dots, \omega'_M\}$,求最小达到门限所需的高斯个数:

$$k = \arg \min_m \left(\sum_{j=1}^m \omega'_j \geq T \right) \quad (2)$$

若当前像素点所在的高斯分布在排序后的序号小于或等于 k ,则当前像素点为背景点,否则为前景点。通常,在实际应用中,目标在图像中有稳定的颜色信息,并且与背景颜色具有较大的对比度,因此前景点所在的高斯的均值与背景高斯均值相差较大,从而可以比较有效地将前景从背景中分割出来。另一方面,由于前背景信息所在的高斯距离比较远,为了抑制背景噪声,通常将可以设置较高的判决门限。

混合高斯模型利用多个高斯模型对像素点的颜色分布进行背景建模,且前背景颜色对比度较大时,利用该方法进行前背景分割时,能够很好地抑制背景图像中快速周期性变化的噪声。但在户外自然场景条件下,目标成像面积小,信息量少,且背景表现为动态特性,图像表现出低信噪比的特性。另外,在远程小目标场景下,目标与背景之间的对比度通常较低,这意味着在利用混合高斯模型进行建模时,某个背景与前景高斯之间均值距离比较近,因此在设置学习模型时,为了能够将前景更好地分离出来,通常让每个高斯的方差保持较小,从而提高背景噪声与前景信号的分辨力。这样最终导致模型对背景的变化敏感,从而导致混合高斯模型对背景图像的代表能力减弱,最终导致前景分割结果中的前景噪声增加。

另一方面,在利用混合高斯模型进行背景建模时,通常选择比较大的判决门限,通过牺牲部分前景信息而达到抑制背景噪声的目的。在小目标场景条件下,目标的面积小,图像前背景对比度低,选择大的判决门限极易将目标判决为背景。因此,利用混合高斯模型在户外远程小目标场景的条件下进行前背景分割时,通常伴随着较强的前景噪声。

从以上分析可以看出,由于小目标图像序列低信噪比的特性,最终导致利用背景模型进行前背景分割

后的图像中存在较强的前景噪声,如果直接利用分割后的图像对跟踪算法进行初始化,则使得跟踪算法将大部分时间都用在对噪声的处理上,并且强噪声容易造成错误的跟踪结果.因此,在利用跟踪算法进行小目标跟踪之前,有必要对分割后的结果作进一步的分析 and 处理,减少前景图像中的前景噪声,提高后续跟踪算法的跟踪性能.

2.2 噪声分析与短时间尺度去噪

混合高斯模型能够较好地描述动态场景,但是由于户外小目标场景的特殊性,导致前景分割结果中仍存在较强的前景噪声.其噪声来源主要体现在以下几个方面:首先,树枝的摇摆、水及雾等造成了环境噪声,这些噪声跟目标的运动模式类似,具有较好的时空一致性,表现为持续噪声;其次,由于环境中存在着丰富的光滑表面,使得环境中的各种物体产生复杂的复合光,最后造成目标亮度的闪烁,导致其颜色在图像中呈现及其复杂的分布,但是时空相关性差,表现为瞬时随机噪声;再次,由地面震动(如爆破、重型载重车辆)等因素所引起的摄像机偶然抖动会引起物体的边缘噪声;最后,还有摄像机内部噪声等系统固有噪声引起的前景噪声.

对于瞬时随机噪声,在图像中同一位置的其帧间相关性较小,最终导致噪声“生存周期”通常比较短.图2对瞬时随机噪声进行在图像序列中同一像素点的噪声“生存周期”进行分析,这里假设帧与帧之间的前景随机噪声相互独立.从图2(a)中可以看出,图像中的前景点噪声浓度较高的情况下,噪声点在同一位置连续出现的时间窗口通常比较小,在8帧噪声图像最右边一列,对上数往下8个像素的时间持续性分析中我们发现,前景点持续时间最长为4帧.我们对帧间独立的椒盐噪声、高斯白噪声和泊松噪声进行了实验,对能够保持在同一位置的前景像素点的持续时间进行统计,结果如图2(b)所示.从结果中可以看出,独立噪声在图像序列中的持续时间呈指数衰减.当前景噪声占图像面积的比例为40%,持续三帧以上(包含三帧)的前景点仅约占所有前景点的15%.因此,理论上可以通过时空联合滤波,滤除前景图像中绝大部分随机前景噪声.

$$Y_i(x, y) = \bigcap_{i=\tau-\tau}^i X_i(x, y) \quad (3)$$

式中 τ 为时间窗口长度, $X_i(x, y)$ 为第 i 帧图像中图像坐标 (x, y) 处像素的前景逻辑值,0 表示背景,1 为前景. $Y_i(x, y)$ 为时空联合滤波的结果.

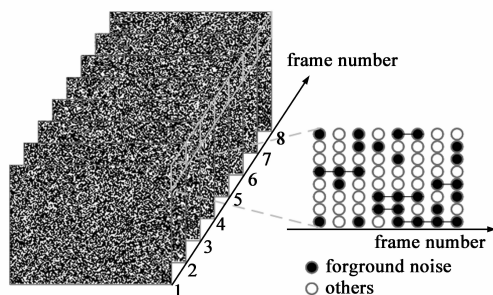
通过前面分析我们知道,随机噪声的时空持续性较差,因此,对于式(3),根据场景的具体情况,选择合适的时间窗口长度,可以滤除大部分的随机噪声.

另外,在户外小目标跟踪场景中,跟踪对象通常为

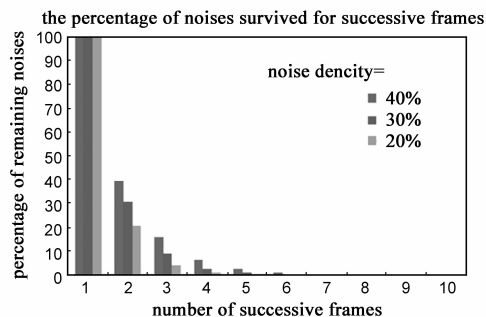
车辆、人等比较规则的目标,其外接矩形长宽比通常比较小,并且在前景分割之后与目标对应的前景区域的外接矩形前景像素点填充率也比较高.最后,由于摄像机抖动及边缘噪声等产生的前景区域与运动目标的前景区域有明显的区别,这些噪声点主要分布在各个区域的边缘处(如图3所示),因此前景连通区域通常比较细长,对于不规则边缘,则表现为其外接矩形内的前景点填充率低.基于这两点,我们可以通过以下规则滤除这些噪声:

(1)前景区域外形:前景区域外形被定义为前景目标区域的长度与平均宽度的比值.当外接矩形的长宽比大于某个值时,判定该前景区域为噪声.

(2)前景像素浓度:前景像素浓度被定义为前景区域面积与前景区域外接矩形面积的比值.当前景像素浓度低于某个值时,将前景判断为噪声.



(a) 短时间尺度去噪原理



(b) 不同噪声浓度下的噪声持续特性

图2 噪声在同一位置出现的持续时长呈负指数分布

图4显示了STS阶段信号处理的结果.从图中可以看出,经过混合高斯模型进行前景分割后的结果中,前景信息大部分为噪声信息(图4(b)),经过LTS去噪处理,大部分噪声已被滤除(图4(c)).

以上提出的短时间信息处理类似于人类视觉在远程小目标检测和跟踪过程中的初始运动目标信息提取的过程,通过初始运动分析,在短时间内通过时空联合信息处理,滤除掉与目标特性区别较大的噪声.剩下的噪声主要为树枝运动、水波运动等引起的持续噪声这一类噪声具有较好的时空连续性,难以在短时间之内将其滤除.但是这类噪声在长时间内所呈现的特性明

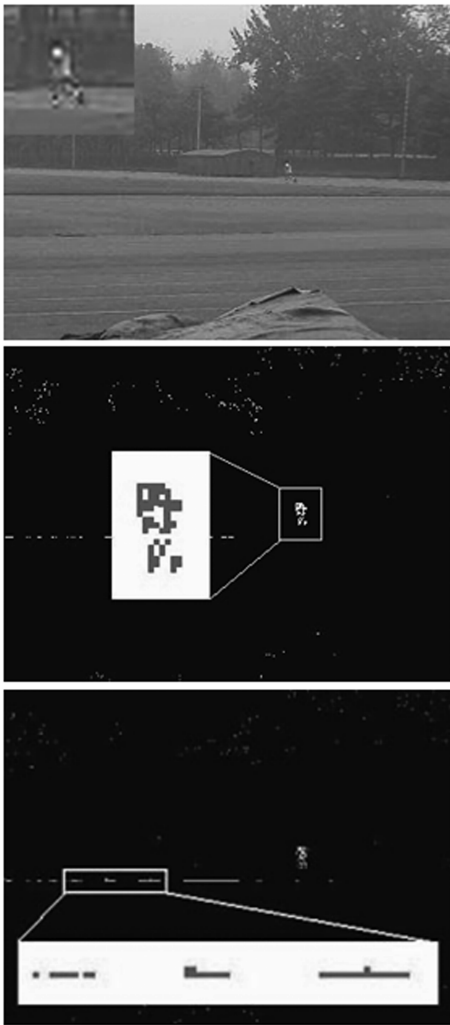


图3 前景分割中的噪声(人为运动目标), 原始图像左上角为放大后的运动目标

显区别于车辆、行人目标. 总得来说, 待跟踪目标的质心运动轨迹、外形及其姿态的变化比较平滑, 而由树枝运动、水的运动等所产生的前景的质心运动轨迹、外形及姿态变化具有较大的随机性. 因此, 通过长时间对 STS 处理后图像中的信息进行分析, 可以将噪声信息与目标信息分离.

2.3 长时间尺度目标检测与跟踪

通过 STS 噪声滤除的处理, 其结果中的前景为目标的位置得到提高. 在 LTS 阶段, 主要集中解决两个问题: 首先, 在候选目标中检测出真实运动目标并得到真实目标在三维时空的运动轨迹; 其次, 还需要处理室外场景中目标遮挡及丢失等问题. 树枝、水波浪等引起的运动干扰未能在 STS 阶段有效滤除, 因此通过 STS 处理之后, 这些噪声仍然存在, 需要分析目标与干扰之间的运动模式, 从而从候选目标中检测出真实目标.

现实世界中, 一段道路总是或直或弯, 汽车、行人等

目标的运动通常以匀速或匀加速运动来近似, 因此其在图像平面投影所形成的运动轨迹在三维时空中是平滑的(图 5), 即小目标在三维时空中的运动轨迹是一次或者二次平滑的. 而树枝等环境运动具有较大的随机性, 从而形成高阶、复杂的运动轨迹曲线; 其次, 由于水波等与周围等颜色信息一致, 因此前景分割之后所形成的前景噪声其外形的随机性比较大, 最终表现为前景区域整体姿态也呈随机变化, 与这类前景噪声相比较, 目标在前景分割后的外形时间一致性较好, 其前景区域的姿态在空间当中的变化比较“平滑”. 因此, 在 LTS 目标检测与跟踪阶段, 对时间观测窗口进行延长, 然后在这一时间窗口内, 对图像中各物体的运动模式等进行分析, 从而滤除背景噪声的干扰, 提取实际的目标信息.



(a) 原始图像



(b) 利用混合高斯模型的前景分割结果



(c) 短时间尺度去噪结果

图4 短时间窗口去噪实验示例

设 X_i 为第 i 帧图像中任意前景连通区域, 我们使用能量函数对前景区域变化的平滑性进行描述:

$$E(X_i, X_{i-1}) = D(X_i, X_{i-1}) - k_1 V(X_i, X_{i-1}) - k_2 \Theta(X_i, X_{i-1}) \quad (4)$$

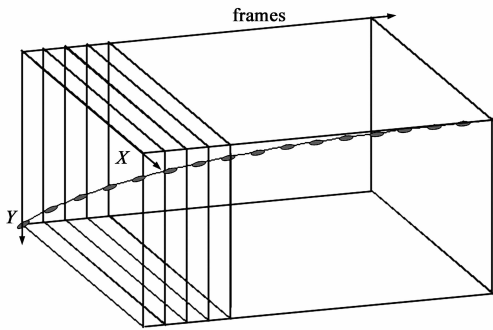


图5 三维时空中的目标轨迹, 椭圆为小目标, 是经STS阶段处理得到的连通的前景点集合的外接椭圆

$E(X_i, X_{i-1})$ 为相邻帧中任意两个前景区域之间的能量函数. 其中, $D(X_i, X_{i-1})$ 为 X_i 与 X_{i-1} 两个前景区域的质心之间的欧氏距离; $V(X_i, X_{i-1})$ 为 X_i 与 X_{i-1} 的速度平滑性约束函数, $\Theta(X_i, X_{i-1})$ 为前景区域 X_i 与 X_{i-1} 的姿态平滑性约束函数.

在实际户外小目标跟踪场景条件下, 两帧之间的时间间隔通常比较短, 在图像中移动的距离很小, 因此, 相邻帧之间同一目标质心欧氏距离通常比较小. 而噪声具有一定的随机性, 在大多数情况下, 他与真实目标之间的距离通常比较大. 另外, 通过前面分析我们知道, 正常目标的轨迹和姿态变化比较平滑. 因此式(4)能较好地对目标平滑性特性进行描述.

式(4)中的速度平滑性函数定义如下:

$$V(X_i, X_{i-1}) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}(\mathbf{v}_i - \boldsymbol{\mu}_{vi})^T \boldsymbol{\Sigma}^{-1}(\mathbf{v}_i - \boldsymbol{\mu}_{vi})\right] \quad (5)$$

其中, \mathbf{v}_i 为当前前景区域的质心运动速度, $\boldsymbol{\mu}_{vi}$ 为已跟踪轨迹在特定时间窗口长度 C 内(本文实验中 $C = 15$ 帧, 对应时间为 1s) X_{i-1} 的质心运动的平均速度. 式(5)中, 当两者运动速度相等时, 速度平滑函数取最大值. 在现实场景中, 目标通常可以近似为匀速或者匀加速运动, 而噪声由于其随机性, 在前景中通常表现为无规律的运动, 前后两阶段速度通常相差比较大. 因此, 在总能量函数中引入 $V(X_i, X_{i-1})$, 可以改善目标检测的性能. 另外, 在多目标情况下, 尤其在目标交叉时, 根据不同运动目标速度的一致性, 引入 $V(X_i, X_{i-1})$ 理论上可以将多个交叉目标准确分离.

式(4)中的姿态平滑函数定义如下:

$$\Theta(X_i, X_{i-1}) = \left|1 - \frac{H}{L}\right| \frac{1}{\sigma_\theta \sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{\theta_i - \mu_{\theta_i}}{\sigma_\theta}\right)^2\right] \quad (6)$$

式中, θ_i 为第 i 帧任意前景区域 X_i 外接椭圆的姿态角, 其定义为外接椭圆长轴在图像平面内相对与图像平面纵轴的旋转角; μ_{θ_i} 为已跟踪轨迹中目标外接椭圆在时

间串口 C 内的平均姿态角, σ_θ 为方差. 由式(6)可以知道, 当前候选目标的外接椭圆姿态角 θ_i 与 μ_{θ_i} 相等时姿态平滑函数取最大值. H 和 L 分别是已跟踪轨迹中目标的长轴和短轴在时间窗口 C 内平均高度和长度. 由于前景分割中存在误差, 外接椭圆宽长比较大时, 外接椭圆姿态角与实际目标姿态角之间可能存在较大的误差, 因此, 姿态能量输出置信度小; 反之应输出较大的能量.

在长时间信号分析阶段, 若观测的时间窗长度为 N , 则在观察时间窗口内其平滑性约束能量函数可以表示为:

$$E(X_1, X_2, \dots, X_N) = \sum_{i=1}^{N-1} E(X_i, X_{i+1}) \quad (7)$$

因为目标运动轨迹平滑, 且目标姿态呈平滑缓慢变化, 因此, 对式(7)求最小化即可得到我们所期望的目标轨迹:

$$E(T_1, \dots, T_N) = \min_{X_1, \dots, X_N} \sum_{i=1}^{N-1} E(X_i, X_{i+1}) \quad (8)$$

上式全局优化问题可以通过 Viterbi 算法快速求解:

$$E(T_1, \dots, T_{i-1}, X_i) = \min_{X_{i-1}} [E(X_{i-1}, X_i) + E(T_1, \dots, T_{i-2}, X_{i-1})] \quad (9)$$

初始条件: $E(X_N) = 0$.

利用 Viterbi 算法对式(8)求解的运算复杂度为 $O(N \times n^2)$, 其中 n 为当前优化问题中每一帧存在最多的候选目标数量. 由于在长时间信号分析之前做了短时间时空联合信号处理, 前景图像中剩下的目标数量已经大大减少, 因此该算法能够满足实时性的要求.

为提高 LTS 目标检测和跟踪的鲁棒性, 还需考虑室外远程监控中通常发生的由于物体遮挡、多目标重叠等引起的目标前景图像暂时性丢失的问题. 为了解决这一问题, 本文将目标分成速度为 0 与速度非 0 两类目标. 如果目标速度为 0, 可以简单地认为目标已经停止运动, 那么在动态规划的过程中并将其从检测结果中排除出去. 如果速度非 0, 那么根据上阶段目标的位置和速度给定一个虚拟目标位置. 若在后续几帧内发现估计位置附近出现目标, 那么可以认为中间所缺失的几帧发生了目标丢失现象.

3 实验及结果分析

为验证我们提出方法的有效性及鲁棒性, 利用本方法对各种室外场景下的小目标进行检测与跟踪实验. 本文算法测试所采用的硬件平台为普通 PC 机, 平台采用主频为 2.8GHz 的 Intel Core i5-2300 双核处理器, 4G DDR2 800 内存, 对于 320×240 的图像进行处理, 动态规划时间窗口长度为 15 帧, 每 5 帧做一次动态规划

处理. 由于在进行动态规划处理前,进行了 STS 处理,在动态规划过程中,每一帧图像中候选目标数量较少,因此在跟踪过程中,动态规划所消耗的资源较少,最后平均处理速度高达 30fps,能够满足实时系统的要求.

在本文的实验中,目标跟踪的轨迹的不同颜色表示不同目标的身份标识.

图 4 显示了前景提取中短时间窗口目标处理的重要性.原始视频通过混合高斯模型背景建模得到的前景图像中不仅包含目标信息,而且包含了较强的背景噪声(图 4(b)所示).由于噪声信号在帧间的独立性,经过本文提出的短时间信息处理方案,可以滤除绝大部分的背景噪声(如图 4(c)所示).

第一组跟踪实验如图 6 所示,对雨天条件下的运动目标进行跟踪.为了便于观察实验结果,我们将目标区域局部放大叠加在图像的左上角.在该视频中,雨天以及地面积水使得前景分割结果中存在强噪声,另外目标运动路径中间有一个黑色灯桩造成对运动目标的短时间遮挡.但是在跟踪结果中可以看出,被跟踪的目标当中没有虚假噪声.另外,根据第 160 帧和第 210 帧的跟踪结果我们可以发现,目标跟踪轨迹的颜色没有发生变化(即两个目标为同一个目标),这说明我们提出的算法在雨天环境条件下对小目标具有较强的跟踪能力,并在目标被短时间遮挡的条件下,算法仍然能持续地跟踪运动目标.



(a) 第160帧跟踪结果



(b) 第210帧图像的跟踪结果

图6 雨天及目标遮挡条件下的跟踪实验,运动目标为骑自行车的人

果可以看出,左起第二个和第三个目标发生运动交叉后,跟踪轨迹的属性没有改变,即运动目标的身份识别号没有改变.从这个实验可以看出,我们的算法能够较好地适应多目标交叉的情形.



(a) 112帧的跟踪结果



(b) 第186帧的跟踪结果

图7 多运动目标交叉情况下的实验结果,运动目标为行人及骑自行车的人

第三组小目标跟踪实验如图 8 所示,对低对比度条件下的远程目标进行跟踪.为了便于观看实验结果,我们将原始视频对应帧的目标区域进行局部放大并进行对比度增强,然后将处理后的局部图片放置于跟踪结果中运动目标的上方.从图中可以看出,即便将目标进行放大和增强,在单帧图像中,人眼仍然很难对目标进行分辨.在实验中,由于人眼难在单帧图像中观测到目标,检查实验效果时,我们通过观察原始视频的视频流对实验结果进行验证.

在该组实验中,要对目标进行跟踪对于人眼来说也是困难的,但从第 125 帧和 200 帧图像的跟踪结果可以看出,利用我们的算法,能够稳定持续地对低对比度运动小目标进行持续跟踪.在这种情况下,当徒步者经过足球门的后面时,本文的算法仍然能够稳定、持续地跟踪目标.

通过以上三组实验我们可以看出,我们所采用的视频数据,目标在图像中所占的面积小,对比度低,人眼在单帧图像中很难分辨出目标.在这种数据条件下,传统的基于特征点^[5,6]、轮廓特征^[9,10]以及纹理特征^[7]等跟踪方法很难适用,而本文提出的方法能够长时间对目标进行正确的持续跟踪.

第二组实验如图 7 所示.从 112 和 186 帧的处理结



(a) 第125帧

(b) 第200帧

图8 低对比度情况下的小目标跟踪结果，
运动目标为运动场上跑步的人

4 结束语

针对户外小目标场景条件下图像信噪比低、目标成像面积小、缺乏稳定目标特征等问题,基于人类视觉目标检测与跟踪的原理,提出一种基于双时间尺度信号分析的目标跟踪方法,在短时间尺度分析中,通过时空联合滤波,对随机瞬时噪声和边缘噪声进行去噪处理,滤除了前景图像中的绝大部分噪声.在长时间尺度信号处理中,分析了目标和持续噪声在三维时空中的运动特征,提出了一种新的平滑性能量约束条件,并用Viterbi算法进行快速求解.实验表明,在复杂户外小目标应用场景下,本文提出的方法能够实时、有效、鲁棒地对远程小目标进行跟踪.

当然,本文提出的全局平滑性能量约束条件在时间尺度上也是有限的,因此在长时间遮挡、或者分割引起的长时间目标信号缺失的条件下,利用本文提出的方法对小目标进行跟踪时,容易引起前后跟踪不一致的问题.

另外,本文提出的目标跟踪算法主要是依赖于速度和姿态的能量平滑性约束条件,因此在多目标跟踪的条件下,可以处理目标交叉等复杂情况.但是,本文通过混合高斯模型的方法提取目标的前景在不同时间帧形态相差较大,从而在跟踪时减弱了目标的区分能力.因此,在下一步工作中,需要对前景检测算法进一步改进,使得前景检测的结果在帧间尽量平滑,使得跟踪算法能够处理更加复杂的情况.

参考文献

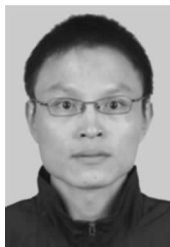
[1] Kim S, Choi B, Kim J, et al. Three plot correlation-based small infrared target detection in dense sun-glint environment for in-

frared search and track[A]. Proceedings of Signal and Data Processing of Small Targets[C]. Baltimore, Maryland, USA: SPIE, 2012. 83930T.

- [2] Sung C. Adaptive and selective multi-target-tracker[A]. Proceedings of Signal and Data Processing of Small Targets[C]. San Diego, California, USA: SPIE, 2011. 81370T.
- [3] Qi B, Wu T, He H, et al. Real-time detection of small surface objects using weather effects[A]. Proceedings of the 10th Asian Conference on Computer Vision[C]. Berlin, Heidelberg: Springer-Verlag, 2011, 6494: 27 - 38.
- [4] Prokaj, J, Z Xuemei, G Medioni. Tracking many vehicles in wide area aerial surveillance[A]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)[C]. USA: IEEE, 2012. 1 - 7.
- [5] Buzzi S, Lops M, Venturino, Ferri M. Track-before-detect procedures in a multi-target environment[J]. IEEE Transactions on Aerospace and Electronic Systems, 2008, 44(3): 1135 - 1150.
- [6] 刘倩,侯建华,牟海军,赵巍,笪邦友.联合生成与判别模型的目标检测与跟踪[J].中国图象图形学报, 2013, 18(10): 1293 - 1301.
- LIU Qian, HOU Jianhua, MOU Haijun, ZHAO Wei, Da Bangyou. Object detection and tracking combining generative and discriminative model[J]. Journal of Image and Graphics, 2013, 18(10): 1293 - 1301. (in Chinese)
- [7] 王欢,王江涛,任明武,杨静宇.一种鲁棒的多特征融合目标跟踪新算法[J].中国图象图形学报, 2009, 14(3): 489 - 498.
- WANG Huan, WANG Jiangtao, REN Mingwu, YANG Jingyu. A new robust object tracking algorithm by fusing multi-features[J]. Journal of Image and Graphics, 2009, 14(3): 489 - 498. (in Chinese)
- [8] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564 - 577.
- [9] Vard A, Jamshidi K, Movahhedinia N. Small object detection in cluttered image using a correlation based active contour model[J]. Pattern Recognition Letters, 2012, 33(5): 543 - 553.
- [10] Tauber C, Batatia H, Ayache A. Quasi-automatic initialization for parametric active contours[J]. Pattern Recognition Letters, 2010, 31(1): 83 - 90.
- [11] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. Fort Collins, USA: IEEE, 1999. 1 - 7.
- [12] Berclaz J, Fleuret F, Fua P. Robust people tracking with global trajectory optimization[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. New York City, USA: IEEE, 2006. 744 - 750.

- [13] Lanz, O. Approximate Bayesianmultibody tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(9): 1436 – 1449.
- [14] 张凯,单甘霖,吉兵,陈海.基于姿态角辅助非线性滤波算法的机动目标跟踪[J].电子学报,2012,40(8):1670 – 1675.
ZHANG Kai, SHAN Gan-lin, JI Bing, CHEN Hai. Maneuvering targets tracking based on pose-aided nonlinear filter algorithm[J]. Acta Electronica Sinica, 2012, 40(8): 1670 – 1675. (in Chinese)
- [15] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564 – 577.
- [16] Yang M, Lv F, Xu W, Gong Y. Detection driven adaptive multicue integration for multiple human tracking [A]. Proceedings of IEEE Conference on Computer Vision [C]. Kyoto, Japan; IEEE, 2009. 1554 – 1561.
- [17] Grabner H, Matas J, Van Gool L, Cattin P. Detection driven adaptive multicue integration for multiple human tracking [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco, USA; IEEE, 2010. 1285 – 1292.
- [18] 刘晨光,程丹松,刘家锋,黄剑华,唐降龙.一种基于交互式粒子滤波器的视频中多目标跟踪算法[J].电子学报,2011,39(2):260 – 267.
LIU Chen-guang, CHENG Dan-song, LIU Jia-feng, HUANG Jian-hua, TANG Xiang-long. Interactiveparticle filter based algorithm for tracking multiple objects in videos[J]. Acta Electronica Sinica, 2011, 39(2): 260 – 267. (in Chinese)
- [19] 王洁,王洪玉,高庆华,张晓云.基于信号特征序列的粒子滤波跟踪算法[J].电子学报,2010,38(10):2297 – 2301.
WANG Jie, WANG Hong-yu, GAO Qing-hua, ZHANG Xiaoyun. Particlefilter algorithm based on signal character sequence for tracking[J]. Acta Electronica Sinica, 2010, 38(10): 2297 – 2301. (in Chinese)
- [20] 杨威,付耀文,潘晓刚,张志勇,黎湘.弱目标检测前跟踪技术研究综述[J].电子学报,2014,42(9):1786 – 1793.
YANG Wei, FU Yao-wen, PAN Xiao-gang, ZHANG Zhi-yong, LI Xiang. Track-before-detect technique for dim targets: An overview[J]. Acta Electronica Sinica, 2014, 42(9): 1786 – 1793. (in Chinese)
- [21] Fortmann T, Bar-Shalom Y, Scheffe M. Sonar tracking of multiple targets using joint probabilistic data association[J]. IEEE Journal of Oceanic Engineering, 1983, 8(3): 173 – 184.
- [22] Sha que K, Shah M. A noniterative greedy algorithm for multiframe point correspondence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(1): 51 – 65.
- [23] Leibe B, Schindler K, Van Gool L. Coupled detection and trajectory estimation for multi-object tracking[A]. Proceedings of IEEE Conference on Computer Vision [C]. Rio de Janeiro, Brazil; IEEE, 2007. 1 – 8.

作者简介



叶喜勇 男,1976年4月出生,浙江省台州人.1994年和2001年于海军航空工程学院通信与信息工程系获得工学学士学位和工学硕士学位.现为在读博士研究生,从事目标检测、跟踪和识别技术的研究.

E-mail: xiyong.ye@qq.com



陶霖密 男.清华大学计算机科学与技术系副教授.1986年于浙江大学获生理学学士学位,1991年在中国科学院生物物理所获认知学硕士学位,2001年于清华大学计算机科学与技术系获计算机应用工学博士学位,主要研究方向为计算机视觉与人机交互,并开展了分布式机器视觉理论与系统、二维、三维非刚体运动估计与跟踪、情感计算、脑机交互等方面的研究.

E-mail: linmi@mail. tsinghua. edu. cn