

# 基于视频分析和多传感器融合的移动式监控系统<sup>①</sup>

叶 锋<sup>1,2,3</sup>, 廖 茜<sup>1</sup>, 汪 敏<sup>1</sup>, 林贵增<sup>1</sup>, 陈超意<sup>1</sup>, 林 晖<sup>1</sup>

<sup>1</sup>(福建师范大学 数学与计算机科学学院, 福州 350007)

<sup>2</sup>(福建星网锐捷通讯股份有限公司 通讯产品研究院, 福州 350002)

<sup>3</sup>(北京邮电大学 多媒体中心, 北京 100876)

**摘 要:** 安防场景下, 大多数传感器系统(视频、红外、烟雾传感)只能实现单一数据采集、简单处理. 提出的移动式智能监控系统以小车为载体, 通过对采集视频进行智能分析处理(包括基于背景差分的入侵检测算法和改进的TLD目标跟踪算法), 并辅助多传感器模块进行多信息融合协同监控, 可以实现入侵检测、移动跟踪监测等功能. 实验表明该系统能提供多种终端形式与用户进行友好交互, 实现真正的智能安防.

**关键词:** 移动式智能监控; 入侵检测; 多信息融合; 自动跟踪

引用格式: 叶锋, 廖茜, 汪敏, 林贵增, 陈超意, 林晖. 基于视频分析和多传感器融合的移动式监控系统. 计算机系统应用, 2017, 26(8): 88-93. <http://www.c-s-a.org.cn/1003-3254/5904.html>

## Mobile Monitoring System Based on Video Analysis and Multi-Sensor Fusion

YE Feng<sup>1,2,3</sup>, LIAO Xi<sup>1</sup>, WANG Min<sup>1</sup>, LIN Gui-Zhen<sup>1</sup>, CHEN Chao-Yi<sup>1</sup>, LIN Hui<sup>1</sup>

<sup>1</sup>(School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

<sup>2</sup>(Fujian STAR-NET Communications Co. Ltd., Fuzhou 350002, China)

<sup>3</sup>(Multimedia Center, Beijing University of Posts and Telecommunication, Beijing 100876, China)

**Abstract:** In security scenarios, most of the sensor systems (video, infrared, smoke sensor) can only achieve simple single data acquisition and processing. A mobile monitoring system based on video analysis and multi-sensor fusion is proposed with car as a carrier. The collected video is analyzed and processed with proposed algorithms (including the background-difference-based the intrusion detection algorithm and the improved TLD algorithm). With collaborative monitoring from multi-sensor, it can achieve intrusion detection, mobile tracking and monitoring. Experiments show that the system can realize the real intelligence security and user friendly interaction

**Key words:** mobile intelligent monitoring; intrusion detection; multi information fusion; automatic tracking

## 引言

虽然传统视频监控与设备经过长期的发展在性能和功能上得到了很大的提高, 但是仍然受到监控时间、人类生理、数据分析、响应时间等诸多因素的限制. 且目前国内使用的视频监控系统大部分均侧重于视频数据的压缩、传输和存储, 其报警功能大都利用红外、烟雾等外接传感器实现, 视频数据到达终端以后还是需要由相关人员来实时观测分析, 或者是存储

起来以备事后查看. 这种系统仍需耗费大量的人力物力, 却只能实现部分的预期功能, 远远无法满足日益增长的对监控系统智能化的需求<sup>[1]</sup>.

针对家庭安防的需求, 本文提出移动式的智能监控系统克服传统摄像头的地域局限性. 基于视频分析, 本系统可以实现对家庭周围环境进行全方位的监控, 加上系统具有行人识别、视频摘要、视频备份、入侵检测和自动跟踪的安防等功能<sup>[2]</sup>, 从而实现住宅智能安

① 基金项目: 福建省引导项目(2016Y0031); 福建省教育厅项目(JA15136); 福建师范大学教学改革研究项目(I201602015); 福建师范大学2015年省级大学生创新训练计划项目(201510394044); 福建省科技厅面上项目(2017J01739)

收稿时间: 2016-12-06; 采用时间: 2017-01-04

防监控. 系统可根据用户设置的特定规则向 PC 端、Andriod 客户端发送预警或其他提示信息, 可以满足用户高效、智能安防的需求.

## 1 系统总体结构分析

提出的系统总体结构框图如图 1 所示. 控制中心(本地服务器)对采集的视频通过视频摘要模块处理压缩后, 将处理完后的视频进行入侵检测分析, 一旦发现可疑人员入侵则启动自动跟踪监测功能. 与此同时将处理完的视频和信息数据上传至云服务器上进行分析整合, 发送预警信息给用户, 用户可通过手机端直接查看并做出相应的处理(如: 对小车发送控制命令); 硬件部分(小兵侦察引擎)主要以小车为主体, 进行移动式视频监控, 并使用传感器进行多信息融合协同监控, 提高系统的性能. 现将各模块的具体功能介绍如下.

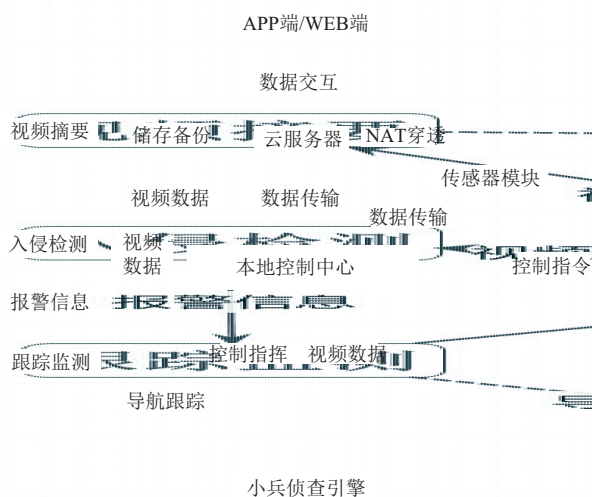


图 1 系统总体结构图

### (1) 视频摘要<sup>[3]</sup>模块

该模块通过对监控视频进行智能分析, 剔除冗余的静态帧, 只保留动态帧, 在不破坏视频的作用的前提下, 最大限度地压缩视频的体积, 大幅降低人工观看视频所带来的麻烦, 为后期快速视频检索提供基础(该模块本文不做具体描述).

### (2) 入侵检测模块

该模块的功能是服务于系统的智能报警通知. 主要实现对环境的监视, 若发现外人入侵, 则进行报警, 并将预警信息发送给用户, 用户可以通过手机端或 WEB 端对此进行相应的处理.

### (3) 跟踪监测模块

该模块的功能是服务于系统的智能导航和追踪监测, 是在入侵检测模块基础上的扩展. 当入侵检测模块检测到有外物入侵时, 则给系统提供导航作用, 对其进行专门的跟踪监测.

### (4) 传感器模块

该模块的功能是多信息融合协同监控, 提高系统的安全性和可靠性. 多种传感器与视频监控相互配合, 从而对环境进行更好的监控. 因为若单一的采用视频智能分析或传感器检测对环境进行监控, 总会存在着不足之处, 而二者融合在一起, 将达到互补的效果. 传感器用于弥补视频监控存在的不足之处, 比如时时检测家中意外(如: 火灾、煤气泄露等)的发生, 做到双重预警的作用. 此外人体红外传感器还能够配合跟踪算法, 使结果更加准确, 让用户更加放心.

### (5) 机器人模块

该模块的功能是为系统提供硬件基础. 此模块选择小车为载体, 主要原因在于其操作方便简单、灵活性强且成本较低. 该小车不仅能够做在主人外出时做到实时巡逻监控, 避免“死角”问题的出现, 更好的守护家居安全.

### (6) WEB 端

该模块的功能在于提供系统与用户的友好交互. 主要用于提升本系统的方便性. 用户不需要下载手机 APP, 便能通过访问网页的形式, 对小车工作环境情况进行实时监控.

### (7) APP 终端

该模块的功能用于丰富系统与用户的交互形式. 该模块用户除了可以通过手机端随时随地查看小兵工作环境情况之外, 还能接收到预警信息, 通过手机端控制小车, 及时对存在着危险隐患做出相应的处理.

### (8) 云服务器

该模块通过对来源于控制中心(本地服务器)的信息整合分析后进行分类保存. 并把处理好的信息(如: 预警信息)发给用户手机端, 方便用户进行相应的处理操作.

## 2 基于背景差分的入侵检测算法模块

入侵检测是指在用户指定的图像区域内, 检测是否有运动目标出现, 若有, 则向用户进行报警. 运动目标检测从场景序列图像中剔除静止的背景区域, 将运

动的目标(前景区域)分割出来,并尽可能地抑制背景噪声和前景噪声,以准确得到感兴趣的运动物体,从而实现识别目标以及跟踪目标.常用的运动目标检测方法一般可以分成三类:帧间差分法、光流法和背景差分法<sup>[4]</sup>.考虑到入侵检测问题的实际应用,其背景较为恒定,采用背景差分法进行入侵检测.主要思路在于利用当前视频序列图像与从视频序列图像中提取的背景图像进行差分处理来检测运动区域.在复杂的应用场景中背景差分法困难在于背景模型的不断建立并不断的更新,以保持与实际背景中精确度近似.

其处理流程图如图2所示.系统将做差后将差分图像转换到灰度空间后,再进行二值化操作.如此,运动的前景就与设置的背景分离开来.在检测过程中只需对前景像素占像素总数的比例进行判断,若高于某一阈值,则判定为入侵事件发生.此处阈值确定方法采用模板学习定值法,具体步骤如下:

- (1) 对已有相同场景下的静态视频进行帧差,计算非零像素点个数占总图像的比例.
- (2) 统计各区段比例,得到直方图,并将数量达到90%以上的值作为视频摘要的阈值.
- (3) 用该阈值对待摘要视频进行摘要.

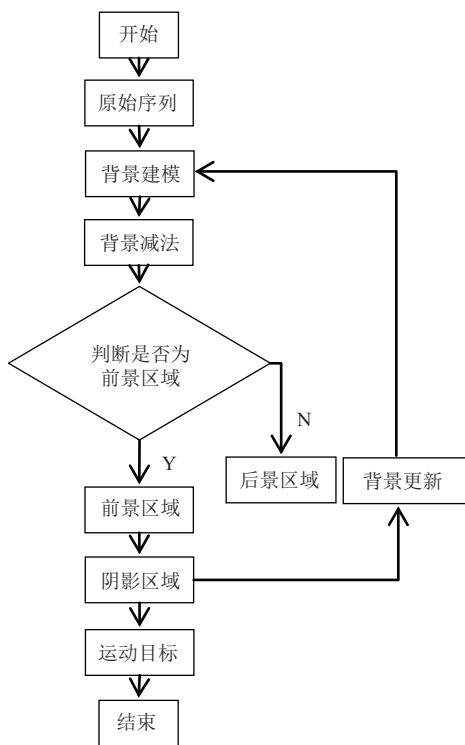


图2 背景差分示意图

### 3 基于改进的 TLD 的跟踪监控模块

当系统检测到有物体入侵之时,跟踪监控模块会发送预警信息,并截图保存入侵画面.此时,用户可选择是否跟踪此物体.系统采用改进的 TLD(Tracking Learning Detection)跟踪算法<sup>[6]</sup>,实现自动控制小车进行自动跟踪物体.TLD 是一种单目标长时间跟踪的算法,该算法将跟踪算法和检测算法相结合来解决被跟踪目标在被跟踪过程中发生的形变、部分遮挡等问题,针对目标旋转变换可能存在的跟踪失败问题,本文引入了 SIFT 算子.通过一种改进的在线学习算法不断地更新跟踪模块的显著特征点以及检测模块的目标模型和相关参数,使跟踪效果更加稳定、鲁棒、可靠.

#### 3.1 改进的 TLD 算法流程框图

基于 TLD 算法的跟踪监测模块算法流程如图3所示:主要由跟踪模块(Tracker)、检测模块(Detector)以及学习模块(Learning)构成.跟踪模块在目标可见情况下用来观察连续帧间的目标动向.检测模块将每张图看成是独立的,然后去定位.学习模块将根据跟踪模块的结果用来对检测模块的错误进行评估,并生成训练样本来对检测模块的目标模型进行更新,避免以后出现类似错误.

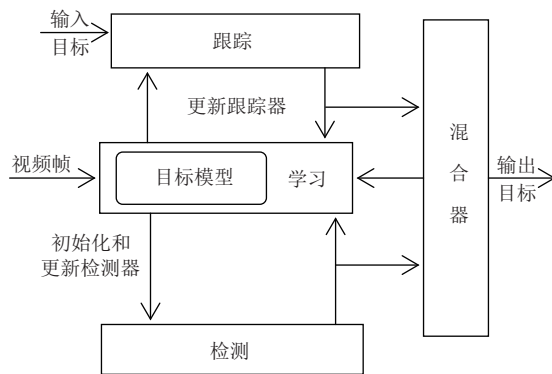


图3 跟踪监测算法模块流程框图

#### 3.2 目标模型

在任意时刻下,被跟踪的目标均可以用其状态属性来表示.目标形状使用图像块(patch) $p$ 来表示,每一个图像块都是从图像目标的矩形框内部采样得到的,并归一化到  $16 \times 16$  像素的大小.两个图像块( $p_i, p_j$ )的相似度计算如式(1),其中 NCC 为归一化相关系数:

$$S(p_i, p_j) = 0.5(NCC(p_i, p_j) + 1) \quad (1)$$

目标模型是一系列正样本  $p^+$ 、和负样本  $p^-$  的集合,

即:  $M = \{P_1^+, P_2^+, \dots, P_m^+, P_1^-, P_2^-, \dots, P_m^-\}$ . 对任意给定的一个图像块  $P$  和目标模型  $M$ , TLD 算法定义了几个量化指标:

- 正样本最近邻相似度

$$S^+(P, M) = \max_{P_i^+ \in M} S(P, P_i^+) \quad (2)$$

- 负样本最近邻相似度:

$$S^-(P, M) = \max_{P_i^- \in M} S(P, P_i^-) \quad (3)$$

- 前 50% 正样本的正最近邻相似度:

$$S_{50\%}^+(P, M) = \max_{P_i^+ \in M \wedge i < \frac{m}{2}} S(P, P_i^+) \quad (4)$$

- 相关相似度:

$$S^r = \frac{S^+}{S^+ + S^-} \quad (5)$$

- 保守相似度:

$$S^c = \frac{S_{50\%}^+}{S_{50\%}^+ + S^-} \quad (6)$$

其中,  $S^c$  的变化范围从 0 到 1, 其取值越大表示图像块的前 50% 越有可能属于正样本.  $S^r$  取值越大表示前 50% 的图像块越有可能是目标区域.

### 3.3 目标检测器和分类器

目标检测器通过扫描窗口来扫描输入图像帧, 然后通过级联分类器判断出每一个图像块中有没有目标. 扫描窗口参数设置为: 缩放比例的步长系数为 1.2, 水平步长为宽度的 10%, 垂直步长为高度的 10%, 最小的扫描窗口大小取 20 个像素. 因此, 考虑到需要处理矩形框的数量巨大, 级联分类器包含三个模块:

#### (1) 图像块方差滤波模块

此分类器模块去除所有像素方差小于被跟踪图相框像素方差 50% 的图相框, 图像框  $P$  方差计算如下:

$$D(P) = E(P) - E^2(P) \quad (7)$$

#### (2) 集成分类器

经过图像块方差检测之后未被去除的图相框进入集成分类器. 此模块利用 2Bit BP 特征, 采用随机森林进行分类. 决策树将目标像素和图像框的像素比较产生一个后验概率  $P_i(y|x)$  ( $y \in \{0, 1\}$ ), 集成分类器把平均后验概率大于 50% 的图相框认定为其包含目标.

#### (3) 最近邻分类器

相似度 ( $S^r, S^c$ ) 被用于指出一个任意的图像块和目标模型中的部分有多大的相似, 如果  $S^r(P, M) > \theta_{NN}$ , 那么图像块  $P$  被分类为正样本, 即被认为包含目标.

### 3.4 学习模块

TLD 采用的学习器是一个半监督在线执行的过程. 针对输入视频流的每一帧, 通过 P-N 学习模块评估当前检测器, 确定出其存在的错误并及时更新以避免将来出现的类似错误. 其主要思想就是让检测器的错误能够被两种类型的专家 (P-experts 和 N-experts) 标识出. P-experts 用来评估错误的负样本 (即现有负样本中可能存在的正样本), N-experts 用来识别错误的正样本. 而他们的相互独立性又能够相互弥补双方的错误.

### 3.5 基于 SIFT 与中值流的跟踪器

TLD 算法的跟踪模块采用的跟踪算法是在中值流跟踪 (Median-Flow tracker) 方法的基础上通过增加跟踪失败检测算法而形成的. 中值流跟踪方法在上一帧的目标框中选择若干个像素点作为特征点, 在下一帧中寻找上一帧中的特征点在当前帧的对应位置. 然后将这若干个特征点在相邻两帧之间的位移变化进行排序, 得到位移变化的中值, 利用该中值, 得到小于中值 50% 的特征点, 将这 50% 的特征点作为下一帧的特征点, 并依次进行下去.

中值流跟踪算法的前提假设是目标是可见的, 所以当目标完全被遮挡或者消失于视野, 则不可避免地出现跟踪失败; 同时 TLD 算法对目标旋转等视觉变换效果较差. 为了能够解决这些问题, 本文在原有算法基础上添加 SIFT 检测<sup>[7]</sup>的跟踪策略, 采用文献[7]提出的 SIFT 特征提取目标上的一些局部外观的与目标的大小、形状、旋转无关的特征点, 具体算法如下:

$ts \leftarrow$  中值检测的阈值

$sift \leftarrow$  采集初始状态的 SIFT 特征

WHILE true

$d[n] \leftarrow n$  个特征点的移动位移(像素)

$dm \leftarrow$  对  $d[n]$  排序, 取中值

$dc = \frac{1}{n} \sum_{i=1}^n (d[i] - dm)$

IF  $dc < ts$  THEN

跟踪成功

ELSE

$flag \leftarrow sift$  与失败时预测目标 SIFT 匹配结果

IF  $flag = true$  THEN

跟踪成功, 目标进行了旋转

$sift \leftarrow$  新的外观的特征值

ELSE

跟踪失败

END IF  
END IF  
END

这个策略能够很稳定地就确定出由剧烈移动或者遮挡所造成的跟踪失败,且对旋转变形目标有较好跟踪效果.

### 3.6 综合器

综合器(Integrator)把检测器和跟踪器得到的目标框予以综合,并作为 TLD 最后的输出.如果跟踪器或者检测器都没有得到目标框,那么就认定当前帧中被跟踪目标没有出现的,否则,综合器将具有最大保守相似度的图像块作为最终的目标框所在位置.

## 4 多信息融合监控模块

为弥补单一采用视频监控带来的不足之处,本系统采用多信息融合监控的方式.将视频智能分析与传感器分析二者相结合,互补不足之处,从而更加有效对安全隐患做出及时预警处理,其实现流程如图 4 所示.

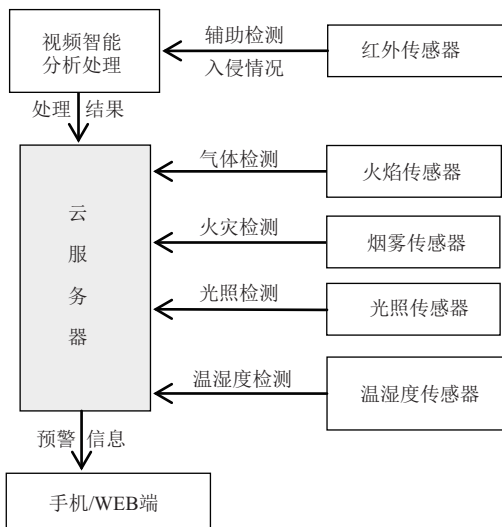


图 4 多信息融合决策算法

## 5 系统应用效果与性能分析

构建的移动式智能监控装置如图 5 所示.为验证多信息融合监控是否能提高本系统性能,做了如下性能测试:

(1) 在安全装置内点燃若干张废纸(模拟烟雾源),将带有 MQ-2 传感器的设备靠近烟雾源,观察现象,发现 LED\_Y 灯闪烁,服务器接收到信息,小车做出相应操作,测试结果成功.

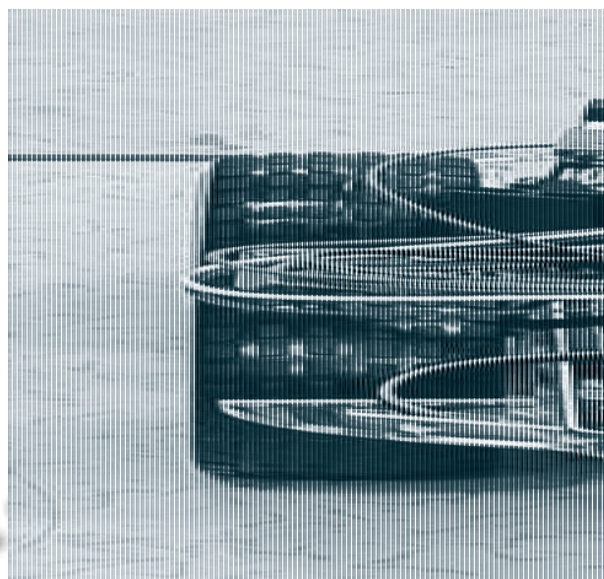


图 5 实体效果图

(2) 用打火机模拟出火焰,将带有火焰传感器的设备靠近烟雾源,观察现象,发现 LED\_H 灯闪烁,服务器接收到信息,小车做出相应操作,测试结果成功(而且,测试火焰距离为 80 cm,且火焰越大,测试距离越远).

(3) 当有外人进入监测范围时,遮盖住摄像头,带有 HC-SR501 人体红外感应模块的设备能识别人体的存在,此时 LED\_R 灯闪烁,服务器接收到信息,小车做出相应操作,测试结果成功.

### 5.1 入侵检测与目标跟踪效果

图 6 和图 7 分别描述了 PC 端和手机端入侵检测效果.系统对输入的视频进行背景差分,并能对入侵的事物进行检测,并发送预警信息通知客户端.用户可以通过客户端启动对目标物体的跟踪(如图 8),且对于倾斜旋转的目标物体具有良好跟踪效果.经测试,采用改进目标跟踪可以持续跟踪对象:对于 3 分钟采集视频测试,平均成功跟踪帧数比例为 96.5%.

### 5.2 多传感融合监控手机客户端效果

视频监控和多传感融合监控不仅可弥补仅视频监控的不足之处,如:无法进行气体检测等,还可辅助视频监控(人体红外),做到双重检测,提高入侵识别的准确率.手机客户端效果图如图 9 所示,软件设计上具有如下的四大功能.

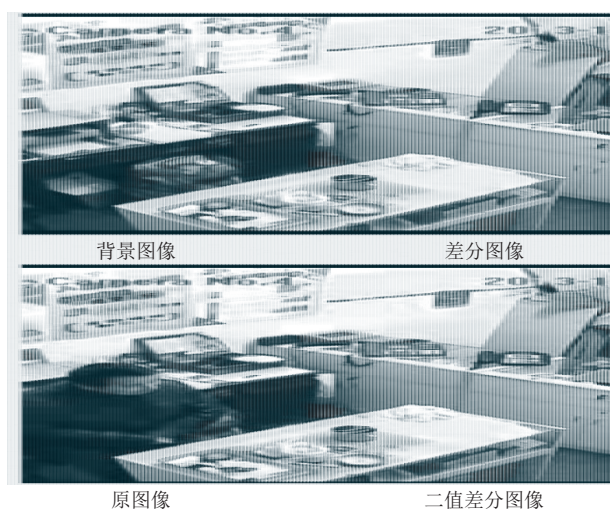


图6 PC端入侵检测效果图



图7 手机端入侵检测效果图



图8 目标跟踪效果



图9 多信息融合效果图(湿度、温度、光照)

- 基于 Android 手机平台的移动视频监控终端的软件, 实现远程视频监控, 随时随地对小车发送命令指示(如控制其走向).
- 能够调节播放视频的大小, 以及进行定格某一幅画面.
- 接收系统发送的预警等信息.

## 6 结语

本文提出的基于视频智能分析和多传感器融合的移动式监控系统, 可对小车所在区域的周围环境进行全方位监控, 加上系统具有的行人识别、视频摘要、视频备份、入侵检测、自动跟踪等功能. 经实验应用证明能够实现真正的智能安防监控. 系统提供多种终端方式与用户进行友好交互.

## 参考文献

- 1 吴群, 王田, 王汉武, 等. 现代智能视频监控研究综述. 计算机应用研究, 2016, 33(6): 1601-1606.
- 2 Davies ER. Computer and machine vision: Theory, algorithms, practicalities. 4th ed. Singapore: Elsevier Pte. Ltd., 2012.
- 3 王娟, 蒋兴浩, 孙铁锋. 视频摘要技术综述. 中国图象图形学报, 2014, 12: 1685-1695. [doi: 10.11834/jig.20141201]
- 4 邬书跃. 基于支持向量机和贝叶斯分析技术的入侵检测方法研究[博士学位论文]. 长沙: 中南大学, 2012.
- 5 亚森江·木沙, 赵春霞. 在线复合模板模型表示的视觉目标跟踪. 中国图象图形学报, 2015, 20(9): 1199-1211. [doi: 10.11834/jig.20150907]
- 6 龚小彪. 基于 TLD 框架的目标跟踪算法研究[硕士学位论文]. 成都: 西南交通大学, 2014.
- 7 曾峦, 王元钦, 谭久彬. 改进的 SIFT 特征提取和匹配算法. 光学精密工程, 2011, 19(6): 1391-1397.