

文章编号:1004-9037(2012)06-0652-06

基于 TMS320DM8168 的视频监控跟踪系统

管 庆 朱 海 王 凯 宋金涛 袁高爽茜

(电子科技大学通信与信息工程学院,成都,611731)

摘要:基于达芬奇系列芯片 TMS320DM8168,开发一个带智能目标跟踪功能的、视频监控与处理系统。系统含有基本的视频监控功能,并加入均值偏移算法和人机交互界面,使目标跟踪与用户操作有机结合在一起。该系统具有体积小,功能高度集成,算法实现高效,良好的用户体验和交互方式等特点。系统可以脱离计算机独立运行,从而能够大大降低设备成本。用户可以通过交互界面指定视频中的目标,启用智能分析算法对用户指定的目标进行跟踪,再由交互界面中绘出目标的运动轨迹,从而为用户提供快速、准确和智能的视频分析业务。本系统通过对均值偏移算法的实现并对其高度优化,解决了遮挡情况下目标的跟踪问题,提高了目标的跟踪能力和目标识别能力。

关键词:智能监控;均值偏移;跟踪

中图分类号:TP911.72

文献标识码:A

Video Surveillance Tracking System Based on TMS320DM8168

Guan Qing, Zhu Hai, Wang Kai, Song Jintao, Yuan Gaoshuangxi

(School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China)

Abstract: The design is based on TMS320DM8168, and it implements a video surveillance and processing system with the function of intelligent object tracking. This system has functions of basic video monitoring, tracking with the mean shift algorithm and human-computer interaction interface. Meanwhile it has several characteristics such as small in size, function highly integrated, algorithm efficient, and good user experience interactive etc. This design can run independently without computers, which can greatly reduce the costs. Users can point out the target in the video by the interaction interface, and then the system will use the intelligent analysis algorithm to track the specific target, and draw the motion path of the target. It offers quick, accurate and intelligent video analysis services. This system can finish the job of target tracking efficiently by using the modified mean shift algorithm which can track the target under occlusion successfully.

Key words: intelligent surveillance; mean shift; tracking

引 言

随着数字信号处理器的迅猛发展,图像处理技术也日益成熟。移动目标的视频检测与跟踪是图像处理、分析应用的一个重要领域,也是当前相关领域的研究前沿。传统的闭路电视只能实现7×24小时不间断录像,造成大量带宽和存储空间的浪费,而且使得检索也非常麻烦。而智能视频跟踪可以实

现报警后才传输或报警触发录像的功能,从而大大缓解了网络资源的紧张,降低了对带宽和系统的要求,录像检索支持时间和报警类型等多种索引方式,真正做到了高效、精准、方便快捷的监控。智能视频跟踪技术有着广泛的用途,例如,它可用于统计进入某区域的人数、评估行走的速度和方向、探测人和物体的可疑移动、识别车牌、确定遗弃物的无人看管时间。

2010年3月, TI公司推出了 TMS320DM8168

达芬奇 (DaVinci™), 将高清多通道系统的所有捕获、压缩、显示以及控制功能完美整合于单芯片之上, 从而不断满足用户对高集成度、高清视频日益增长的需求。

本系统是一个基于 TMS320DM8168 的、带目标跟踪功能的视频监控与处理系统。基于 McFW 框架, 构建了视频输入、处理、输出通路; 在 DSP 核心上有效地实现背景提取, 目标识别和目标跟踪, 从而可以根据用户需求检测多种违规行为, 如贵重物品丢失, 跨越警戒线, 区域内徘徊等等。本系统改进了均值偏移算法^[1], 从而能更加准确有效地提供智能视频分析服务。

1 系统方案

1.1 系统指标

系统硬件指标如下:

- (1) 核心处理器 TMS320DM8168 (C674x 工作于 400 MHz, ARM Cortex-A8 工作于 1 GHz)。
- (2) 同时支持 4 路智能视频分析 (最多支持 16 路)。
- (3) 支持智能视频分析算法。
- (4) 支持 PAL/NTSC 标清输入。
- (5) 支持 QT 用户交互界面。
- (6) DDR 最大工作频率可达 1 066 MHz。
- (7) 支持 1 GB 的 NAND FLASH。
- (8) 核心板大小 9.5 cm × 10.5 cm。
- (9) 系统平均功耗 8 W, 待机 6 W, 算法处理时 10 W。

系统软件指标如下:

- (1) 支持全天候的智能视频监控、分析和实时处理。
- (2) 标清输入下均值偏移跟踪算法的帧率达到 25 帧/s。
- (3) 采用多通道视频处理框架 (Multichannel framework, McFW)。
- (4) 支持 QT 图形界面库。

1.2 系统方案

本系统的目的是实现一个完整的智能视频分析系统, 系统框图如图 1 所示。

2 系统硬件设计

TMS320DM8168 智能视频分析系统的硬件平台设计包括系统硬件平台整体结构, 各个模块的功能, 器件的选型关键环节。系统硬件框图如图 2

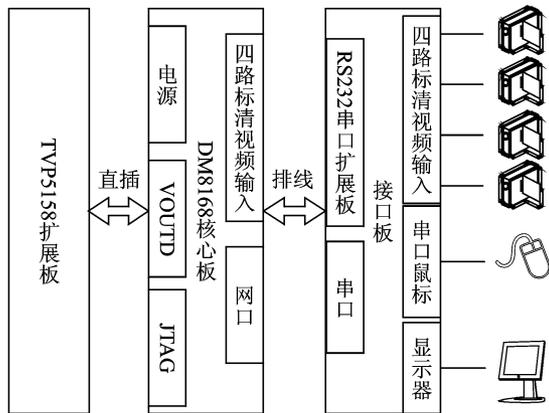


图 1 系统结构框图

所示, 主要分为 4 个模块: 视频处理模块、存储模块, 电源模块和接口模块。

目前视频处理模块可以接收 4 通道模拟信号, 将其通过 TVP5158 视频解码器解码成 YUV 格式。视频输入信号会被同时传送给 TMS320DM8168, 经过视频处理后, TMS320DM8168 芯片内带有视频数模转换器 (Digital analog converter, DAC), 通过运算放大器芯片 OPA361 就可以把视频图像直接输出到视频插孔, 连接到显示器上显示。

TMS320DM8168 的 ARM Cortex-A8 和 DSP C6747 主频都为 1 GHz, 处理能力强大。TMS320DM8168 内带有的 Cortex-M3 核负责视频处理子系统 VPSS, 其对来自 TVP5158 的视频信号进行接收, 编解码, 以及视频输出; A8 负责整个 SoC 的控制; C6747 运行算法对视频进行处理。

存储模块包括的存储设备有 1 GB 的 DDR3 内存、1 GB 的 NAND FLASH 以及 4 MB 的 SPI FLASH。DDR3 内存作为操作系统和系统软件的运行空间, 其单片大小为 256 MB, 主频最高可达 1 600 MHz, DDR3 连接在 TMS320DM8168 的 EMIF0 和 EMIF1 总线上。1 GB 的 NAND FLASH 用于存储 Boot loader、启动环境变量、Linux 内核、文件系统等数据。

系统电源部分采用 5 V 直流电源输入, 通过 TPS65232 的可调 BUCK 控制器模块产生 5.0 V 电压, 5.0 V 电压通过 TPS65232 的两个电源转换模块, 转换成 3.3 V 和 1.0 V_CON 电压; 5 V 同时输入到 TPS56420 产生 1.5 V 电压; 用 5 V, 3.3 V 和 1.0 V_CONN 通过 TPS65001 产生 0.9 V 和两路 1.8 V 电压; 3.3 V 通过 TP40041 产生 1.0 V, 称之为 1.0 V_AVS 电压; 3.3 V 和 1.5 V 通过 TPS51200 产生 0.7 V 电压。

接口模块包括了 4 路视频输入的视频插孔,1 路视频输出插孔;为了方便调试和以后的开发,主要接口有 JTAG,10 M/100 M 以太网口,RS232 串口,其中网口用了 PHY 芯片 ET1011 和 RJ45 插座,以及已经扩充存储的 GPMC 扩展接口。

3 系统软件设计

本系统是一个基于 TMS320DM8168 的、带目标跟踪功能的视频监控与处理系统。基于 McFW 框架^[2],构建了视频输入、处理、输出通路;在 DSP 核心上实现了均值偏移算法,在 Cortex-A8 核心提供了用户界面与文件系统,使目标跟踪与用户操作有机地结合在一起。

3.1 TMS320DM8168 软件结构

如图 3 所示,TMS320DM8168 片内共有 4 个处理核心,系统软件需要使 TMS320DM8168 的 4 个核协同工作起来。其中 Cortex-A8 为 Host,运行

linux 系统,可以动态地为其他核心装载程序,管理其他核程序的运行。同时,linux 带有文件系统、TCP/IP 协议栈、外设驱动,其上运行有 Qt 图形用户界面以及 FTP 服务器,处理与用户的交互。

Target 部分由 3 个核心构成,运行 BIOS6 操作系统。其中 VPSS-M3 主要管理视频的输入输出,完成视频 buffer^[3]的捕获、存储搬移和显示处理、翻译处理;Video-M3 负责完成视频的编解码^[4];C674x DSP 负责运行均值偏移算法,实现对目标的跟踪。各个处理核心通过核间通信(Inter-process communication,IPC)来完成数据的传递以及视频 Buffer 的共享。

3.2 TMS320DM8168 视频处理通路

多通道视频处理框架(Multi channel framework,McFW)针对视频的捕获、处理、显示,为开发者创建了一系列的处理节点。这些节点能够自动地管理对应的硬件资源或者实现相应的软件功能。

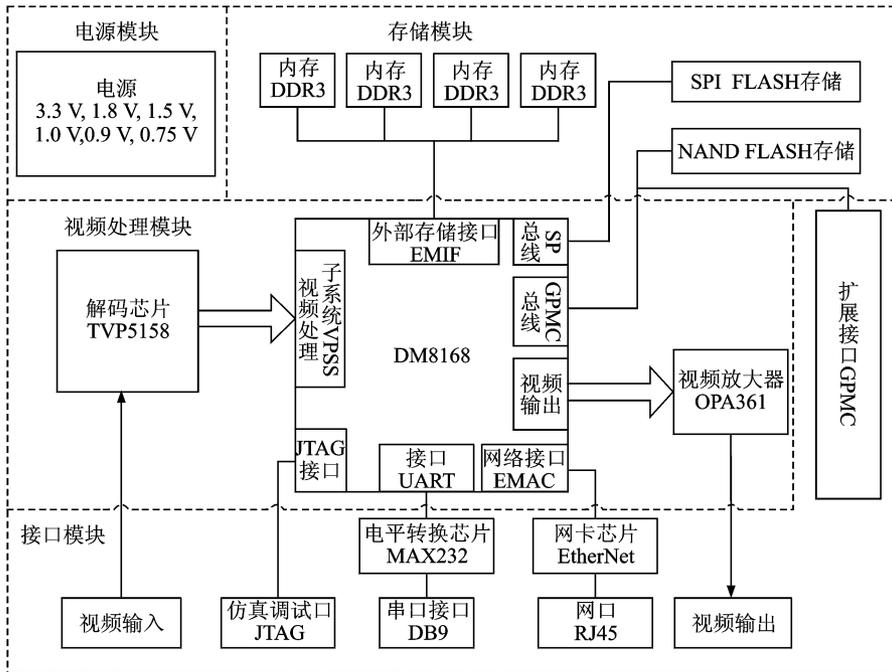


图 2 系统硬件结构框图

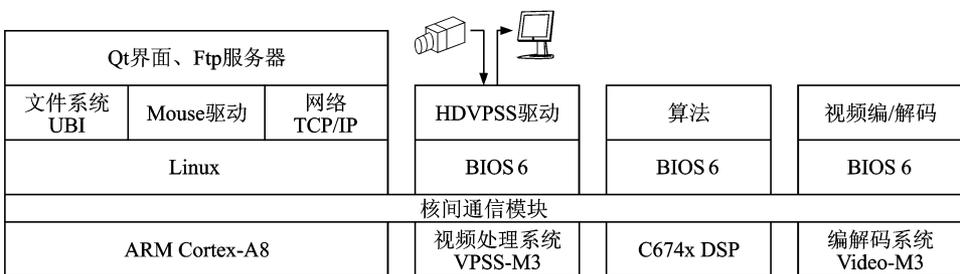


图 3 TMS320DM8168 软件结构图

开发者可以自由地将节点连接起来,构建自己的视频处理流。

本系统的视频处理流程如图4所示。在VPSS-M3通过CaptureLink获得一路D1格式的信号,同时通过NullSrcLink产生一路空白的视频信号,将它们融合成同一视频队列后传递给Cortex-A8。用户通过图形用户界面指定待跟踪的目标,将目标的信息放入视频Buffer中appData指针指向的结构体。C674x DSP根据指定的目标信息对目标进行跟踪,并将成功跟踪目标的视频放入NullSrcLink产生的那一路;SWMS Link可以将多路视频信号拼在同一画面中显示。最后在输出显示之前共有3路视频信号,包括输入的原始视频、成功跟踪上目标的视频,以及将前两者同时显示在同一画面上的视频,用户可以自由切换显示。

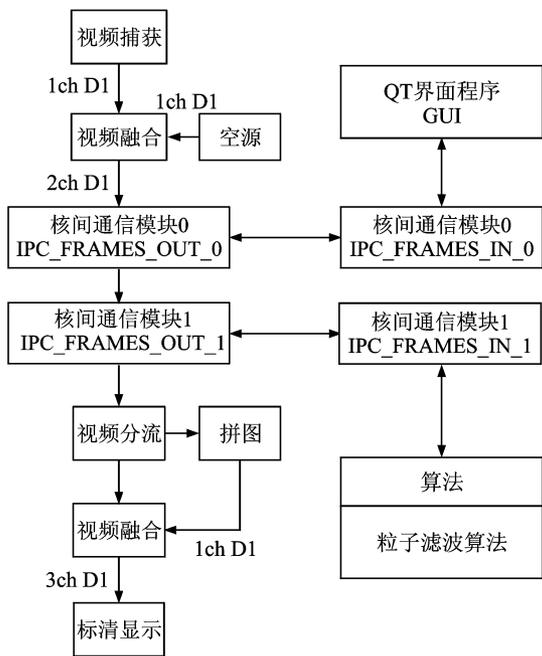


图4 视频处理流程软件结构图

3.3 均值偏移算法

对目标进行跟踪,首先要对目标进行表示,即要提取出有效的特征来表征目标。然后就是采用一定的算法逐帧地或者每隔一定的帧数对目标进行定位。选取RGB直方图作为目标的特征,采用Mean Shift算法^[4]对目标进行定位和跟踪。

以该区域为目标模板,在后续的帧中找到与模板最为相似的区域,就是通过这个过程实现对目标的跟踪。

首先提取区域中的特征,设特征空间为 m 维,

则目标模板和候选的目标区域都可以表示为这个 m 维特征空间中的一个向量,即

$$\hat{q} = \{\hat{q}_u\}_{u=1, \dots, m} \quad \sum_{u=1}^m \hat{q}_u = 1$$

$$\hat{p}(y) = \{\hat{p}_u(y)\}_{u=1, \dots, m} \quad \sum_{u=1}^m \hat{p}_u = 1 \quad (1)$$

式中 \hat{q} 为模板的特征向量, $\hat{p}(y)$ 为位置为 y 的候选区域的特征向量^[5]。在本文中,特征向量实际上就是颜色直方图,直方图的量化区间被划分为 m 个bin,选取 $m=8 \times 8 \times 8=512$,即RGB每种颜色的量等级为8。

设目标的初始位置为 x_0 ,按下式计算模板区域的特征向量

$$\hat{q}_u = C \sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right) \delta[b(x_i) - u] \quad (2)$$

式中 $\{x_i\}_{i=1, \dots, n}$ 为目标模板区域内的像素点的位置, h 为控制核函数的作用范围,函数 $b: R^2 \rightarrow \{1, \dots, m\}$ 将像素点的位置映射到对应像素点所在的颜色区间。 $k(x)$ 为核函数,本文选取Epanechnikov函数

$$k(x) = \begin{cases} 1-x & x \leq 1 \\ 0 & \text{其他} \end{cases} \quad (3)$$

核函数对不同像素点进行了加权,中心的像素点权重较大,处于边缘的像素点权重较小。 C 为归一化参数,为

$$C = \frac{1}{\sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right)} \quad (4)$$

同理,也可以计算候选区域的特征向量

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_k} k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \delta[b(x_i) - u] \quad (5)$$

式中各参数的意义与模板区域计算式相类似。

采用Bhattacharyya系数 $\hat{\rho}(y)$ 衡量模板与候选区域的相似性,其计算公式如下

$$\hat{\rho}(y) \equiv \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (6)$$

如前面所述,跟踪的过程就是在当前帧中找到位置 y ,使 $\hat{\rho}(y)$ 最大。如果从 y_0 处开始搜索目标的位置,那么可以将 $\hat{\rho}(y)$ 在 y_0 处进行泰勒级数展开,取其线性近似即

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} + \frac{1}{2} \sum_{u=1}^m \hat{p}_u(y) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \quad (7)$$

将 $\hat{p}_u(y)$ 表达式代入,可将式(7)改写为

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} +$$

$$\frac{C_h}{2} \sum_{i=1}^{n_h} w_i k \left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \quad (8)$$

其中 w_i 为对应像素点的权重, 故有

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\mathbf{y}_0)}} \delta[b(\mathbf{x}_i) - u] \quad (9)$$

观察上式, 第 1 项只与 \mathbf{y}_0 有关, 为了使 $\hat{\rho}(\mathbf{y})$ 达到最大, 实际上需要使第 2 项取得最大值。形式如第 2 项的式子就可以使用 Mean Shift 算法逐步迭代到其最大值。Mean Shift 算法主要计算式为

$$\mathbf{y}_1 = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i w_i g \left(\left\| \frac{\mathbf{y}_0 - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_h} w_i g \left(\left\| \frac{\mathbf{y}_0 - \mathbf{x}_i}{h} \right\|^2 \right)} \quad (10)$$

其中 $g(x) = -k'(x)$ 。采用式(10)可以计算出一个新的位置 \mathbf{y}_1 , 使得对目标定位向与模板更为相似的区域移动。此过程实际上就是由一个旧的位置 \mathbf{y}_0 计算得出一个新位置 \mathbf{y}_1 , 使新的位置更接近目标所在的位置, 如此迭代数次就能够到达目标所在的位置。

综上所述, 整个 Mean Shift 算法的流程如下。第一帧开始, 首先根据目标框(由用户自己设定)计算模板区域的特征向量 $\hat{\mathbf{q}}$; 第二帧以后每输入一帧新的数据就按以下步骤处理, 进入步骤(1)时将 \mathbf{y}_0 设为上一帧目标的位置。

(1) 计算 $\hat{\mathbf{p}}(\mathbf{y}_0)$ 和 $\hat{\rho}(\mathbf{y}_0)$;

(2) 计算 \mathbf{y}_1 ;

(3) 计算 $\hat{\mathbf{p}}(\mathbf{y}_1)$ 和 $\hat{\rho}(\mathbf{y}_1)$;

(4) 当 $\hat{\mathbf{p}}(\mathbf{y}_1) < \hat{\mathbf{p}}(\mathbf{y}_0)$ 时, 令 $\mathbf{y}_1 = (\mathbf{y}_1 + \mathbf{y}_0)/2$, 返回步骤(3)。如果返回次数超过 3 次, 则不再返回, 执行步骤(5);

(5) 令 $\mathbf{y}_0 = \mathbf{y}_1$, 返回步骤(1), 循环 3 次。

按上述步骤所得 \mathbf{y}_1 就是当前帧的目标位置, 如此逐帧地得到目标的位置就实现了目标的跟踪。当目标走出画面时停止跟踪模块, 等待用户的下一次输入。

为了使算法在 DSP 上高效、实时地实现, 对算法做了如下优化:

(1) 考虑到核函数的计算比较费时, 可以预先计算好核函数的值, 在需要时通过查表获得结果。

(2) 将算法改为定点实现。

(3) 算法的计算主要涉及到一个小的区域而不是涉及到整个画面, 所以将区域内的数据拉直, 读入 cache, 以减少存取时间。

(4) 尽量减少循环的复杂度, 一个循环只实现几种简单的运算。这样就可以展开循环, 进一步减

少了计算的时间。

通过以上处理, 整个模块的运行时间比原来减少了上百倍, 实现了目标的实时跟踪。

均值偏移跟踪算法在目标被遮挡的情况下, 仍然可以对目标进行准确的跟踪。如图 5(a)所示。

在初始帧选取目标, 一辆小车用均值偏移跟踪算法对其进行跟踪。小在经过大树的过程中, 依次经历了部分遮挡和完全遮挡。在这个过程中, 均值偏移的跟踪椭圆一直在锁定并跟踪目标, 说明均值偏移在遮挡情况下, 仍然可以对目标进行正确跟踪, 具有很好的鲁棒性。对于图 5(b), 行人在经历众多的遮挡后, 跟踪椭圆也仍然可以锁定目标, 准确地画出其运动轨迹, 这个过程充分体现了均值偏移跟踪算法的优越性。大部分跟踪算法, 如卡尔曼滤波^[6], 都不能解决遮挡情况下目标的跟踪问题, 而均值偏移跟踪算法, 在短时间部分遮挡及完全遮挡情况下, 可以准确地跟踪到目标, 上述实验结果说明了这一点。均值偏移跟踪算法在解决遮挡情况下目标的跟踪问题, 具有无可比拟的优越性。



(a) 跟踪效果1



(b) 跟踪效果2

图 5 均值偏移跟踪算法

3.4 评测结果

系统测试结果如表 1 所示。

表 1 系统测试结果

被测项目	系统指标	实测结果 (平均值)
支持标准制式	PAL/NTSC	PAL/NTSC
系统平均功耗/W	8	7.5
目标跟踪的准确率/%	>90	80>
系统处理帧率/(帧·秒 ⁻¹)	25	27
DDR 工作频率/MHz	1 066	800
核心板尺寸/(cm×cm)	9.5×10.5	9.2×10.2

4 结束语

本系统设计实现了一款基于 TI DaVinci 系列 TMS320DM8168 的智能视频监控跟踪系统。系统体积小,使用简单,能提供稳定可靠的智能视频监控与跟踪服务。本智能视频分析系统应用领域极为广泛,涉足安防、交通、金融、电力、远程教育等多行业,可应用于交通管理,车流量或人流量统计,自动柜员机的实时监控,数控机床的边缘检测等方案,市场前景广阔。

参考文献:

[1] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift[C]//Proceedings IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, USA: IEEE, 2000:142-149.

- [2] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C]//Proc IEEE CVPR Conf. Fort Collins, USA: IEEE, 1999.
- [3] Brasnett P, Mihaylovat, Canagarajah N, et al. Particle filtering with multiple cues for object tracking in video sequences [C]//Proceedings of SPIE's 17th Annual Symposium on Electronic Imaging. San Jose, California, USA: SPIE, 2005:16-20.
- [4] Horprasert T, Harwood D, Davis L S. A statistical approach for real-time robust background subtraction and shadow detection [C]//IEEE ICCV' 99 FrameRate Workshop. California, USA: IEEE, 1999.
- [5] Comaniciu D, Ramesh V. Mean shift and optimal prediction for efficient object tracking[J]. International Conference on Processing, 2000,13:70-73.
- [6] Maggio E, Cavallaro A. Hybrid particle filter and mean shift tracker with adaptive transition model [C]//IEEE International Conference on Acoustics, Speech, and Signal Processing. [S. l.]: IEEE, 2005:221-224.

作者简介:管庆(1968-),男,副教授,研究方向:数字信号处理,E-mail:qguan@uestc.edu.cn;朱海(1989-),男,硕士研究生,研究方向:数字信号处理;王凯(1986-),男,硕士研究生,研究方向:数字信号处理;宋金涛(1988-),男,硕士研究生,研究方向:数字信号处理;袁高爽茜(1987-),女,硕士研究生,研究方向:数字信号处理。