

文章编号:1000-582X(2010)06-115-06

## 结合均值偏移和多特征的自动人头识别

赵敏 孙棣华 张路 何恒攀

(重庆大学 自动化学院 重庆 400044)

**摘要:**为提高固定单目垂直摄像方式下人头目标识别的正确率,提出一种新的头部目标区域获取方法。首先给出基于 Mean-shift 的人头目标分割算法,由于综合考虑了像素点在空间信息和色彩信息的联系,能够较为完整地分割出人头部目标候选区域。在此基础上,基于运动人头区域的轮廓具有近似圆形以及人头发色具有聚类性 2 个关键特征,提出并建立了基于发色信息的头部区域评价模型和基于连通域边缘轮廓的头部目标评价模型来实现人头部目标区域的识别。实验结果表明,提出的算法能有效抑制光照的影响和消除与发色分布类似的伪目标,静态图像检测正确率约为 89.4%。

**关键词:**均值偏移;人头识别;发色分布;单目视觉;图像分割;目标识别

**中图分类号:** TP317.4

**文献标志码:** A

## Automatic head recognition by integrating mean shift with multiple features

ZHAO Min, SUN Di-hua, Zhang Lu, HE Heng-pan

(College of Automation, Chongqing University, Chongqing 400044, P. R. China)

**Abstract:** To improve the head detection accuracy in video sequences captured with fixed vertical monocular camera, a novel method of head recognition based on mean shift and multiple features is proposed. Firstly, mean shift-based image segmentation algorithm with color information and spatial information is suggested to derive the candidate head components in images. Furthermore, by referring to two features that the contour of human head regions are approximate round and the hair color distribution is clustered, the evaluation models based on the contour information and inside color information of candidate head components are presented for head recognition. The experimental results show that the proposed algorithm can effectively reduce the light interfere and eliminate fake target whose color information is similar to hair color distribution. The detection rate for static images can reach about 89.4%.

**Key words:** mean shift; head detection; hair color distribution; monocular vision; image segmentation; object recognition

基于视觉的人体检测可以通过包括头部、躯干和四肢在内的完整人体进行识别。由于图像采集设备安装位置受限或为了尽可能避免人体间的相互遮

挡,通常只能获取检测区域的俯视图像(vertical view image)。在这类图像中,人体只有部分可见,难以通过基于完整人体的检测方法实现人体的识

**收稿日期:** 2009-01-02

**基金项目:** 国家 863 计划资助项目(2006AA04A124); 重庆市科技计划攻关资助项目(CSTC, 2005AC6037); “211 工程”三期建设项目(S-09108)

**作者简介:** 赵敏(1980-),女,重庆大学博士研究生,主要从事图像处理 and 计算机视觉方向研究。

孙棣华(联系人),男,重庆大学教授,博士生导师,(Tel)023-65106953;(E-mail)d3sun@163.com。

别<sup>[1-2]</sup>。在俯视图像中,人体的头部(尤其是头顶)通常体现得较为完整,因此,头部的俯视信息(如头部轮廓,头部区域颜色分布及相关统计特性等)可作为人体的局部特征来实现对俯视图像中的人体进行检测<sup>[1]</sup>。目前,在涉及到人体检测的应用中,基于人的头部特征进行人体检测(特别是固定单目垂直视觉拍摄方式下),被认为是检测人活动最有效的手段之一<sup>[1-4]</sup>。

现有单目视觉下的头部检测方法可以分为基于头部轮廓模型的方法和基于肤色/发色统计模型的方法<sup>[4]</sup>。在基于头部轮廓的方法中,最具代表性的是头部轮廓域彩色模型<sup>[5-6]</sup>和基于 Hough 变换的头部检测方法<sup>[1-8]</sup>,头部轮廓域彩色模型<sup>[6]</sup>对于正面提取人头效果明显,但对于侧面和倾斜旋转的人头不能有效检测。基于 Hough 变换的头部轮廓特征提取方法计算量较大且容易造成类圆区域或类圆目标的误检测<sup>[1]</sup>。而肤色/发色方法利用人头肤色/发色的统计特性来进行检测,主要有基于肤色的混合高斯模型<sup>[10]</sup>和头部发色色度概率分布模型<sup>[9-10]</sup>。肤色检测方法在图像信息通过垂直摄像方式取得时,该检测方法将失效。基于头部发色色度概率分布模型的头部检测方法易受光照的影响,难以完整地分割出头部目标区域和非头部目标区域<sup>[9-10]</sup>。

由于基于单一特征的头部目标检测方法容易造成目标区域的误检测与漏检测,因此基于多特征的头部目标检测方法是当前研究的热点。文献[1]采用改进的 Hough 变换结合灰度判决器的算法具有较好的检测成功率,但对深色着装乘客的肩部或背部部分区域形成的类圆区域容易造成误检测。文献[10]将颜色检测器和轮廓检测器作为 2 个级联的检测器,基于肤色与发色的颜色检测器容易受光照的影响,难以完整地分割出头部目标区域,容易造成头部目标漏检测,且当图像中存在与发色分布类似的区域时,同样也存在非头部目标误检测。

综上所述,基于头部特征提取的单目视觉头部识别方法目前尚需解决:1)如何在光照变化的情况下完整地分割出头部目标区域,2)如何消除与头部灰度相近的区域(如行人携带的深色物品以及深色着装行人的背部、肩部等)的干扰,提高识别的准确率。

针对上述问题,提出了一种新的结合均值偏移彩色图像分割算法和多特征的人头目标检测方法。首先通过引入 mean-shift 算法分割出人头部目标候选区域,然后建立人头部发色评价模型消除与发色

分布不同的伪目标区域。在此基础上,进一步结合头部候选区域的轮廓特征,建立连通域边缘轮廓与圆相似程度的评价模型,将与发色分布类似的一些区域从候选目标区域中去除,实现人头部目标的正确识别。

## 1 人头部目标识别算法的总体框架

按照“分割-评价”的检测模式,将基于 mean-shift 的人头分割算法与基于发色信息的头部区域评价模型和基于连通域边缘轮廓的头部目标评价模型结合起来,算法流程如图 1 所示,由 2 个步骤构成

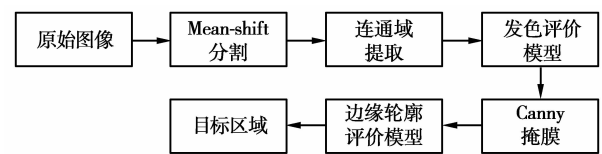


图 1 头部目标识别算法流程

在分割阶段,由于光照影响对人头部区域发色信息产生干扰,易于导致人头目标漏检测。对此,本文提出采用 mean-shift 算法进行人头区域分割。其较传统发色分割方法主要区别在于综合考虑了像素点在空间信息和色彩信息的联系,能够抑制光照变化的影响。它不仅能够将受到光照影响的人头部区域完整分割出来,还能将深色着装的背部区域及其它与发色分布类似的区域分割为连通域整体,便于通过后续处理将其去除。

在识别阶段,基于获得的候选目标区域,建立了基于级联的发色检测器和轮廓检测器。基于发色的检测方法速度快,在不考虑光头、白发或佩戴帽子的情况下,发色检测可将所有头部区域检测出来,不会造成漏检,并可出色度分布与发色分布不同的伪目标清除,但对于与发色分布类似的一些区域如背景等可能造成误检。为此,在前述发色检测基础上,以连通域边缘轮廓与圆的相似程度作为另一判别标准,排除与发色分布类似的伪目标区域干扰,提高检测准确性。

## 2 候选人头目标区域建立

### 2.1 基于 Mean-shift 算法的人头部目标分割

Mean-shift 算法是一种非参数密度梯度估计过程,最早由 Fukunaga 和 Hosteler 提出,而后由 Cheng 以及 Meer 等进一步发展<sup>[11-12]</sup>。目前,Mean-shift 算法已成功运用于特征空间的分析、图像平

滑、图像分割和目标跟踪中。文献[12]已证明 Mean-shift 向量为局部均值的向量,具有变步长梯度上升和轨迹光滑的特性。Mean-shift 算法所具有的这一重要特性是其被引入到人头部分割的另一关键因素。

将基于无参概率密度估计的 Mean-shift 算法应用到图像分割中,必须对其做适当的修改,以适用于图像中的应用<sup>[13-14]</sup>。一个典型图像中的任何像素点可由二维的空间坐标和  $p$  维的矢量值来表示。当  $p=1$  时,为灰度级图像,当  $p=3$  时为彩色图像。在这 2 个空间中,引入欧几里德距离,结合像素的空间信息和值域矢量信息,将图像中的像素用  $2+p$  维的联合空间-值域矢量来统一表示。定义在此假设情况下的联合矢量所对应的核函数为

$$K_{h_s, h_r}(x) = Ck\left(\left\|\frac{x^s}{h_s}\right\|^2\right)k\left(\left\|\frac{x^r}{h_r}\right\|^2\right), \quad (1)$$

其中:  $C$  为归一化常量;  $x^s$  和  $x^r$  分别为联合矢量的 2 维空间分量和  $p$  维值域分量,  $h_s$  和  $h_r$  则为各自相对应的核函数带宽。这样,针对联合空间的 Mean-shift 偏移轨迹表示如下

$$x_{new}^r = \frac{\sum_{i=1}^n x_i^r k\left(\left\|\frac{x_{old}^s - x_i^s}{h_s}\right\|^2\right) g\left(\left\|\frac{x_{old}^r - x_i^r}{h_r}\right\|^2\right)}{\sum_{i=1}^n k\left(\left\|\frac{x_{old}^s - x_i^s}{h_s}\right\|^2\right) g\left(\left\|\frac{x_{old}^r - x_i^r}{h_r}\right\|^2\right)}, \quad (2)$$

$$x_{new}^s = \frac{\sum_{i=1}^n x_i^s k\left(\left\|\frac{x_{old}^r - x_i^r}{h_r}\right\|^2\right) g\left(\left\|\frac{x_{old}^s - x_i^s}{h_s}\right\|^2\right)}{k\left(\left\|\frac{x_{old}^r - x_i^r}{h_r}\right\|^2\right) g\left(\left\|\frac{x_{old}^s - x_i^s}{h_s}\right\|^2\right)}, \quad (3)$$

其中,在 2 个空间中选用相同的核函数,即

$$K_N(x) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}x\right), \quad (4)$$

且:  $g(x) = -k(x)$ 。

根据文献[9-10],LUV 和 Lab 彩色空间最接近人对色彩的感官认识。而对于常用的 RGB 彩色空间,3 个彩色分量并非线性独立。基于这种原因,选用 LUV 作为 Mean-shift 分割的目标彩色空间,因此,  $p=3$ ,加上二维的空间坐标信息,建立的联合空间维数为 5 维。在此基础上采用聚类算法,就可以把空间和颜色距离都相近的点归为一类。采用以下几步完成

令  $x_i$  和  $z_i, i=1, \dots, n$ , 分别为 5 维联合空间中的输入图像像素和输出图像像素。对于每个像素点

①令  $j=1$  和  $y_{i,1} = x_i$ ;

②根据式(2)、(3)计算  $y_{i,j+1}$ , 并计算  $\text{dist} = \|x_{new}^s - x_{old}^s\|$ 。如果  $\text{dist} < \lambda$ , 则进入③, 否则重复②;

③令  $z_i = (x_i^s, y_{i,c}^r)$ 。

当所有点都完成上述处理之后,将坐标空间距离小于  $h_s$  和值空间小于  $h_r$  的点进行合并,形成一个整体的连通域,并将像素值小于 20 的连通域合并到其相邻的连通域中,为后期的目标识别阶段降低运算复杂度。通过上述步骤对图像进行分割后的结果见图 2(b)所示。

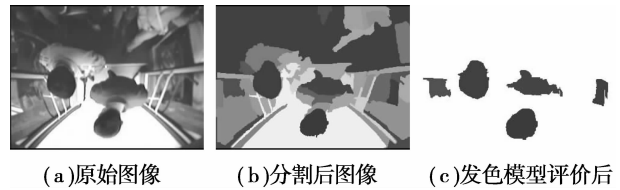


图 2 Mean-shift 分割结果

## 2.2 候选目标区域提取

对经过 mean-shift 分割后的图像进行扫描,提取和标识图像中的连通域,并根据各标识连通域的面积大小、长宽比进行连通域筛选,可获得候选人头目标,得到的候选目标区域可能包括伪目标,如人的背部、背景区域等。

## 3 基于发色模型和轮廓信息的人头目标识别

### 3.1 基于发色模型的人头目标识别

文献[9]结果表明,发色采样点在 YCbCr 彩色空间具有较好的聚类现象,发色在 YCbCr 彩色空间的聚类区域为

$$R_{Cb} = [R_{Cbmin}, R_{Cbmax}] = [111, 141].$$

$$\text{以及: } R_{Cr} = [R_{Crmin}, R_{Crmax}] = [115, 143]. \quad (5)$$

为了识别人头部候选区域,建立了基于发色信息的候选区域属于头部区域的评价模型,来表征该候选区域属于人头目标的可能性或者概率大小。

考虑在 YCbCr 空间中,令  $z$  为候选连通域,  $x_i \in z, i=1, \dots, n$ , 即  $x_i$  为该连通域中的点,且  $x_i = (Y_i, Cb_i, Cr_i)$ , 有

$$f(z) = \frac{1}{n} \sum_{i=1}^n g(x_i). \quad (6)$$

其中

$$g(x_i) = \begin{cases} 0, & \text{others;} \\ 1, & \text{if } R_{Cbmin} < Cb_i < R_{Cbmax}; \\ & R_{Crmin} < Cr_i < R_{Crmax}. \end{cases} \quad (7)$$

上式表明,  $f(z)$  越大, 该区域属于人头部区域的可能性就越大, 反之则越小。当  $f(z) > f_T$  时, 则认为该连通区域为人头部目标区域, 否则为伪人头部目标区域并将其清除。

根据式(5)、(6), 可以将图像中的头部区域全部检测出来, 但当图像中存在颜色信息恰巧与人头部颜色信息十分相近时, 将引入伪目标区域。图 2(c) 为对图 2(b) 的分割结果用发色模型进行头部区域评价的结果, 可以看到, 利用发色模型, 大部分颜色分布与发色分布不同的伪目标连通区域都能被有效地清除, 但部分区域如背景中的一些区域及右一乘客背部部分区域, 在 YCbCr 空间颜色分布与头部颜色分布类似, 因此, 仅通过发色检测不能将其有效滤掉。

### 3.2 基于轮廓信息的人头目标识别

通过对采集的大量原始视频数据分析和观察发现, 一般情况下, 当头部没有佩带较大遮挡物时, 例如帽子等, 人头部区域形状接近圆形, 即人头部区域的轮廓具有近似圆形的特征。连通域形状和圆形的相似性越大, 该连通域属于人头部区域的可能性就越大。为此, 选择各个连通域的轮廓信息作为识别人头部区域的另一标准, 通过建立评价模型来表征连通域相对于圆形的相似性。

利用 canny 算子对发色模型评价后得到的连通域图像进行掩膜处理, 提取出各个连通域的边界后, 令  $z$  为图像分割后任意一个连通域,  $e_i = (x_i, y_i)$ ,  $i = 1, \dots, n$  为该连通域的边界点。假设该连通域的中心为  $R_0 = (x_0, y_0)$ , 选取欧几里德距离作为测度, 则中心到边界点的距离平方为

$$\epsilon_i^2 = \|R_0 - e_i\|^2, \quad i = 1, \dots, n.$$

取目标函数如下

$$\begin{aligned} \text{Min Obj} &= \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n \|R_0 - e_i\|^2, = \\ &= \sum_{i=1}^n (x_0 - x_i)^2 - (y_0 - y_i)^2. \end{aligned} \quad (8)$$

变量约束为:  $R_0 = (x_0, y_0) \in z$ 。

于是, 问题转化为最优化问题。采用牛顿最速下降法进行求解, 得到连通域中心位置  $R_0$ , 进而求出式(8)目标函数的最小值。为了对不同的连通域实施统一的度量, 目标函数进行归一化处理, 获得连通域边缘轮廓与圆相似程度的评价模型

$$\text{Obj}_{\text{Normal}} = \frac{1}{n} \sum_{i=1}^n \frac{\epsilon_i^2}{\epsilon_{\max}} = \frac{1}{n} \sum_{i=1}^n \frac{\|R_0 - e_i\|^2}{\epsilon_{\max}}. \quad (9)$$

其中  $\epsilon_{\max} = \max\{\epsilon_i^2\} | i = 1, \dots, n$ 。

当  $\text{Obj}_{\text{Normal}} < T$ , 该连通域属于人头部目标区域, 否则为伪人头部目标区域并将其清除。  $T$  为阈值, 可以通过大量实验获得。

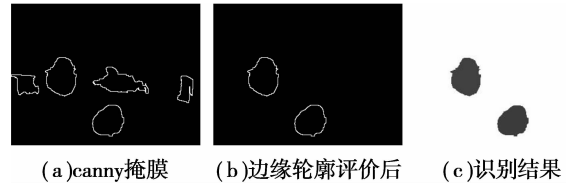


图 3 边缘轮廓评价结果

图 3 为对图 2 发色模型评价后得到的连通域进行边缘轮廓评价的结果。首先利用 canny 算子对发色模型评价后的连通域进行掩膜处理, 得到包括人背部在内的 5 个连通域的边缘图像, 如图 3(a) 所示。然后将连通域边缘轮廓进行圆相似程度评价, 成功消除 3 个伪人头部目标, 正确检测出图像中的 2 个人头部目标, 如图 3(b) 所示。

## 4 实验结果分析

原始实验视频序列通过安装在公交车车门正上方的摄像头采集, 图像大小为  $320 \times 240$ , 采样频率为 30 帧/s。

### ① 分割算法在抑制光照影响方面的作用

在此实验中, 将分割算法与文献[10]做比较, 识别结果如下图 4 所示。

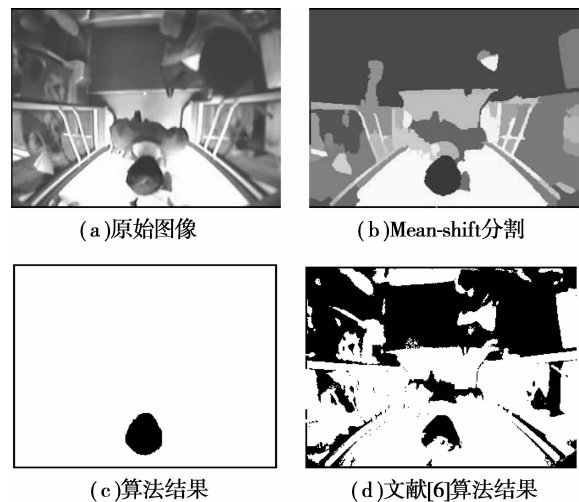


图 4 分割算法在抑制光照影响方面的作用

图 4 中, 当人头部区域靠近车门外部时, 受外部光照影响, 导致相应像素色彩信息偏离正常范围。当采用色度信息分割方法, 如文献[9]的方法, 使得靠近车门外的部分头部区域丢失, 导致去真类型识

别错误,见图 4(d)。分割算法由于考虑了像素空间位置信息和色彩信息,因此在抗光照影响方面有良好的性能。在分割结果中可以看到,人头区域被分割成一个完整的整体,识别结果见图 4(b)和 4(c)。

②算法在消除与发色发布类似区域伪目标方面的作用

将人头目标识别算法与文献[1]采用改进的 Hough 变换结合灰度判决器的人头识别算法进行比较,实验结果如图 5 和图 6 所示:

1)在消除背景中与发色分布类似的类圆区域方面

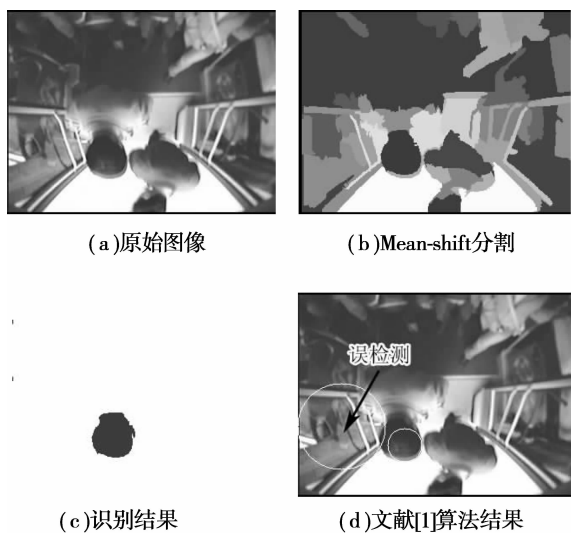


图 5 算法与文献[1]算法在消除背景中的类圆区域比较

从图 5(a)可以看出,当车门上的边缘点与灰度接近头部的背景块叠加容易形成类圆区域,采用文献[1]的方法无论通过 Hough 变换检测或是后续的灰度判决均不能将该伪目标滤出,因此存在背景中的类圆区域误检测如图 5(d),同样的误检测结果可参阅文献[15]。而算法考虑了像素颜色和位置的影响,能将灰度类似但色彩信息不同的物体分别分割出来,有效地避免了背景中的类圆区域的误检测,识别结果如图 5(c)所示。

2)在消除深色着装乘客的肩部或背部形成的类圆区域方面

从图 6(d)可以看出,由于人的背部边缘接近圆形,因此通过 Hough 变换能将其检出,而又由于乘客深色着装,因此文献[1]不能通过后期的灰度判决器将其滤出,造成误检测如图 5(e)所示,同样的误检测结果可参阅文献[15]。而算法将背部分割为一个整体,通过轮廓评价模型将其滤出,识别结果如图

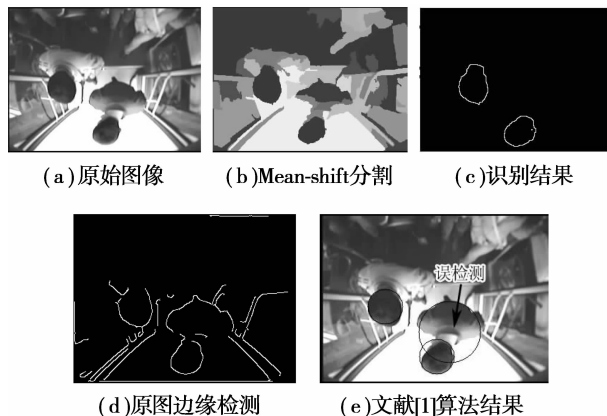


图 6 算法与文献[1]算法在消除深色着装乘客的肩部或背部形成的类圆区域比较

6(c)所示。

③实验结果统计分析

为了表明检测算法的普适性,另选在出入口通道处,以单目垂直方式安装摄像头采集到的图像作为另一组实验场景,所拍摄图像分辨率为  $176 \times 144$ ,结果如图 7 所示。

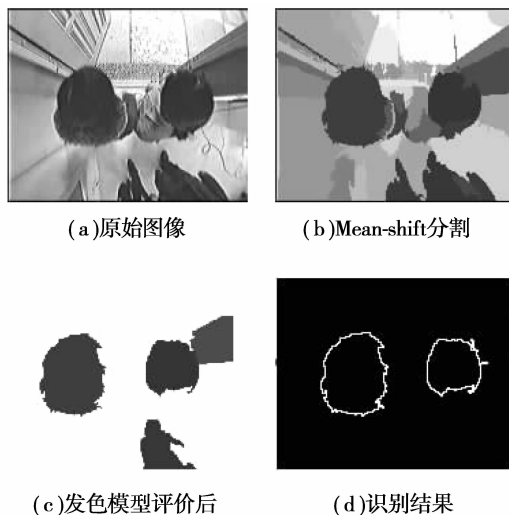


图 7 通道入口处实验结果

图 7 说明检测算法对垂直摄像方式下采集的通道入口处的行人头部目标检测同样适用,并能正确检测出存在女士披肩发情况的头部目标。

将此方法在 PC 机上运行,并以垂直单目方式下采集的公交车乘客上下车序列和出入通道口行人序列进行测试,共计检测 120 帧图像,真实人头目标为 179,漏检测目标 5 个,误检测目标 14 个,检测正确率约为 89.4%,与文献[1]的检测率接近。同时,给出的检测结果是静态图像的检测结果,还可通过后续的目标跟踪来降低目标漏检测和误检测。

## 5 结 论

在固定单目垂直摄像条件下,针对光照变化和头部灰度相近的区域干扰对人头识别准确率的影响,分别采用基于均值偏移的彩色图像分割和基于级联的发色检测及轮廓检测,能够在抑制光照变化影响的基础上,完整地提取出头部目标候选区域有效地消除与发色发布类似区域的伪目标,提高了头部目标检测的正确率。

提出的基于单目视觉的自动人头识别算法可以广泛地应用于行人检测、乘客计数等智能监控场合。由于 Mean-shift 算法计算量较大,对于识别实时性要求较高的场合,尚需进一步改进算法的效率。

### 参考文献:

- [1] 于海滨,刘济林. 应用于公交客流统计的机器视觉方法[J]. 中国图象图形学报, 2008, 13(4): 716-722.  
YU HAI-BIN, LIU JI-LIN. A vision-based method to estimate passenger flow in bus [J]. Journal of Image and Graphics, 2008, 13(4): 716-722.
- [2] 潘浩,高枝宝,何小海,等. 基于计算机视觉的公交系统人流量检测算法[J]. 计算机工程, 2007, 33(11): 216-218.  
PAN HAO, GAO ZHI-BAO, HE XIAO-HAI, et al. Pedestrian flow detection algorithm in public transport system based on computer vision [J]. Computer Engineering, 2007, 33(11): 216-218.
- [3] ZHAO TAO, NECATIA R. Segmentation and tracking of multiple humans in crowded environments [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2008, 30(7): 1-14.
- [4] KELLY P, CONNOR N E, SMEATON A F. Robust pedestrian detection and tracking in crowded scenes[J]. Image and Vision Computing, 2009, 27(10): 1445-1458.
- [5] JUSTIN G, NILES V L, et al. Automatic detection of heads in colored images [C] // Proceedings of the Second Canadian Conference on Computer and Robot Vision, May 9-11, 2005, Victoria, BC, Canada. [S. l]: IEEE, 2005: 276-281.
- [6] StanBirchfield. Elliptical head tracking using intensity gradients and color histograms[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 23-25, 1998, Santa Barbara, CA, USA. [S. l]: IEEE, 1998:232-237.
- [7] 卢湖川,张继霞,张明修. 基于 Hough 变换头部检测与跟踪的方法研究[J]. 系统仿真学报, 2008, 20(8): 2127-2132.  
LU HU-CHUAN, ZHANG JI-XIA, ZHANG MING-XIU. Head detecting and tracking based on hough transform[J]. Journal of System Simulation, 2008, 20(8):2127-2132.
- [8] ZOU W, LI Y, YUAN K, et al. Real-time elliptical head contour detection under arbitrary pose and wide distance range[J]. Journal of Visual Communication and Image Representation, 2009, 20(3):217-228.
- [9] ZHAO M, SUN D H, HE H P. Hair-color modeling and head detection[C]// Proceedings of the 7th World Congress on Intelligent Control and Automation, WCICA '08, June 25-27, 2008, Chongqing, China. [S. l]: IEEE, 2008: 7768-7772.
- [10] 王长军,朱善安. 基于颜色和变形模版的实时人体检测[J]. 中国图象图形学报, 2006, 11(6): 861-866.  
WANG CHANG-JUN, ZHU SHAN-AN. Color and deformable templates based real-time pedestrian detection [J]. Journal of Image and Graphics, 2006, 11(6): 861-866.
- [11] 周芳芳,樊晓平,叶榛. 均值漂移算法的研究与应用[J]. 控制与决策, 2007,22(8):841-847.  
ZHOU FANG-FANG, FAN XIAO-PING, YE ZHEN. Mean shift research and applications [J]. Control and Decision, 2007,22(8):841-847.
- [12] COMANICIU D, MEER P. Meanshift: a robust approach toward feature space analysis [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [13] TAO W B, HAI J. Color image segmentation based on mean shift and normalized cuts[J]. IEEE Transactions on systems, man, and cybernetics, 2007, 37(5): 1382-1389.
- [14] 伊怀锋,黄贤武. 基于均值偏移的彩色图像分割算法[J]. 计算机应用, 2006, 26(7):1605-1608.  
YI HUAI-FENG, HUANG XIAN-WU. Color image segmentation algorithm based on mean shift [J]. Computer Applications, 2006, 26(7):1605-1608.
- [15] 于海滨. 基于头部特征提取的人体检测与跟踪及其应用[D]. 浙江: 浙江大学信息学院, 2007:45-46.

(编辑 侯 湘)