

中图法分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)07-1283-06

论文索引信息: 安国成, 陈建军, 张凤军, 戴国忠. 改进权值计算的均值移动目标跟踪 [J]. 中国图象图形学报, 2011, 16(7): 1283-1288

改进权值计算的均值移动目标跟踪

安国成¹⁾, 陈建军²⁾, 张凤军¹⁾, 戴国忠¹⁾

¹⁾(中国科学院软件研究所, 人机交互实验室, 北京 100190)

²⁾(东南大学, 信息科学与工程学院, 南京 210096)

摘要: 针对基于 Bhattacharyya 相似度的均值移动跟踪算法精度较差的问题, 提出一种基于直方图交集思想的新型颜色分量加权方法, 该方法利用参考模板与候选模板归一化颜色概率密度对应颜色分量的比值作为均值移动算法的加权系数。新权值计算方法在目标快速运动, 有场景相似颜色干扰等情况下具有很好的适用能力, 从而提高目标的跟踪精度。另外处理跟踪过程中, 因摄像机抖动、光照变化等因素导致跟踪线索变化的情况, 利用基于辅助模板的目标更新机制, 有效地解决了目标短暂遮挡以及更新过程中的累积误差问题。通过多组对比实验结果可以看出, 算法具有更强地抑制背景干扰以及特征自适应的能力, 从而提高了均值移动跟踪算法的鲁棒性。

关键词: 复杂环境; 均值移动; 直方图交集; 模板更新

Target tracking based on mean shift with improved weights

An Guocheng¹⁾, Chen Jianjun²⁾, Zhang Fengjun¹⁾, Dai Guozhong¹⁾

¹⁾(Intelligence Engineering Laboratory, Institute of Software Chinese Academy of Sciences, Beijing 100190 China)

²⁾(School of Information Science and Engineering, Southeast University, Nanjing 210096 China)

Abstract: Aimed to solve the poor tracking accuracy problem of mean shift based on Bhattacharyya similarity, a novel weighting method is proposed based on the enhanced histogram intersection. The paper uses the ratio between the color components of the normalized color probability densities of the candidate model and that of the reference model as the weight coefficient. Results show that the new color weighting coefficient method has good adaptability under fast object moving or color similarity in scene, and the target tracking precision is enhanced effectively. In addition, a template updating method based on auxiliary model mechanism is used to solve the problems of camera vibration, illumination and object changing, therefore the cumulative error problem is also effectively alleviated. Through groups of contrast experiments, the new tracker shows its good adaptive ability and its robustness.

Keywords: complex environments; mean shift; histogram intersection; model updating

0 引言

视频序列中的目标跟踪以数字图像处理和计算

机视觉为理论依据, 以动态图像处理技术为基础。它是智能领域中的热点研究之一, 具有广泛的应用前景。均值移动是一种基于密度梯度无参数估计的方法。Comanicu 将其应用到视频跟踪以及图像分

收稿日期: 2010-07-09; 修回日期: 2010-12-13

基金项目: 中国博士后科学基金项目(20100470588); 国家重点基础研究发展计划(973)基金项目(2009CB320804); 国家高技术研究发展计划(863)基金项目(2009AA01Z337, 2008AA01Z303); 国家自然科学基金项目(60673188, U0735004)。浙江大学 CAD&CD 国家重点实验室开放基金项目(A1009)。

第一作者简介: 安国成(1979—), 男, 博士后。主要研究领域为人机交互、视频图像跟踪等。
E-mail: anguocheng99@gmail.com。

割中,并且取得了很好的效果^[1]。Hu 等人^[2]利用目标的空间信息和颜色信息对原有均值移动算法进行增强。Liu 等人^[3]提出利用多特征线索对目标进行跟踪,并采用目标运动信息进行补偿,从而达到信息互补的作用。Ren 等人^[4]提出利用双学习方法来处理目标因光照变化而导致参考模板与候选模板不匹配的问题。

本文对经典均值移动算法在基于 Bhattacharyya 相似度计算的权值系数进行分析,经过推导,指出其在跟踪迭代过程中的缺陷,即对跟踪目标主要颜色分量具有抑制作用,对次要颜色分量具有放大作用。在此基础上本文提出利用直方图交集思想的改进权值计算方法,从而保持了不同颜色分量在参考模板中的作用恒定,使均值移动算法在复杂的跟踪场景下具有更强的鲁棒性以及跟踪精度。

1 经典均值移动算法

被跟踪目标按照下式进行核直方图建模:

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

式中: x_0, x_k 分别表示目标中心像素坐标和第 k 个像素坐标; n 为目标像素总数目; m 为颜色分量总数目; $K(\|x\|^2)$ 是核函数; h 表示核函数的带宽; $\delta(\cdot)$ 是 Kronecker delta 函数;函数 $b: \mathbf{R}^2 \rightarrow \{1, 2, \dots, m\}$ 的作用是计算位置在 x_k 处的像素所属颜色分量; C 为归一化常数,使得 $\sum_{u=1}^m \hat{q}_u = 1$ 。在跟踪的后续过程中,也采用类似的方法,从而得到被跟踪目标在位置 y 的候选模板核概率密度

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

式中, C_h 为归一化因子,使其满足 $\sum_{u=1}^m \hat{p}_u(y) = 1$ 。Comaniciu 采用 Bhattacharyya 相似度来度量两者之间的距离^[1],即

$$d(y) = \sqrt{1 - \rho[\hat{p}(y), \hat{q}]} \quad (3)$$

然后对 $\rho[\hat{p}(y), \hat{q}]$ 进行泰勒展开得到权值

$$\omega_k = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(x_k) - u] \quad (4)$$

跟踪过程就是按照式(5),根据初始位置 y_0 计算新位置。

$$y_1 = \frac{\sum_{k=1}^n x_k \omega_k K'\left(\left\|\frac{y_0 - x_k}{h}\right\|^2\right)}{\sum_{k=1}^n \omega_k K'\left(\left\|\frac{y_0 - x_k}{h}\right\|^2\right)} \quad (5)$$

然后将 $y_0 \leftarrow y_1$, 经过若干次迭代,直到得到目标的最终跟踪位置。

2 经典均值移动算法的缺陷

首先对基于 Bhattacharyya 相似度权值计算方法所具有的不足之处进行定性分析。从上面可以看出, ω_k 的计算方法以及数值的大小将直接影响目标位置求取。为了不失一般性,关注其中一个颜色分量 j 的计算过程。假设在参考模板中颜色分量 j 含有 n_q^j 个像素而在位置 y_0 处的候选模板中含有 n_p^j 个像素,而且假设参考模板一共包含 N_q 个像素,而候选模板包含 N_p 个像素。如果核导数为常数,即参考模板中不对像素的空间位置进行加权,那么式(5)计算的具有颜色分量 j 的像素位置权值为

$$\omega_j = \sqrt{\frac{\hat{q}_j}{\hat{p}_j(y_0)}} = \sqrt{\frac{\frac{n_q^j}{N_q}}{\frac{n_p^j}{N_p}}} = \sqrt{\frac{n_q^j N_p}{n_p^j N_q}} \quad (6)$$

当参考模板和候选模板所包含的像素个数 N_q 与 N_p 相等时,式(6)可以化简为 $\omega_j = \sqrt{\frac{n_q^j}{n_p^j}}$,所以颜色分量 j 在一次迭代过程中的总作用为

$$n_p^j \omega_j = n_p^j \sqrt{\frac{n_q^j}{n_p^j}} = \sqrt{n_q^j n_p^j} \quad (7)$$

如果 $n_p^j < n_q^j$,说明候选区域中的颜色分量 j 不足,并且其在候选区域中所起到的总体作用只有 $\sqrt{n_q^j n_p^j}$ 小于 n_q^j 。如果 $n_p^j > n_q^j$,说明候选区域中的颜色分量 j 受到干扰,并且其在迭代过程中所起到的作用 $\sqrt{n_q^j n_p^j}$ 大于 n_q^j 。当目标尺度变化时,可以近似认为 n_p^j 与 N_p 成线性变化,并不影响所有的分析。

从上面可以看到,基于 Bhattacharyya 相似度权值计算,在跟踪过程中,放大了受干扰颜色分量的整体作用,却减弱了目标主要颜色成分的定位能力,从而降低了跟踪算法抑制背景干扰的能力以及跟踪精度。这种情况的产生是对 Bhattacharyya 相似度进行泰勒展开存在误差导致的。

3 基于直方图交集的相似度系数

视频目标跟踪,也就是在候选区域内对像素位置进行加权,使得经均值移动迭代计算出来的最终

位置为目标中心位置。所以一般来说,只要能够达到此目的的加权方法,都是可以应用到均值移动算法之中的^[5]。

直方图交集^[6]是由 Swain 提出的,他将其应用到基于颜色的图像检索之中,并且取得了很好的效果。直方图交集主要用来比较两个给定直方图 q 、 p 的相似度,具体如下式:

$$\sum_{u=1}^m \min(q_u, p_u) \quad (8)$$

其中 m 为颜色分量总数。然后基于直方图比的概念, Swain 提出后映射直方图 (histogram backprojection) 目标搜索定位算法。首先对每个颜色分量 u , 定义直方图对应颜色分量比 $R_u = \min\left(\frac{q_u}{p_u}, 1\right)$, 然后将图像区域内空间位置的像素数值替换为其对应颜色分量的 R_u , 最后利用与目标大小相等的求和模板对整个图像权值区域进行卷积操作, 从而得到权值之和最大的区域即为目标区域。算法的关键是对不同像素位置加权所使用权值 R_u 的计算。算法使用的是整个图像的直方图 p_u , 对于大多数非目标, 主要颜色分量 p_u 远大于 q_u , 所以大部分权值 R_u 是一个远小于 1 的数; 而目标具有的主要颜色分量 q_u 对应的权值 R_u 才接近或等于 1。这种权值计算方法并不适合均值移动跟踪算法, 对此提出以下改进。

首先使用归一化核直方图 \hat{q}_u 、 \hat{p}_u 替代直方图 q_u 、 p_u , 这样不仅可以避免因参考目标与候选目标尺度大小不一致所带来的权值波动问题, 还可以通过核函数的选择利用目标的空间信息。根据均值移动跟踪算法, \hat{p}_u 应为候选区域的核直方图。一般来说候选区域与参考目标空间大小相近, 由于目标的无规则运动, 可能导致候选区域 \hat{p}_u 包含的目标主要颜色分量较少, 此时计算的 $\frac{\hat{q}_u}{\hat{p}_u}$ 远大于 1, 但是由 R_u 的定义只能令其为 1, 从而严重抑制了目标主要分量在跟踪过程中所起到的积极定位作用。另外, 由于候选区域比整个图像小很多, 其包含的一部分非目标主要分量的 \hat{p}_u 比较小, 按照 $\frac{\hat{q}_u}{\hat{p}_u}$ 计算的权值 R_u 并不是远小于 1, 所以候选区域像素对应的权值都较接近或等于 1, 从而导致目标定位精度很差。

其次, 为了较为合理地计算不同颜色分量所对应的权值, 也就是候选区域内不同位置像素所对应的权值, 对 R_u 的定义进行改进, 为了符号统一, 这里用 ω_k 来表示权值大小, 即

$$\omega_k = \begin{cases} 0 & \hat{p}_u = 0 \\ \frac{\hat{q}_u}{\hat{p}_u} \delta[b(x_k) - u] & \text{其他} \end{cases} \quad (9)$$

由式(9)所定义的权值, 得到颜色分量 j 在一次迭代过程中的所起到的总作用为

$$n_p^j \omega_j = n_p^j \frac{n_q^j}{n_p^j} = n_q^j \quad (10)$$

由此可见, 每一种颜色分量在候选模板中所起到的作用与其在参考模板中所起到的作用相同, 这里称之为作用恒定性质。此种权值计算有效地避免了基于 Bhattacharyya 相似度权值计算所具有的缺陷。

最后, 为了进一步提高权值计算的鲁棒性与合理性, 权值可以修正为

$$\omega_k = \begin{cases} D & \hat{p}_u \geq T_E \hat{q}_u \text{ 或} \\ & \hat{p}_u < T_N \\ \frac{\hat{q}_u}{\hat{p}_u} \delta[b(x_k) - u] & \text{其他} \end{cases} \quad (11)$$

式中 T_E 、 T_N 、 D 为调节因子。式(11)主要是对候选模板中的颜色分量做进一步区分。参数设置主要考虑以下几个要素。对于参数 T_E 的设置主要考虑到候选区域某些颜色分量超过参考模板对应颜色分量, 其主要原因是目标尺度变化。对于这种情况, 因为使用的是归一化直方图, 所以其作用已经被削弱, 最为主要的一种情况是候选模板中存在干扰, 也就是作为背景的颜色分量进入到候选模板。这时虽然使用归一化核直方图, 但是其并不能削弱它们的作用, 所以对 T_E 的设置要兼顾算法容错能力与跟踪精度两个方面。 T_N 主要是降低非主要颜色分量的作用, 防止奇异权值出现。 \hat{p}_u 在满足式(11)的两个条件下, 权值可以直接设置为 D , 对于 D 的设置可以基于不同的考虑, 例如可以按照颜色分量在参考模板中的作用恒定准则, 即退化到式(9)。此时认为当条件满足时, 该颜色分量已经受到严重污染, 所以在跟踪过程中不再考虑对应颜色分量的作用, 令 $D=0$ 。

4 辅助模板更新策略

通常, 在跟踪过程中, 目标的特征很可能会因转动^[7]、光照、摄像机角度而发生变化。对此提出利用目标的统计特性设计一个辅助模板, 利用其对参考模板更新进行指导。

采集 N 幅在不同光照、不同姿态下目标的多幅图像进行辅助模板 B 的构造。对其中第 i 帧目标图

像,按照式(1)构造归一化核直方图,所有目标图像的核直方图为 $\{\hat{q}_{i,u}\}_{i=1,2,\dots,N;u=1,2,\dots,m}$,然后求和得到归一化组合核直方图 $\hat{Q}_u = \frac{1}{N} \sum_{i=1}^N \hat{q}_{i,u}$ 。辅助模板 B 如下式构造:

$$B_u = \begin{cases} 1 & \hat{Q}_u \geq h \\ 0 & \hat{Q}_u < h \end{cases} \quad u = 1, 2, \dots, m \quad (12)$$

式中, h 为克服样本采集过程中的干扰而设置的阈值,其作用是控制辅助模板 B 中可更新颜色特征分量的数目。在跟踪过程中,基于辅助模板 B 的模板更新具体过程如图 1 所示。

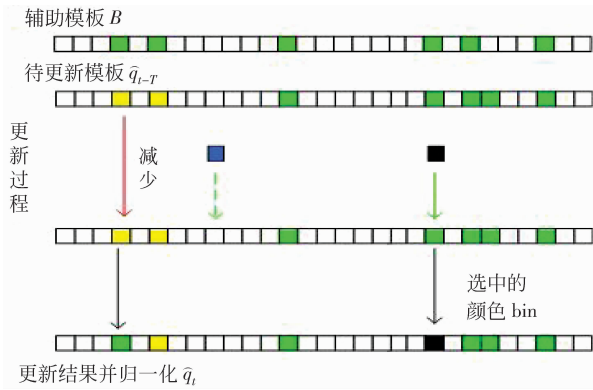


图 1 基于辅助模板 B 更新机制

Fig. 1 The mechanism of the new model update based on B

在图 1 中,辅助模板 B 中绿色代表可更新颜色特征分量,白色代表不可更新颜色分量,黄色代表待更新参考模板 \hat{q}_{i-T} 中存在的颜色分量。在更新过程中,进入模板的颜色对应辅助模板 B 中不可更新的颜色分量,则不对参考模板 \hat{q}_{i-T} 中相应颜色特征分量进行更新,如蓝色方块所代表的颜色分量,反之,如果进入模板的颜色对应辅助模板 B 中可更新颜色分量,则进行相应的更新,如黑色方块所对应的颜色分量。具体按照下式进行:

$$q_i = (1 - \alpha) \hat{q}_{i-T} + \alpha \cdot \hat{q}_i \quad (13)$$

式中, \hat{q}_{i-T} 是第 $t - T$ 帧更新后的归一化参考模板, T 是更新所隔帧数; q_i 为当前第 t 帧经过辅助模板筛选的新参考模板,对其进行归一化得到 \hat{q}_i ; α 为更新速率。由于采用了这种基于辅助模板的更新机制,从而避免了模板更新时机选择的难题,有效解决了目标特征线索变化的问题。整个算法流程如下:

1) 跟踪之前根据具体应用情况构造可更新颜色分量的辅助模板,手动获得初始帧 $i = 1$ 跟踪目标区域以及初始位置 y_0 ,建立参考模板核直方图 \hat{q}_u ;

2) 读入下一帧 $i = i + 1$ 图像,在位置 y_0 处按照式(2)计算当前图像的候选模板颜色核直方图 $\hat{p}_u(y_0)$;

3) 按照式(11)计算候选区域对应像素权值大小;

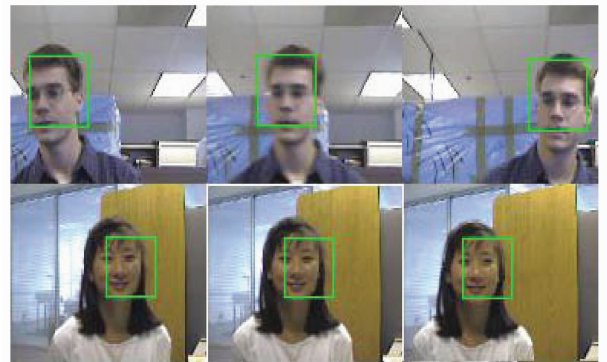
4) 按照式(5)得到更新窗口位置 y_1 ;

5) 如果 $\|y_1 - y_0\| < \varepsilon$,其中 ε 为阈值,其通常小于 1,迭代结束,返回步骤 2),否则 $y_0 \leftarrow y_1$,继续迭代;

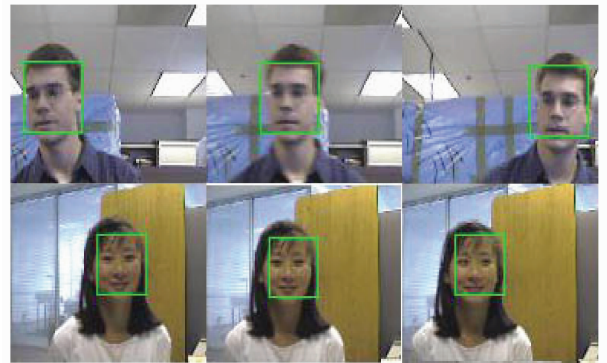
6) 每隔 T 帧,进行模板更新,计算当前跟踪位置的候选区域核直方图,然后参照辅助模板 B 对参考模板中不同的颜色分量进行更新处理,对可更新颜色分量其更新方法按照式(13)进行。

5 实验结果

为了证明本文算法的跟踪能力,我们进行了大量的对比视频跟踪实验,首先验证基于改进直方图交集权值计算的均值移动算法的有效性。实验所使用的图像是斯坦福大学所提供的。图 2 是两个场景的对比实验效果,参数设置为 $T_E = 1.2, T_N = 0.001$,



(a) 经典算法跟踪结果



(b) 新权值跟踪结果

图 2 对比跟踪结果

Fig. 2 Comparison of tracking results

迭代终止条件是 $\epsilon = 1$, 初始跟踪目标为手动选择。

对于第 1 个场景, 跟踪目标作快速的左右摇摆运动, 这时经典均值移动算法在跟踪精度上明显存在偏差, 说明基于 Bhattacharyya 相似度的权值计算对目标由于快速运动所引起的目标模糊的区分能力不如本文算法, 所以导致算法陷入局部极值点。对于第 2 个场景, 是一名女子在类似肤色的家具周围运动, 可以看出在一般较为理想的情况下, 两个算法性能相似, 但是当目标运动到家具范围内, 新算法具有明显的抑制背景干扰的作用。

为了验证整个跟踪系统的有效性, 我们也进行了自拍场景实验。摄像机为手持方式, 所以存在明显的抖动现象。视频图像的分辨率为 208×176 ; 更新速率 $\alpha = 0.1$; 更新时机 $T = 7$

从图 3 可以看出, 当跟踪目标从较为明亮的区域向灯光较为阴暗的区域移动时, 由于光照的变化使得参考模板在候选区域内找不到较为精确的位置, 从而产生偏差, 如第 18 帧。这种偏差随着时间的推移, 在经典算法中并没有得到纠正。但是在新算法中, 由于采用了基于辅助模板更新的机制, 所以从跟踪结果可以看出, 参考模板自适应学习了目标的特征变化, 但是并没有受到背景干扰, 如图 3 (b) 第 18 帧、第 43 帧。从第 108 帧可以看出, 经典算法在跟踪过程中已经产生了明显的偏差, 而且在第 149 帧中, 目标已经丢失。新算法却可以得到较为满意的结果。从图 4 也可以看出, 在整个视频的跟踪过程中, 改进算法一般都具有较小的偏差, 而且有效避免了目标丢失现象。

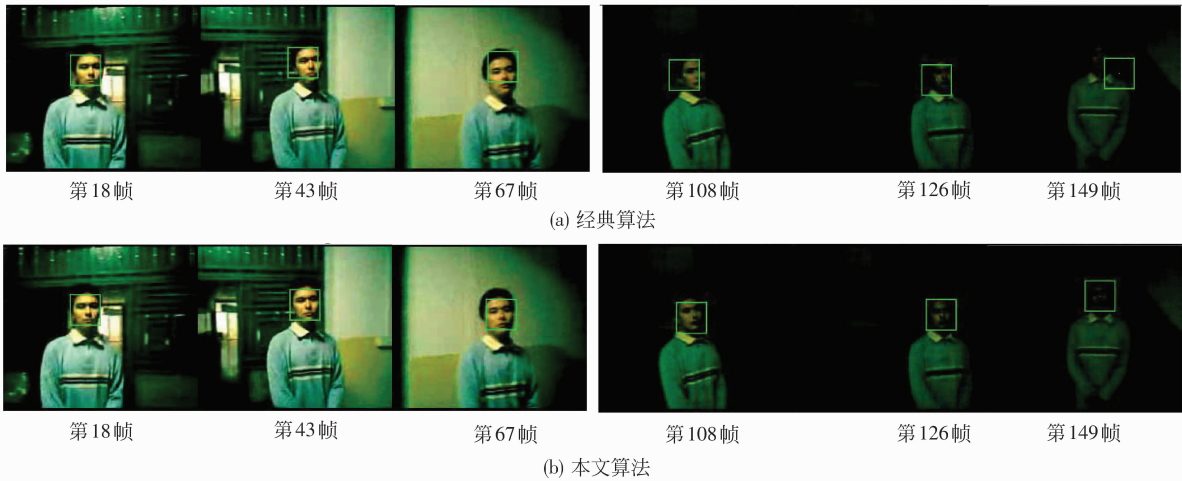


图 3 跟踪对比

Fig. 3 Comparison of tracking results

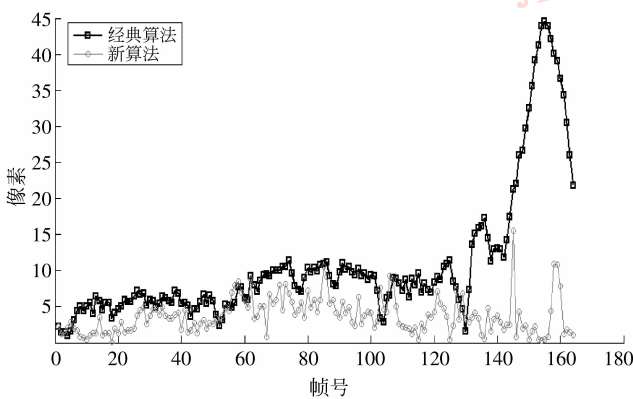


图 4 性能比较

Fig. 4 Performance comparison

6 结 论

提出一种基于改进相似度权值计算的均值移动算法, 增强了均值移动算法在目标跟踪过程中的自适应能力, 从而实现感兴趣目标的连续跟踪。并且分析经典算法以及常规模板更新机制所存在的不足, 然后详细介绍了辅助模板更新机制, 最后从多个真实场景对比实验结果可以看出, 该算法可以很好地解决由目标快速运动、光照变化等因素所导致目标特征变化而引起目标跟踪性能下降, 甚至丢失的问题。另外本文算法经过适当调整, 还可以应用到其他系统之中^[8-10]。

参考文献 (References)

- [1] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [2] Hu J S, Juan C W, Wang J J. A spatial-color mean-shift object tracking algorithm with scale and orientation estimation [J]. Pattern Recognition Letters, 2008, 29(16): 2165-2173.
- [3] Liu H, Yu Z, Zha H B, et al. Robust human tracking based on multi-cue integration and mean shift [J]. Pattern Recognition Letters, 2009, 30(9): 827-837.
- [4] Ren Y, Chua C S. Bilateral learning for color-based tracking [J]. Image and Vision Computing, 2008, (26): 1530-1539.
- [5] Cheng Y Z. Mean shift, mode seeking, and clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790-799.
- [6] Swain M J, Ballard D H. Color indexing [J]. International Journal of Computer Vision, 1991, 7(1): 11-32.
- [7] An G C, Wu Z Y. Head tracking based on prior knowledge model update [J]. Opto-Electronic Engineering, 2008, 35(6): 10-14. [安国成, 吴镇扬. 基于先验知识模板更新的头部跟踪算法 [J]. 光电工程, 2008, 35(6): 10-14.]
- [8] Li Z, Tang Q L, Sang N. Improved mean shift algorithm for occlusion pedestrian tracking [J]. Electronics Letters, 2008, 44(10): 622-623.
- [9] Zimmermann K, Matas J, Svoboda T. Tracking by an optimal sequence of linear predictors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(4): 677-691.
- [10] Zimmermann K, Svoboda T, Matas J. Anytime learning for the NoSLLiP tracker [J]. Image and Vision Computing, 2009, 27(11): 1695-1701.