

文章编号: 1000-5013(2010)03-0282-06

# 一种融合 LBP 纹理特征的多姿态人脸跟踪方法

陈远, 陈锻生

(华侨大学 计算机科学与技术学院, 福建 泉州 362021)

**摘要:** 提出一种改进的 Camshift 算法,它融合目标人脸的局部二值模式(LBP)纹理特征的  $T$  分量,以及肤色的 HSV 色彩空间的  $H$  分量的统计直方图来生成概率分布图像,实现纹理与肤色特征的有效融合;然后,利用 Kalman 滤波器来预测目标人脸的运动信息,快速地跟踪到目标人脸.实验表明,在复杂的跟踪条件下,这种算法比原始的仅采用颜色直方图信息的 Meanshift 和 Camshift 算法,在跟踪速度和精度上有显著的提高.

**关键词:** 多姿态人脸;连续 AdaBoost 算法;特征查找表;局部二值模式

**中图分类号:** TP 391.41

**文献标识码:** A

人脸检测与跟踪对于自动人脸处理系统的研究具有重要的意义,在信息安全,智能监控,虚拟现实等领域都具有广泛的应用前景. Rowley 等<sup>[1]</sup>提出利用多个人工神经网络(ANN)来检测多姿态人脸.为了检测平面内旋转的人脸,他们将检测图像进行不同角度的旋转后得到一组检测图像,依次进行检测,以提高平面内旋转的人脸检测精度,但其检测速度降低了. Feraud 等<sup>[2]</sup>采用 CGM (Constrained Generative Model) 检测器分别针对左右摇摆 $[0^\circ, 20^\circ]$ 和 $[20^\circ, 40^\circ]$ 的人脸. Schneiderman 等<sup>[3]</sup>将人脸划分为左侧面、正面和右侧面,并训练了 3 个基于积分图的检测器,分别处理这 3 种情况. 文[4]用 SVM 方法进行多视角人脸的检测和识别;文[5-6]将层叠型检测器扩展为金字塔型检测器,用来检测多视角人脸. 本文提出一种多姿态人脸检测与跟踪算法,并在此基础上构建了一个实时人脸检测与跟踪系统.

## 1 多姿态人脸检测

针对多视角人脸检测系统的低效率问题,在 Real AdaBoost 算法<sup>[7]</sup>的基础上,提出融合肤色特征和利用特征查找表(FST)输出进行姿态预估的 Real AdaBoost 改进算法,并构建出一个多姿态人脸检测系统,如图 1 所示.

### 1.1 离线训练过程

首先,收集人脸样本和非人脸样本. 在预处理阶段,计算各样本的积分图,以便于计算样本的类 Haar 特征;然后,从这两类样本集中提取所有类 Haar 特征,生成特征空间. 在这个特征空间中,应用连续的 AdaBoost 算法进行训练,即对特征空间中的每一个特征构造一弱分类器.

根据归一化因子最小化原则,选择一个最优的弱分类器,并根据其输出更新样本权重,进入下一次的迭代,直到分类误差达到要求为止. 最后,整合所有弱分类器,输出一强分类器,并将其加入到 3 层金字塔检测结构的分类器中. 应用当前离线训练好了的分类器对非人脸样本集进行过滤,去除那些可以

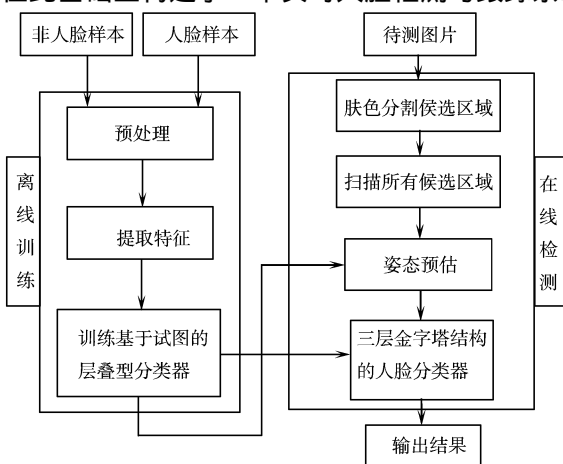


图 1 多姿态人脸检测系统的框架

Fig. 1 System framework for multi-view face detection

收稿日期: 2008-12-25

通信作者: 陈锻生(1959-),男,教授,主要从事数字图像处理与模式识别的研究. E-mail: dschen@hqu.edu.cn.

基金项目: 福建省科技计划重点项目(2008I0021)

被正确分类的样本. 如果非人脸样本集不为空,继续应用连续的 AdaBoost 训练新的强分类器;如此循环,直到所有的非人脸样本集用完为止.

多姿态人脸检测模块分类器的训练吸取了 3 层金字塔检测结构的思想,按照不同的角度范围训练 8 个层叠分类器. 由这 8 个分类器组合成一个完整的多姿态人脸检测模块分类器,如图 2 所示.

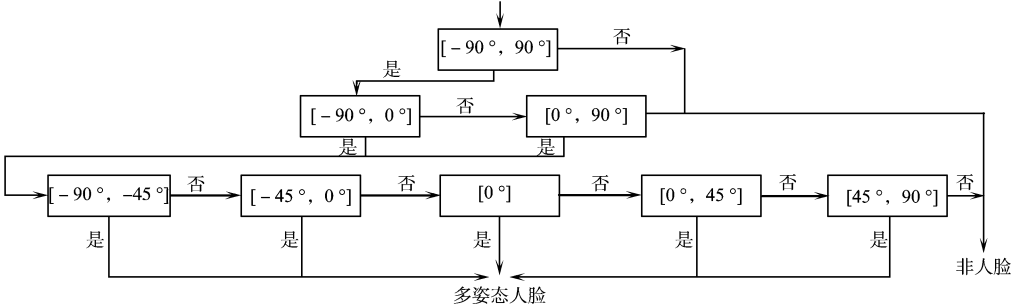


图 2 金字塔检测器的结构框图

Fig. 2 Framework of pyramidal detectors

1.2 在线检测过程

首先,利用 YCbCr 空间的肤色模型分类器分割出人脸候选区域;其次,利用姿态预估和多分辨率搜索进行在线检测的加速. 在进行姿态预估的同时,利用不同尺寸的检测子扫描所有人脸候选区域. 对于每一个候选区域,应用训练好的 8 个层叠分类器组成的金字塔型分类器(包括  $[-90^{\circ}, 90^{\circ}]$ ,  $[-90^{\circ}, 0^{\circ}]$ ,  $[-90^{\circ}, -45^{\circ}]$ ,  $[0^{\circ}, -45^{\circ}]$ , 右侧的 3 个检测器由左侧的镜像生成)判断是否为人脸,输出矩形区域;对所有的矩形区域,合并重叠的子窗口并输出人脸矩形区域.

在检测多姿态人脸时,最后的人脸检测结果也就是那些通过了最后一层的检测子窗口. 最后一层含有 5 个检测器,只要检测子窗口通过了任何一个检测器,都意味着被检测成人脸窗口. 那么,最终待检测图片中的人脸窗口就是那些分别通过了不同检测器的子窗口的总和.

2 融合局部二值模式纹理特征的人脸跟踪

在 CamShift 跟踪算法中,目标模式的表示方法对跟踪结果有着重要的影响. 当前主要使用颜色直方图表示跟踪目标,但是它在复杂条件下跟踪效果并不理想. 局部二值模式(LBP)能够自动提取跟踪目标的关键模式点,抑制表示光滑背景和噪音点等次要模式对目标的影响. 在目标特征空间的选择上,LBP 用更小的特征空间准确地表示跟踪目标. LBP 纹理模型与传统的颜色直方图一样,都属于一种点样本估计方式.

因此,如果能将跟踪目标的 LBP 直方图嵌入 Meanshift 算法,便可解决纹理模型与 Meanshift 结合的难题,得到一种基于纹理的目标跟踪算法. 在应用过程中,对基于彩色直方图统计的 Camshift 算法进行改进,提出融合纹理和彩色特征的 Camshift 人脸跟踪算法,其框架如图 3 所示.

2.1 建模过程

首先,将经过人脸检测模块检测到的人脸区域作为初始化样本,并规格化成同样大小的图片,同时将彩色的样本转换成灰度图. 然后,对得到的灰度图进行伽马校正,去除一定的光照影响,再利用 LBP 编码算法,选择平移不变模式(LBP)、旋转不变模式(LBP<sub>R</sub>)和旋转不变 Uniform 模式(LBP<sub>R,U</sub>),对该人脸灰

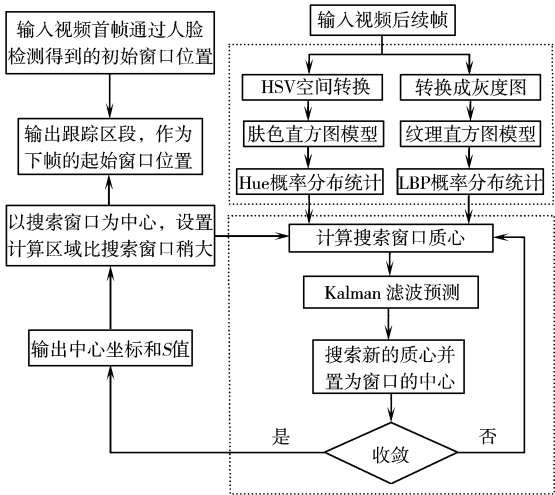


图 3 融合 LBP 特征的人脸跟踪模块框架

Fig. 3 Modular structure of face tracking syncretized LBP feature

度图的每个像素进行 LBP 特征描述

将得到的 LBP 特征进行直方图统计,直方图向量描述了图像中各个微小特征分布情况,并对该直方图进行归一化处理.最后,根据统计得到 LBP 特征直方图运行反向投影操作,可以得到灰度空间的人脸目标 LBP 特征的概率分布图.

不同的 LBP 算子得到直方图向量的维数是不一样的.以  $P=8, R=1$  为例,如果采用  $LBP(8,1)$ ,得到的直方图向量的维数是 256 维;如果采用旋转不变模式  $LBP_R(8,1)$ ,得到的直方图向量是 36 维;而如果采用旋转不变 Uniform 模式  $LBP_{R,U}(8,1)$ ,得到的直方图向量则是 9 维.

3 种不同模式的 LBP 算子对人脸进行建模的实验对比,如图 4 所示.图 4(a),(b),(c)中,第 1 列是将彩色静态图片上检测到的人脸区域转成灰度图,第 2 列是经过 LBP 编码的目标区域,第 3 列是 LBP 直方图统计的结果显示,最后一列是将统计直方图在原来彩色静态图片进行反向投影的结果.

由图 4 可见,旋转不变的 LBP 算子能够表现具有一定旋转不变性的纹理,但是它也就丧失了原图像的一些方向信息.有些时候方向性也会包含很多信息,而此时旋转不变的 LBP 算子并不是一个合适的选择.另外,标准旋转不变的 LBP 算子可以有效地减少直方图 bin 的个数,从而使直方图更加紧致.充分考虑了以上 3 种算子的分类能力和算法复杂度之后,采用  $LBP_{R,U}(8,1)$  模式来进行目标人脸的纹理建模,得到目标模型.



(a)  $LBP(8,1)$



(b)  $LBP_R(8,1)$



(c)  $LBP_{R,U}(8,1)$

图 4 LBP 算子人脸建模对比

Fig. 4 Comparison of different LBP based face modeling

## 2.2 融合颜色与纹理的直方图统计算法

设  $\{r_i\}_{i=1, \dots, n}$  为目标人脸区域中的点,  $Hist(H, T)$

是目标直方图特征空间矩阵,  $H=40$  (将常用的肤色 Hue 空间定义域设在  $90 \sim 130$ ), 分别表示 Hue 通道的量化, 设  $T=9$ , 表示  $LBP_{R,U}(P, R)$  纹理模式 bin 的个数. 这样, 所选取的特征空间维数为  $40 \times 9 = 360$  (仅 360 个向量). 初始化  $Hist(H, T) = 0$ .

(1) 依次取出目标人脸区域中的一个点, 计算该点的  $LBP_{R,U}(P, R)$  值, 量化为 9 个等级; 然后, 将该点的 Hue 通道颜色分别按要求量化为 40 个等级, 得到相应的特征向量, 并根据目标从当前跟踪窗口的中心位置  $\hat{y}_0$  移动到新的窗口中心位置  $\hat{y}_1$  的迭代方程计算该点的权值. 将权值导入到直方图  $Hist$  矩阵相应的 bin 中, 按同样的方法统计出目标人脸区域所有点的模式.

(2) 归一化目标直方图  $Hist$ , 得到目标人脸模型  $\{\hat{p}_u\}_{u=1, \dots, m}$ . 候选区域的人脸模型  $\{\hat{p}_u(\hat{y}_0)\}_{u=1, \dots, m}$  计算过程与目标人脸模型一致.

## 2.3 基于颜色纹理联合直方图的跟踪算法过程

跟踪算法过程主要有如下 6 个步骤.

(1) 将利用多姿态人脸模块标定出来的椭圆形人脸区域作为初始化的搜索窗口, 迭代次数  $k$  初始化, 即  $k=0$ .

(2) 在当前帧中, 计算以  $\hat{y}_0$  为中心的目标人脸候选区域的颜色纹理直方图, 并建立相应的候选模型  $\{\hat{p}_u(\hat{y}_0)\}_{u=1, \dots, m}$ ; 然后, 经过反向投影变换, 转化为融合颜色与纹理的概率分布图像 (CTPDI).

(3) 根据目标从当前跟踪窗口的中心位置  $\hat{y}_0$  移动到新的窗口中心位置  $\hat{y}_1$  的迭代方程, 计算其权重值.

(4) 利用 Meanshift 算法, 对每一帧 CTPDI 图像搜索目标匹配的最优区域, 并根据搜索区域的不变矩估算人脸目标的质心  $\hat{y}_1$  和大小  $S$ . 如果  $\hat{y}_1 - \hat{y}_0 < \epsilon$ , 停止迭代, 保存当前帧搜索结果, 作为下一帧图像初始化搜索窗口, 进入 (5); 否则, 返回 (2), 继续当前帧的跟踪.

(5) 将当前帧 Camshift 的测量值送入 Kalman 滤波进行质心方向的预测, 然后将预测结果反馈给 Camshift 算法再进行搜索.

(6) 当前帧如果是最后一帧, 则跟踪过程结束; 否则, 读取下一帧作为当前帧, 以  $y_0$  为目标候选窗口的初始中心位置, 返回 (1), 在新来的视频帧中定位目标人脸。

3 结果与分析

3.1 样本的收集

多姿态人脸样本主要考虑人脸绕水平轴旋转(俯仰)和绕垂直轴旋转(摇摆), 其姿态范围: 摇摆为  $[-90^{\circ}, +90^{\circ}]$ , 俯仰为  $[-20^{\circ}, +20^{\circ}]$ . 使用 VidTIMIT 人脸库, 包括 43 个人(其中 19 个女性的, 25 个男性的)头部上、下、左、右转动的人脸样本, 手工裁剪出 5 000 幅归一化为  $24 \times 24$  的彩色半侧面人脸样本, 5 000 幅全侧面人脸, 以及约 5 000 正面人脸样本, 构成多视角人脸样本训练集合. 利用“自举”的方法增加负样本, 即将 1 幅  $1\,024 \times 768$  的非人脸灰度图按 0.9 缩放处理后形成 33 幅; 然后用  $24 \times 24$  的窗口穷举, 得  $(1\,024 - 24) \times (768 - 24) \times (1 + 0.81 + 0.81^2 + \dots + 0.81^{32}) = 3\,911\,172$  个非人脸样本. 这可解决非人脸样本不足的问题。

3.2 多姿态人脸检测

利用上述的算法, 离线训练了一个多姿态金字塔检测器, 其主要参数如表 1 所示. 表 1 中,  $N_1$ ,  $N_0$  分别为训练集人脸样本数和训练集初始非人脸样本数,  $\gamma$  为检测率,  $\eta$  为误检率。

每个角度范围都是独立训练的。

训练时, 使用了约 1 000 幅分辨率高的非人脸图像, 通过不断采样生成的足够多的负样本. 当然, 由于训练集样本有限, 并不是所示角度范围内的所有情况都能包含到, 仅以一些易获得的角度作为代表. 如在训练  $45^{\circ} \sim 90^{\circ}$  对应的分类器时, 人脸样本大多数是 900 和 600 的姿态情况, 只有少量的其他角度。

训练后得到的第 1 层检测器含有 2 层强分类器, 分别含有 5, 26 个特征, 检测正确率达到 99%, 误检率为 22%. 这里的误检率稍低了一些, 主要是因为训练集里的初始非人脸数较少. 为了在第 1 层达到拒绝半数非人脸样本的目标, 将误检率设为 22%, 这也意味着拒绝了半数的非人脸. 第 2 层检测器平均含 5 层强分类器, 正确率可以达到 96%, 误检率降低到 2%. 第 3 层检测器更为精确, 平均含有 9 层强分类器, 正确率可以达到 92%, 而误检率可以下降到 0.001%。

为了对比彩色图像的多姿态人脸检测的实验性能, 采用独立于训练集合之外的 UCD 彩色人脸库 (<http://dsp.ucd.ie/~prag/>) 进行测试, 对比数据如图 5 所示. 图 5 中, 曲线 1, 2, 3 分别代表采用文[3]、文[8]和改进的算法(增加肤色分割预处理和姿态预估的 3 层金字塔型分类器)。

根据实验结果的 ROC 曲线分析可看出, 将粗略检测确定的人脸区域作为细致定位的检测区域, 不但大大减小后续处理阶段的搜索区域, 加快检测速度, 而且有助于提高整个检测算法的准确率. UCD 测试集和 CMU-PIE 人脸库上的部分检测结果, 如图 6 所示. 由于角度分类时允许样本交叉, 故会出现左侧面跟正面, 或右侧面跟正面矩形标识框重叠. 系统在 UCD 彩色的人脸图像库上进行测试, 取得较好的实验结果. 对 94 幅 299 张 UCD 人脸库的图像(其中包含 26 个侧面、182 个正面、50 个旋转人脸)进行检测, 检测出人脸数 269 个, 检测率为 90.2%, 误检的窗口数仅有 56 个。

表 1 训练姿态检测器的主要参数

Tab. 1 Main parameters for pose detector training

层次	角度范围	$N_1$	$N_0$	训练器设定目标	
				$\gamma$ / %	$\eta$ / %
第 1 层	$[-90^{\circ}, 90^{\circ}]$	10 000	8 000	99	22
第 2 层	$[-90^{\circ}, 0^{\circ}]$	6 000	7 000	96	2
第 3 层	$0^{\circ}$	3 000	6 000	92	0.001
	$[10^{\circ}, 45^{\circ}]$	6 000	6 200		
	$[45^{\circ}, 90^{\circ}]$	1 500	2 000		

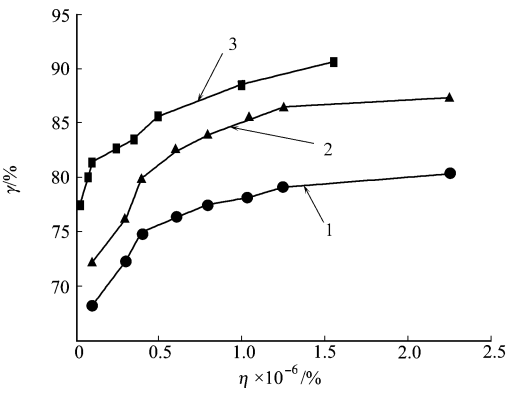


图 5 UCD 彩色人脸检测结果  
Fig. 5 Detection results on  
UCD color face database

综上,通过 UCD 库的实验结果分析,改进方法的检测率与误检率都优于前两种方法,而且收敛速度能够达到实时效果.最后,将多姿态人脸检测模块加入到实时的人脸检测与跟踪系统中,取得非常好的检测效果.实时图像序列人脸多姿态检测的结果,如图 7 所示.



图 6 静态图像检测结果

Fig. 6 Detection results on face images



图 7 实时图像序列检测结果

Fig. 7 Detection results on face video

### 3.3 人脸跟踪实验

对仅采用彩色特征统计直方图的 Meanshift 算法、Camshift 算法,以及改进的融合 LBP 纹理特征的 Camshift + Kalman 算法,在同一个视频图像序列的质心点漂移的坐标变化曲线,如图 8 所示.分析几种算法的各自跟踪性能,其中 Meanshift 算法和 Camshift 算法都在被手机遮挡的时候发生了漏跟的情况.由此可见,加入了 LBP 和运动信息特征,可以明显提高对运动中遮挡人脸跟踪的准确度.

在跟踪过程中,发现如在手经过人脸时会使人脸跟踪窗口发生变形,甚至漂移到非人脸的区域.这是因为 Camshift 算法只是通过查询目标的颜色直方图模型计算目标区域内像素是人脸像素的概率,而没有考虑人脸固有的形状信息.

在跟踪过程中,限制长轴和短轴的比例在 1.0 ~ 1.5 的范围内,不但可以减少搜索的迭代次数,加快跟踪的速度,而且还可以有效地防止类肤色物体对跟踪效果的干扰,提高跟踪准确率,结果如图 9 所示.

采用原始的彩色特征直方图的 Camshift 算法,出现两个人脸交叉形成连通域的情况后,再分开会导致原来较小的后面人脸跟踪失败.改进的算法融合 LBP 纹理特征和 Kalman 滤波后,可以预测到原始小人脸的运动信息,所以能够在人脸区域交叉时,跟踪小人脸成功,结果如图 10 所示.



图 9 肤色干扰实验结果

Fig. 9 Experiment on skin color noise



图 10 融合 LBP 特征的实验结果

Fig. 10 Experiment based on LBP featured

综上,LBP 纹理特征在目标特征空间的选择上,具有用更小的特征空间准确地表示跟踪目标的优势.实验表明,在复杂的跟踪条件下,LBP 的鲁棒性强,在跟踪速度和跟踪性能上,均比当前主流的颜色直方图表示法都有显著提高.

## 4 结论

在正面检测的基础上,利用连续 AdaBoost 算法的思想构造了多姿态人脸分类器,并实现了实时的

多姿态人脸检测模块.通过在UCD库上测试,得出该算法构造的分类器在多姿态人脸检测上具有一定的可行性和实用性.

提出了利用融合LBP特征和脸形约束的Camshift+Kalman的改进算法,以进行实时人脸跟踪模块的开发.通过实验对比,分析了几种算法在实现人脸实时跟踪方面的性能差异,并且对影响人脸跟踪效果的因素做了详细地讨论和分析.结果表明,融合LBP特征和脸形约束的Camshift+Kalman的改进算法,能够更好地实现实时复杂的跟踪条件下的运动人脸的跟踪.

#### 参考文献:

- [1] ROWLEY H, BALUJA S, KANADE T. Rotation invariant neural network-based face detection[C]. Proc IEEE Conf Computer Vision and Pattern Recognition. Washington D C: IEEE Computer Society, 1998: 38-44.
- [2] FERAUD R, BERNIER O J, JEAN-EMMANUEL VIALLET, et al. A fast and accurate face detector based on neural networks[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2001, 23(1): 42-53.
- [3] SCHNEIDERMAN H, KANADE T. A statistical method for 3D object detection applied to faces and cars[C]. IEEE Conf Computer Vision and Pattern Recognition. South Carolina: [s. n.], 2000: 746-751.
- [4] LI Yong-min, GONG Shao-gang, LIDDELL H. Support vector regression and classification based multi-view face detection and recognition [C]. Proc 14th IEEE International Conference on Automatic Face and Gesture Recognition. Grenoble: [s. n.], 2000: 300-305.
- [5] LI S Z, ZHU L, ZHANG Z Q, et al. Statistical learning of multi-view face detection[C]. Proc 7th European Conference on Computer Vision: Part ( ). London: Springer-Verlag, 2002: 67-81.
- [6] LI S Z, ZHU L, ZHANG Z Q, et al. Learning to detect multi-view faces in real-time[C]. Proc 2nd International Conference on Development and Learning. Washington D C: IEEE Computer Society, 2002: 172.
- [7] SCHAPIRE R E, SINGER Y. Improved boosting algorithms using confidence-rated predictions[J]. Machine Learning, 1999, 37(3): 297-336.
- [8] 武勃, 黄畅, 艾海舟, 等. 基于连续 Adaboost 算法的多视角人脸检测[J]. 计算机研究与发展, 2005, 42(9): 1612-1621.

## A Multi-View Face Tracking Method Syncretized LBP Texture Feature

CHEN Yuan, CHEN Duan-sheng

(College of Computer Science and Technology, Huaqiao University, Quanzhou 362021, China)

**Abstract:** We proposed a new improved Camshift algorithm, the hue in an HSV color space of skin color and the texture in the local binary patterns (LBP) texture future of target face is used to produce probability distribution image (PDI), to realize the fusion of the LBP and skin color information. And we also used Kalman filter to predict motion information of target face and find it fast. The experiments show that compare with the original Meanshift and Camshift algorithm using only the color histogram, this algorithm can great improve the tracking speed and precision in complicated tracking conditions.

**Keywords:** multi-view face; real AdaBoost algorithm; feature search table; local binary patterns

(责任编辑: 鲁斌      英文审校: 吴逢铁)