

文章编号: 1004-4353(2018)01-0063-06

# 基于特征组合的多人脸跟踪算法

骆绍烨<sup>1</sup>, 刘丽桑<sup>2</sup>

( 1. 莆田学院 信息工程学院, 福建 莆田 351100;  
2. 福建工程学院 信息科学与工程学院, 福建 福州 350118 )

**摘要:** 提出了一种融合肤色特征、ORB 特征和运动状态估计的多人脸跟踪算法. 该算法以多线程跟踪为基础, 根据不同跟踪算法的适用特点, 在未受肤色干扰时依靠基于分块加权的改进 Camshift 算法跟踪, 在受干扰时则结合包含尺度变化的 ORB 特征匹配算法进行跟踪. 算法同时利用 Kalman 滤波器修正跟踪误差, 以提高跟踪效果. 实验表明, 基于特征组合的多人脸跟踪算法具有较好的跟踪准确性和实时性.

**关键词:** 多人脸跟踪; 特征组合; Camshift; ORB; Kalman 滤波

**中图分类号:** TP391.41

**文献标识码:** A

## Multi-face tracking algorithm based on feature combination

LUO Shaoye<sup>1</sup>, LIU Lisang<sup>2</sup>

( 1. College of Information Engineering, Putian University, Putian 351100, China; 2. College of Information Science and Engineering, Fujian University of Technology, Fuzhou 350118, China )

**Abstract:** In this paper, a multi-face tracking algorithm which integrates skin-color feature, ORB feature and motion state estimation is proposed. This algorithm creates an independent tracking thread which is based on multi-threading. And according to the characteristics of application of different tracking algorithms, it tracked the faces relied on the improved Camshift algorithm tracking which was based on block weighting with no interference of skin color. However, when disturbed by the skin color, this algorithm tracked the faces by using ORB feature matching algorithm with scale change. And it uses the Kalman filter to improve the effect of tracking and correct the tracking error. The experiment shows that the multi-face tracking algorithm based on feature combination has better tracking accuracy and higher real-time performance.

**Keywords:** multi-face tracking; feature combination; Camshift; ORB; Kalman filter

## 0 引言

人脸检测跟踪技术在智能监控、信息安全等领域有着广泛的应用. 目前, 人脸的正面检测技术在较理想的图像条件下, 已相对较为成熟, 但在姿态变化频繁、干扰遮挡等因素的影响下, 人脸的跟踪效果还难以令人满意. 常用的人脸跟踪算法按照思想的不同可以分为基于肤色特征、基于局部

特征和基于运动估计等 3 类. 基于肤色特征的典型算法主要是 Camshift 算法<sup>[1]</sup>, 该算法具有良好的适应性和实时性, 对人脸的姿态变化不敏感, 但容易受到类肤色的干扰. 基于局部特征的跟踪算法主要有基于主动轮廓模型的 Snake 模型算法<sup>[2]</sup>、基于 Harris 角点的 SIFT 特征匹配算法<sup>[3]</sup>和 SURF 特征匹配算法<sup>[4]</sup>、基于 FAST 角点的 ORB 特征匹配算法<sup>[5]</sup>等, 其中 Snake 模型算法由

收稿日期: 2017-12-18

作者简介: 骆绍烨(1982—), 男, 讲师, 研究方向为计算机视觉、信息检索.

基金项目: 国家自然科学基金资助项目(81373552); 莆田市科技计划项目(2015G2014); 莆田学院科研创新专项基金资助项目(2017024)

于主要考虑整体轮廓特征而使得其抗形变能力较差; SIFT 特征匹配算法跟踪效果相对较好, 但计算量大; SURF 特征匹配算法的效率比 SIFT 算法虽有了大幅度提升, 但其计算复杂度依然高于 ORB 特征匹配算法. 基于运动估计的跟踪算法主要有 Kalman 滤波算法<sup>[6]</sup>和粒子滤波算法<sup>[7]</sup>, 其中 Kalman 滤波通常用于跟踪中的修正, 较少单独用于人脸跟踪, 而粒子滤波跟踪算法则存在计算复杂、粒子退化等问题.

相对于单人脸跟踪, 多人脸跟踪由于各人脸之间的干扰影响以及跟踪环境更为复杂, 因此单一的特征匹配或运动估计都难以适用. 常用的多人脸跟踪算法主要有基于人脸检测模型算法和特征组合算法两种. 基于人脸检测模型算法的跟踪效果主要是依赖于检测模型, 但检测效果和时间效率之间往往难以平衡, 如 S. Duffner 等提出的一种基于进出场人脸模型的跟踪算法<sup>[8]</sup>. 特征组合跟踪算法通常是将两种特征简单地结合起来进行跟踪, 如 W. P. Choi 等提出的一种基于形状和纹理信息的多人脸跟踪算法<sup>[9]</sup>, 粮龙亚等将 SURF 特征和 KLT 匹配算法相结合进行多人脸跟踪研究<sup>[10]</sup>, 孟繁静利用 Kalman 滤波器来校正 Camshift 算法的跟踪结果<sup>[11]</sup>. 在利用特征组合方法研究多人脸跟踪时, 目前大多研究者都是将组合特征跟踪直接用于所有的视频场景, 未能完全发挥不同特征跟踪适用场景不同的优势. 基于此, 本文提出了一种融合肤色特征、ORB 特征匹配和运动状态估计的多人脸跟踪算法. 该算法对每张人脸建立各自独立的跟踪线程, 在无严重的肤色干扰时采用基于肤色特征的改进 Camshift 算法跟踪人脸, 当被跟踪人脸受到较大干扰或遮挡时, 则采用 ORB 特征匹配算法进行跟踪; 且在整个跟踪过程中, 利用 Kalman 滤波器进行修正, 以获得更好的跟踪效果.

## 1 基于特征组合的多人脸跟踪算法

### 1.1 基于分块加权的改进 Camshift 跟踪算法

Camshift 算法是一种通过爬升概率密度函数梯度来获得最近主峰的非参数估计方法<sup>[12]</sup>. 算法的步骤为: ①将视频图像转换至 HSV 颜色空间; ②计算目标人脸区域 Hue 分量的概率直方

图, 并依此将原跟踪图像转换为基于目标颜色直方图的颜色概率分布图像; ③利用 MeanShift 算法进行迭代比较目标人脸和候选区域的相似度. 目标人脸和候选区域的相似度通过评价边缘检测的巴氏系数来衡量, 目标人脸区域  $q$  与在  $y$  处的候选区域  $p$  之间的颜色直方图相似度可表示为

$$\rho[p(y), q] = \sum_{b=1}^m \sqrt{p_b(y)q_b}. \quad (1)$$

由于传统的 Camshift 算法抗干扰能力较弱, 因而本文采用基于分块加权模型的改进 Camshift 算法<sup>[13]</sup>, 即对目标模型和候选模型进行分块加权后再比较相似度. 目标模型和候选模型的颜色直方图相似度则变为

$$\rho[p(y), q] = \sum_{j=1}^s \rho[p_j(y), q_j] w_j. \quad (2)$$

其中:  $\rho[p_j(y), q_j]$  为对应分块的相似度;  $w_j$  为该分块的权重, 由其所包含的人脸结构细节的多少决定;  $s$  为总的分块数.

通过均值偏移向量计算出相似度的极大值, 以此获得新的人脸候选位置  $y_1$ . 计算新的候选人脸区域与目标人脸的相似度时, 若新位置的相似度小于旧位置相似度, 则执行  $y_1 = (y_1 + y)/2$ , 重复判断和  $y_1$  的设置直到条件不成立为止; 若新位置的相似度大于旧位置的相似度, 则重复执行位置偏移计算, 直到新旧位置间的距离小于规定的阈值. 当相似度迭代比较中止后, 即可得到该帧图像的人脸中心位置和区域. 将该结果作为下一帧算法搜索窗的初始值, 继续后续帧的人脸跟踪. 改进的 Camshift 算法通过分块加权加强了眼睛、眉毛等面部器官的颜色特征描述, 提高了算法的抗干扰性能, 但由于其跟踪依然依赖于人脸肤色直方图, 所以在受到较为严重的肤色干扰时, 仍然会发生跟踪失败或跟踪误差较大的问题.

### 1.2 包含尺度变化的 ORB 跟踪算法

针对上述改进的 Camshift 算法存在的问题, 本文算法利用包含尺度变化的 ORB 跟踪算法进行解决. ORB 跟踪算法在 FAST<sup>[14]</sup> 特征检测和 BRIEF<sup>[15]</sup> 特征描述子的基础上, 通过对 FAST 特征的不带方向性和 BRIEF 算法不具有旋转不变性这两大弱点进行改进, 不仅能够获得与 SIFT、SURF 算法相类似的匹配效果, 而且还可以大大

提升算法的实时性<sup>[15]</sup>.

FAST 角点被定义为在角点邻域内,灰度值大于或小于该角点的像素点数量必须超过数值  $N$  (ORB 算法中  $N=9$ ). ORB 算法采用 Harris 角点的评价函数来解决 FAST 特征点边缘响应过大的问题. Harris 角点响应函数  $R$  为

$$R = \det(\mathbf{A}(x, y)) - \alpha (\text{trace} \mathbf{A}(x, y))^2, \quad (3)$$

式中  $\det(\mathbf{A}(x, y))$  是结构矩阵的行列式的值,  $\text{trace} \mathbf{A}(x, y)$  是矩阵  $\mathbf{A}(x, y)$  的迹,  $\alpha$  通常取值为 0.04<sup>[16]</sup>. 为了使 FAST 算子具有方向性, ORB 算法采用待测点的灰度值和其邻域的质心之间的偏移向量来增加方向信息. 若角点邻域内的矩为

$$m_{pq} = \sum_{x, y} x^p y^q I(x, y), \quad (4)$$

则 FAST 特征点的方向为

$$\theta = \arctan(m_{01}, m_{10}). \quad (5)$$

由于 FAST 本身不具有尺度不变性, 因此本文算法借鉴了 SIFT 算法的图像金字塔思想<sup>[17]</sup>, 在计算特征点的过程中构建了多尺度的特征点金字塔, 以此实现 FAST 角点的多尺度性. ORB 算法中使用 BRIEF 描述子对检测到的 FAST 特征点进行描述, 以便用相对少量的强度对比来表示图像邻域<sup>[18]</sup>. BRIEF 描述子的提取过程为: 在一个图像邻域  $P$  中, 任意选取  $n_d$  (通常为 256) 个像素点对, 且定义准则  $\tau$  为

$$\tau(P; x, y) = \begin{cases} 1, & p(x) < p(y); \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

式(6)中  $p(x)$  表示滤波处理后邻域  $P$  在点  $x = (u, v)$  的灰度值. 每个点对确定一个  $\tau$ ,  $n_d$  个点对构成了  $n_d$  维的 BRIEF 描述子, 即

$$f_{n_d} = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(P; x, y). \quad (7)$$

ORB 算法通过比较邻域窗口中的两个随机子窗口像素和的方式进行编码, 这种滤波方式相对于 BRIEF 原有的高斯滤波算子具有更好的抗噪声敏感性能. 另外, 为了使 BRIEF 具备旋转不变性, ORB 算法将每个 FAST 特征点所产生的像素点对进行旋转后判别, 得

$$\mathbf{S}_\theta = \mathbf{R}_\theta \mathbf{S}, \quad (8)$$

其中  $\mathbf{R}_\theta$  为 FAST 特征点邻域方向  $\theta$  的对应旋转矩阵,  $\mathbf{S}$  为随机像素点矩阵. 然后, 再利用积分图像进行二进制编码, 这样就得到具有旋转校正的

BRIEF 描述子向量:

$$\mathbf{g}_{n_d} = f_{n_d}(p) \mid (x_i, y_i) \in \mathbf{S}_\theta. \quad (9)$$

为了得到相关性低的点对, ORB 算法利用贪婪搜索的方法对上述描述子向量进行筛选, 然后使用 Hamming 距离和阈值来完成图像的特征点匹配, 最后通过匹配出的特征点来实现目标的跟踪定位.

### 1.3 融合运动预测估计

为了进一步提高跟踪的准确性, 本文在跟踪过程中融入了 Kalman 滤波. Kalman 滤波器通过线性系统状态方程来计算状态变量和输出信号的估计值, 然后利用系统观测数据校正状态变量的估计值, 最后采用使状态变量估计误差的均方差最小的估计值作为更接近真实值的结果<sup>[19]</sup>. 设  $k$  时刻系统状态向量为  $\mathbf{X}_k = (x_k, y_k, v_{kx}, v_{ky})^T$ , 观测向量为  $\mathbf{Z}_k = (x_k^z, y_k^z)^T$ , 其中  $(x_k, y_k)$  为该时刻的质心坐标,  $v_{kx}$  和  $v_{ky}$  是质心在横向和纵向上的速度分量.  $\mathbf{A}$  和  $\mathbf{H}$  为系统状态转移矩阵和系统观测矩阵,  $\mathbf{W}_t$  和  $\mathbf{V}_t$  是互不相关且均值都为 0 的系统噪声和观测噪声, 则 Kalman 滤波器的系统状态方程和系统观测方程为:

$$\begin{cases} \mathbf{X}_t = \mathbf{A} \mathbf{X}_{t-1} + \mathbf{W}_t, \\ \mathbf{Z}_t = \mathbf{H} \mathbf{X}_t + \mathbf{V}_t. \end{cases} \quad (10)$$

跟踪预测过程中, 利用前一帧的质心位置和速度 (初始帧的质心位置由人脸检测得到, 速度为 0) 来预测当前帧的系统状态, 通过改进的 Camshift 算法和 ORB 算法得到的质心坐标作为测量值, 用以修正系统状态方程产生的系统估计值, 从而获得最终的质心位置和状态.

## 2 干扰的判断和算法步骤

### 2.1 干扰的判断

多人脸跟踪下的肤色干扰主要有两大来源: 类肤色背景和其他被跟踪人脸. 针对这两种情况, 本文分别采用不同的判别依据.

1) 类肤色背景干扰. 在使用 Camshift 算法跟踪人脸的过程中, 虽然人脸在视频中可能由于姿态变化而产生跟踪区域的改变, 但姿态改变而造成的变化通常比较和缓. 若跟踪遇到大面积的类肤色背景干扰时, 跟踪区域往往会出现突然急剧

变化的情况,这种区域变化通常表现为跟踪区域的单条边或两条边的突然增大.因此,若前后相邻两帧同一人脸区域的高或宽的增长率超过阈值(本文设定为 40%),则认为该跟踪人脸区域受到严重的类肤色干扰.

2)其他干扰.在多人脸跟踪过程中,各个被跟踪的人脸由于其位置的变化经常出现人脸互相接近或交叉遮挡的情况,此时基于 Camshift 算法的跟踪往往会将互相接触的跟踪人脸区域合并,导致后续的人脸跟踪错误或失败.针对这一问题,本文采用主动检测的方法来帮助跟踪算法做出提前调整.对于每帧图像在跟踪完成后进行如下判断:设中心分别在  $a$  点  $(x_a, y_a)$  和  $b$  点  $(x_b, y_b)$  的两个人脸区域的横向宽度为  $w_a$  和  $w_b$ ,纵向的高度为  $h_a$  和  $h_b$ ,则人脸边缘之间的横向距离  $l_x$  和纵向距离  $l_y$  为

$$\begin{cases} l_x = |x_a - x_b| - (w_a + w_b)/2, \\ l_y = |y_a - y_b| - (h_a + h_b)/2. \end{cases} \quad (11)$$

若  $l_x > l_m$  且  $l_y > l_m$ ,则认为两个被跟踪的人脸未互相干扰,其中  $l_m$  为设定人脸间边缘最小距离的阈值,否则判定两个人脸之间将产生肤色跟踪干扰.

## 2.2 算法步骤

基于特征融合的多人脸实时跟踪算法采用多线程的跟踪方法,在检测到人脸后为其建立一个独立的跟踪线程.为了提高跟踪的准确性和效率,各个人脸跟踪线程采用不同的特征组合跟踪算法进行跟踪,以适应跟踪过程中人脸所处的不同环境和状态.各线程的跟踪算法步骤如下:

1)利用人脸检测工具获得跟踪人脸的中心和窗口.

2)创建各人脸跟踪线程,初始化各跟踪线程的目标人脸模板和 Kalman 滤波器.

3)读取下一视频帧,采用基于分块加权的 Camshift 跟踪算法计算出各个候选人脸跟踪窗口.

4)对各个跟踪人脸进行肤色干扰判定.若某一跟踪人脸满足肤色干扰判断中的任意一个条件,则该线程执行步骤 5),否则执行步骤 6).

5)使用包含尺度变化的改进 ORB 算法进行跟踪.将初始帧中的该人脸目标与当前帧图像进行匹配,得到新的候选人脸窗口.

6)将各个候选目标人脸位置作为系统观测值,利用 Kalman 滤波器预测修正,以此获得各人脸位置的最优估计值.

7)将上一步骤中得到的人脸区域设置为新的目标人脸模板,并回到步骤 3)继续执行,直至跟踪结束.

## 3 实验结果及分析

为了验证本文提出的基于特征组合的多人脸实时跟踪算法的效果,本文设计了仿真实验,将本文算法与经典 Camshift 算法、ORB 算法进行仿真实验对比.仿真实验采用 Microsoft Visual Studio 2010 作为实验开发平台,结合 Intel 建立的开源计算机视觉库 OpenCV(Open Source Computer Vision Library,版本:2.4.10)编程实现.实验计算机的主要配置:处理器为 Intel Core i5-3470,主频为 3.2 G,内存为 12 GB,操作系统为 Windows 7.实验所用视频均由本地摄像头录制获得.实验中人脸初始位置和大小均利用相同的人脸检测函数在视频的初始帧获得.3 种算法的部分多人脸跟踪效果分别如图 1—图 3 所示.



图 1 经典 Camshift 算法的跟踪效果





图 2 ORB 算法的跟踪效果



图 3 本文算法的跟踪效果

由图 1 可以看出,经典 Camshift 算法在人脸交错时,两个人脸跟踪区域出现了融合的现象(图 1b);这两个人脸分离后,人脸跟踪区域仍然重叠在面积较大的人脸上,另外一个跟踪线程出现了跟踪失败的情况(如图 1c). 由图 2 和图 3 可以看出,ORB 算法和本文算法能够较好地处理人脸间交错时的跟踪问题.

为了进一步量化 Camshift 算法、ORB 算法和本文算法的跟踪效果,对 3 种算法的跟踪处理时间 & 准确率进行了统计,统计结果如表 1 所示. 设定当跟踪区域包含对应人脸的整个可见面部区域,且人脸与整个跟踪区域的面积之比超过  $2/3$  时,视为跟踪成功.

表 1 3 种跟踪算法的准确率和时间效率

算法	总帧数	算法准确跟踪的帧数				平均跟踪准确率/%	平均每帧处理时间/ms
		人脸 1	人脸 2	人脸 3	平均值		
Camshift 算法	376	331	253	148	244	64.9	13
ORB 算法	376	334	326	317	326	86.6	132
本文算法	376	356	342	325	341	90.7	38

由表 1 可以看出,本文算法的平均跟踪准确率显著优于 Camshift 算法,且略高于 ORB 算法;平均处理时间高于 Camshift 算法,但显著低于 ORB 算法.

4 结论

经实验表明,本文提出的融合肤色特征、局部

角点特征和运动状态估计的多人脸跟踪算法,不仅能够保证实时性,而且具有较好的多人脸跟踪效果;因此,本文方法有助于提高复杂场景下的人脸跟踪效率,可为安全监控、人机交互等应用提供相应的支持. 在现实环境中,由于干扰多人脸跟踪的因素众多,使得跟踪算法的准确性和实时性仍需进一步提高;因此,今后将继续研究各人脸特征

跟踪对复杂场景的适应特性,通过优化特征选取、改进组合方式等方法来提高多人脸跟踪的适应性和效率.

## 参考文献:

- [1] Bradski G R. Computer vision face tracking as a component of a perceptual user interface[C]//Proceedings of IEEE Workshop Applications of Computer Vision. Princeton, NT: IEEE, 1998:214-219.
- [2] Kass M, Witkin A, Terzopoulous D. Snake: active contour moduels[J]. International Journal of Computer Vision, 1987,1(4):321-331.
- [3] Lowe D G. Object recognition from local scale-invariant features[C]//Proceeding of the 7th IEEE International Conference on Computer Vision. Corfu, Greece: IEEE, 1999.
- [4] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up Robust Features [M]. Berlin Heidelberg: Springer, 2006.
- [5] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]//IEEE International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011.
- [6] Kalman R E. A new approach to linear filtering and prediction problems[J]. Transaction of the ASME-Journal of Basic Engineering, 1960,82(Series D): 35-45.
- [7] Michael Isard, Andrew Blake. Condensation-conditional density propagation for visual tacking[J]. International Journal of Computer Vision, 1998, 29(1):5-28.
- [8] Duffner S, Odobez J M. A track creation and deletion framework for long-term online multi-face tracking[J]. IEEE Transactions on Image Processing, 2013,22(1):272-285.
- [9] Choi W P, Lam K M. An effective shape-texture weighted algorithm for multi-view face tracking in videos[J]. Congress on Image & Signal Processing, 2008,4:156-160.
- [10] 粮龙亚,钱雪忠. 基于 SURF 特征点的多人脸跟踪方法研究[J]. 计算机应用与软件, 2015, 32(2): 178-181.
- [11] 孟繁静. 基于视频的实时多人脸检测跟踪与优选方法研究[D]. 沈阳:东北师范大学, 2016.
- [12] 杨超,蔡晓东,王丽娟,等. 一种改进的 Camshift 跟踪算法及人脸检测框架[J]. 计算机工程与科学, 2016,38(9):1863-1869.
- [13] 骆绍烨,刘丽桑. 基于改进 Camshift 的人脸跟踪算法[J]. 延边大学学报(自然科学版), 2017,43(2): 144-149.
- [14] Rosten E, Drummond T. Machine learning for high-speed corner detection [M]. Berlin Heidelberg: Springer, 2006.
- [15] Calonder M, Lepetit V, Strecha C, et al. BRIEF: Binary Robust Independent Elementary Features [M]. Berlin Heidelberg: Springer, 2010.
- [16] 孟凡清. 基于背景差分法和 ORB 算法的运动目标检测与跟踪算法研究[D]. 北京:北京印刷学院, 2014.
- [17] 葛山峰,于莲芝,谢振. 基于 ORB 特征的目标跟踪算法[J]. 电子科技, 2017,30(2):98-100.
- [18] 许宏科,秦严严,陈会茹. 基于改进 ORB 的图像特征点匹配[J]. 科学技术与工程, 2014,14(18):105-109.
- [19] 王丽,郝晓丽. 基于 Kalman 滤波器和改进 Camshift 算法的双眼跟踪[J]. 微电子学与计算机, 2016,33(6):109-112.