

基于在线学习机制的人脸持续跟踪方法^①

蔡丽仪^② 吴怀宇^③ 陈镜宇 陈 洋

(武汉科技大学信息科学与工程学院 武汉 430081)

摘 要 针对目标人脸短暂离开画面后不能继续跟踪的问题,提出将增量分层判别回归方法(IHDR)与改进的核相关滤波(KCF)跟踪算法相结合,以解决人脸持续跟踪问题。首先,提取人脸光照不变特征,增量构建人脸特征 IHDR 树。然后,检索 IHDR 树识别目标人脸,通过循环矩阵获取人脸正负样本,训练岭回归分类器对人脸进行跟踪。在人脸短暂离开画面时,重新识别目标人脸,使用识别结果重新初始化跟踪器,实现对面脸的持续跟踪。此外,针对 KCF 跟踪器的跟踪框尺度不能自适应的问题,对 KCF 跟踪器进行了改进,设置 3 个尺度不同的模板区域并计算响应,以响应最大的区域的尺度为跟踪框的尺度。最后,在不同光照下进行了动态人脸识别实验,识别率达到 97.84%。与传统跟踪算法进行对比,所提方法能够在尺度上自适应跟踪目标人脸,并满足实时性要求。在人脸短暂离开画面的视频中,该方法亦能实现对目标持续跟踪。

关键词 人脸持续跟踪,光照不变特征,在线学习,尺度自适应,核相关滤波器(KCF)

0 引 言

随着公共安全越来越受到人们的重视,人脸跟踪作为管理公共人员的重要方法之一,引起了科研人员的广泛关注,并被应用于各个领域。如 Degadwala 等人^[1]使用 HOG^[2]特征、SVM^[3]特征分类器和 Viola-Jones^[4]面部检测框架,构建监控视频人脸跟踪系统,保护商业、银行和政府等重要场所活动安全。Ren 等人^[5]通过跟踪人脸,计算某人停留在敏感区域的时间来判断该人员是否可疑。Elrefaei 等人^[6]提出一种应用于手机的刑事侦查框架,可以帮助警察识别犯罪者或嫌犯的面貌。此外,人脸跟踪还应用于人机交互^[7]和无人机跟随^[8]等领域。

人脸跟踪是指通过帧序列或视频中的运动,估计和分析帧平面上人脸的轨迹。现有的人脸跟踪方法主要分为三种:基于特征、基于模型以及两者相结

合的方法。基于特征的方法将面部视为组件的集合,提取特征,通过逐帧检测和匹配,确保每时刻的人脸跟踪,这种方法跟踪精度高,但逐帧的人脸检测导致内存消耗极大,实现实时的跟踪是该方法改进的方向。Ramkumar 和 Logashanmugam^[9]提出一种基于复合掩膜变化的人脸跟踪方法,复合掩膜算法即高斯滤波和自适应中值滤波,通过标准化亮度和增加对比度,模糊图像并消除噪声和细节。文中使用 Haar^[10]特征和 SURF^[11]级联检测人脸,针对光照问题,提出了一种改进的梯度对数场(MGLF)对人脸进行跟踪。作者通过对视频帧的下采样与逐帧检测,实现了低分辨率视频中的人脸实时跟踪。

基于模型的方法将面部表示为整体单元,在跟踪过程中不断学习和更新模型,适应目标在运动过程中的形态、光照、比例等变化,这种方法跟踪速度快、抗干扰能力强,但当目标人脸被遮挡时,跟踪器被遮挡物污染,最终导致轨道偏离目标轨迹。Khur-

① 国家重点研发计划专项(2017YFC0806503)和湖北省科技支撑计划(2015BAA018)资助项目。

② 女,1994 年生,硕士生;研究方向:视觉跟踪;E-mail: 1123579750@qq.com

③ 通信作者,E-mail: wuhy@wust.edu.cn

(收稿日期:2018-09-05)

shid 等人^[12]提出一种基于多模型字典的在线学习方案,组合重建和分类字典,从正例和负例中学习每个字典,增强被跟踪的目标与背景之间的区别。该方案的优点在于对面部表情的变化有较强的鲁棒性,可运用于非刚性物体。

综合基于特征和基于模型的方法存在的优缺点,研究人员致力于将两种方法结合起来,优势互补。Ren 等人^[13]使用卷积神经网络(convolutional neural network, CNN)^[14]和卡尔曼滤波器对人脸进行跟踪。首先用 CNN 逐帧检测视频中的人脸,当面部被严重遮挡时,使用卡尔曼滤波算法估计面部位置,定位目标,最终实现高精度的人脸跟踪。

针对上述方法中存在的计算速度慢、手动选取跟踪目标、不能变尺度跟踪以及不能持续跟踪(目标短暂离开摄像头后仍能继续跟踪)的问题,本文提出一种基于在线学习机制的人脸持续跟踪方法。结合光照不变特征(illumination invariant features, IIF)^[15],通过建立增量分层判别回归(incremental hierarchical discriminant regression, IHDR)^[16]树对人脸特征进行在线增量学习,准确地识别目标人脸,并用识别结果初始化核相关滤波器(kernelized correlation filters, KCF)^[17],对目标人脸进行跟踪。针对 KCF 不能变尺度以及不能持续跟踪的问题,提出一种改进的 KCF 方法,当目标离开视频边缘时,启动 IHDR 树检索功能,识别目标人脸,以识别结果框重新初始化 KCF 跟踪器,实现对目标人脸的持续跟踪。

1 增量学习光照不变特征

本文提出的基于在线学习机制的人脸跟踪方法,通过结合检测、识别与跟踪,实现非特定人脸的尺度自适应持续跟踪。首先,进行人脸检测,提取人脸 IIF 特征;其次,使用 IHDR 算法增量学习人脸特征与其姓名标签之间的映射关系,建立 IHDR 树;然后,检索 IHDR 树,识别目标人脸;最后,使用目标人脸识别结果框初始化跟踪框,启动改进的 KCF 跟踪器,在跟踪过程中结合识别与跟踪算法,实现对目标人脸的持续跟踪。

1.1 图像预处理

在进行人脸特征提取和人脸识别之前,都要先进行人脸检测。人脸检测作为人脸跟踪的第一步,其检测正确率与检测速度影响着整个人脸跟踪系统。本文选用开源的人脸检测库 libfacedetection 对人脸进行检测。为了提高人脸识别率,减少图像噪声以及面部偏转对人脸识别带来的影响,先对检测出的人脸图像进行人脸摆正、椭圆掩膜修正以及双边滤波等预处理。

(1) 人脸摆正

人脸摆正即将偏转了一定角度的人脸进行摆正,使左右两只眼睛在同一水平线上,如图 1 所示。

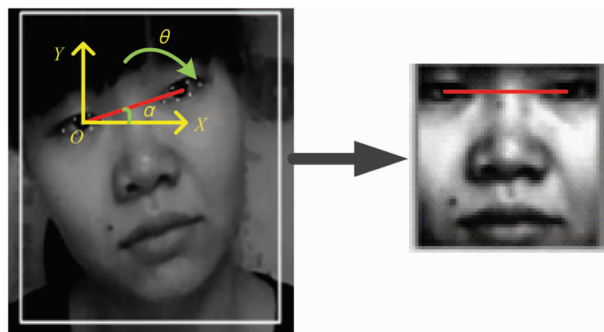


图1 人脸摆正示意图

图 1 中,首先使用 OpenCV 自带的人眼检测算法,定位左右两只眼睛的中心坐标,以左眼中心坐标为原点,建立直角坐标系;然后,将两只眼睛中心坐标点连线,计算该线段与 x 坐标轴的夹角 $\alpha \in (0, 180)$ (单位 $^\circ$),角 α 则为人脸的偏转角度,当 $\alpha < 90^\circ$ 时,表示人脸向左偏转,当 $\alpha \geq 90^\circ$ 时,表示人脸向右偏转;最后,将人脸图像旋转 $\theta \in (-90, 90)^\circ$ (默认逆时针旋转为正),完成人脸摆正,其中,当 $\alpha < 90^\circ$ 时, $\theta = -\alpha$; 当 $\alpha \geq 90^\circ$ 时, $\theta = 180 - \alpha$ 。

(2) 椭圆掩膜修正

图像掩膜是提取图像中感兴趣区域的一种技术,本文使用椭圆掩膜来提取椭圆形人脸区域,如图 2 所示。

图 2 中,首先根据人脸摆正后的图片大小,生成一张大小一致的空掩膜,空掩膜中的每个像素值均为 0;其次根据人脸图片中脸部的大小及位置,在空掩膜中拟合近似的椭圆,将椭圆内部像素值置为

255;然后生成一张输出图像,将像素值置为 128;最后,按照掩膜上的椭圆形状,将原图像复制到输出图像中,其中椭圆掩膜上像素值为 0 所对应的原图中

的像素不复制,像素值不为 0 的像素则复制。使用椭圆掩膜修正人脸,可以剔除背景信息,突出面部特征。

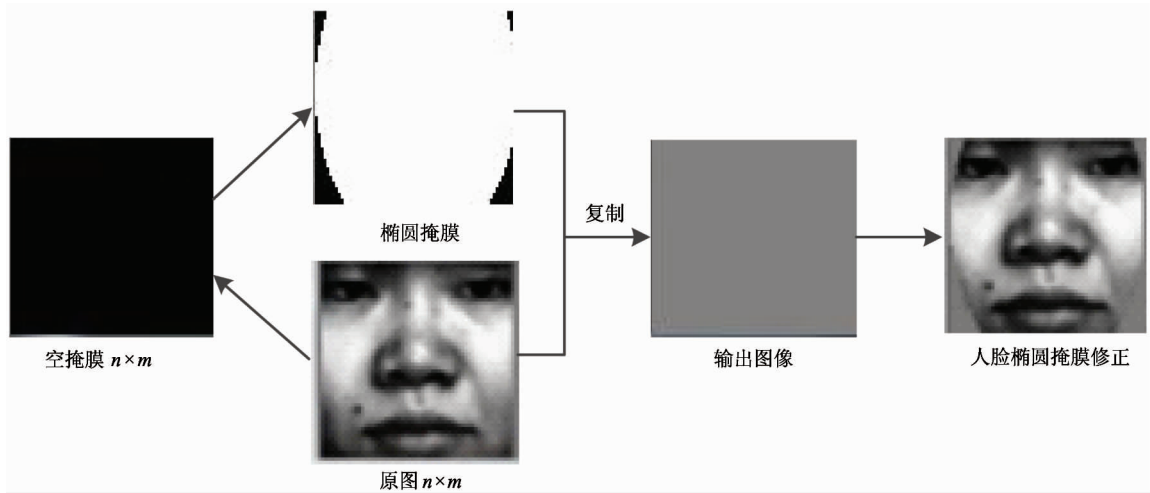


图 2 人脸椭圆掩膜修正

(3) 双边滤波

双边滤波是一种非线性滤波,其权重同时考虑了图像的空间邻近度和像素范围域中的辐射差异。在像素值变化平缓的区域,像素范围域权重接近 1,此时起主导作用的是空间邻近度;当像素处于图像边缘时,像素值变化明显,像素范围域起主导作用,能够较好地保留图像的边缘信息。

$$g(i, j) = \frac{\sum_{k,l} f(k, l) \omega(i, j, k, l)}{\sum_{k,l} \omega(i, j, k, l)} \tag{1}$$

式(1)中, $g(i, j)$ 是像素 (i, j) 的去噪强度, $\omega(i, j, k, l)$ 为权重系数,其定义如式(2)所示。

$$\omega(i, j, k, l) = e^{-\frac{(i-k)^2 + (j-l)^2}{2\sigma_d^2}} \cdot e^{-\frac{\|f(i, j) - f(k, l)\|^2}{2\sigma_r^2}} \tag{2}$$

式(2)中, σ_d 和 σ_r 是平滑参数, $f(i, j)$ 和 $f(k, l)$ 分别是像素 (i, j) 和 (k, l) 的强度。

1.2 光照不变特征

经过上述预处理后,人脸图片剔除了背景,滤去了部分噪声,突出了面部特征。针对跟踪环境中通常会有光照变化的问题,本文提取对光照不敏感的人脸 IIF 特征,训练人脸特征 IHDR 树,以削弱在跟踪过程中,光照变化对人脸识别的影响。

IIF 特征对光照不敏感的原因是在光照变化的情况下,亮度值在某一区间上的像素数是不变的。如式(3)所示。

$$L_p = \sum_{b=1}^B \exp\left(-\frac{(b - b_p)^2}{2\max(k, r_p)^2}\right) \cdot H_p^E(b) \tag{3}$$

式(3)中, L_p 是亮度值在区间 $[1, B]$ 中的像素数, B 是亮度值量化的区间的总数, b_p 代表像素点的亮度值 I_p 所对应的量化 bin, $k = 0.1$ 是一个经验常量,参数 $r_p = 4$ 通常用来控制区间的长度, H_p^E 是在像素点 p 处计算得到的局部敏感直方图。

使用标准库 FERET 中的图片验证 IIF 特征的光照不变性,其结果如图 3 所示。

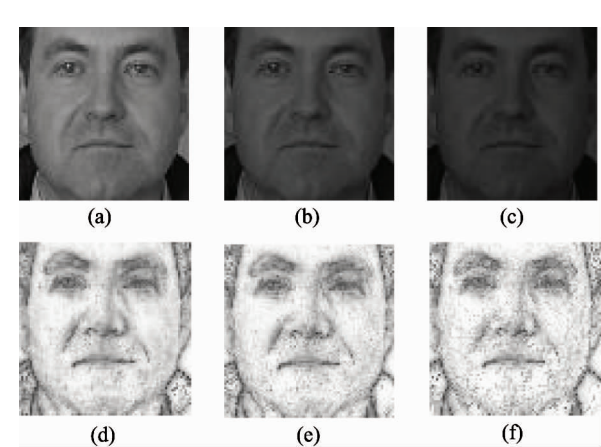


图 3 光照不变特征

图 3 中,子图(a) ~ (c) 是光照强度不同的输入图片,子图(d) ~ (f) 是相应的特征输出。可以看

出,随着光照的不断变化,用该方法提取出来的特征变化不大,即该特征提取方法对光照具有鲁棒性,能有效减少光照变化对人脸识别的影响。

1.3 IHDR 树

将提取的人脸 IIF 特征与对应的姓名标签之间的映射关系作为输入,增量构建人脸特征 IHDR 树。IHDR 树的特征:IHDR 树分为 X 空间和 Y 空间,其中 X 空间存储输入的人脸特征向量,针对输入的不同人脸特征序列,生成相应的 x 聚类子(x 节点),此过程叫作聚类。 Y 空间存储虚拟标签,虚拟标签与 X 空间中的 x 聚类子相对应,存在映射关系 $h:X \rightarrow Y$ 。同时,每个虚拟标签指向对应 x 聚类子的姓名标签。IHDR 树的结构如图 4 所示。

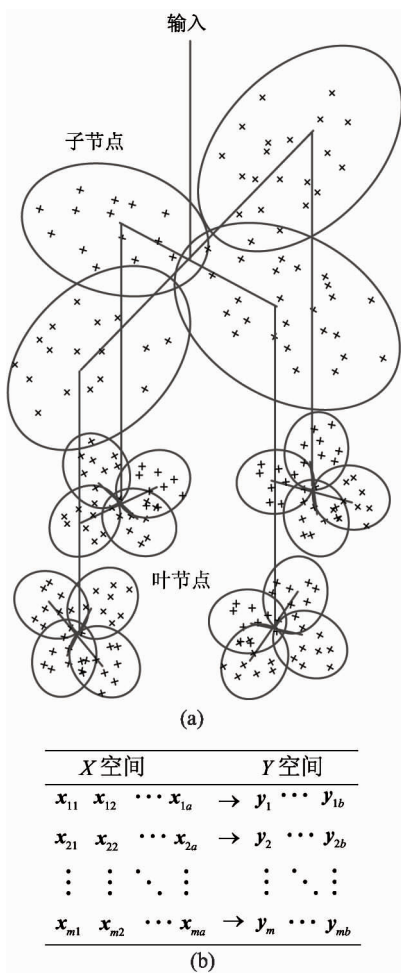


图 4 IHDR 树的结构与 X 、 Y 空间的映射关系

步骤 1 更新 y 聚类集。给定样本 y 和聚类 $c = \{y_i | i = 1, 2, \dots, n\}$ 集合, q 是内部节点最大子节点数, $\delta_y > 0$ 是输出空间 Y 的灵敏度:

步骤 1.1 利用式(4)找到最近邻 y_j :

$$j = \operatorname{argmin}_{1 \leq i \leq n} \{ \|y_i - y\| \} \quad (4)$$

式(4)表示使 $\|y_i - y\|$ 达到最小值的参数 y_j , $\|y_i - y\|$ 表示样本与聚类中心的欧氏距离;

步骤 1.2 若 $\|y_j - y\| < \delta_y$, 则将 y 归为 y_j 类, 求 y 与 y_j 的均值作为 y_j 的聚类中心, 更新 y_j ;

步骤 1.3 若 $n < q$ 且 $\|y_j - y\| > \delta_y$ (防止相似或相同的样本形成不同的聚类), 则将 n 加 1, 设置新的聚类 $y_n = y$, 将 y_n 添加到 c 中;

步骤 1.4 若 $n = q$ 且 $\|y_j - y\| > \delta_y$, 则将 y 归为 y_j 类, 以 y 与 y_j 的均值更新 y_j 聚类中心。

步骤 2 更新 x 聚类集。给定样本 (x, y) 和聚类 $c = \{(x_i, y_i) | i = 1, 2, \dots, n\}$ 集合, $\delta_x > 0$ 是输入空间 X 的灵敏度:

由于 X 空间和 Y 空间存在映射关系, 根据步骤 1 中求得的样本 (x, y) 的虚拟标签, 确定样本中 x 部分更新哪个 x 聚类子 (如 y 属于 Y 空间中第 k 个聚类子 y_k , 则 x 属于 X 空间中第 k 个聚类子 x_k)。

步骤 2.1 计算样本 x 与聚类 x_k 的马氏距离 $\|x_k - x\|$;

步骤 2.2 若 $\|x_k - x\| < \delta_x$, 则求 x 与 x_k 的均值, 更新 x_k ;

步骤 2.3 若 $\|x_k - x\| > \delta_x$, 且 x_k 不是叶子节点, 则在 x_k 的 p 个子节点中找到最近邻 x_j :

$$j = \operatorname{argmin}_{1 \leq i \leq p} \{ \|x_i - x\| \} \quad (5)$$

步骤 2.4 若 $\|x_j - x\| < \delta_x$, 则用 x 更新 x_j ;

步骤 2.5 若 $\|x_j - x\| > \delta_x$ 且 x_k 的子节点数 $p < q$, 则 $p + 1$, 为 x_k 生成新的子节点 $x_{p+1} = x$;

步骤 2.6 若 $\|x_j - x\| > \delta_x$ 且 $p = q$, 则继续向下搜索 x_j 的子节点, 循环步骤 2.3 ~ 2.5, 直到 x_j 为叶子节点。

步骤 2.7 为 x_j 生成子节点 $x_n = x$;

步骤 2.8 IHDR 树更新完成。

步骤 3 检索 IHDR 树。给定待检测样本 x , 检索精度 k , 敏感系数 ε :

步骤 3.1 以从上往下的顺序, 从第一层开始检索, 计算每一层的每个节点 x_i 与 x 的马氏距离 $\|x_i - x\|$, 选出前 k 个距离最小的节点作为活跃节点;

步骤 3.2 若这 k 个节点中存在 x_j , 使得 $\|x_j -$

$x \parallel < \varepsilon$, 返回 x_j 对应的 y_j , 检索结束;

步骤 3.3 若 $\|x_n - x\| > \varepsilon, n = 1, 2, \dots, k$, 以这 k 个节点为根节点, 继续向下搜索, 若这 k 个节点有子节点, 则取消该节点的活跃标志, 计算其子节点与 x 的马氏距离, 选出前 k 个距离最小的节点继续作为活跃节点;

步骤 3.4 重复步骤 3.2 和步骤 3.3, 直到所有活跃节点均为叶子节点, 检索结束, 并未学习过该样本序列。

2 KCF 跟踪器

KCF 跟踪器首先使用循环矩阵获得正负样本, 提取 HOG 特征, 训练岭回归分类器, 再通过计算目标响应获得目标位置, 在目标位置附近继续循环采样, 更新分类器, 从而实现对目标的跟踪。

2.1 循环矩阵

KCF 的训练样本由目标样本循环位移得到。对于二维图像而言, 循环位移的结果是在目标样本的上下左右移动取样, 如图 5 所示。

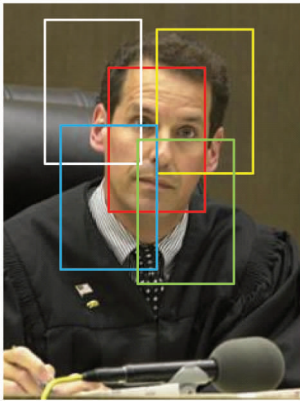


图 5 循环位移取样

图 5 中, 贴合人脸区域的框为目标样本, 其余的框为目标样本通过循环位移得到的结果。

2.2 岭回归分类器

岭回归分类器的训练目标是找到一个函数 $f(z) = w^T z$, 使得样本 x_i 与其回归目标 y_i 的平方误差最小, 即:

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \tag{6}$$

式(6)中 λ 用于控制结构复杂性, 防止过拟合。写成矩阵形式有:

$$\min_w \|Xw - y\|^2 + \lambda \|w\|^2 \tag{7}$$

式(7)中数据矩阵 X 每行 x_i 表示一个样本, y 的每个元素对应一个样本标签, 这个最小化问题有闭式解:

$$w = (X^T X + \lambda I)^{-1} X^T y \tag{8}$$

由于通常处理的都是非线性问题, 非线性问题在低维空间中不可分, 将其通过一个核函数:

$$\varphi^T(x) \varphi(x') = k(x, x') \tag{9}$$

映射到高维空间, 便变得线性可分:

$$w = \sum_i \alpha_i \varphi(x_i) \tag{10}$$

式(10)中, $\varphi(x_i)$ 表示将 x 映射到高维空间的函数, 此时在原空间(低维)中求最小 w 的问题, 变成在对偶空间中求最小 α 的问题:

$$\alpha = (K + \lambda I)^{-1} y \tag{11}$$

式(11)中, K 是所有训练样本的核相关矩阵 $K_{ij} = k(x_i, x_j)$, α 是系数 α_i 的向量, 代表双空间中的解。

2.3 尺度自适应跟踪

KCF 跟踪器的跟踪过程包括检测视频序列中的目标以及更新岭回归分类器。对目标进行检测, 检测样本为预测区域和将其位移得到的待测样本集合 $z_j = P^j z$, 则使式(12)最大的样本即为检测出的新目标区域。

$$f(z_j) = \alpha^T \phi(X) \phi(z_j) \tag{12}$$

为了实现尺度自适应跟踪, 以检测出的新目标区域为模板, 按相同比例缩放模板长宽, 分别设置一个比模板大(模板的 1.05 倍)和一个比模板小(模板的 0.95 倍)的矩形区域, 用式(12)计算 3 个区域的响应, 响应最大的区域为适应尺度的新目标区域, 从而实现对目标的尺度自适应跟踪。

3 实验及结果分析

本实验的运行环境为 Intel Core i5-3210 2.5GHz PC, Win7 64 位操作系统。运行平台为 Visual Studio 2013, 使用版本为 2.4.9 的开源 OpenCV 库。

在本文提出的识别与跟踪相结合的方法中, 人

脸识别作为校准装置,其识别率决定着整个跟踪系统的稳定性。而跟踪系统中的人脸识别是在在线视频中动态进行的,所以本文使用录制的7段人脸视

频进行实验,每段视频的长度为30 s,帧速率为30 帧/s,共900 帧,帧分辨率为640×480,其人脸识别结果如图6 所示。



图6 基于 IHDR 的人脸识别结果

图6 中展示了人员1 到人员7 的人脸识别结果,其中第(1)行为有灯光下的识别结果,第(2)行为无灯光下的识别结果,不管在有无灯光的情况下,系统都能正确地识别人脸,表明人脸识别系统对光照具有一定的鲁棒性;第(3)和(4)行为在人脸遮挡、偏转和有表情变化下的人脸识别,从人员6 的(4)图中可以看出,在人脸具有表情变化的时候,系统仍能正确识别人脸;人员7 的(4)图中,半边人脸离开画面,相当于在人脸遮挡情况下,系统仍能正确识别人脸;人员1 到人员5 的(4)图中展示了识别失败的图片,在这些图片中,因人脸偏转过度而导致人脸识别失败的占绝大部分,尽管如此,(3)行中展现的在一定偏转范围内系统正确识别人脸的图片,说明了人脸识别系统对人脸形变仍然具有一定的鲁棒性。

表1 中展示了各视频的识别率以及系统总识别率。

分析表1 可知,人员1 到人员7 的识别正确率都大于94%,符合人脸识别系统的要求。表中人员3 和4 的识别率相较其他人员要低,可能与佩戴眼镜且头部偏转过度有关,总体来说,所有人员的正确

识别率达到97.84%,能够保证系统在目标短暂离开画面后重新定位人脸。

表1 基于 IHDR 的人脸识别结果

视频	总帧数 (帧)	正确识别帧数 (帧)	正确率 (%)
人员 1	900	890	98.89
人员 2	900	878	97.56
人员 3	900	868	96.44
人员 4	900	852	94.67
人员 5	900	876	97.33
人员 6	900	900	100
人员 7	900	900	100
所有人员	6300	6164	97.84

为了验证本文提出方法的实时性及尺度自适应效果,现将改进的 KCF 跟踪方法与以下几种方法进行了对比: KCF 跟踪方法、基于时空上下文学习 (spatio-temporal context learning,STC) 的跟踪方法以及基于压缩感知 (compressive sensing,CS) 理论的压缩跟踪 (compressive tracking,CT) 方法。分别将以上方法在人脸视频 david_indoor 上进行测试,视频 david_indoor 的帧分辨率为320×240。在实验过程

中,在第 300 帧时手动选取人脸区域,对人脸进行跟踪,选取视频中第 400 帧到 499 帧之间的图像进行

跟踪效果分析,实验结果如图 7 所示。



图 7 标准库人脸跟踪算法对比

由图 7 可知,选取的跟踪算法都能对人脸进行跟踪,相较之下,CT 的跟踪效果较差,经过第 458 帧的人脸偏转之后,在第 477 帧出现轨迹漂移。将 KCF 与改进 KCF 的跟踪效果对比可发现,在目标人脸逐渐远离摄像头时(第 418 帧~第 458 帧),改进 KCF 的跟踪框能够不断适应人脸的大小,具有更好的跟踪效果。改进的 KCF 跟踪算法在原 KCF 跟踪算法上增加了跟踪框尺度自适应的功能,势必会增加计算时间,为了说明改进的 KCF 算法仍然具有实时性,现给出各算法在第 400 帧到第 499 帧之间图像上的运行耗时,如表 2 所示。

表 2 各跟踪算法运行耗时

算法	平均帧耗时(s)
STC	0.03796
CT	0.03404
KCF	0.01602
改进 KCF	0.03498

处理一张图像耗时 0.04 s 时,代表每秒处理 25 帧图像,故以处理每帧图像的平均耗时低于 0.04 s 为程序能实时运行的标准。表 2 中展示的 4 种算法平均帧耗时均低于 0.04 s,均能实现实时跟踪。4 种算法中,KCF 的跟踪速度最快,STC 的速度最慢,而改进的 KCF 算法虽增加了帧处理时间,但仍能达到实时性的要求。各算法耗时曲线如图 8 所示。

图 8 中,虽然改进的 KCF 算法、CT 算法和 STC 算法的平均帧耗时相差不大,但是改进的 KCF 算法的耗时曲线相对而言更加平稳,这也保证了跟踪系统的稳定性。

由于标准库中没有能够展示持续跟踪效果的测试视频,本文使用 PC 自带的摄像头录制了一段视频,来展示本系统的增量学习人脸特征、智能选择感兴趣目标以及尺度自适应持续跟踪的效果,如图 9 所示。

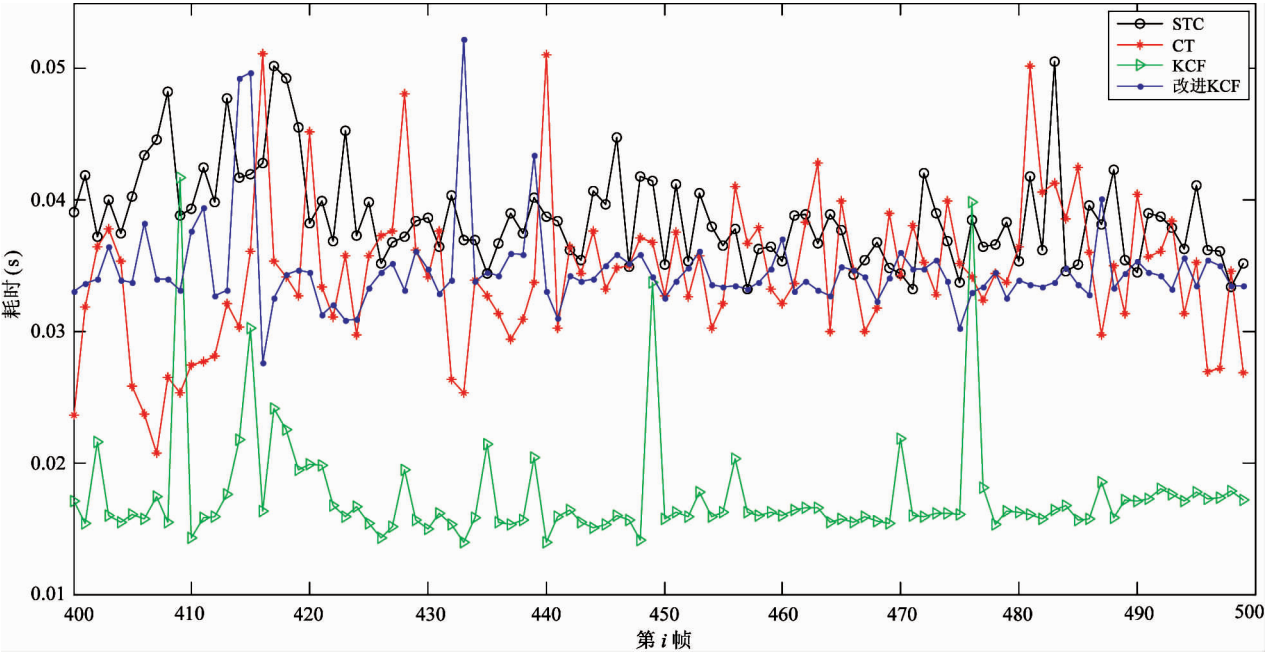


图 8 算法运行耗时曲线



图 9 持续跟踪效果图

图 9 中展示了结合识别与跟踪的人脸跟踪系统的整个实现过程。首先,检测视频中的人脸分别对多个人脸进行在线学习,如图 9(t1)和(t2)所示。其中图 9(t1)为人员 1(yh)的人脸特征学习过程,图

9(t2)为人员 2(cly)的人脸特征学习过程;然后,检测与识别人脸,并选定感兴趣人脸进行跟踪,如图 9(t3)~(t5)所示。其中图 9(t3)为人脸识别结果,图 9(t4)和(t5)中选定了目标人脸(cly)进行跟踪,

先利用人脸识别结果框初始化跟踪框,关闭人脸检测与识别程序(未继续对人员 1 进行人脸检测和识别),再启动改进的 KCF 跟踪器,实现了对感兴趣人脸的自动框选并跟踪;当感兴趣人脸逐渐远离摄像头时,如图 9(t4)和(t5)所示,随着人脸在画面中越来越小,跟踪框也越变越小,适应人脸尺度;当感兴趣人脸短暂离开摄像头画面,如图 9(t5)~(t9)所示,在人脸离开画面到达一定程度时,停止跟踪,重新启动人脸检测与识别程序,在连续 3 帧识别出感兴趣人脸后,重新定位感兴趣人脸(将识别出的人脸姓名标签与输入的感兴趣人脸姓名作对比,仅显示感兴趣人脸的检测与识别结果,未显示非感兴趣人脸(人员 1)的人脸检测和识别结果)。其中图 9(t8)为对感兴趣人脸进行重新识别定位,图 9(t9)为用识别框重新初始化跟踪框,启动改进的 KCF 跟踪器,实现了对人脸的持续跟踪。

本文所提方法能够在目标人脸短暂离开画面后重新跟踪目标人脸。为了探究人脸短时离开的时间长度对跟踪效果的影响,本文在不同离开时长下进行了跟踪实验,实验结果如图 10 所示。



图 10 人脸短时离开的时间长度容忍分析图

图 10 中分别给出了目标人脸离开画面 10 s、20 s、1 min 和 5 min 后的重新跟踪结果图。由图可知,在不同的离开时长下,系统都能准确地跟踪目标人脸。因为在目标离开画面后,需重新识别目标人脸,再用人脸识别框初始化跟踪框,从而实现再跟踪。而人脸重新识别用到的人脸特征检索库(IHDR 树)在进行人脸跟踪前已构建完成,并不随着人脸跟踪过程而更新,即人脸特征检索库不随着时间的推移而改变,人脸识别效果不随时间而改变,故人脸短时离开的时间长度不影响重新识别与跟踪效果。但对于许多应用场景而言,无限期的等待并没有太大的意义,故设定等待时间阈值 T (T 的大小根据实际情况而定),当目标离开时间 $t > T$ 时,关闭人

脸检测与识别系统,此时可进入下一个感兴趣人脸的跟踪。

4 结 论

本文提出了一种将人脸识别与跟踪相结合的方法,实现了对人脸的持续跟踪。针对人脸识别易受光照变化和人脸形变的影响,先对人脸进行对齐和椭圆掩膜修正,再提取人脸光照不变特征。针对以往的人脸跟踪大多是手动选取跟踪目标,既不方便也不精准的问题,提出用人脸识别框初始化跟踪框,自动选取感兴趣目标。针对 KCF 跟踪器不能自适应跟踪框的问题,提出了一种改进的 KCF 跟踪器,在计算出的下一个目标位置处取 3 个大小不同的区域,计算它们的响应,以确定跟踪框的大小。针对目标短暂离开画面后不能重新跟踪的问题,提出在跟踪过程中加入识别装置,在目标离开画面后关闭跟踪器,启动人脸识别装置,在连续 3 帧识别出目标后重新定位目标位置,重新初始化跟踪框,继续对目标进行跟踪。通过实验证明,本系统能够实现人脸持续跟踪,虽然改进的 KCF 跟踪器比 KCF 跟踪器更为耗时,但是改进的 KCF 跟踪器能够实现跟踪框尺度自适应,在能满足实时性的要求下具有更好的跟踪效果,对之后的人脸跟踪研究具有重要的意义。

参考文献

- [1] Degadwala S, Pandya S, Patel V, et al. A review on real time face tracking and identification for surveillance system[C]. In: Proceedings of the International Conference on Recent Trends in Engineering, Science & Technology, Hyderabad, India, 2016. 1-5
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005. 886-893
- [3] Ukil A. Support vector machine[J]. Computer Science, 2002, 1(4):1-28
- [4] Viola P A, Jones M J. Detecting pedestrians using patterns of motion and appearance in videos[P]. US patent: 20040258307, 2004
- [5] Ren Z, Zhang X, Yang S. A real-time suspicious stay detection system based on face detection and tracking in monitor videos[C]. In: Proceedings of the 2017 10th International Symposium on Computational Intelligence and

- Design, Hangzhou, China, 2017. 264-267
- [6] Elrefaei L A, Alharthi A, Alamoudi H, et al. Real-time face detection and tracking on mobile phones for criminal detection[C]. In: Proceedings of the 2017 2nd International Conference on Anti-Cyber Crimes, Abha, KSA, 2017. 75-80
 - [7] Sanjaya W S M, Anggraeni D, Zakaria K, et al. The design of face recognition and tracking for human-robot interaction[C]. In: Proceedings of the 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering, Yogyakarta, Indonesia, 2017. 315-320
 - [8] Flores-Delgado J I, Martínez-Santos L G, Lozano R, et al. Embedded control using monocular vision: face tracking[C]. In: Proceedings of the 2017 International Conference on Unmanned Aircraft Systems, Miami, USA, 2017. 1285-1291
 - [9] Ramkumar G, Logashanmugam E. An effectual face tracking based on transformed algorithm using composite mask[C]. In: Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research, Chennai, India, 2017. 1-5
 - [10] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]. In: Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition, Kauai, USA, 2001. 511-511
 - [11] Bay H, Tuytelaars T, Gool L V. SURF: speeded up robust features[C]. In: Proceedings of the European Conference on Computer Vision, Graz, Austria, 2006. 404-417
 - [12] Khurshid A, Scharcanski J. Incremental multi-model dictionary learning for face tracking[C]. In: Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, Houston, USA, 2018. 1-6
 - [13] Ren Z, Yang S, Zou F, et al. A face tracking framework based on convolutional neural networks and Kalman filter[C]. In: Proceedings of the IEEE International Conference on Software Engineering and Service Science, Beijing, China, 2017. 410-413
 - [14] Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: a convolutional neural-network approach[J]. *IEEE Transactions on Neural Networks*, 1997, 8(1):98-113
 - [15] He S, Yang Q, Lau R W H, et al. Visual tracking via locality sensitive histograms[C]. In: Proceedings of the Computer Vision and Pattern Recognition, Portland, USA, 2013. 2427-2434
 - [16] Weng J, Hwang W S. Incremental hierarchical discriminant regression. [J]. *IEEE Transactions on Neural Networks*, 2007, 18(2):397
 - [17] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(3):583-596

Face continuous tracking method based on online learning mechanism

Cai Liyi, Wu Huaiyu, Chen Jingyu, Chen Yang

(School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081)

Abstract

Aiming at the problem that the target face can not be continuously tracked after leaving the screen for a short time, the incremental hierarchical discriminant regression method (IHDR) is combined with the improved kernelized correlation filters (KCF) tracking algorithm to solve the face continuous tracking problem. Firstly, the face illumination invariant features is extracted, and the face feature IHDR tree is incrementally constructed. Then, IHDR tree is retrieved to identify the target face, the positive and negative samples of the face are obtained through the cyclic matrix, and the face regression is tracked by the training ridge regression classifier. When the face briefly leaves the screen, the target face is re-identified, and the tracker is re-initialized using the recognition result to achieve continuous tracking of the face. In addition, for the problem that the tracking frame size of the KCF tracker is not adaptive, the KCF tracker is improved. Three template areas with different scales are set and the responses are calculated, and the scale of the area with the largest result is the scale of the tracking frame. Finally, the dynamic face recognition experiment is carried out under different illuminations, and the recognition rate reaches 97.84%. Compared with the traditional tracking algorithm, the proposed method can adaptively track the target face on the scale and meet the real-time requirements. In the video where the face briefly leaves the screen, the method can also achieve continuous tracking of the target.

Key words: face continuous tracking, illumination invariant feature, online learning, scale adaptation, kernelized correlation filter(KCF)