

在线修正的人脸姿态跟踪算法及交互应用*

胡银焕⁺, 安国成, 张凤军, 戴国忠

(中国科学院软件研究所 人机交互技术与智能信息处理实验室,北京 100190)

Face Tracking Based on Online Modifying under Complex Postures and Interactive Applications

HU Yin-Huan⁺, AN Guo-Cheng, ZHANG Feng-Jun, DAI Guo-Zhong

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

+ Corresponding author: E-mail: huyinhuan@126.com

Hu YH, An GC, Zhang FJ, Dai GZ. Face tracking based on online modifying under complex postures and interactive applications. *Journal of Software*, 2010,21(Suppl.):1-11. <http://www.jos.org.cn/1000-9825/10001.htm>

Abstract: Because human face tracking is prone to be affected under complex postures (largely tilt, rotate, occlusion and color interference), this paper proposes a new face tracking algorithm based on online modifying, which includes two modules: detection module and tracking module. Tracking module is effective when detection is failed, and in the process of face tracking, in order to reduce the continuous accumulation of errors, we use real-time face detection to fix the parameters (including tracking the location and scale) of the tracking module, which made use of face detection and tracking with their respective advantages. Experiments show that the algorithm is able to accurately track a face under largely tilt and rotate and is also robust to color interfere and occlusion. In addition, the algorithm is used to control commercial games, which provides a new interactive way for HCI.

Key words: online modify; complex postures; occlusion; tracking

摘要: 针对人脸在跟踪过程中可能存在大幅度的倾斜、旋转、遮挡以及肤色干扰等问题,提出一种基于在线修正的人脸跟踪算法.该算法当人脸检测失效时,人脸跟踪模块将用于提取目标参数;而在人脸跟踪过程中,为减小由连续跟踪造成的累积误差,利用人脸实时检测机制新检测到的人脸目标参数来修正跟踪模块的参数,包括跟踪窗口的位置和尺度,从而利用了人脸检测和人脸跟踪各自的优点.通过实验,其结果表明,该算法能够精确地跟踪复杂姿态下的人脸目标,并且能够解决肤色干扰和遮挡的问题,具有很好的适应性和鲁棒性.另外,将在线修正的跟踪方法应用于娱乐游戏控制,为人机交互提供了新的方式.

关键词: 在线修正;复杂姿态;遮挡;干扰;跟踪

随着计算机硬件的发展和计算机视觉技术的进步,越来越多的研究者开始从事基于摄像头的用户界面

* Supported by the National Basic Research Program of China under Grant No.2009CB320804 (国家重点基础研究发展计划(973)); the National Natural Science Foundation of China under Grant Nos.U0735004, 60673188 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant Nos.2009AA01Z337, 2008AA01Z303 (国家高技术研究发展计划(863)); the China Postdoctoral Science Foundation under Grant No.20100470588 (中国博士后科学基金资助)

Received 2009-07-15; Accepted 2010-07-09

(camera based user interface,简称 CBUI)的研究.基于摄像头的用户界面通常使用一个或者多个摄像头跟踪人的运动情况,并将运动信息转化为对应的语义来控制操作.基于摄像头的用户界面为人们提供了更自然方便的交互方式,它可以应用在视频会议、计算机娱乐游戏、机器人控制等,并能为残疾人和老年人使用计算机提供方便的交互界面^[1].

1 相关工作

CBUI 作为新一代的人机交互方式之一,通常用人脸、手势、运动姿态、目光和表情等自然的方式和机器进行交互,它具有场景适应能力强、方便快捷、自然和谐的优点,使用户在一定程度上摆脱了鼠标和键盘的束缚,并且不需要任何其他可穿戴设备,减轻了人们的负担.Yang^[2]等人使用 DDLE 无监督学习方法对人脸表情进行分类,通过计算人脸的神经张力能量跟踪识别表情,并将该方法应用到人与智能服务机器人的交互中,使机器人不仅能够进行人脸检测,而且能够对人脸的表情和运动准确地识别.Dmitry^[3]则尝试用鼻子进行人机交互.为了能够准确的提取鼻子的位置,他们使用两个普通的摄像头进行检测和跟踪,从而使用户摆脱了传统的手持鼠标进行交互的方式.

在基于摄像头的交互界面中,人脸检测和跟踪起着重要的作用.Bradoski^[4]将人脸跟踪用于 3D 图形的控制,头部的运动被转换为 3D 图形摄像头视点的变动.作者在中国紫禁城和夏威夷岛这两个三维模型环境中漫游,使用人脸跟踪定位头部的朝向,进而控制在虚拟环境中的漫游方向.Fu^[5]等人则使用人脸跟踪技术模拟鼠标,利用人脸信息实现了一系列鼠标功能,并且用人脸跟踪操纵游戏,取得了良好的效果.Wang^[6]等人指出在交互系统中仅使用运动检测具有自身的局限性,系统并不能真正意识到用户的存在,不能确认用户的身份,故他们设计了两个游戏原型,在游戏中,利用实时人脸检测和跟踪技术解决了以上两个问题,增强了游戏的真实性和用户沉浸感.

可以看到,人脸检测和跟踪是基于摄像头的人机界面的核心技术之一,它有着广泛的实际应用.过去的十多年间,人脸检测和跟踪技术取得了很大的进展.Viola 与 Jones^[7]使用 AdaBoost 检测人脸,这种方法比先前已有的检测技术在精确度和速度上都有很大程度的提高,尽管如此,他们的方法只能检测人的正脸姿态,当人脸倾斜或者旋转时,该方法就会失效,因此具有很大的局限性.很多研究人员针对人脸跟踪,提出了不同的解决方案.Boltz^[8]等人提出基于高维特征的目标跟踪算法,他们利用目标区域的像素信息、梯度信息以及图像片信息构成表示目标的高维特征,然后对该特征进行简化处理,从而实现基于高维特征空间的目标跟踪.而 Han^[9]等人则是利用低维空间的跟踪模块融合成目标高维空间的状态参数估计算法来进行目标跟踪的.Bradoski^[4]的 CAMSHIFT 算法基于肤色和概率论对人脸区域进行跟踪,算法速度快,并被后来很多学者利用,但是该方法受肤色干扰的影响.

现有的检测算法对正脸姿态的检测率高,速度快,对于倾斜和旋转姿态下的人脸检测性能有所下降.已有的跟踪算法也存在需要解决的问题:人脸运动速度较快可能导致图像模糊,降低跟踪器的性能;在跟踪过程中,人脸容易被干扰,如手、颈部、手臂等都会影响跟踪器的性能;人脸目标在跟踪过程中可能存在自身遮挡、场景遮挡等现象,这些都会造成跟踪不精确的现象.为解决上述问题,本文利用检测和跟踪各自的优点(检测对于正脸姿态的适应性强,检测精确,而跟踪可以处理倾斜和旋转下的实时跟踪),提出一种新的基于在线修正的跟踪方法.该算法首先利用检测模块在正脸姿态下获取的人脸的位置和尺度信息初始化跟踪模块,然后利用检测模块的实时检测结果在线修正跟踪模块的位置和尺度参数,使得新的跟踪算法能够处理肤色干扰、遮挡等问题;另外,当检测模块失效时,使用跟踪模块跟踪人脸,以解决检测模块因人脸姿态发生变化而导致性能下降的问题.

2 算法概述

本文检测模块基于 Viola 和 Jones^[7]的人脸检测算法对视频帧进行人脸检测,并在此基础上针对如何精确地确认人脸作了进一步的研究;跟踪模块基于 CAMSHIFT^[4]算法,提出了一种在线修正的跟踪方法,使跟踪的精确

度得到了很大程度的提高.算法的流程图如图 1 所示.

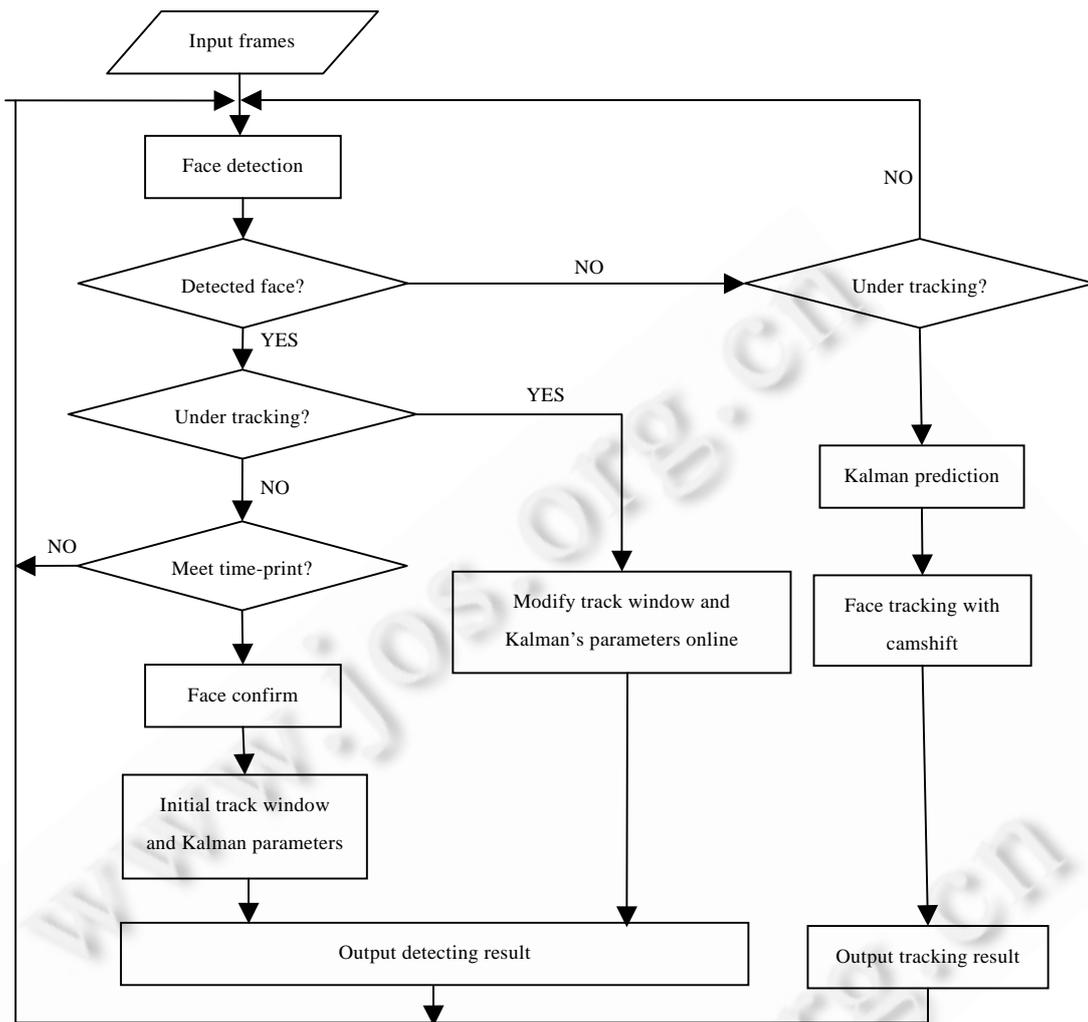


Fig.1 Algorithm flow chart

图 1 算法流程图

算法采用一种相互补偿的机制.首先,当人脸处于正脸姿态时,检测模块检测人脸,记录人脸状态的位置、尺度等信息,并利用这些信息对跟踪模块进行初始化,从而达到跟踪模块自动初始化的目的.当人脸倾斜或者旋转时,检测模块可能失效,这时跟踪模块被激活,为人脸的连续跟踪提供必要的参数,从而实现目标的连续跟踪;其次,在跟踪过程中,由于连续跟踪产生累积误差,造成跟踪结果发生偏移.为了减小跟踪模块产生的误差,提高跟踪结果的精确度,利用检测模块的实时检测结果在线修正跟踪模块的位置、尺度等参数,使已经偏离真实位置的跟踪窗口重新定位于真实的(或接近真实的)人脸位置.

3 基于时间印的人脸检测

3.1 AdaBoost人脸检测原理

目前基于 AdaBoost 的人脸检测方法在实际中得到广泛的应用.本文人脸检测模块正是采用 Viola^[7]的基于

AdaBoost 的学习方法进行人脸检测的. Viola 人脸检测方法是一种基于积分图、级联检测模块和 AdaBoost 算法的方法, 它是一种有监督的统计模式识别方法. 该方法首先使用 Harrlike 特征表示人脸, 使用“积分图”实现特征数值的快速计算; 然后利用 AdaBoost 方法选出一些最能代表人脸的弱分类器, 按照加权的方式将弱分类器构造为一个强分类器; 最后将训练得到的若干强分类器串联组成一个级联结构的层叠分类器, 级联结构能有效地提高分类器的检测速度和精确率.

尽管 Viola 的人脸检测方法速度快, 检测率高, 但仍存在一些问题: (1) 该检测方法是基于矩形特征的, 检测窗口在待检测图像上的移动步长为一个像素, 这就导致检测结果窗口的重叠现象, 或者存在无效的检测结果窗口; (2) Viola 的人脸检测方法仅对正脸姿态产生很好的检测效果, 而当人脸处于倾斜或者旋转姿态时, 该方法的适应性降低, 经常检测不到人脸, 产生人脸丢失的现象. 为此, 本文将在第 3.2 节和第 4 节分别对以上两个问题提出解决方案.

3.2 有效窗口合并

通过 Viola 的人脸检测方法得到的检测结果存在的问题之一是: 检测窗口的重叠. 本文提出一种解决窗口重叠和去除无效窗口的方法: 合并有效检测窗口. 作者将可能为人脸的窗口记为“有效检测窗口”. 要合并有效检测窗口, 首先要寻找到有效检测窗口, 然后才能进行合并. 具体过程如下:

首先寻找有效检测窗口. 设某一帧检测窗口序列为 $Sequence[n] = \{R_1, R_2, \dots, R_n\}$, n 为检测窗口的数目, R 为检测窗口, $R \rightarrow x, R \rightarrow y$ 表示检测窗口的中心.

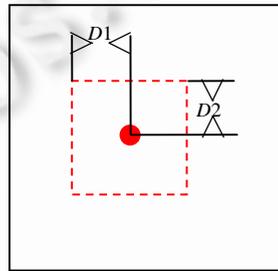


Fig.2 Effective detection window

图 2 有效检测窗口

在图 2 中, 黑色矩形表示一个检测窗口, 实心圆表示该结果窗口的中心位置 $R(x, y)$, 虚线小矩形表示预先设定的一个以 R 为中心的邻域矩形. $D1, D2$ 分别表示虚线矩形宽和高的一半. 假设检测窗口序列中的任意一个窗口 R_i, R_j 是除 R_i 外的检测窗口序列中的任意一个, $count$ 为整数, 初始值为 0; N 为整数, 它和检测窗口序列中窗口的个数 n 存在一个比例系数 $\alpha, N = n \times \alpha, \alpha \in (0, 1)$ 其中, N 为判断 R_i 是否为有效检测窗口的阈值, n 为检测窗口的总数, 本文中 $\alpha = 0.5$. 若: $|R_j \rightarrow x - R_i \rightarrow x| \leq D1$, 且 $|R_j \rightarrow y - R_i \rightarrow y| \leq D2$, 则: 认为 R 处于 R_i 规定的阈值范围内, $count$ 增加 1. 同样, 分别计算其他窗口是否也在 R_i 的阈值范围内, 累计 $count$ 的值. 若 $count \geq N$, 则 R_i 为有效检测窗口, 否则, 该为无效检测窗口.

其次, 合并有效检测窗口(如图 3 所示, 视频序列图像采用斯坦福大学提供的人脸图像). 合并有效检测窗口的方法是: 计算有效检测窗口的平均窗口, 将平均窗口作为该检测帧的最终检测窗口. 计算方法如下: 设 w, h 分别为有效检测窗口的宽和高, W, H 为最终检测窗口的宽和高, (x, y) 为最终检测窗口的中心坐标, 则 X, Y, W, H 的计算公式如(1)所示:

$$X = \frac{\sum_{i=1}^N R_i \rightarrow x}{N}, Y = \frac{\sum_{i=1}^N R_i \rightarrow y}{N}, W = \frac{\sum_{i=1}^N R_i \rightarrow w}{N}, H = \frac{\sum_{i=1}^N R_i \rightarrow h}{N} \quad (1)$$

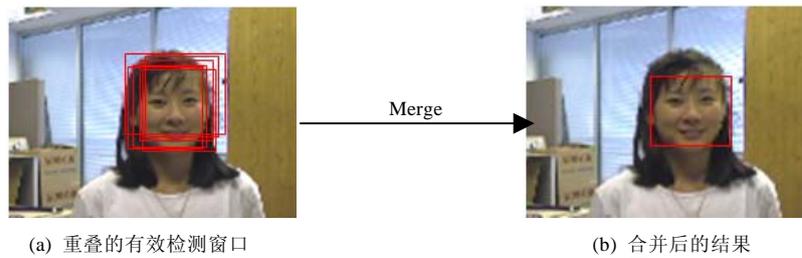


Fig.3 Schematic diagram of the merger effective detection window

图3 合并有效检测窗口示意图

3.3 基于时间印的人脸确认

尽管对一帧中的检测结果窗口序列进行了处理,但并不能完全消除无效检测窗口对检测结果的影响.为尽量减小无效检测窗口对检测结果的影响,本文采用基于时间印^[15]的机制确认检测模块最终的人脸检测结果.时间印的基本思想是:在连续帧中,分析每一帧中的检测结果的位置相似度,若该相似度在设定的阈值范围内,则计算所有连续帧的检测结果窗口的平均窗口,平均窗口即为人脸检测模块最终确认的人脸结果.

假设时间印要通过连续 M 帧的检测结果来确认人脸,由第 3.1 节和第 3.2 节可以得到 M 帧的检测结果窗口,用 $S_i(X, Y, W, H)$, $i \in [1, M]$ 表示,其中, (X, Y) 表示检测结果窗口的中心位置, (W, H) 表示检测结果窗口的宽和高.首先,求得 N 帧检测结果窗口的平均窗口 $A(X, Y, W, H)$, 其中,

$$A \rightarrow X = \frac{\sum_{i=1}^M S_i \rightarrow X}{M}, A \rightarrow Y = \frac{\sum_{i=1}^M S_i \rightarrow Y}{M}, A \rightarrow W = \frac{\sum_{i=1}^M S_i \rightarrow W}{M}, A \rightarrow H = \frac{\sum_{i=1}^M S_i \rightarrow H}{M} \quad (2)$$

然后,采用第 3.2 节中寻找有效检测窗口的方法判断 M 帧的最终检测结果的位置是否在设定的平均窗口的阈值内,若是,则将 M 帧的平均窗口作为检测模块最终确认的检测结果.

4 在线修正的人脸跟踪

本文基于 CAMSHIFT^[4] 算法,实现人脸跟踪模块.由于 CAMSHIFT 算法是以颜色为跟踪线索的,在人脸跟踪过程中,容易受到非目标肤色区域的干扰,如手、颈部等,并且也容易受到目标遮挡的影响,而且当人脸目标快速运动时,其跟踪精度也会降低.为了解决 CAMSHIFT 算法所存在的问题,本文提出了在线修正的方法.

4.1 CAMSHIFT 原理

CAMSHIFT^[4] 是一种运动跟踪方法.它主要通过视频图像中运动物体的颜色信息来达到目标跟踪的目的.该算法首先计算被跟踪目标的色彩直方图,在各种色彩空间中,只有 HIS 空间的 H 分量可以表示颜色信息,所以在具体计算过程,需将其他色彩空间的值转化到 HIS 空间,使用其中的 H 分量作直方图计算,根据获得的色彩直方图将原始图像转化成色彩概率分布图像;然后根据色彩概率分布图像,通过二阶矩阵的方法计算目标跟踪区域的重心,调整跟踪窗口的位置到重心,这样就在帧中找到了目标跟踪区域;CAMSHIFT 是在连续图像序列中重复上述步骤,从而实现视频目标的连续跟踪.最初的 CAMSHIFT 方法搜索窗口的初始化是通过手动完成的,本算法实现了跟踪目标自动初始化的功能.

4.2 在线修正机制

CAMSHIFT 方法虽然速度很快,但在跟踪人脸时,一旦人脸运动速度很快,或者受到非目标肤色的干扰,该算法很难找到人脸的精确位置,会发生很大偏移,甚至将非人脸区域作为人脸,造成目标人脸丢失.为此,采用 Kalman^[10] 滤波对人脸运动状态进行估计,以提高人脸跟踪的精确度. Kalman 滤波是基于最小均方误差准则的线性系统最优估计方法,它的主要作用是在观测噪声和过程噪声的作用下,利用预测值和输出值估计出系统的真值.

已有的研究工作表明将 CAMSHIFT 和 Kalman 滤波组合使用在一定程度上提高了目标跟踪的精确度.但是跟踪人脸面临特殊的挑战,原因在于人脸与颈部、手等具有相近的颜色特征,在跟踪过程中人脸不可避免会受到它们的干扰或者遮挡,产生累积误差.这种累积误差会使跟踪窗口发生偏移,CAMSHIFT 和 Kalman 滤波结合并没有纠正这种累积偏移的能力,依然会导致人脸目标的丢失.为此,本文利用检测模块对跟踪模块的参数进行实时修正.

在线修正的主要思想是:算法的检测模块和跟踪模块都在实时运行,在连续跟踪的过程中,尽管进行了人脸状态估计,但是跟踪结果的精确度也难免受其他接近肤色的物体(颈部、手、胳膊等)遮挡和干扰的影响,跟踪窗口必然会发生偏移,而且这个偏移量随着跟踪的继续会越来越大,从而导致跟踪窗口与真实人脸位置的偏差逐渐变大.由于检测模块在实时运行,而且检测方法的速度特别快,因此,利用新检测到的人脸的位置、尺度等信息去修正跟踪模块的参数,使搜索窗口的位置调整到人脸中心或者接近人脸中心的位置.由于检测模块速度快,精确率高,所以会在跟踪结果的偏移量不是很大时就做出调整.

参数在线修正包括两个部分(如图 4 所示):(1) 修正 Kalman 滤波位置参数;(2) 修正搜索窗口的尺度.本文算法的检测模块除了要自动初始化跟踪模块,还担负着另外一个重要的任务:那就是为跟踪模块提供在线修正的位置、尺度等参数.设检测窗口的状态为 $DetectWin(x, y, with, height)$, Kalman 滤波的位置参数为 $Position(x, y)$, 搜索窗口为 $TrackWin(x, y, with, height)$, 则使用 $DetectWin(x, y, with, height)$ 对 $Position(x, y)$ 和 $TrackWin(x, y, with, height)$ 进行在线修正如下:

$$Position \rightarrow x = DetectWin \rightarrow x, Position \rightarrow y = DetectWin \rightarrow y \quad (3)$$

$$TrackWin \rightarrow with = DetectWin \rightarrow with, TrackWin \rightarrow height = DetectWin \rightarrow height \quad (4)$$

得到修正后的位置和尺度信息后,重新计算搜索窗口和设置新的 ROI.然后根据新的搜索窗口和 ROI 计算颜色直方图、计算反向投影图像,找到搜索窗口的重心,移动搜索窗口到搜索重心,从而找到要跟踪的人脸目标.

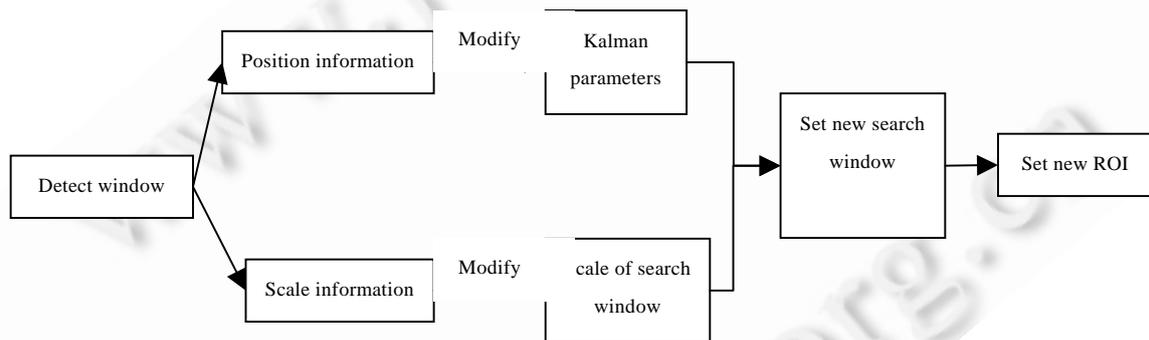


Fig.4 Online modifying diagram

图 4 参数在线修正示意图

5 实验分析

为验证在线修正的人脸跟踪算法的性能,本文进行了自拍场景实验.实验场景是在实验室内拍摄的人脸各种姿态的视频.在整个视频中,人脸的姿态和尺度随意变化,使用普通的 USB 摄像头,视频图像的分辨率为 320×240 ,共有 768 帧.具体实验结果如下.

5.1 跟踪算法性能验证

下面是验证算法稳定性的部分实验结果,如图 5 所示.其中右上角有圆斑标记的帧代表检测结果,帧中矩形窗口表示检测到的人脸区域,左上角有圆斑标记的帧代表跟踪结果,帧中矩形表示跟踪到的人脸区域.图 5(a)中,实验者由正脸姿态逐渐转变到倾斜姿态,在 31 帧检测模块失效,跟踪模块马上被激活,准确地跟踪到人脸.图 5(b)显示的是实验者由正脸姿态变为仰脸姿态,在第 144 帧检测模块失效时跟踪模块起到了补偿作用,而在第

145 帧检测模块又重新检测到了人脸,在第 146 帧的状态下检测模块又一次失效,而跟踪模块又一次跟踪到人脸,在该过程中,人脸检测和跟踪结果交替出现,增强了算法的鲁棒性和稳定性.图 5(c)是人脸逐渐被完全遮挡时的跟踪性能,在第 486 帧,人脸被遮挡,特征丢失,检测模块失效,而跟踪模块依然及时地避免了人脸丢失.由实验结果可以看出,无论从正脸姿态转化为倾斜姿态,还是在人脸特征被逐渐遮挡的情况下,都未发生人脸丢失的现象,而且跟踪结果偏差不大,验证了基于在线修正的人脸跟踪算法的鲁棒性.

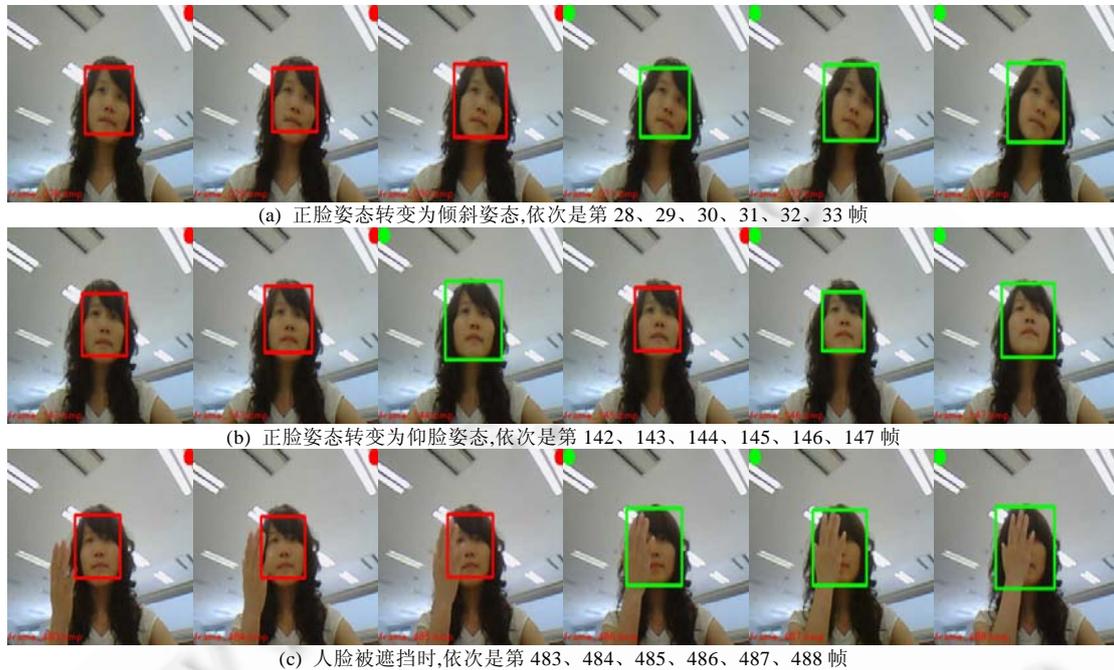


Fig.5 Mutual compensation between detection module and tracking module

图 5 检测模块和跟踪模块相互补偿

图 6 是 CAMSHIFT 算法、CAMSHIFT 和 Kalman 滤波相结合算法以及基于本文算法跟踪结果的比较,是在人脸未被遮挡情况下进行的.自拍视频中实验者的颈部和人脸肤色非常接近,是对跟踪结果影响最大的因素.图 6(a)是 CAMSHIFT 算法人脸跟踪结果,可以看到跟踪结果受颈部的影响非常大,跟踪窗口完全将颈部区域也作为人脸区域,跟踪结果误差较大.图 6(b)是 CAMSHIFT 和 Kalman 相结合的跟踪结果,虽然比 CAMSHIFT 方法本身的跟踪结果在精确度上有所提高,但是受肤色影响依然比较明显,如第 21、149 帧还是将部分颈部当作了人脸区域.而图 6(c)是基于在线修正跟踪方法的跟踪结果,从跟踪结果的图序列中可以看到,颈部并没有对跟踪结果造成很大影响,跟踪窗口的位置更加接近人脸区域.

图 6 中序列是在脸部无遮挡情况下的性能比较,而且肤色干扰区域仅限于颈部,图 7 中所做的实验情况相对比较复杂,视频序列中实验者的脸部部分或者全部被遮挡,并且颈部、手、手臂和人脸肤色非常接近,也是影响跟踪结果的重要因素.图 7(a)是 CAMSHIFT 算法的跟踪结果.由图中可以看出,跟踪窗口远远超出了人脸范围,受颈部和手臂等肤色干扰的影响相当大,跟踪结果几乎完全失效.图 7(b)是 CAMSHIFT 和 Kalman 相结合的跟踪结果,跟踪的精确度有了提高,但是人脸的中心位置和人脸区域相对真实的人脸位置仍然存在较大误差.图 7(c)中是本文算法的跟踪结果,尽管手臂和颈部在其中既遮挡脸部又产生肤色干扰因素,但是跟踪窗口依然紧紧围绕在人脸的区域.充分显示了本文跟踪方法的精确性和鲁棒性.

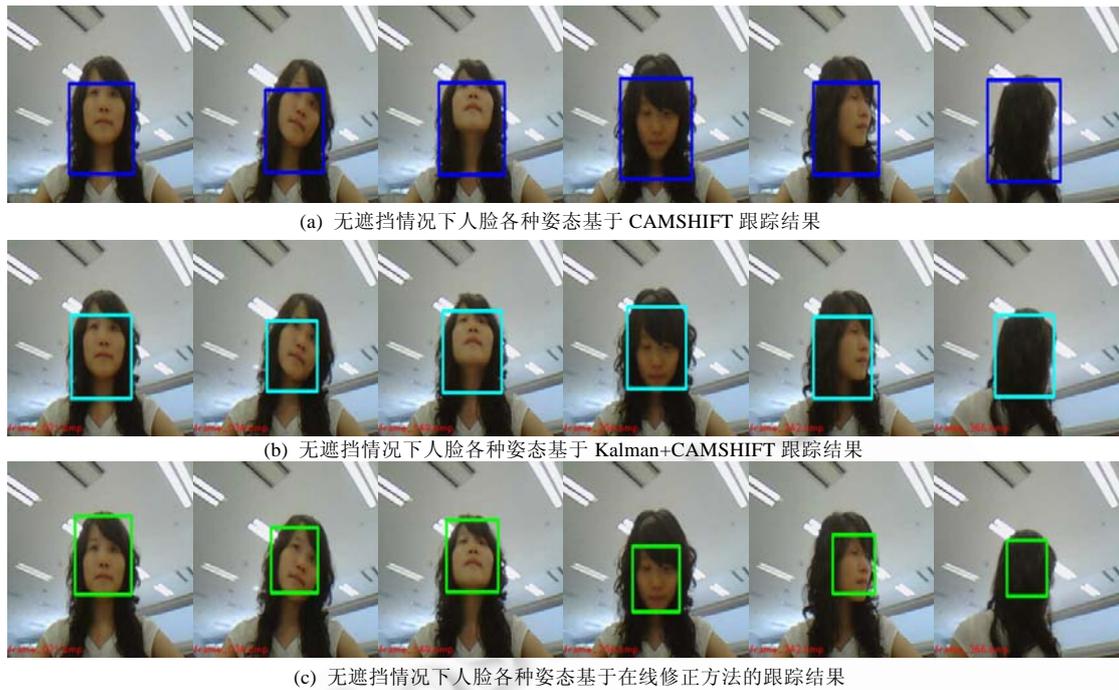


Fig.6 Face tracking results comparison without occlusion

图 6 无遮挡情况下,人脸各种姿态跟踪结果的比较
(依次是第 21、34、149、204、242、366 帧)

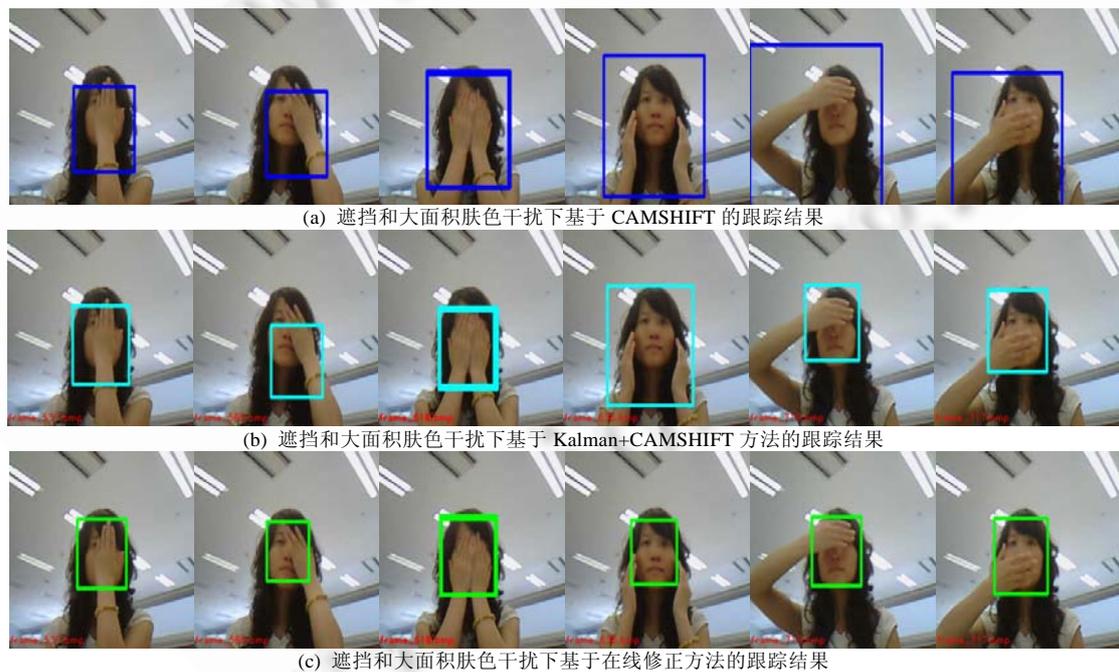


Fig.7 Face tracking results comparison with occlusion and color interfere

图 7 有遮挡和面积肤色干扰下跟踪结果的比较
(依次是第 537、585、616、636、713、717 帧)

图 8 显示了 CAMSHIFT 算法、CAMSHIFT 和 Kalman 相结合方法以及在线修正的跟踪算法的性能图.横坐标表示帧数,纵坐标表示跟踪结果与真实位置之间的欧式距离(用像素计算),真实位置是手动标记的.图中蓝色三角形曲线表示 CAMSHIFT 算法的性能,红色圆形曲线表示 CAMSHIFT 和 Kalman 滤波相结合方法的性能,绿色十字曲线表示在线修正的跟踪算法的性能.由图 8 可以看出在线修正的跟踪算法在大部分情况下,跟踪结果和真实位置之间的欧式距离比较小,跟踪性能比较稳定,而其他两种算法的跟踪性能相对很低,产生的误差也相对较大.在 200 和 300 帧之间,人由正脸姿态逐渐转变为侧脸,再转变为背对摄像头,由于肤色信息逐渐减弱,前两种算法的误差偏大,而在线修正的跟踪算法误差较小.在 400 和 600 帧之间,人脸被遮挡,并且受到手和手臂大面积近似人脸肤色的干扰,前两者的跟踪误差急剧增大,而本文的算法则表现相对比较稳定,误差较小,受到遮挡和干扰的影响较小.表 1 给出了这 3 种跟踪方法相对于真实值的均值和标准差.可以看出本文方法的均值和标准差相对 CAMSHIFT 方法分别缩小了 3.7 倍和 1.5 倍,说明本文算法具有很强的稳定性.

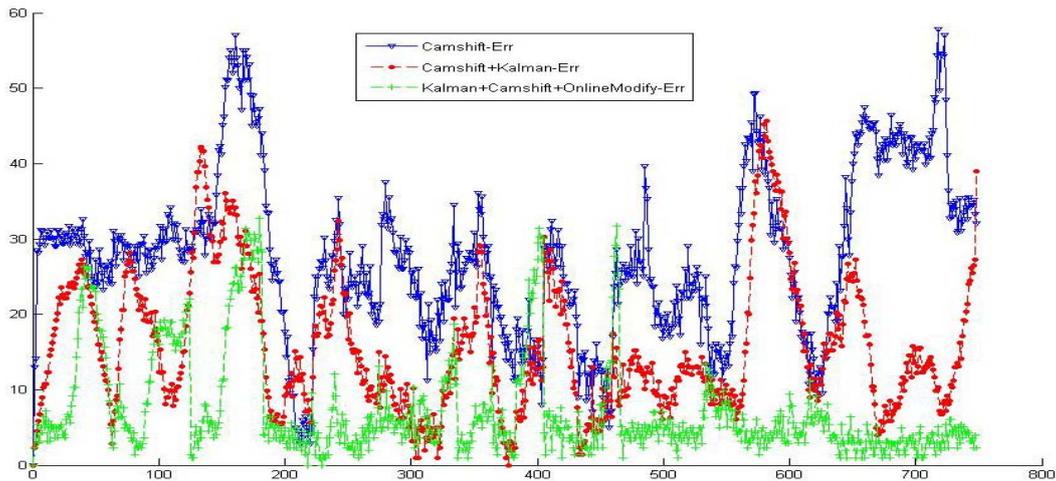


Fig.8 Performance comparison of three tracking algorithms

图 8 3 种跟踪方法的性能对比图

Table 1 Mean and standard deviation between the three methods and the true value

表 1 3 种方法和真实值之间的均值和标准差

Name	CAMSHIFT	Kalman+CAMSHIFT	Online modify
Mean	27.722 7	15.834 3	7.358 4
Standard deviation	10.824 8	9.195 4	6.882 2

5.2 交互设计

目前基于人脸跟踪技术的互动娱乐系统已被开发出来,并受到人们的喜爱.它不仅为人们提供了可选择的便捷的交互方式,而且为老人和残疾人提供方便可用的交互界面.为了验证本文跟踪算法的性能和实用性,作者将算法应用到两个互动娱乐系统中,并且利用该算法,设计了合理、有效、自然的交互方式.下面是对两个娱乐系统交互设计的详细描述.

(1) 赛车游戏

在该游戏中,用户可以使用头部的运动来控制赛车的运动(左转、右转、加速、减速等).该交互设计的主要原理是:根据跟踪窗口的尺度和位置信息,系统预先设定好相应的左转、右转、加速和减速的阈值,在规定的阈值范围内驱动赛车做出相应的运动.本文利用跟踪窗口的中心位置的横坐标值控制赛车的左右转向,利用跟踪窗口的尺度信息来控制赛车加速和减速.假设控制赛车方向的阈值为 P ,控制赛车加速和减速的尺度阈值为 S ,那么当跟踪窗口中心位置的横坐标 $X > P$ 时,赛车左转;若 $X < P$ 时,赛车右转.当跟踪窗口的宽度 $width > S$ 时,赛车加速行驶,反之赛车则减速.在本文实验中 $X=170, S=75$,用户根据实际交互情况来重新设定阈值.实验结果表明,

使用人脸控制此类游戏能够获得更多的沉浸感,交互更加自然.图 9 是交互游戏的效果图,效果图的上部分是游戏场景,下部分是人脸跟踪系统,实时和游戏场景进行交互.



Fig.9 Play racing game using the face

图 9 使用人脸控制赛车游戏

(2) 接礼物游戏

本文根据人脸运动的特点,设计开发了一款适合使用人脸跟踪技术作为交互方式的游戏“接礼物”.该游戏中,一只小松鼠是玩家的化身,当玩家的头部左右运动时,小松鼠随之做出相应的运动.游戏里从天空随机落下各种各样的水果,玩家要不停地左右移动才能接到更多的水果,如果在规定的时间内完成了接礼物的任务,则闯入下一关,游戏难度将增加,即礼物落下的速度会更快.游戏中作者用玩家的脸部替代了松鼠本身的脸部,进一步提高了游戏的沉浸感.

该互动娱乐游戏的设计原理是:玩家在摄像头前左右运动,人脸的位置也随之移动,系统使用本文提出的跟踪算法实时跟踪人脸位置,将其转化为化身在游戏场景中的位置信息,进而驱动化身做出相应的运动.当化身移动到某一礼物下方时,礼物就会落到化身的口袋里,便赢得相应的成绩.该游戏将跟踪窗口的中心位置映射为玩家在游戏场景中的位置,中心位置的横坐标控制化身左右运动,中心位置的纵坐标控制化身上下跳跃运动,玩家的人脸移动的速度增加,化身也在场景中加速运动,反之,化身也放慢运动速度.图 10 为该游戏的真实效果图.

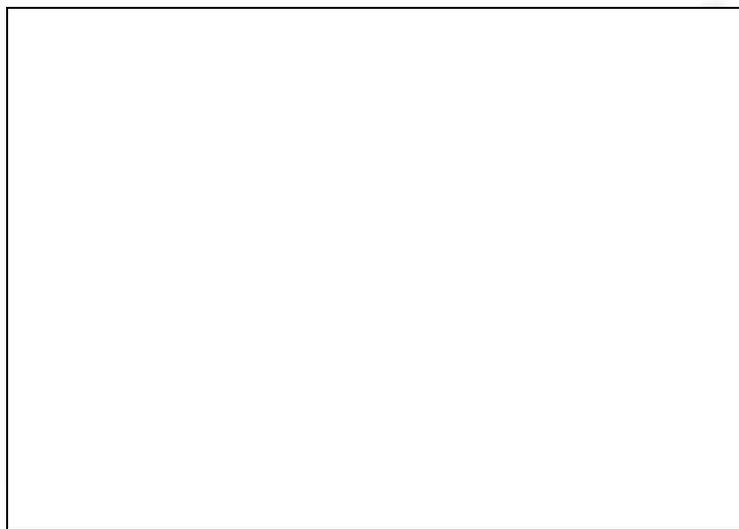


Fig.10 Receive gift game

图 10 接礼物

6 结 论

本文利用人脸检测和人脸跟踪在各种人脸姿态下分别具有的优缺点,提出了一种新的基于在线修正的跟踪算法,并将其应用到游戏控制中,为人机交互提供了新的方式.通过实验证明,在线修正的跟踪方法不仅可以处理人脸检测时人脸丢失的问题,而且,该方法能够解决人脸在大幅度倾斜、旋转和受干扰遮挡等情况下的跟踪问题,跟踪精确,速度快,为进行和谐自然的人机交互提供了技术保障.

References:

- [1] Betke M, Gips J, Fleming P. The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Trans. on NSRE*, 2002,10(1):1-10.
- [2] Yang Y, Ge SS, Lee TH, Wang C. Facial expression recognition and tracking for intelligent human-robot interaction. *Intel Serv. Robotics*, 2008,1:143-157.
- [3] Gorodnichy DO, Roth G. Use your nose as a mouse's perceptual vision technology for hands-free games and interfaces. *Image and Vision Computing*, 2004,22:931-942.
- [4] Bradski GR. Real time face and object tacking as a component of a perceptual user interface. In: *Proc. of the IEEE Workshop on WACV'98*. 1998. 214-219.
- [5] Fu Y, Huang TS. hMouse: Head tracking driven virtual computer mouse. In: *Proc. of the IEEE Workshop on Applications of Computer Vision*. 2007. 30.
- [6] Wang S, Xiong XC, Xu Y, Wang C, Zhang WW, Dai XF, Zhang DM. Face tracking an augmented input in video games: Enhancing presence. In: *Proc. of the CHI 2006, Role-Playing and Control, Selecting and Tracking*. 2006. 22-27.
- [7] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2001.
- [8] Boltz S, Debreuve E, Barlaud M. High-Dimensional statistical measure for region-of-interest tracking. *IEEE Trans. on Image Processing*, 2009,18(16):1266-1283.
- [9] Han BH, Davis LS. Probabilistic fusion-based parameter estimation for visual tracking. *Computer Vision and Image Understanding*, 2009,113:435-445.
- [10] Kalman RE. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 1960,82(Series D):35-45.



胡银焕(1984—),女,硕士生,主要研究领域为计算机视觉.



张凤军(1971—),男,博士,副研究员,主要研究领域为人机交互,虚拟现实,CAD.



安国成(1979—),男,博士,主要研究领域为人机交互,视频图像跟踪.



戴国忠(1944—),男,研究员,博士生导师,CCF高级会员,主要研究领域为人机交互,计算机图形学.