

基于交互式随机动作的人脸活体检测

黄叶珏

(浙江经贸职业技术学院 信息技术系, 浙江 杭州 310018)

摘要:针对人脸识别可以通过照片和视频进行攻击的问题,提出一种交互式随机动作的活体检测方法。对认证用户发出随机化的动作指令,根据对视频中的人脸进行检测和对齐跟踪来判断用户是否按要求完成动作,从而实现活体检测。利用随机化序列方式有效地抵抗了照片和视频认证的攻击。

关键词:人脸识别;人脸检测;活体检测;交互式随机动作;认证

DOI: 10.11907/rjdk.1511355

中图分类号: TP306

文献标识码: A

文章编号: 1672-7800(2015)012-0026-02

0 引言

人脸识别作为一项成熟的生物识别技术,目前已被广泛应用于金融、公安、社会服务、电子商务等领域。然而人脸很容易用视频或照片等进行复制,从而达到攻击的目的,活体检测是人脸识别能否有效应用的前提,目前对活体检测方法的研究有很多。大多数活体检测方法基于特征提取与训练的方式,由于应用环境的复杂性和不可控性,基于训练的方法的准确性不可控^[1-6]。另一类方法要求用户做转头、摇头、眨眼或者张嘴等动作,这类方法对于视频的防欺骗性不高,比如北京智慧眼科技有限公司的活体检测只要求用户眨眼即可;腾讯微众银行通过语音识别技术识别用户是否正确认读数字并结合用户嘴巴是否张开作为辅助技术来确保用户是活体,但是这种方法需要声音接收设备,比常规的视频活体检测的应用场景限制更严格。本文克服现有技术的不足,提供一种交互式随机动作的活体检测方法,采用人脸对齐技术,使得检测的健壮性更高,能准确判断出当前检测者是否为活体真人,解决了现有活体检测技术中存在的照片或视频欺骗问题。采用对视频中的人脸进行检测、跟踪和对齐,对用户发随机化的动作指令,对人脸特征位置进行分析来确定用户所做的动作是否按要求完成,从而实现活体检测。通过随机化的序列方式有效地抵抗了照片和视频的攻击。

1 人脸检测与人脸对齐算法

在人脸活体检测认证的应用场景中,一般要求人脸图像清晰且人脸正面朝向镜头。在该应用场景中,人脸检测

算法的要求不高,可采用 OpenCV 中的 AdaBoost 人脸检测器实现人脸定位。获取视频中的人脸位置后,可利用人脸对齐算法^[7]定位人脸的特征点位置。人脸对齐算法可得到比较准确的人脸各个特征点的位置,给定图像中 n 个特征点位置 $P_i \in P$,通过提取对应位置的 SIFT 特征^[8]、HOG 特征^[9]或者其它特征,根据式(1)求最小化实现人脸特征点对齐:

$$F(P^0 + \Delta P) = || H(P^0 + \Delta P) - H(P^*) || \quad (1)$$

其中, P^0 为根据人脸检测器的定位以及平均人脸对齐特征点计算得到的初始位置配置, P^* 为训练集中手工标注的人脸特征点位置, ΔP 为初始位置配置与 P^* 的位移量, $H(\cdot)$ 为 SIFT 特征或者 HOG 特征提取函数。

在训练阶段,先采用人工标注的方式对训练数据集人脸各个特征点的位置进行标注。使用人脸检测器对训练数据集中的人脸进行定位,根据检测到的人脸框,计算各个特征点在人脸框中的局部坐标,对训练数据集中所有特征点的坐标计算其平均值,就可以得到一个训练集的平均人脸特征点位置,将它作为人脸特征点位置的初始配置。人脸对齐模型训练通过计算每个训练数据的人工标注特征点与初始配置特征点的差异,得到迭代方向和步长生成对齐模型。为了提高算法的健壮性,需要对训练数据进行增强,即在原始训练数据对的基础上,通过对人脸检测框进行旋转、平移和缩放来增加初始配置的多样性。

训练得到人脸对齐模型后,就可以根据对齐模型和人脸检测框通过迭代的方式来得到各个人脸标注特征点位置。对于视频的每一帧使用人脸检测算法得到人脸框,人脸检测算法可能会获得多个人脸框。由于人脸活体检测认证应用中,被认证者是端坐在摄像机前的,所以取最大的 1~2 个人脸框作为候选人脸即可。然后根据人脸框使

作者简介:黄叶珏(1978—),女,浙江嵊州人,浙江经贸职业技术学院信息技术系副教授,研究方向为计算机图形图像与网络安全。

用平均人脸特征点位置作为初始配置,再根据人脸对齐模型进行迭代获得人脸特征点的位置。

为了进一步提高算法的健壮性,引入一个评分模型来降低误判。首先使用 OpenCV 的人脸检测器对 FDDB 人脸数据库^[10]进行检测,检测时降低阈值以达到 99% 以上召回率。该人脸检测器就可产生很多的误报人脸窗口,将这些输出结果通过人工筛选方式分成人脸和非人脸两类。训练时,利用前面得到的人脸对齐模型对人脸和非人脸都进行对齐迭代,根据各个特征点的位置分计算 32×32 的 128 位 SIFT 特征,然后将这些 SIFT 特征连接起来,输入到线性 SVM 中进行训练。这样就可以根据线性 SVM 的得分为阈值来评判人脸对齐的结果。本文实验中阈值为 0.3,小于 0.3 的就认为该人脸对齐结果不符合要求,如图 1 所示。

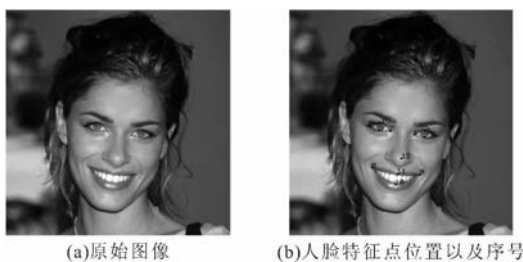


图 1 人脸对齐算法产生的人脸特征点位置

2 基于有限状态机的张嘴和眨眼动作识别

对视频进行人脸检测和人脸对齐后,就可以得到如图 1(b)所示的人脸对齐特征点位置。可以根据各个特征点位置来估计人脸的姿态,只需判断张嘴和眨眼动作即可。以右眼为例,右眼的左右眼角的特征点位置标号分别为 0 和 2,上下眼皮的特征点位置标号分别为 1 和 3,那么可以计算当前眼睛的宽和高,并根据下式可以判断右眼的状态:

$$J(P) = \frac{||P_0(x) - P_2(x)||}{P_3(y) - P_1(y)} \quad (2)$$

其中, $P_i(x)$ 和 $P_i(y)$ 分别特征点 i 的 x 和 y 的坐标值, $J(P)$ 为右眼的宽和高的比例。根据眼睛的宽和高的比例,就可以确定当前眼睛是处于睁开状态还是闭眼状态,同理可以估计左眼和嘴巴的状态。

对于张闭嘴和眨眼动作,它们是一个序列,张闭嘴动作表示从开始的闭嘴状态转换到张嘴状态再到闭嘴状态来完成整个动作,同理眨眼动作也是由睁眼状态转换到闭眼状态再转换到睁眼状态。通过有限状态来完成张闭嘴和眨眼动作的检测。以张闭嘴动作为例,首先算法进入初始状态(0),如果当前为张嘴状态则继续在原状态循环,如果检测到闭嘴状态,则转入到闭嘴状态(1);在闭嘴状态(1)时,如果检测嘴巴是闭合,则继续在原状态,如果检测到嘴巴张开,则跳转到张嘴状态(2);在张嘴状态(2)时,如果检测到嘴巴还张着,则继续在原状态,如果检测到嘴巴闭上,则跳转到张嘴动作完成状态(3),此时表示判断用户

完成了张闭嘴的整个动作,返回动作成功的结果,如图 2 所示。

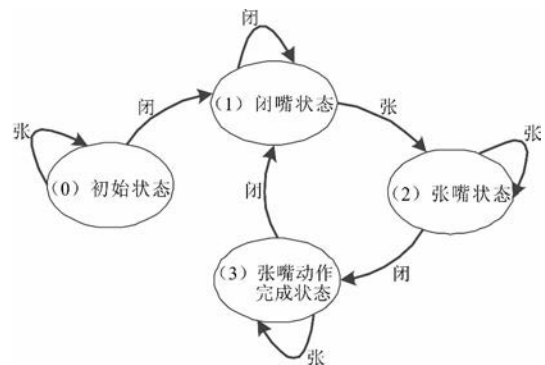


图 2 张闭嘴动作检测的有限状态机

3 基于交互式随机动作序列的活体检测

识别视频中的人脸动作后,就可以根据需要发出指令让用户来完成相应的动作。本文设计了“张闭嘴动作”、“左眼睁开右眼闭上”和“右眼睁开左眼闭上”3 种动作,由于眨眼动作是无意识的,让用户强行控制会不自然,所以没有采用。在活体认证时,计算机检测到人脸后,并进行人脸对齐,这样可以根据特征点位置来分析人脸的动作。然后计算机从上述的 3 个动作中随机选取一个,要求用户在 5 秒钟内完成,计算机根据特征点位置来分析用户是否完成动作。交互式动作认证重复 8 次,计算机根据这 8 次的识别结果判断用户是否按要求完成动作,以确定摄像机中的用户是否是活体。在整个认证过程中,计算机对用户人脸进行跟踪,以防止用户认证过程中切换视频,如果计算机发现跟踪中断,认证过程中止。

由于每一次的动作是从 3 个动作中随机选取,那么 8 次认证所产生的可能总数为 $= 6561$ 种,所以使用图片或者预先录制的视频通过认证的概率是非常低的。在实际的活体认证应用场景中,一般会限制用户进行认证的次数,对于异常的认证结果会引起人工核查,所以在实际的使用中通过图片或者预先录制的视频通过认证几乎是不可能的。活体认证系统采用 Visual C++ 实现,在主流的计算机上运行速度可以达到实时。

4 结语

本文提出了一种基于交互式随机动作序列的活体检测方法,通过视频人脸检测和人脸对齐算法定位人脸特征点,根据人脸特征点位置估计人脸动作。设计了 3 种随机动作进行交互式认证,根据概率认证了该方法可以很好地防止照片和视频对活体认证攻击。

参考文献:

- [1] 杨健伟. 面向人脸识别的人脸活体检测方法研究. [D]. 北京: 北京邮电大学, 2014.

核技术专业计算程序二次开发技术研究

王思力

(中国原子能科学研究院, 北京 102413)

摘要:结合核技术应用专业技术人员在科研生产中遇到的问题,针对专业计算程序二次开发中的典型问题给出了具体的技术解决方案。以放射性后果分析程序和射线屏蔽计算程序为例,分别采用 PAVAN、PAVDOS 编制了放射性事故后果分析计算程序 PandP,采用 LabWindows/CVI、VC++ 设计实现了加速器屏蔽计算程序 AcDose,对专业计算程序的二次开发进行了详细分析。

关键词:核技术应用;二次开发;FORTRAN;VC;LabWindows/CVI

DOI:10.11907/rjdk.1511375

中图分类号:TP303

文献标识码:A

文章编号:1672-7800(2015)012-0028-03

1 核技术专业程序应用中存在的问题

核技术应用领域涉及大量专业计算,科研人员进行分析计算时需要专业的程序。然而,实际应用中这些程序存在以下问题:

(1)多数采用 FORTRAN 语言开发,源代码不宜阅读。程序开发人员在编制程序时,注释的内容比较少,同时程序中经常会用到经验公式,很多参数采用了 FORTRAN 程序的默认变量声明,造成源程序非常难懂^[1]。

(2)计算程序没有操作界面,输入输出都是基于文本文件,采用“卡片”(就是一行有特定要求和意义的字符串)式操作,需要对计算结果进行二次处理,过程繁琐。“卡片”编写较慢、不直观,经常需要查阅手册;如果分析对象太复杂,输入卡就会非常复杂,容易出错;输出结果中包含多种信息,分析结果提取比较繁琐。

(3)专业计算程序基本上都是由核工程专业技术人员编制和修改,他们既是开发者又是使用者。专业计算程序用户集中且总数较少,他们关注的重点不是程序的实现形式,而是程序的算法和数学模型,导致这些计算程序在数据录入和结果输出方面功能比较单一。

(4)大部分专业程序没有源代码,只有可执行程序,不能直接在源代码的基础上进行修改,虽然程序比较权威,但功能单一,需要进行大量的假设和后续数据处理。

2 技术解决方案

在核技术科研生产中,经常会出现一项任务涉及多个计算任务,每一种计算任务又从属于不同的研究方向。如加速器设计过程中,就会涉及到热工(冷却系统)、力学(结构设计)、辐射防护(屏蔽计算)、加速器物理(磁场计算)、自动控制(PLC编程)等。尽管有专业划分和多人协作,

- [2] 罗浩. 人脸识别中的活体检测方法研究. [D]. 长沙:湖南师范大学, 2015.
- [3] 刘华成. 人脸活体检测关键技术研究. [D]. 宁波:宁波大学, 2014.
- [4] PINTO A, ROBSON SCHWARTZ W, PEDRINI H, et al. Using visual rhythms for detecting video-based facial spoof attacks[J]. Information Forensics and Security, 2015, 10(5):1025-1038.
- [5] X TAN, Y LI, J LIU, L JIANG. Face liveness detection from a single image with sparse low rank bilinear discriminative model[C]. Proc. 11th Eur. Conf. Comput. Vis. 2010.
- [6] W R SCHWARTZ, A ROCHA, H PEDRINI. Face spoofing detection through partial least squares and low-level descriptors. [C] Proc. Int. Joint Conf. Biometrics, 2011.

- [7] XIONGET F, DE LA TORRE. Supervised descent method and its applications to face alignment[J]. CVPR, 2013.
- [8] DAVID G LOWE. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004(2): 91-110
- [9] NAVNEET DALAL, BILL TRIGGS. Histograms of oriented gradients for human detection[C]. International Conference on Computer Vision & Pattern Recognition, 2005.
- [10] JAIN V, LEARNED-MILLER, E FDD. A benchmark for face detection in unconstrained settings[M]. Tech. Rep. UM-CS-2010-009, University of Massachusetts, Amherst, 2010.

(责任编辑:陈福时)

作者简介:王思力(1981—),女,黑龙江佳木斯人,硕士,中国原子能科学研究院工程师,研究方向为核技术应用。