

# 基于局部异常行为检测的欺骗识别研究

夏凡,王宏

(清华大学 智能技术与系统国家重点实验室,北京 100084)

**摘要:**基于局部异常行为检测的欺骗识别研究是计算机视觉与情感计算领域中的一个新兴课题,其核心是利用计算机视觉技术从视频中检测、跟踪人的局部动作并进行分析,得出检测对象欺骗的可能性.它在安全监控、辅助面试、情感合成等领域均有着广阔的应用前景.从视觉技术的发展水平和常用处理方法入手,对欺骗识别研究的几个典型问题的研究现状进行了综述,并归纳了欺骗识别研究所涉及的相关理论与技术方案,最后结合课题研究给出了该研究的难点、急需解决的若干重要问题及未来的发展趋势.

**关键词:**情感计算;欺骗识别;计算机视觉;智能系统

**中图分类号:** TP391.4 **文献标识码:** A **文章编号:** 1673-4785(2007)05-0012-08

## Methodologies for deception detection based on abnormal behavior

XIA Fan, WANG Hong

(State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China)

**Abstract:** Research on deception detection based on visual cues for abnormal behavior is a new project in the area of computer vision and affective computing research. This research aims at detecting and tracking an individual's behavior via computer vision technology and analyzing the probability of it being deceptive. This has many promising applications in areas such as security surveillance, automated interview assistants, emotion synthesis, etc. This paper primarily focuses on overall methods and general characteristics of computer vision related to the detection process. A discussion on research challenges and future directions is also provided at the end of the paper.

**Keywords:** affective computing; deception detection; computer vision; intelligent systems

欺骗检测(deception detection)是指通过记录并分析观察对象的各种状态数据,判断出其欺骗的可能性.传统的欺骗检测,如安全部门使用的测谎仪,主要是基于生理数据的检测,因为人在欺骗状态下会在心理和生理上产生异常反应,包括呼吸变化、心跳加快、血压升高、皮肤出汗等,由于这些反应都直接受控于植物神经系统,一般不受人的意识控制,因此无论被测者如何回答或缄默,仪器都将如实记录下被测试人的血压、脉搏、呼吸、皮电阻等生理参数的变化,并作为判定被测者是否有欺骗嫌疑的重要依据<sup>[1]</sup>.

但科学界人士普遍认为传统的测谎仪的可信度有限,部分原因是因为其有效性与询问者的询问技

巧密切相关.传统测谎仪所测试的是人撒谎时的压力值,通过心率加快、呼吸加速、血压升高、汗液增多等表征来判断被测者是否说谎,而那些对自己的罪行并不感到内疚的人以及那些能够在压力下克制反应的人可以轻易地打败测谎仪<sup>[2]</sup>.

近年来,计算机视觉技术得到了迅速的发展,运动物体的视觉分析技术<sup>[3]</sup>也有了长足的进步,并已逐渐应用到如手势识别<sup>[4]</sup>、道路监控<sup>[5]</sup>等相关领域,因此如何将视觉技术应用于欺骗检测已成为视觉领域的一个具有挑战性的新课题.已有大量研究证明,对检测对象肢体动作的观察是欺骗检测中一种有效的方法,因为在撒谎和诚实的情况下人的表现通常是不一样的<sup>[6-8]</sup>.这些肢体动作的不同包括头部运动的变化<sup>[9]</sup>,说话时阐述性动作的减少<sup>[10]</sup>等.很多人都会认同欺骗者的外在动作表现会和诚实的人有

收稿日期:2006-12-18.

基金项目:国家自然科学基金资助项目(60433030).

所不同,但他们很难确定具体哪些动作是和欺骗真正关联的,就连经过训练的安全专家也会被一些线索误导<sup>[11]</sup>. 基于局部异常行为检测的欺骗识别的目的就是利用计算机视觉技术,通过记录并分析被测者的肢体活动来判断检测对象是否处于欺骗状态.

## 1 相关理论与主要研究内容

与欺骗相关联的一些行为动作通常可分为 2 类:过度控制和不安<sup>[12]</sup>. 不安通常表现为紧张或害怕,包括语速变快、声音变大以及烦躁等. 但是不安与欺骗之间的联系仍是一个很有争议的话题,大量的实验数据表明欺骗和不由自主的不安之间确实有重要的联系<sup>[13-14]</sup>. 另一方面,在很多情况下说谎者很清楚那些会让他露出破绽的外在行为表现,因而主动的控制自己,表现出镇静的一面,这时候就有可能出现过度控制.

有 2 种广泛提及的理论对通过异常行为检测进行欺骗识别的研究具有指导意义. 这 2 种理论分别是人际欺骗理论 (interpersonal deception theory)<sup>[15]</sup>和期望背离理论 (expectancy violations theory)<sup>[16]</sup>. Buller 以及 Burgoon 的人际欺骗理论指出欺骗是一种动态的过程,欺骗是根据信息接收者对信息的信任或怀疑来调整发送的信息,以获得信息接收者的信任的过程. 仅仅通过测试过程中被试处于紧张或过度控制是无法判断被试是否处于欺骗状态. 期望背离理论指出,可以根据被试在正常情况下的状态对其在不安及过度控制时的状态进行估计,如果被试的实际表现与期望的表现过分背离,则认为被试处于欺骗状态中或者欺骗的可能性很大.

人类的姿势可以被划分为 3 种范畴:表征 (emblems)、阐述 (illustrators) 和表演 (adaptors). 表征是在同一个文化团体中有确定意义的手的姿势或肢体动作,例如点头代表同意或确定. 阐述是与说话内容相协调的姿势,此类行为能够强调正在说明的话题或者表征一个复杂的空间概念,如果脱离说话内容则一般无具体含义,譬如说口头描述方向的同时做右转弯的手势. 表演是为完成某一个物理的或心理的需要而表现出来的姿势,譬如说擦拭、赞扬、捡东西和坐立不安等.

已进行的研究显示撒谎者和说真话者姿势有所不同. 人们普遍认同的是撒谎者其坐立不安或者踢脚的表演动作会显著增加. 不过许多的研究也表明表演动作和欺骗之间通常没有必然的联系. 然而阐述动作的减少却和欺骗有着密切的联系. 据显示心

理负担的增加可能是阐述动作减少的一个原因,撒谎者可能更关注如何编织一个真实的谎言而忽略了阐述动作的表述.

基于以上理论,研究者们进行了一系列的验证性实验,包括模拟偷窃和机场测试<sup>[17-18]</sup>. 前者的目的是描述能被用在欺骗检测中的表征. 在这个实验中,一些参与者饰演小偷的角色而其他人在偷窃事件时均在场,所有的参与者都分别接受没有训练过和训练过的面试官连续的面试,并记录下视频、声音和一些其他信息;而机场测试实验则是雇佣 4 个演员来帮助进行概率证明的研究,以便决定通过姿势和身体动作来区分行为状态的可行性. 他们参与进入一个模拟机场屏幕检测的场景,这个场景包括坐在椅子上、站立、排队以及运动等场面,在每个场面中,每个演员都被要求演示 3 种状态:放松、激动以及过度控制. 在这些实验数据的基础上,美国 Arizona、Rutgers 等大学的研究者们应用视觉技术,对被测者的肢体运动及手的姿势变化进行了检测和异常行为分析,并由此得出了被测者的欺骗可能性<sup>[19-22]</sup>.

Shan<sup>[22]</sup>等人进行的局部区域 (blob) 分析与欺骗检测的研究是一个较为深入的工作. 他们首先建立一个头部和手部的训练集,得到一个肤色的分布直方图并通过阈值判断某像素点是否为肤色点. 这种方法方便快捷,但容易受到光照及位姿变化的影响,而且不同的人种对应的肤色分布相差很大. Shan 等采取了对每个人种都训练一组直方图,并用 RGB 归一化的方法,在一定程度上解决了光照和位姿变化的问题. 然后用粗搜索细分类的思想对划分出的肤色区域进行手部和头部的查找,即首先用椭圆匹配的方法找到 blob 的候选区域,其次用 KL 变换进行特征匹配,找到正确的手部和头部的位置,再用 Kalman 滤波的方法进行跟踪. 在跟踪的过程中记录各个 blob 的速度、移动频率及相互间的距离,最后形式化为被试状态与上述特征的关系表达式,达到欺骗检测的目的. 在 2 个场景 5 个片断的测试中正确率为 100 %.

Enrica<sup>[23]</sup>则针对手指运动这样的局部细微特征进行了分析. 他首先利用训练集得到肤色分布直方图,并用贝叶斯分类器进行手部区域的查找,然后用一种特征跟踪的方法对手部进行跟踪. 具体做法是确定 2 个手部区域的质心,并根据该区域是处于运动还是静止标上标签,再用贝叶斯网络进行预测跟踪,达到了较好的效果. 随后 Enrica 对手部区域用

一族特殊的小波函数进行分解,用分解系数表示手指的相位,也取得了不错的效果.如果能够将正常与异常状态下手指的相位变化规律进行对比分析可能会得出更好的结果.

上述方法都是基于普通摄像机的视频数据. Buddharaju<sup>[24]</sup>等人则利用红外摄像机的数据进行了研究.目前文献中提到的只是利用伪彩色技术对红外视频进行增强,可以直观显示局部皮肤的温度变化,以及利用椭圆匹配查找头部及手部,为进一步的分析打下基础.

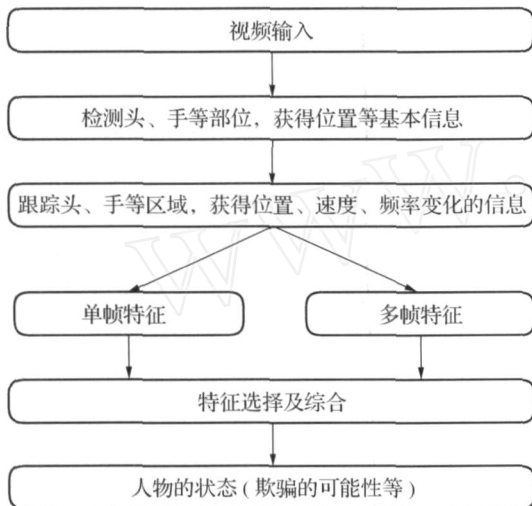


图1 主要处理流程

Fig. 1 Main process flow

## 2 主要技术方案比较分析

利用计算机视觉技术进行局部异常行为检测的欺骗识别的研究,目前主要集中在国外的一些大学和研究机构,国内这方面目前几乎没有可供查阅的文献.基于视觉技术的欺骗检测通常遵循如下处理流程:首先利用肤色检测技术从视频中分离出皮肤区域,再用特征匹配等方法检测出感兴趣的身体部位(如头部、手部等),并对其进行跟踪,记录这些身体区域的运动特征(如速度、频率、相互间的距离等),最后再从中抽取出能够代表与欺骗相关的异常的特征并进行形式化表达,建立异常特征与欺骗状态间的联系.下面将就以上所述处理流程的各步骤,对计算机视觉所常用的一些技术进行归纳和比较.

### 2.1 人体肤色区域分割

人体肤色区域分割主要分为2种类型:一种是像素级的分割,另一种是区域级的分割.

像素级的人体肤色区域的分割主要涉及到颜色空间变换和肤色建模.颜色空间有色度学的CIE -

RGB 真实三原色系统和CIE - XYZ 虚拟三原色系统等<sup>[25]</sup>,另外还有一些常用于肤色检测的颜色空间如YUV<sup>[26]</sup>和LUX<sup>[27]</sup>等,关于颜色空间更为详尽的分析可以参考文献[28].目前常用的静态肤色建模基本方法有3种:阈值化、参数化和非参数化方法.

阈值化方法即用数学表达式明确规定肤色范围,是一种简单的肤色建模方法,例如在YUV空间中根据Y范围和UV平面中的色度范围划定肤色范围,如果输入像素的颜色落入限定的区域,就认为是属于肤色像素.这种简单明确的判断运行起来快速高效,特别是在实时系统中更具有吸引力.规定肤色范围虽然快捷,但要取得好的效果需要解决2个问题:一是如何选择合适的颜色空间,二是如何确定规则中的参数.

参数化方法是用函数衡量输入像素与肤色的相似度.常用的模型函数有高斯密度函数,可以用单峰高斯模型SGM或高斯混合模型GMMs<sup>[29]</sup>.单峰高斯模型是用椭圆高斯联合概率密度函数来衡量某一像素属于肤色的可能性,也可以通过高斯分布参数计算输入像素颜色与像素均值的马氏距离,来表示该像素与肤色模型的接近程度.高斯混合模型通常对肤色和非肤色分别用一个高斯混合模型表示,用标准的期望最大化EM算法训练这2个高斯混合模型.高斯肤色模型通常选用二维色度平面,在色度平面上的肤色分布是否符合高斯分布,决定了本建模方法的好坏.

非参数化方法最常用的模型是直方图模型.利用肤色样本的直方图统计训练出肤色概率图SPM,在离散化的颜色空间中查找输入像素对应的概率值,判断其为肤色的可能性.利用SPM检测肤色像素主要有2种方法:规则化查找表和贝叶斯分类器<sup>[30]</sup>.贝叶斯分类器学习正反样本,很可能得出较好的分类器.规则化查找表只学习肤色样本,但需足够数量的有代表性的样本以及适当的量化级数.关于像素级肤色分割的更详细的讨论可以参考文献[31].

区域级的人体肤色分割以像素级肤色区域分割为基础,再加上区域、边缘或者纹理信息进行进一步准确的分割. Son Lam Phung<sup>[32]</sup>在他的工作中首先利用简单的阈值化方法得到候选的肤色区域,再用Sobel或者Canny算子进行边缘提取(Sobel算子对于不同质的颜色区域之间的边缘有很好的提取效果,而Canny算子则对同质颜色区域的内部的细小边缘比较敏感),最后利用边缘信息在各候选区域内

容对肤色分布进行判断得出实际皮肤区域. Solar<sup>[33]</sup> 先在 RGB 空间用高斯混合密度函数做基于像素的检测,然后用可信度高的像素作为种子进行“散射”处理,其他像素要确认是否肤色必须在其某邻域内有确认过的肤色像素. Chen Yu<sup>[34]</sup> 在 LUV 颜色空间中利用高斯混合模型分别训练出皮肤颜色/非皮肤颜色、皮肤纹理/非皮肤纹理分布的概率密度函数模型,再对输入的像素对颜色、纹理 2 个条件分别进行判断,均满足时判定为肤色像素.

2.2 肤色区域的跟踪

跟踪等价于在连续的图像帧间建立基于位置、速度、形状、纹理、色彩等有关特征的对应匹配问题,常用的数学工具有 Kalman 滤波<sup>[35]</sup>、条件概率密度传播<sup>[36]</sup>及动态贝叶斯网络<sup>[37]</sup>等. 其中 Kalman 滤波是基于高斯分布的状态预测方法,能够最优化自回归数据处理,不能有效地处理多峰模式的分布情况;条件概率密度传播算法是以因子抽样为基础的条件密度传播方法,结合可学习的动态模型,可完成鲁棒的运动跟踪.

按照跟踪方法来划分,主要有如下 4 类<sup>[38]</sup>.

1) 基于模型的跟踪. 将人的运动看成是骨骼的运动,用线图法、二维轮廓或立体模型等几何模型来近似身体的各个部分,该方法要求参数更新和匹配过程中更大的计算量.

2) 基于区域的跟踪. 将人体看作由头、躯干、四肢等身体部分所对应的局部区域所组成,利用高斯分布建立人体和场景的模型,通过跟踪各个局部区域来完成整个人的跟踪. 基于区域的跟踪方法目前已有较多的应用,例如利用区域特征进行室内单人的跟踪. 基于区域跟踪的难点是处理运动目标的阴影和遮挡,如能结合纹理、彩色及形状等特征进行跟踪将会有助于上述问题的解决.

3) 基于活动轮廓的跟踪. 基于活动轮廓的跟踪<sup>[39]</sup>思想是利用封闭的曲线轮廓来表达运动目标,并且该轮廓能够自动连续地更新. 相对于基于区域的跟踪方法,轮廓表达降低了计算复杂度,如果初始化时能够合理地区分每个运动目标并实现轮廓初始化,既使在部分遮挡存在的情况下也能连续地进行跟踪. 然而初始化通常是很困难的.

4) 基于特征的跟踪. 基于特征的跟踪包括特征提取和特征匹配 2 个过程. Shan 的工作<sup>[22]</sup>就是一个很好的区域特征跟踪的例子,文中将检测对象的头部和双手各用一个矩形框加以标识,将运动和静止的矩形框分别加上标签,作为跟踪的特征.

2.3 行为特征与欺骗状态之间关系的形式化表达

基于视觉技术进行欺骗检测的最终目的是要将异常的行为与欺骗状态进行关联,简单说来就是把检测对象的行为模式与欺骗状态之间的关系进行分析和编码,并用自然语言等进行描述. 将检测对象的行为特征与欺骗状态之间的关系进行形式化表达,可以简单地认为是时变数据的匹配分类问题,即将测试序列与预先标定的代表典型欺骗状态下行为的参考序列进行匹配. 由此可见,关键问题是如何从学习样本中获取参考序列,并且学习和匹配的行为序列能够处理在相似的运动模式下空间和时间尺度上轻微的特征变化.

匹配时变数据的技术通常有:

1) 动态时间规整 DTW<sup>[40]</sup>. DTW 具有概念简单、算法鲁棒的优点,早期被应用于语音识别中,并且最近才被用于匹配人的运动模式<sup>[41-42]</sup>;对 DTW 而言,即使测试序列模式与参考序列模式的时间尺度不完全一致,只要时间次序约束存在,它仍能较好地完成测试序列和参考序列之间的模式匹配.

2) 隐马尔可夫模型 HMMs<sup>[43]</sup>. HMMs 是更加成熟的匹配时变数据的技术,它是随机状态机. HMMs 的使用涉及到训练和分类 2 个阶段,训练阶段包括指定一个隐马尔可夫模型的隐藏状态数,并优化相应的状态转换和输出概率,便于输出符号与特定行为模式下的特征相匹配. 对于每一个行为模式,一个 HMM 是必须的. 匹配阶段涉及到一个特定的 HMM,产生相应于所观察行为特征的测试符号序列的概率计算. HMMs 在学习能力和处理未分割的连续数据流方面比 DTW 有更好的优越性,当前被广泛应用于人的运动模式匹配中<sup>[44-45]</sup>.

3) 神经网络. 同样也是目前较好的匹配时变数据的方法,例如 Rosenblum 等<sup>[46]</sup>使用径向基函数网络从运动中识别人物的情感.

表 1 各类研究方法比较  
Table 1 The comparison of methods

	肤色分离	跟踪方法	行为表达
Gabriel <sup>[21]</sup>	颜色表查找	Kalman 滤波 + 形状信息	HMMs
Shan Lu <sup>[22]</sup>	颜色表查找	Kalman 滤波 + 形状信息	自定义模型
Enrica <sup>[23]</sup>	贝叶斯分类	Kalman 滤波 + 区域质心	Wavelet

此外 Kojima 等<sup>[47]</sup>提出了一种新的方法,选择一组运动词语或短句来报告场景中检测目标的行为,用于产生实时视频图像中人的行为的自然语言描述:首先,从每帧图像中提取出代表整个人体的头部区域,并利用基于模型的方法,估计其三维姿势和位置;然后,这些参数轨迹被分解为单一运动的基元,评估每个基元的诸如姿势和位置的变化度、与环境中其他一些目标的相对距离等概念特征,并且选择最合适的词语和其他语义元素;最后,利用机器翻译技术,产生检测对象行为描述的自然语言文本。

### 3 欺骗识别的应用

#### 3.1 辅助审讯

在刑事调查中刑侦人员经常要面对一些老奸巨滑的犯罪嫌疑人,传统的利用生理数据进行辅助测谎的设备会令嫌疑人先警觉并有心理防备,使得测试结果大打折扣。利用视觉技术的异常行为欺骗检测可以在嫌疑人不知情的情况下进行,进而分析出那些具有潜在可疑性的对象供进一步调查,给审讯带来便利。

此外,每天都有数以万计的游客从机场或是海关出入境,在当今国际恐怖犯罪不断加剧的形势下,如何在出入境处阻止这些危害极大的危险分子已成为安全专家们的一个重要任务。通常他们不得不从监控录像或是其他途径选择可疑的人进行检查,在每次的检查过程中专家们必须判断被测对象是否在诚实地回答问题,这会是一项非常艰难的工作。而基于视觉技术的欺骗检测可以从海量的视频数据中,提取出具有一定欺骗可能性的潜在嫌疑人的录像片断,这将会给专家们提供非常有益的帮助。

#### 3.2 机要部门人员的筛选

冷战以来,间谍人员的渗透与反渗透已经成为国家安全和机要部门一个非常重要的课题,对于新工作人员招募时的测试和内部工作人员的定期检查也逐渐制度化,测谎也成为必不可少的一环。以美国为例,其国家安全局、中央情报局(CIA)、联邦调查局(FBI)在招募新雇员时都要进行测谎检查。但利用传统的生理仪器进行这一工作时,往往会在不同的测试中得出完全相悖的结果。其原因是身体上佩戴过多的传感器会对被测者造成额外的心理压力,影响结果的准确性。而基于视觉技术的欺骗检测可以达到遥测的目的,在一定程度上缓解被试者的压力,有助于取得更准确的结果。

#### 3.3 自动柜员机的安全监控

随着银行卡应用的逐步普及,自助设备也被人

们广泛接受。犯罪分子随之将黑手伸向了这一领域,伪装吞卡、仿制假卡、窥探密码、封堵出钞口等作案手段层出不穷,令银行防不胜防。据 Gartner 研究公司所做的最新报告,在 2005 年 12 月,美国网络用户共损失 27.5 亿美元,这些钱都与自动提款机和银行借记卡有关。防范来自 ATM 的犯罪不仅关系到银行的利益和信誉,也直接关乎人民群众的生活和社会的安定。基于计算机视觉的欺骗检测技术可以给自动柜员机增加一道安全保护层,在预防和检测相关犯罪活动方面提供有力的工具。

在 2006 中国国际金融(银行)技术暨设备展览会上,已经有自动柜员机生产公司提出要在柜员机上配置监控摄像头,提高自动柜员机的安全性。如果将基于计算机视觉的欺骗检测技术应用到自动柜员机上,可以分析出当前操作者是否具有明显异常的欺骗行为,进而采取措施给出警告或加以阻止。

#### 3.4 辅助情感合成研究

目前人工合成情感的研究主要集中在语音合成、面部表情合成方面,是通过音调、音色、说话的频率及节奏等合成具有情感表现力的语音;面部表情的合成则是通过人脸肌肉模型,利用嘴唇、眼睛、眉毛的变化来表现。如何利用人的肢体动作来合成并表现情感,目前所涉及的文献甚少。在基于视觉技术的欺骗检测的研究过程中,可以挖掘人的肢体动作和情感表达之间的关系,进而应用到情感合成中来。

#### 3.5 人的异常行为的辅助检测

当前的视觉技术领域,异常检测已经成为热门研究课题之一,而人的异常行为检测也成为其中重要的组成部分。Boiman<sup>[48]</sup>在他的论文中提出的方法,能够在没有训练集的前提下有效的从视频数据中分析出人的异常行为。这些研究通常是将人体看作一个整体来进行异常检测,而基于局部异常行为的欺骗检测的研究,对于其中的肢体行为的细节与异常状态之间联系的分析,无疑是对异常行为检测的一个有益补充。

## 4 研究难点及发展方向

前面介绍了近年来基于视觉技术的欺骗检测的研究概况和主要的技术路线,从这些内容可以看出,该研究目前还处于起步阶段,尚有许多难点和问题需要研究者们共同探索。

#### 4.1 获取具备真实性的数据

基于局部异常行为的欺骗检测最大的困难在于训练数据集的获取。完好的数据集应是被试者真实

的欺骗过程的记录,但在实验中很难事先知道被试在接下来的过程中是否有欺骗行为,如果能事先知道的话,也就不能称之为欺骗了。

一种做法是表演欺骗,例如在 Arizona 大学的一个实验中<sup>[17]</sup>,就是用演员饰演小偷以及偷窃现场的其他角色。尽管这只是由演员进行表演,但考虑到平常人对影视作品中演员表演出的情感,与现实中出现的相同情感,二者之间并不能从人物的非言语行为上进行区分,因此这种方法仍然具有一定的真实性。第2种做法是设计巧妙的实验,控制被试在实验中的行为,使其达到自然欺骗的状态。作者曾经利用“杀人游戏”<sup>[49]</sup>设计过一个该类型的实验。在这个游戏中,“杀手”必须欺骗其他的游戏者才能达到最终获胜的目的,因此欺骗的真实性可以保证。由于游戏不会在现实世界中造成严重的后果,对欺骗者的心理压力较小,效果并不十分明显。Gabriel<sup>[21]</sup>在实验中使用了真实的犯罪嫌疑人(这些人最后都被确认有罪)的标准的测谎仪采访的视频磁带,由这些视频分析所得的实验结果显然更具有说服力。

#### 4.2 数据挖掘算法的运用

数据挖掘是从大量的数据中抽取潜在的、不为人知的有用信息、模式和趋势。在欺骗识别的检测过程中,会产生大量的数据,这正是数据挖掘技术适用的场合,如何合理运用各种数据挖掘的技术,从海量的数据中提取出与被检测对象心理状态相关的信息,是值得大家去研究的。

在欺骗识别的一般检测过程中,会对人体的各个部分进行跟踪检测,得到各种帧内和帧间的检测结果数据,如头部、手部的位置、大小,相互间的距离,运动的速度,双手交叉的频率,头部摆动的频率等等,这些数据通常被表达成高维的向量,如果合理的运用数据挖掘的算法和技术,是可以从中获取大量信息的。作者目前正在进行一些相关研究工作,运用数据挖掘的相关算法将被检测人的运动状态进行分类,并从中获得与其心理状态相关的数据,以期达到欺骗识别的目的。

#### 4.3 分析动作细节

在通向应用的过程中,基于视觉技术的欺骗检测最基本的要求是能够处理实际情况下的欺骗状况。以辅助审讯为例,审讯官需要面对的往往都是那些既十分老练,又善于在压力下克制自我反应的人。这些人在进行欺骗时,很少会有异常的不安或是过度控制的表现,而现实中审讯官也是从他们在面对压力问题时的一些细微反应观察到异常,这些细微

反应通常包括手指和嘴唇的微动、眼球速率变快等。以现有的研究方法,希望从他们的肢体动作的速度频率的变化上找出欺骗的痕迹是非常困难的。

Enrica<sup>[23]</sup>在他的研究中采用了一种不同的方法,他将注意力集中到考查对象的手部,通过皮肤分割技术分离出手部区域,再用一族特殊的小波函数来表达手指的重叠状态和方向,能够反映出手指一些突然的细小抖动。

#### 4.4 性能评估方法的建立

一般说来,鲁棒性、实时性和准确性是衡量一个欺骗检测系统的主要标准。由于基于局部异常行为的欺骗检测技术应用于诸如刑侦审讯之类的场合时,它的任务是协助专业人士的工作,并要求能够连续的工作,从海量的视频数据中抽取可疑片段供专家进一步分析,因此,鲁棒性必不可少。实时性的要求对于自动柜员机监控这类需要实时分析并响应的系统是最基本的要求。为此,如何选择有效的技术方案来提高系统性能、降低计算代价也是个特别值得考虑的问题。至于准确性的衡量则相对复杂,何种状态应该被定义为具有欺骗可能性,是仅仅从肢体的运动上就能判断,还是需要借助其他衡量工具,则是一个值得深入研究的问题。

### 5 结束语

基于局部异常行为的欺骗检测已经成为计算机视觉中一个重要的研究课题,它在刑侦审讯、智能监控等方面的应用前景引起了科研人员的浓厚兴趣和关注,特别是视觉技术在处理人的肢体运动方面的发展,更促使许多研究者将其推向更加智能的应用场合。本文从基于视觉技术的欺骗检测的原理和一般处理流程出发,归纳了常用的技术方法和路线,同时阐明了研究难点和未来的发展趋势,以期能对相关领域深入研究提供参考和借鉴。

### 参考文献:

- [1] GU GAS C. The silent witness [M]. Prentice-Hall, 1979.
- [2] MASCHKE GW, SCALABRINI GJ. The lie behind the lie detector[M]. 4th ed. AntiPolygraph.org, 2005.
- [3] GAVRILA D. The visual analysis of human movement: A survey[J]. Computer Vision and Image Understanding, 1999, 73(1):82-98.
- [4] FREEMAN W, WEISSMAN C. Television control by hand gestures [A]. Proc International Conference on Automatic Face and Gesture Recognition[C]. Zurich,

- Switzerland, 1995.
- [5] REMAGNINO P, TAN T, BAKER K. Multi-Agent visual surveillance of dynamic scenes[J]. Image and Vision Computing, 1998, 16(8): 529 - 532.
- [6] DEPAULO B, LINDSA Y J, MALONE B, et al. Cues to deception[J]. Psychological Bulletin, 2003, 129: 74 - 118.
- [7] EKMAN P, FRIESEN W V. Nonverbal leakage and clues to deception[J]. Psychiatry, 1969, 32: 88 - 106.
- [8] ZUCKERMAN M, DRIVER R. Telling lies: verbal and nonverbal correlates of deception[A]. Nonverbal Communication: An Integrated Perspective [C]. Hillsdale, NJ: Erlbaum, 1985.
- [9] BULLER D B, BURGOON J K, WHITE C, et al. Interpersonal deception: VII. behavioral profiles of falsification, equivocation and concealment [J]. Journal of Language and Social Psychology, 1994, 13: 366 - 395.
- [10] VRIJ A, EDWARD K, ROBERTS K, et al. Detecting deceit via analysis of verbal and nonverbal behavior[J]. Journal of Nonverbal Behavior, 2000, 24: 239 - 263.
- [11] VRIJ A. Detecting lies and deceit: the psychology of lying and its implications for professional practice [M]. Chichester, UK: John Wiley, 2000.
- [12] EKMAN P. Telling lies: Clues to deceit in the marketplace, politics, and marriage [M]. New York: WW Norton and Company, 2001.
- [13] LINDSA Y J, MALONE B, MUHLENBRUCK L, et al. Cues to deception[J]. Psychological Bulletin, 2003, 129: 74 - 118.
- [14] ZUCKERMAN M, DEPAULO B, ROSENTHAL R. Verbal and nonverbal communication of deception[J]. Advances in Experimental Social Psychology, 1981, 14: 1 - 59.
- [15] BULLER D, BURGOON J. Interpersonal deception theory[J]. Communication Theory, 1996, 6: 203 - 242.
- [16] GEORGE J, BIROS D P, BURGOON J K, et al. Training professionals to detect deception [A]. NSF/NIJ Symposium on Intelligence and Security Informatics [C]. Tucson, AZ, 2003.
- [17] MATTHEW L J, THOMAS O M, JOHN K, et al. Identification of deceptive behavioral cues extracted from video[A]. Proc IEEE International Conference on Intelligent Transportation Systems [C]. Vienna, Austria, 2005.
- [18] BURGOON J K, BLAIR J P, MOYER E. Effects of communication modality on arousal, cognitive complexity, behavioral control and deception detection during deceptive episodes[A]. Annual Meeting of the National Communication Association [C]. Miami Beach, USA, 2003.
- [19] BURGOON J K, ADKINS J, KRUSE M, et al. An approach for intent identification by building on deception detection[A]. Proc Hawaii International Conference on System Sciences [C]. Hawaii, USA, 2005.
- [20] THOMAS O M, MATTHEW L J. Deception detection through automatic, unobtrusive analysis of nonverbal behavior[J]. IEEE Intelligent Systems, 2005, 20(5): 36 - 43.
- [21] TSECHPENAKIS G, METAXAS D, ADKINS M, et al. HMM - based deception recognition from visual cues [A]. IEEE International Conference on Multimedia and Expo [C]. Amsterdam, Netherlands, 2005.
- [22] SHAN L, TSECHPENAKIS G, METAXAS D N. Blob analysis of the head and hands: a method for deception detection[A]. Proc Hawaii International Conference on System Sciences [C]. Hawaii, USA, 2005.
- [23] DENTE E, NGJ, VRIJ A, et al. Tracking small hand movements in interview situations [J]. IEEE International Symposium on Imaging for Crime Detection and Prevention, 2005, 35(2): 55 - 60.
- [24] BUDDHARAJU P, DOWDALL J, TSAM YRTZIS G, et al. Automatic thermal monitoring system (ATHEMOS) for deception detection[A]. Proc IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. Washington, DC, USA, 2005.
- [25] WYSZECKI G, STILES W. Color science: concepts and methods, quantitative data and formulae [M]. 2nd ed. Chichester: John Wiley & Sons, 2000.
- [26] YAO H, LIU M, GAO W, et al. Method of face locating and tracking based on chromatic coordinates transformation of color images [J]. Chinese J Computers, 2000, 23(2): 158 - 165.
- [27] LIEVIN M, LUTHON F. Nonlinear color space and spatiotemporal MRF for hierarchical segmentation of face features in video [J]. IEEE Transactions on Image Processing, 2004, 13(1): 63 - 67.
- [28] FORD A, ROBERTS A. Colour space conversions [R]. Westminster University, London, 1998.
- [29] HSU L, ABDEL M, JAIN A. Face detection in color images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 696 - 706.
- [30] LEE J Y, YOO S I. An elliptical boundary model for skin color detection [A]. Proc International Conference on Imaging Science, Systems and Technology [C]. Las Vegas, USA, 2002.
- [31] VEZHNEVETS V, SAZONOV V, ANDREEVA A. A survey on pixel - based skin color detection techniques [A]. Proc Graphicon - 2003 [C]. Moscow, Russia, 2003.



- 2003.
- [32] PHUNG S L, BOUZERDOUM A, CHAID. Skin segmentation using color and edge information[A]. Proc Int Symposium on Signal Processing and its Applications [C]. Paris, France, 2003.
- [33] RUIZ DEL S, VERSCHAE R. Skin detection using neighborhood information[A]. Proc IEEE International Conference on Automatic Face and Gesture Recognition [C]. Seoul, Korea, 2004.
- [34] CHEN Y, CHEN F. Skin detection based on color and texture[J]. Computer Aided Engineering, 2006, 15(2):
- [35] WELCH G, BISHOP G. An introduction to the Kalman filter[R]. SIGGRAPH Course Notes, Los Angeles, California, USA, 2001.
- [36] ISARD M, BLAKE A. Condensation-conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1):5-28.
- [37] PAVLOVIC V, REHG J, CHAM T J, et al. A dynamic bayesian network approach to figure tracking using learned dynamic models[A]. Proc IEEE International Conference on Computer Vision [C]. Corfu, Greece, 1999.
- [38] WANG L, HU W, TAN T. A survey of visual analysis of human motion[J]. Chinese Journal of Computers, 2002, 25(3):225-237.
- [39] PARAGIOS N, DERICHE R. Geodesic active contours and level sets for the detection and tracking of moving objects[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2000, 22(3):266-280.
- [40] MYERS C, RABINIER L, ROSENBERG A. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition[J]. IEEE Trans Acoustics, Speech, and Signal Processing, 1980, 28(6):623-635.
- [41] BOBICK A, WILSON A. A state-based technique for the summarization and recognition of gesture[A]. Proc International Conference on Computer Vision[C]. Cambridge, 1995.
- [42] TAKAHASHI K, SEKI S, et al. Recognition of dexterous manipulations from time varying images[A]. Proc IEEE Workshop on Motion of Non-Rigid and Articulated Objects[C]. Austin, 1994.
- [43] PORITZ A. Hidden Markov models: a guided tour [A]. Proc IEEE International Conference on Acoustics, Speech and Signal Processing[C]. New York, 1988.
- [44] RABINIER L. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2):257-285.
- [45] STARNER T, PENTLAND A. Real-time american sign language recognition from video using hidden Markov models[A]. Proc International Symposium on Computer Vision[C]. Coral Gables, Florida, 1995.
- [46] ROSENBLUM M, YACOOB Y, DAVIS L. Human emotion recognition from motion using a radial basis function network architecture[A]. Proc IEEE Workshop on Motion of Non-Rigid and Articulated Objects [C]. Austin, 1994.
- [47] KOJIMA. Generating natural language description of human behaviors from video images[A]. Proc IEEE International Conference on Pattern Recognition[C]. Barcelona, 2000.
- [48] BOIMAN O, IRANI M. Detecting irregularities in images and in video[A]. Proc IEEE International Conf on Computer Vision[C]. Beijing, China, 2005.
- [49] KILLBAR. Rules of mafia game [EB/OL]. [2007-1-10]. <http://www.killbar.com/intro.htm>.

#### 作者简介:



夏 凡,男,1982年生,硕士研究生,主要研究方向为计算机视觉、情感计算。

E-mail: jf@mails.tsinghua.edu.cn.



王 宏,男,1954年生,副教授,主要研究方向为移动机器人、智能感知与人机交互,发表学术论文多篇。