(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110998598 A
(43)申请公布日 2020.04.10

(21)申请号 201880053716.4

(22)申请日 2018.06.29

(30)优先权数据
1710560.2 2017.06.30 GB

(85)PCT国际申请进入国家阶段日
2020.02.18

(86)PCT国际申请的申请数据
PCT/EP2018/067692 2018.06.29

(87)PCT国际申请的公布数据
WO2019/002602 EN 2019.01.03

(71)申请人 挪威科技大学
地址 挪威特隆赫姆

(72)发明人 K·B·拉贾 R·拉玛钱德拉 S·文卡提什 C·布希

(74)专利代理机构 北京信诺创成知识产权代理有限公司 11728
代理人 刘金峰

(51)Int.Cl.
G06K 9/00(2006.01)
G06K 9/46(2006.01)
G06K 9/62(2006.01)

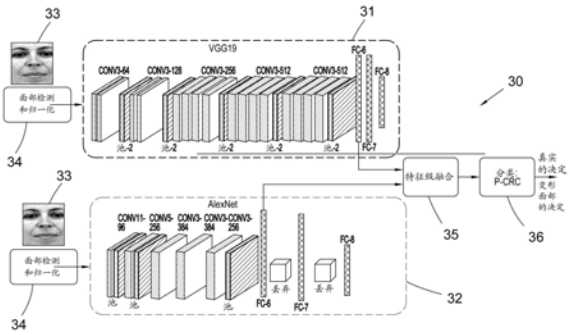
权利要求书3页 说明书16页 附图7页

(54)发明名称

对被操纵图像的检测

(57)摘要

用于检测变形的或平均化的图像的装置(30),其中变形的或平均化的图像是合成生成的图像,包括来自两个或更多个不同的源图像的信息,两个或更多个不同的源图像对应于两个或更多个主体。装置包括:特征提取模块,用于接收输入图像(33)并输出该图像的一组描述符特征特性;分类器模块(36),被配置为,基于描述符特征而将输入图像认定为表明图像已经变形或平均化的第一类型或表明图像没有变形或平均化的第二类型。特征提取模块包括多个神经网络(31、32),神经网络向分类器模块提供互补性的描述符特征。装置进一步包括融合模块(35),用于将来自每个神经网络的描述符特征数据组合并将融合的特征数据发送给分类器模块。分类器模块包括机器学习系统,机器学习系统被训练为使用训练数据集将图像分类,训练数据集包括变形的或平均化的图像和没有变形或平均化的图像。



CN 110998598 A

1.一种用于检测变形的或平均化的面部图像的装置,其中所述变形的或平均化的图像是合成生成的图像,所述合成生成的图像包括来自两个或更多个不同的源图像的信息,所述两个或更多个不同的源图像对应于两个或更多个主体,所述装置包括:

特征提取模块,其用于接收输入图像并输出该图像的一组描述符特征特性;以及

分类器模块,其被配置为,基于所述描述符特征而将所述输入图像认定为第一类型或第二类型,所述第一类型表明图像已经变形或平均化,所述第二类型表明图像没有变形或平均化;

其中所述分类器模块包括机器学习系统,所述机器学习系统被训练为使用训练数据集对单个图像进行分类,所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

2.根据权利要求1所述的装置,其中所述特征提取模块包括机器学习系统,并且所述描述符特征取决于参数,所述参数由对包括图像的训练数据集的使用而确定。

3.根据权利要求2所述的装置,其中所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

4.根据前述任意一项权利要求所述的装置,其中所述变形的或平均化的图像是变形的或平均化的面部图像。

5.根据前述任意一项权利要求所述的装置,进一步包括图像预处理模块,所述图像预处理模块被布置为从所述图像提取并归一化感兴趣的区域(例如,人的面部),并将预处理后的图像发送给所述特征提取模块。

6.根据前述任意一项权利要求所述的装置,其中所述特征提取模块包括一组滤波器,所述滤波器与所述输入图像的斑块进行卷积,以提供一组描述符特征。

7.根据权利要求6所述的装置,其中所述一组描述符特征包括一串从卷积导出的二值化量,例如BSIF,正如此处所描述的。

8.根据前述任意一项权利要求所述的装置,其中所述分类器模块包括线性支持向量机或概率协同表示分类器。

9.根据前述任意一项权利要求所述的装置,其中所述特征提取模块包括至少一个卷积神经网络。

10.根据权利要求9所述的装置,其中所述特征提取模块包括多个实质上统计独立的神经网络,所述神经网络向所述分类器模块提供互补性的描述符特征。

11.根据权利要求9或10所述的装置,其中所述神经网络包括深度卷积神经网络,优选地具有三个或更多个卷积层。

12.根据权利要求11所述的装置,其中所述描述符特征从每个深度卷积神经网络的第一完全连接层提取。

13.根据权利要求10至12中任意一项所述的装置,其中所述装置进一步包括特征级融合模块,用于将来自每个神经网络的描述符特征数据组合并将融合的(例如,级联的)特征数据发送给所述分类器模块。

14.根据权利要求9至13中任意一项所述的装置,其中使用包括变形的或平均化的图像和没有变形或平均化的图像的一组图像来对所述神经网络分别训练,以便训练它们的滤波器,从而提供适于确定一图像是否已经变形或平均化的描述符特征。

15.一种用于检测变形的或平均化的图像的装置，其中所述变形的或平均化的图像是合成生成的图像，所述合成生成的图像包括来自两个或更多个不同的源图像的信息，所述两个或更多个不同的源图像对应于两个或更多个主体，所述装置包括：

特征提取模块，其用于接收输入图像并输出该图像的一组描述符特征特性；以及

分类器模块，其被配置为，基于所述描述符特征而将所述输入图像认定为第一类型或第二类型，所述第一类型表明图像已经变形或平均化，所述第二类型表明图像没有变形或平均化；其中

所述特征提取模块包括多个神经网络，所述神经网络向所述分类器模块提供互补性的描述符特征；

该装置还包括融合模块，用于将来自每个神经网络的描述符特征数据组合并将融合的特征数据发送给所述分类器模块；并且

所述分类器模块包括机器学习系统，所述机器学习系统被训练为使用训练数据集将图像分类，所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

16.根据权利要求15所述的装置，其中所述神经网络是不同的深度卷积神经网络。

17.根据权利要求15或16所述的装置，其中所述神经网络实质上是统计独立的。

18.根据权利要求15、16或17所述的装置，其中所述变形的或平均化的图像是变形的或平均化的面部图像。

19.根据权利要求15所述的装置，进一步包括权利要求2至14中任意一项的特征。

20.一种检测变形的或平均化的图像的方法，其中所述变形的或平均化的图像是合成生成的图像，所述合成生成的图像包括来自两个或更多个不同的源图像的信息，所述两个或更多个不同的源图像对应于两个或更多个主体，所述装置包括：

接收输入图像；

生成该图像的一组描述符特征特性；

基于所述描述符特征，通过将该图像认定为第一类型或第二类型而将该图像分类，所述第一类型表明图像已经变形或平均化，所述第二类型表明图像没有变形或平均化；

其中分类步骤包括使用机器学习系统，所述机器学习系统被训练为使用训练数据集对单个图像进行分类，所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

21.根据权利要求20所述的方法，其中生成所述描述符特征的步骤包括使用多个优选为实质上统计独立的神经网络，所述神经网络向所述分类器模块提供互补性的描述符特征；并且来自每个神经网络的描述符特征数据在分类前被组合。

22.根据权利要求20或21所述的方法，进一步包括训练在所述分类步骤中使用的分类器的步骤。

23.根据权利要求20至22中任意一项所述的方法，包括训练特征提取模块，所述特征提取模块使用包括变形的或平均化的图像和没有变形或平均化的图像的一组图像来生成描述符特征，以便训练所述模块，从而提供适于确定一图像是否已经变形或平均化的描述符特征。

24.根据权利要求20至23中任意一项所述的方法，包括使用根据权利要求1至18中任意一项所述的装置。

25.一种软件产品，用于编程或以其他方式配置计算机，以执行权利要求20至24中任意一项所述的方法。

对被操纵图像的检测

技术领域

[0001] 本发明涉及对被操纵图像的检测，尤其涉及为了蒙蔽身份检查和相关安全措施而操纵的图像。本发明特别但不限于涉及：当例如在护照或其他身份证件、此类证件的申请、边境管制应用或银行应用中使用时，对已变形的或平均化的面部图像的检测。

背景技术

[0002] 面部识别是一种广泛使用的生物测定方法，已成为日常生活的一部分。对个人的自动识别(观察到面部生物特征)，尤其是在受限条件下，可产生非常高的准确性。这一事实使面部生物测定在国际边境管制中发挥了重要作用。面部识别系统是基于过去40年从信号和模式识别算法中收集到的知识而建立的，产生了准确且可靠的面部识别算法。这种性能的提升使面部生物测定可用于从取证、监视、物理和逻辑访问控制到电子商务和电子政务应用的各种应用中。

[0003] 生物面部参考图像已成为电子护照和签证申请的重要部分，在十年的引进期内已发行了近8亿本这样的护照。选择面部识别作为边境管制方案的原因之一是，在假的(false)否定系统决定的情况下，边境管制官员可以进行视觉比较。与其他的生物测定方法(例如，指纹识别)相比，这是一个明显的优势。这些因素证明了面部识别在自动边境管制(ABC)电子门中的适用性。在典型的ABC系统中，通过将现场获取的面部图像与存储在电子机读旅行证件(eMRTD)护照中的面部参考图像进行比较，可自动验证eMRTD与护照持有人(向边境警卫出示eMRTD的个人)之间的关联。这有助于高度可靠且准确的边境管制过程。实际上，国际民航组织(ICAO)已决定要求将面部图像作为eMRTD的主要标识符。因此，面部图像是全球所有电子护照中唯一的生物测定参考。

[0004] 随着ABC系统的广泛采用，其面部识别子系统对各种攻击的脆弱性得到了更多关注。这些攻击大致可分为两类：对ABC系统的攻击和对eMRTD生物测定参考的攻击。

[0005] 前一类的攻击通常是通过呈现面部伪像而在护照门的获取设备(或照相机)上进行的。此类攻击称为面部欺骗或呈现攻击。但是，这些攻击在生成面部伪像(即，人工的呈现攻击工具)并将其呈现给ABC电子门时需要付出很大的努力。此外，这种攻击只有在攻击者可以访问丢失或被盗的eMRTD护照时(这将允许他准备类似于eMRTD护照中存在的面部照片的面部伪像)才能成功。

[0006] 后一类的攻击包括可能操纵存储在(被盗)护照的逻辑数据结构(LDS)中的生物测定数据，目的是替换参考图像。但是，这种攻击应当容易被发现，因为它们将涉及更改在面部图像数据组上计算出的哈希值。因此，一种更有效的方法是利用护照申请和签发协议中的缺陷。在本文中进一步讨论的正是这类攻击，并且与本发明的实施例有关。

[0007] 这种攻击很容易进行，因为大多数护照签发程序都会在申请过程中接受(打印的)面部照片。此外，一些国家将接受上传到网络门户的数字照片，用于更新护照和签证申请。无论在哪种情况下，这都为攻击者提供了将被操纵的面部照片提交给护照发行机构的机会，进而获得了具有所有物理和电子安全特征并包含被操纵照片的真实eMRTD护照。虽然使

用生物测定亭可以解决此问题,但仅在少数护照申请办公室使用它们。

[0008] 适合进行对执行此类攻击来说所必需的图像操纵的软件是可免费获得的。例如,在线社交媒体经常提供允许图像修饰,特别是面部图像修饰的软件。各种不同类型的面部图像改变(例如,几何、纵横比和美化)都属于被修饰图像的类别,并且可以用于这种攻击。实际上,人们已经认识到由被修饰图像所引起的这种攻击的风险。Bharati、Singh、Vatsa和Bowyer在IEEE信息取证和安全事务(Transactions on Information Forensics and Security)(美国,IEEE)2016年9月,第11卷,第9期,第1903-1913页的"使用受指导的深度学习来检测面部修饰(Detecting Facial Retouching Using Supervised Deep Learning)"的研究文章证明了数字变更对自动面部识别性能的影响。该文章还介绍了一种检测已修饰图像的算法。

[0009] 面部变形是另一种形式的图像操纵,它已成为对eMRTD护照和签证签发协议的最严重攻击,这些协议始终依赖于申请人提供的图像。同时,如下文所述,这会导致一个复杂的问题:由于大量免费可用的变形软件产品,创建变形的面部图像的复杂性相当低。例如,GIMP和GAP工具可以用最少的工作生成高质量的变形的面部图像。

[0010] 面部变形(经由整个面部变形或平均化,或基于部分的面部变形或平均化来实现)的目的是使用可从两个或更多个不同的源面部图像(对应于两个或更多个不同主体)获得的唯一信息来生成新的合成生成的面部图像。因此,面部变形与对面部图像的修饰不同,因为后者依赖于仅可从与单个主体相对应的单个源面部图像获得的唯一信息。创建的变形面部图像将实质上构成面部外观特征,所述面部外观特征与对变形的面部有贡献的多个数据主体相对应。也许令人惊讶的是,对人类专家(例如,边防警卫)而言,所得到的图像在视觉上非常类似于这两个主体,并且在自动生物面部识别系统的相似性评分方面也是如此。

[0011] 这为任何攻击者(例如,已知的罪犯)提供了一个将他的面部图像变形为他人(例如,没有犯罪记录的人)的面部图像并申请可由这两个人使用的eMRTD护照的机会。由于eMRTD护照现在可广泛地与用于边境管制的ABC系统一起使用,因此,攻击者无需伪造护照证件即可执行此攻击。因此,需要减轻这种攻击,以确保边境管制过程的安全性。

[0012] 已经在商用面部识别算法上证明了注册程序对于面部变形攻击的脆弱性。而且,已经在实验上证明了人类检测变形图像的困难——即使面部识别专家也未能检测到变形的面部图像。

[0013] 进一步的复杂性是:与大多数提交打印图像的国家所使用的护照申请程序一致,可以打印变形的面部图像并且接着按照ISO/IEC标准以通常300dpi的分辨率再次扫描(在护照办公室处),用于生成ID证件。变形图像的重新数字化版本丢失了像素级信息,该像素级信息原本可帮助检测变形图像。同样,可以对数字变形的图像进行修饰以增强视觉质量,并在打印和提交之前消除任何重影外观,以使变形的图像类似于任何主体为了护照/签证签发而提交的高质量图像。已经证明,ABC系统特别容易受到对包含此类图像的eMRTD护照的攻击。

[0014] 迄今为止,除了发明人的工作以外,尚未报道有方法可以基于单个图像自动且可靠地检测变形的面部图像。

发明内容

[0015] 根据本发明的第一方面，提供了一种用于检测变形的或平均化的图像的装置，其中所述变形的或平均化的图像是合成生成的图像，包括来自两个或更多个不同的源图像的信息，所述两个或更多个不同的源图像对应于两个或更多个主体，所述装置包括：特征提取模块，其用于接收输入图像并输出该图像的一组描述符特征特性；以及分类器模块，其被配置为，基于所述描述符特征而将所述输入图像认定为第一类型或第二类型，所述第一类型表明图像已经变形或平均化，所述第二类型表明图像没有变形或平均化；其中所述分类器模块包括机器学习系统，所述机器学习系统被训练为使用训练数据集对单个图像进行分类，所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

[0016] 因此，本发明提供了一种装置，其通过机器学习(使用一组训练数据来训练分类器)来检测变形或平均化的图像(基于部分或整个图像)，例如，变形或平均化的照片。特别地，它能够使用单个输入图像来做到这一点，受过训练的装置对所述单个输入图像进行分析，以确定其是否已变形或平均化。训练数据集可以包括一组变形的或平均化的图像以及没有变形或平均化的图像。

[0017] 尽管该装置能够通过仅分析单个输入图像本身来进行确定，但是其当然可以用于以这种方式顺序地或同时地分析多个这样的图像。

[0018] 优选地，所述提取模块还包括机器学习系统，使得所述描述符特征取决于参数，所述参数由对包括图像的训练数据集的使用而确定。

[0019] 尽管本发明总体上可以用于检测变形的或平均化的图像，但是鉴于上面讨论的关于护照和身份证件照片的变形的已知问题，该装置优选地被配置为检测变形的或平均化的面部图像，因此分类器模块确定单个输入图像是否已变形或平均化。

[0020] 尽管可以将单个原图像输入到装置中，但是非常优选的是将输入图像尽可能地标准化。因此，该装置优选地还包括图像预处理模块，该图像预处理模块被布置为提取并归一化图像中感兴趣的区域(例如人的面部)，并将预处理后的图像发送给特征提取模块。

[0021] 为了提取人的面部，可以使用任何适宜的面部检测器算法，例如，Viola-Jones算法。预处理还可以包括将图像定向，例如通过仿射变换和/或将图像裁剪为合适大小以输入到系统中。

[0022] 尽管原则上可以采用任何适宜的特征提取系统，但是特征提取模块可以有用地包括一组滤波器，这些滤波器与输入图像的斑块卷积以提供一组描述符特征。斑块可以是不重叠的(即，可将图像铺开)或者可以以任何适宜的“步幅”(即，它们的中心之间的像素数)重叠。图像斑块和滤波器可以被视为通常具有相同大小的矩阵，并且卷积可涉及将矩阵的对应元素相乘并可涉及对结果求和，以提供每个滤波器的单个标量。

[0023] 优选地，多个这样的滤波器与每个斑块卷积，以提供一组输出值。

[0024] 上述一组描述符特征可以包括一串从卷积导出的二值化量，例如二值化统计图像特征(BSIF)。如本文所述，这涉及将二进制值1或0指定到上述每个标量值。

[0025] 本发明可以使用滤波器大小，例如是3×3、5×5、7×7、9×9、11×11、13×13、15×15和17×17，并具有8种不同的位长，例如5、6、7、8、9、10、11和12。例如，就精度而言，与其他滤波器相比，大小为11×11、长度为12位的滤波器可提供最佳性能。

[0026] 分类器模块可以是任何适宜的类型。例如，其可以包括线性支持向量机或概率协

同表示分类器。

[0027] 在本发明的优选形式中，特征提取模块包括至少一个卷积神经网络。这样的网络包含一个或多个卷积层(级)，每个卷积层均具有一组可学习的过滤器(也称为“核”)。优选地，神经网络是深度卷积神经网络(D-CNN)。术语“深”表示提供了多个这样的滤波器。优选地，提供3个或更多个卷积层，但是最优选地，存在5个或更多个，甚至更优选地是7个或更多个。

[0028] D-CNN的第一卷积层接收一组图像斑块作为其输入。这些斑块可以重叠，每个斑块的中心之间的像素数称为“步幅”。对于黑白图像，每个斑块的尺寸为m×n，对于输入图像为彩色的，每个斑块的尺寸为m×n×3(对于RGB)，m和n为所选斑块的行和列大小。

[0029] 滤波器通常在尺寸上对应于斑块，并且通过将每个矩阵的对应元素相乘并将结果相加来计算每个卷积。换句话说，每个过滤器可以在输入体积(斑块)的宽度和高度上卷积，以提供激活图。所有滤波器的图的堆栈提供了深度尺寸。因此，该层输出的矩阵的尺寸可以对应于斑块的数量和所应用的滤波器的数量。

[0030] 卷积层通常串联排列，使得作为来自给定斑块的输出而被提供的矩阵形成下一级的输入。由于这些卷积层优选地独立地(即，并行地)处理每个斑块，因此它们仅是“局部连接的”，而非“完全连接的”。

[0031] D-CNN还可具有“池化”层，在其中提取有意义的特征，具有“简化”数据的效果，例如，通过“最大池化”(在其中，将一组j×k个值替换为对应于该组中最大者的单个值)来实现。

[0032] 尽管装置可以包括单个神经网络，但是优选地，特征提取模块包括向分类器模块提供互补性的描述符特征的多个(至少实质上)统计独立的神经网络。

[0033] 如下所述，这提供了协同效果，因为它使得分类器能够考虑独立的描述符特征，从而有效提供了“交叉检查”，大大提高了系统的可靠性。这是基于发明人所认识到的：即使神经网络是在相同数据上训练的，它们具有实质上不同的体系结构，这也将带来神经网络所提供的独特和互补的描述符特征。

[0034] 该概念被认为是特别且独立作出的发明。因此，从另一方面看，提供了一种用于检测变形的或平均化的图像的装置，其中所述变形的或平均化的图像是合成生成的图像，包括来自两个或更多个不同的源图像的信息，所述两个或更多个不同的源图像对应于两个或更多个主体，所述装置包括：特征提取模块，其用于接收输入图像并输出该图像的一组描述符特征特性；以及分类器模块，其被配置为，基于所述描述符特征而将所述输入图像认定为第一类型或第二类型，所述第一类型表明图像已经变形或平均化，所述第二类型表明图像没有变形或平均化；其中所述特征提取模块包括多个(优选为实质上统计独立的)神经网络，所述神经网络向所述分类器模块提供互补性的描述符特征；该装置还包括融合模块，用于将来自每个神经网络的描述符特征数据组合并将融合的特征数据发送给所述分类器模块；并且所述分类器模块包括机器学习系统，所述机器学习系统被训练为使用训练数据集将单个图像分类，所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

[0035] 术语“实质上统计独立的”是指它们提供互补的描述符特征，从而与使用单个神经网络相比，显著提高了检测的准确性。

[0036] 如上所述，神经网络优选是D-CNN。尽管在原则上，可从D-CNN的任何层获得描述符特征，但优选是从每个深度卷积神经网络的第一个完全连接层提取描述符特征。该方面的发明可以包括针对于其他方面描述的任何优选或可选的特征。特别地，它可以适于检测变形的面部图像。

[0037] 由于该装置使用来自两个神经网络的输出，因此优选地，该装置还包括特征级融合模块，用于组合来自每个深度卷积神经网络的描述符特征数据，并将融合的(例如，级联的)特征数据发送给分类器模块。

[0038] 可以针对这些目的，基于正常图像而对旨在做图像识别/对象检测和识别的神经网络进行预训练。由于分类器被分别训练以检测变形的或平均化的图像，因此无需进一步训练就可以在本发明中使用这样的网络。然而，优选的是，使用包括变形的或平均化的图像和没有变形或平均化的图像的一组图像(例如，变形的和正常的)来对神经网络分别训练，以便训练它们的滤波器，从而提供适于确定一图像是否已经变形或平均化的描述符特征。

[0039] 尽管本发明中可以使用任何适宜的D-CNN，但是特别是一旦在适宜的图像数据集上对其进行了训练(或微调)，本发明优选地使用两个流行的预训练D-CNN，即VGG19和AlexNet。发明人已经确定这些D-CNN实质上是统计独立的。

[0040] VGG19使用非常小的(即3×3)卷积滤波器，具有大量的(16-19)层。另一方面，AlexNet使用较大的滤波器(例如，11×11，第一层的步幅为4)和八层，其中五层为卷积。它还包括某些“丢弃”层，其中某些输出设置为零。

[0041] 本发明还扩展到相应的方法，因此从另一方面看，本发明提供了一种检测变形的或平均化的图像的方法，其中所述变形的或平均化的图像是合成生成的图像，包括来自两个或更多个不同的源图像的信息，所述两个或更多个不同的源图像对应于两个或更多个主体，所述方法包括：接收输入图像；生成该图像的一组描述符特征特性；基于所述描述符特征，通过将该图像认定为第一类型或第二类型而将该图像分类，所述第一类型表明图像已经变形或平均化，所述第二类型表明图像没有变形或平均化；其中分类步骤包括使用机器学习系统，所述机器学习系统被训练为使用训练数据集对单个图像进行分类，所述训练数据集包括变形的或平均化的图像以及没有变形或平均化的图像。

[0042] 如上所述，生成所述描述符特征的步骤优选地包括使用多个实质上统计独立的神经网络(优选地，D-CNN)，所述神经网络向所述分类器模块提供互补性的描述符特征；并且来自每个神经网络的描述符特征数据在分类前被组合。

[0043] 同样地，所述方法可以进一步包括，在使用前，对在所述分类步骤中使用的分类器的进行训练的步骤。其还可以包括训练特征提取模块，所述特征提取模块使用包括变形的或平均化的图像和没有变形或平均化的图像的一组图像来生成描述符特征，以便训练所述模块，从而提供适于确定一图像是否已经变形或平均化的描述符特征。

[0044] 事实上，所述方法可以进一步涉及对应于以上关于本发明的装置所讨论的任何或全部优选或可选特征的步骤。

[0045] 如上所讨论的，本发明的一种特殊应用是检测变形的或平均化的图像，该图像可以呈现给护照/签证签发办公室，或者实际上呈现在自动护照门口，用于犯罪或欺诈目的。因此，本发明扩展到一种通过检测与护照或其他身份证件有关的变形图像来防止这种活动的方法，并且扩展到用于该目的的自动装置，包括包含这种装置的自动护照门。

[0046] 通常通过使用计算机或其他适宜的数据处理装置来实现本发明，所述计算机或其他适宜的数据处理装置通常涉及存储器、处理器和数据存储。因此，本发明还扩展到软件产品，该软件产品用于对这种装置进行编程或以其他方式配置，以提供本发明的装置和/或方法。因此，本发明的另一方面提供了这样的软件产品。

[0047] 尽管本发明的上述方面特别集中在对变形的或平均化的图像的检测上，但是将认识到，以上讨论的装置和方法可以同样地用于检测已经以某种其他方式(例如，修饰)操纵的图像。因此，本发明可以更一般地用于检测被操纵的图像。

[0048] 因此，根据本发明的另一方面，提供了一种用于检测被操纵的图像的装置，所述装置包括：特征提取模块，其用于接收输入图像并输出该图像的一组描述符特征特性；以及分类器模块，其被配置为，基于所述描述符特征而将所述输入图像认定为第一类型或第二类型，所述第一类型表明图像已经被操纵，所述第二类型表明图像没有被操纵；其中所述分类器模块包括机器学习系统，所述机器学习系统被训练为使用训练数据集对单个图像进行分类，所述训练数据集包括被操纵的图像以及没有被操纵的图像。在本发明的又一方面，还提供了一种相应的方法。将会认识到，本发明的这些方面同样得益于以上结合本发明的上述那些方面所讨论的可选特征和/或步骤。

附图说明

[0049] 现在仅通过示例的形式并参照附图，对本发明特定的优选实施例进行描述，其中：

[0050] 图1是示出了根据本发明第一实施例的变形面部图像检测系统的框图；

[0051] 图2是示出了图1的实施例的BSIF阶段的定性结果的图，其中(a)是一组归一化的面部图像，(b)是对应的一组示出了BSIF特征的图像，(c)是那些BSIF特征的直方图，而(d)是同一直方图的放大部分；

[0052] 图3是用于产生变形图像的一对主体的照片；

[0053] 图4是一组成对主体的归一化图像，以及对应的、用于产生变形图像数据库的变形图像；

[0054] 图5是示出了将来自数据库的变形图像与创建变形图像的主体进行比较而获得的分散的比较分数的图；

[0055] 图6是示出了根据本发明第二实施例的变形面部图像检测系统的框图；

[0056] 图7示出了在第二实施例中使用的深度卷积神经网络的卷积层中的滤波器的权重，具有3×3的滤波器大小，分别来自(a)已知的AlexNet和(b)已知的VGG19网络；

[0057] 图8是一组成对主体的归一化图像，以及对应的、与第二实施例一起使用的、用于产生变形图像数据库的变形图像；

[0058] 图9是将IG-SVM方法(a)和第二实施例的方法(b)的性能进行比较的一对曲线图。

具体实施方式

[0059] 首先参照图1，示出了用于健壮的变形面部图像检测的系统的第一实施例10的框图。其具有两个主要的阶段：预处理11和变形面部检测12。

[0060] 第一步骤是针对图像13，该图像将受到检测，将被输入到系统中。

[0061] 接着，预处理阶段11提取归一化的面部区域。这首先涉及检测图像中的面部。这是

使用Viola-Jones算法执行的,Viola-Jones算法是一种众所周知的对象检测框架,该框架是健壮的并且在现实世界场景中表现良好。(可以使用任何其他的面部检测器,例如,基于D-CNN的面部检测器。)在下一步骤中,使用仿射变换对图像13进行归一化以补偿旋转,最后,将归一化后的图像大小调整为x乘y个像素(此处为120×120个像素)。这使得输出标准大小的归一化图像14。

[0062] 概况来说,变形的面部检测阶段12基于对从归一化面部图像14提取的微纹理特征的分析。这使用二值化统计图像特征(BSIF)滤波器15来完成,将输出(一组检测到的特征)传递到线性支持向量机(SVM)16进行分类。这将图像分类为“正常面部”17或“变形面部”18(即,被拒绝)。

[0063] BSIF滤波是一种提供图像描述符的已知技术(参见J.Kannala和E.Rahtu,“BSIF:二值化统计图像特征(BSIF:Binarized statistical image features)”,第二十一届国际模式识别会议(International Conference on Pattern Recognition)(ICPR),2012年,第1363-1366页)。首先对灰度图像进行归一化,使得0表示平均强度(即,表示归一化图像的矩阵包括负值和正值),然后将其划分为l×l个像素的正方形斑块(即,每个斑块由l×l的实数矩阵表示)。

[0064] 然后,每个斑块与一系列的n个线性滤波器进行卷积,每个线性滤波器也是一个l×l的实数矩阵,以提供滤波器响应。每个滤波器的滤波器响应是通过将每个矩阵的相应元素相乘并将结果相加得出的。然后,如果总和大于0,则通过提供输出1来对总和进行二进制化,否则,则通过提供输出0来对总和进行二进制化。由于应用了一组n个滤波器,因此提供了n个二进制位的串,其值构成图像的BSIF描述符(特征),可用于对其进行分类。

[0065] 返回图1,在阶段15从图像获得特征。所使用的特定滤波器组是开源组。这最初是由机器学习训练阶段使用从13个不同的自然风景图像中随机采样的50,000个图像斑块来确定的(参见A.Hyvearinen、J.Hurri和P.O.Hoyer,自然图像统计学(Natural Image Statistics),第39卷,Springer,2009年)。机器学习未受指导,并使用独立成分分析方法(参见J.H.van Hateren和A.van der Schaaf,与主要视觉皮层中的简单细胞相比,自然图像的独立成分滤波器(Independent component filters of natural images compared with simple cells in primary visual cortex),伦敦皇家学会会议录(Proceedings of the Royal Society of London),系列B:生物科学(Biological Sciences),265(1394):359-366,1998)。所得的滤波器在统计上高度独立。因此,该组滤波器被优化以用于图像识别,但是在该实施例中,它们并不针对于变形图像检测而被专门训练。

[0066] 由于BSIF滤波器是基于无指导的学习生成的,因此,可以使用任意数量的具有不同大小的滤波器。发明人评估了8种不同的滤波器大小,例如3×3、5×5、7×7、9×9、11×11、13×13、15×15和17×17,并具有8种不同的位长,例如5、6、7、8、9、10、11和12。基于开发数据集上的实验,就精度而言,选择大小为11×11,长度为12位的滤波器是最优选的(请参见下文)。

[0067] 图2示出了使用优选的滤波器大小在正常的和变形的面部图像上获得的BSIF特征的定性输出。第一和第二主体20、21的归一化图像与相同的两个主体的归一化变形图像22一起在2(a)处示出。

[0068] 应用滤波器的结果如图2(b)所示,其中BSIF特征(即每个斑块的描述符)被示出

为:在与其相关的斑块的位置处被转换为灰度。此外,图2(c)所示的直方图示出了线性面元序列中每个斑块的描述符的值。由于使用了12个滤波器,因此,每个描述符的位长为12,进而每个描述符的尺寸为1×4096(即,2^{12})。

[0069] 值得注意的是,BSIF直方图特征(图2(c))示出了正常的和变形的面部图像之间的直方图轮廓变化。如图2(d)所示,通过将所有三个面部图像的对应面元中的直方图轮廓放大,可以进一步看到差异。

[0070] 返回图1,接下来将描述符数据(即,图2(c)中所示的数据)传递到线性SVM分类器16。这先前已经以公知的方式进行了训练,使用大量肯定的(正常面部)样品和否定的(变形的)样品(例如,参见下面的实验),从而能够以高度的可靠性区分它们。因此,SVM分类器使用描述符数据来确定所呈现的面部图像是属于正常类型(附图标记17)还是变形类型(附图标记18),并将相应的输出提供给用户。在使用该系统的地方,例如在护照/VISA签发办公室,将图像认定为变形类型将导致图像被拒绝。

[0071] 在第一实施例的变型中,BSIF特征提取系统被单个深度卷积神经网络(D-CNN)取代,正如在下面描述的第二实施例中所使用的——可以使用所描述的任一D-CNN。其也可以使用第二实施例的分类器系统,该分类器系统可以从D-CNN的第一完全连接层接收其输入,正如同样针对第二实施例所进行的描述(并不需要第二实施例的特征级融合)。

[0072] 实验

[0073] 发明人构建了一个新的大规模的变形面部数据库,该数据库由450个变形图像组成,这些变形图像使用来自110个数据主体的不同面部图像组合而生成。数据收集的第一步是按照eMRTD护照规范中定义的ICAO获取标准获取面部图像。在此程度上,他们首先在具有统一照明、统一背景、中性姿势和正常面部表情的工作室中收集了正面图像。使用安装在三脚架上并位于与拍摄主体相距2米的Canon EOS 550D DSLR相机获取图像。图3示出了所获取的两个数据主体的高质量面部图像的示例。

[0074] 然后使用免费的GNU图像处理程序v2.8(GIMP)和GIMP动画包(GAP)工具生成变形的面部图像。将待变形的两个面部图像手动对齐,并作为输入提供给GAP工具。然后,GAP工具生成图像帧的序列,显示一个主体到另一个主体的转变。最终的变形图像是通过确认其与对变形过程作出贡献的主体面部的相似性来手动选择的。图4示出了三个变形面部图像的示例,这些示例是分别使用两个不同的主体获得的,尽管变形图像也是使用三个主体以相似的方式生成的。左列显示第一主体的图像,第二列显示第二主体的图像,而第三列显示相应的变形图像。

[0075] 为了充分评估变形面部数据库并对变形面部检测算法基准化,450个变形图像的整个数据库被分为三个独立的子集:训练集、开发集和测试集。训练集包括200个变形图像,这些变形图像专门用于训练SVM分类器。开发集包括50个变形图像,这些变形图像用于调整所提出方案的参数,尤其是关于选择BSIF滤波器的大小和长度。测试集包括200个变形图像,这些变形图像仅用于报告变形面部检测算法的结果。

[0076] 使用可从Neurotechnology获得的可商用的Verilook面部识别SDK对变形的面部图像数据库进行了脆弱性分析(请参见http://www.neurotechnology.com/verilook.html)。通过将变形的面部图像注册到Verilook并使用探针样本(对应于用于创建变形图像的数据主体之一)进行分析。因此,对于每个注册的变形面部,取决于用于创建

变形面部图像的主体数量，可以获得两个或三个不同的比较分数。

[0077] 图5示出了使用Verilook面部SDK在包括450个变形图像的整个数据库上所获得的分散的比较分数。这些分数的有效性是根据FRONTEX(欧盟成员国对外边境的运营合作管理的欧洲机构)的准则进行评估的。FRONTEX建议：以验证模式运行的ABC系统中的面部验证算法应提供性能，使得错误接受率(FAR)为0.001(0.1％)或更佳。在此配置下，FRR不应高于0.05(5％)。在实验中，使用了SDK提供的阈值分数，对于给定的FAR＝0.1％，分数为36。图5中的竖直红线表示此验证阈值，对应于目标FAR＝0.1％。因此，将大于36的比较分数视为成功验证。从图5可以看出，针对该阈值，所有变形图像都已成功匹配，从而表明了变形面部图像在现实世界应用中的攻击潜力。

[0078] 下面讨论使用第一实施例进行自动的变形面部检测的定量结果，以及对当前四种不同的特征提取方案的比较评估：图像质量分析(IQA)、局部二值模式(LBP)、局部相位量化(LPQ)和2D快速傅立叶变换(2DFFT)。选择这些比较器是考虑到它们与问题的相关性，以及它们在最近的呈现攻击检测(Presentation Attack Detection)工作中的准确性。这些特征提取方案的分类是使用线性SVM进行的，以与提出的方案一致。

[0079] 定义了两个不同的性能评估指标，以量化结果：(i)分类为变形面部图像的正常面部图像(NFCM)，即，分类为变形面部图像的正常面部图像的比率；(ii)分类为正常面部图像的变形面部图像(MFCN)，即，分类为正常面部的变形面部图像的比率。可使用平均分类错误率(ACER)来测量总体准确性，其被定义为ACER＝(NFCM+MFCN)/2。

[0080] 表1给出了第一实施例(表中的“提出的方法”)的定量结果以及这项工作中采用的四种不同的基线算法。

[0081]

算法	MFCN（%）	NFCM（%）	ACER（%）
图像质量[7]-SVM	1.73	73.37	37.55
LBP [16]-SVM	37.66	13.20	25.43
LPQ [2]-SVM	29.00	11.47	20.23
2DFFT [15]-SVM	61.03	37.22	49.12
提出的方法	3.46	0	1.73

[0082]

[0083] 表1

[0084] 基于所获得的结果，可以观察到以下内容：

[0085] ●注意到了，第一实施例(“提出的方法”)的使用具有最佳性能，ACER为1.73％。

[0086] ●注意到了，图像质量分析特征具有最佳的MFCN为1.73％，但是此设置的NFCM值很高，为73.37％，这在现实世界情况下不适用。

[0087] 因此，与常规的特征提取技术相比，基于BSIF滤波器的统计图像特征的使用显示出最佳性能。所获得的结果证明了第一实施例对于自动的变形面部检测的适用性。

[0088] 现在将参照其余附图讨论本发明的第二实施例。它特别适合于识别经过了打印-扫描过程，因而比“数字”变形图像更难检测的变形图像。所述打印-扫描过程对应于使用最广泛的护照申请过程。

[0089] 图6示出了第二实施例的系统30的框图。正如下文更加详细的讨论，它基于两个预训练的深度卷积神经网络(D-CNN)31、32的特征级融合，以检测变形的面部图像。已经知道所采用的神经网络用于图像识别且为此目的而被预先训练。

[0090] 卷积神经网络包括一个或多个卷积层(级)，每个卷积层都具有类似于先前实施例中所使用的一组可学习滤波器(也称为“核”)。术语“深”表示提供了多个这样的滤波器。

[0091] 就像第一实施例中使用的BSIF特征提取系统一样，D-CNN的第一卷积层接收图像的一组斑块作为其输入。所述斑块可以重叠，每个斑块的中心之间的像素数称为“步幅”。对于黑白图像，每个斑块的尺寸为m×n，当输入图像是彩色的时，每个斑块的尺寸为m×n×3(对于RGB)，m、n为图像斑块的行和列。每个滤波器在尺寸上对应于这些斑块，并且通过将每个矩阵的对应元素相乘并将结果相加来计算每个滤波器的滤波器响应(即，卷积)。换句话说，每个滤波器在输入体积(斑块)的宽度和高度上卷积，以提供激活图。所有滤波器的图的堆栈提供了深度尺寸。因此，该层输出的矩阵的尺寸对应于斑块的数量和所应用的滤波器的数量。

[0092] 卷积层串联排列，使得作为给定斑块的输出而被提供的矩阵形成下一级的输入。由于这些卷积层独立地(即，并行地)处理每个斑块，因此它们仅是“局部连接的”，而不是“完全连接的”。

[0093] 深度CNN还具有“池化”层，在其中提取有意义的特征，具有“简化”数据的效果，例如，通过“最大池化”(在其中，将一组j×k个值替换为对应于该组中最大者的单个值)来实现。

[0094] 在多个卷积层(通常是池化层)之后，将存在一个或多个完全连接的层，这些层基于整个图像的特性来接收数据，并包括一个分类级，该分类级提供输出，该输出将输入图像认定为一些类型中的一个类型。

[0095] 在该实施例中使用的两个已知的深度CNN(被设计用于图像识别)中的每一个都具有大量的卷积和池化层，其后是一些完全连接的层。对它们进行了预训练以用于图像识别，其方式与关于BSIF系统所讨论的方式大体相似——即，在机器学习过程中使用了训练数据集，从而学习到了用于图像识别的最佳滤波器。

[0096] 返回图6，使用三个主要的功能块来构造第二实施例的系统，这三个主要的功能块包括：(i)在块34处预处理图像33，(ii)在块35处两个D-CNN 31、32的特征级融合，以及(iii)在块36处的分类。下面讨论每个块的功能。

[0097] 预处理块34中的每一个大致类似于关于第一实施例所描述的相应块11。因此，从面部图像33开始，使用Viola-Jones算法进行面部检测，使用仿射变换对检测到的面部区域进行归一化以补偿旋转，并且将图像重新调整大小为227×227像素。(使用此大小是因为它适合于D-CNN输入层的大小，该D-CNN也已使用此大小的图像进行了预训练。)

[0098] 相同的预处理输出接着被提供到两个不同的D-CNN中的每一个。

[0099] 该系统使用两种流行的预训练D-CNN，称为VGG19(附图标记31)和AlexNet(附图标记32)——分别参见(i)K.Simonyan和A.Zisserman“用于大规模图像识别的非常深的卷积网络(Very deep convolutional networks for large-scale image recognition)”，arXiv预印本，arXiv:1409.1556，2014年和A.Krizhevsky、I.Sutskever;和(ii)G.E.Hinton“具有深度卷积神经网络的图片网络分类(Imagenet classification with deep

convolutional neural networks)”,神经信息处理系统的进展(Advances in neural information processing systems),第1097-1105页,2012年。

[0100] VGG19使用非常小的(即3×3)卷积滤波器,具有大量的(16-19)层。另一方面,AlexNet使用较大的滤波器(例如11×11,第一层中的步幅为4)和八个层,其中五层为卷积。(它还包括一些“丢弃(dropout)”层,其中某些输出设置为零。)每个D-CNN的卷积和池化层在图中示意性地显示。

[0101] 两个D-CNN都在相同的大型ImageNet数据库上训练。但是,由于它们基于不同的配置(即,体系结构),所以尽管在相同的数据库上进行了训练,但它们提供了适合检测变形面部图像的互补功能。

[0102] 在使用变形面部数据库(见下文)之前,会对D-CNN进行微调。

[0103] 这两个网络的组合发生在特征级,其中使用一组卷积层将提取自每个网络的特征进行融合,例如通过级联。从图6可以看出,从两个网络31、32的第一个完全连接层(FC-6)提取了输出。选择该层是为了利用全尺寸特征,尤其是在AlexNet中采用的丢弃层之前。然后,在特征级融合块35中,将这些特征级联起来以形成单个向量,然后将该向量输入到分类块36中。

[0104] 类似于第一实施例的分类器,先前已经用已知方式训练的分类块36使用特征向量来确定图像是正常的还是变形的(因此应当被拒绝)。所使用的分类器是概率协同表示分类器(P-CRC),该分类器使得测试样本连同其他类型的似然比最大化,以执行分类。

[0105] 与第一实施例一样,第二实施例的分类器36使用来自变形面部图像数据库(包含正常图像和变形图像)的训练数据集进行训练。表2列出了数据库的详细内容:

[0106]

类别	样本数量		
	训练集	测试集	总数
真实的图像	206	146	352
变形的图像	225	206	431

[0107] 表2

[0108] 但是,除此之外,此训练数据集用于独立地微调VGG19和AlexNet网络,这些网络在图像识别中进行了预训练,但与变形图像无关。微调涉及训练每个单独的D-CNN,以在将图像如上所述地组合使用之前将其分类为正常图像或变形图像。

[0109] 在对网络进行微调时,与网络中的其他层相比,在最后一层应用了高值的学习率。使用的学习参数是权重学习率因子10和偏向学习率因子20。

[0110] 分析

[0111] 为了将经过微调的VGG19和AlexNet的网络特征可视化,图7示出来自这两个D-CNN的第三卷积层(Conv3)的特征。

[0112] 图7(a)示出了来自经微调的AlexNet的Conv3特征,卷积滤波器大小为3×3,长度为384。有趣的是,观察到了每个卷积滤波器都表现出面部特征和纹理特征。此外,突出显示的区域(以不同的颜色表示)表明由经微调的AlexNet所学习的面部特定特征。

[0113] 从图7(b)中也可以看到类似的观察结果,它是具有Conv3层的经微调的VGG19的对应图示,卷积滤波器大小为3×3,长度为256。图7(b)中突出显示的区域显示了来自由经微

调的VGG19所学到的面部特征的要素。

[0114] 通过下式对相关的互相关系数(CC)进行计算,对来自两个经微调的D-CNN网络的特征提供互补性信息的程度(因此产生了协同效应)进行了研究:

[0115] $$CC=\frac{\sum_m\sum_n(F_{Alex}-\bar{F}_{Alex})(F_{VGG}-\bar{F}_{VGG})}{\sqrt{\left(\sum_m\sum_n(F_{Alex}-\bar{F}_{Alex})^2\right)\left(\sum_m\sum_n(F_{VGG}-\bar{F}_{VGG})^2\right)}} \quad (1)$$

[0116] 其中F_{VGG}和F_{Alex}表示每个D-CNN的描述符特征的值,而$\bar{F}_{VGG}$和$\bar{F}_{Alex}$是它们的平均值。平均值分别对应于跨行和列(m和n)的像素强度。CC的值越低,表明数据越互补。

[0117] 下面的表3示出了结果:

[0118]

数据类型	特征	CCC
数字的	训练 Alex-训练 VGG	0.047
	测试 Alex-测试 VGG	0.045
打印-扫描 (HP)	训练 Alex-训练 VGG	0.024
	测试 Alex-测试 VGG	0.009
打印-扫描 (RICOH)	训练 Alex-训练 VGG	0.007
	测试 Alex-测试 VGG	0.012

[0119] 表3

[0120] 将会注意到,CC值(表中称为"CCC")都非常低,因此表明从两个D-CNN获得的特征本质上是互补的——即,它们组合在一起具有协同效果。

[0121] 为了检查分类器的变形检测性能,令从AlexNet提取的FC-6特征为F_A,而令从VGG19提取的FC-6特征为F_V。然后,通过将提取到的特征进行级联来组合它们,以形成用于训练P-CRC的单个特征向量$Tr_F=[F_A||F_V]$(也可以使用其他等效的分类器,例如SVM、Random Forest)。然后,将测试面部图像F_{Te}独立地投影到AlexNet和VGG19网络的FC-6层上,以获得相应的特征,并令其为Fte_A和Fte_V。然后,使用特征级联将这些特征组合起来,以形成单个向量$Te_F=[Fte_A||Fte_V]$,将其用作测试特征向量,特别是与P-CRC一起,以获取变形检测分数。这项工作中使用的P-CRC利用学习到的特征向量与探针特征向量的正则最小二乘回归(LSR)来表示:

[0122] $$\widehat{F}=argmin_\alpha\|Te_F-\mathscr{D}\alpha\|_2^2+\lambda\|\alpha\|_2^2 \quad (2)$$

[0123] 其中Te_F是测试图像的特征向量,$\mathscr{D}$是使用Tr_F学习到的协作子空间字典,α是系数向量,而λ是正则化参数。所获得的距离用作变形检测分数,以获得变形的面部检测性能——参见下面实验中对图9的讨论。

[0124] 实验

[0125] 在变形的面部图像数据库上对第二实施例进行了评估,该数据库是使用包括104个主体的可公开获得的面部数据库创建的。使用免费的软件包(如GNU图像处理程序v2.8

(GIMP)和GIMP动画软件包(GAP)工具)生成变形的面部图像,并通过手动干预来对齐界标点,以获取高质量的变形图像。然后,为了模拟护照签发程序的实际情况,生成了数字变形图像的打印-扫描版本。

[0126] 因此,数字变形图像中的每一个都使用高质量的相纸(300克)和打印分辨率为1200dpi的HP Photosmart 5520进行打印。下一步,使用两种不同的扫描仪对打印的图像进行扫描:(i)HP照片扫描仪(HP Photo Scanner)和(ii)RICOH办公扫描仪(RICOH office scanner)。按照ICAO的关于ePass中面部图像参考的规范,将打印的照片扫描为具有300dpi。

[0127] 图8示出了数字变形图像40(顶部行)、使用HP扫描仪打印-扫描的图像41(中间行),以及使用RICOH扫描仪打印-扫描的图像42(底部行)的示例。图8的每一行中的中间图像44对应于使用两个不同的主体43、45生成的变形的面部图像,这两个不同的主体在变形的面部图像的两侧示出。

[0128] 从图8可以看出,随着打印-扫描过程的进行,感知图像质量下降。这种效果使检测更具挑战性。此工作中使用的数据库具有352个真实的面部图像和431个变形的面部图像(对于所有三个版本,例如数字的、使用HP进行打印-扫描的,以及使用RICOH进行打印-扫描的),这些图像被分为两个不相交(或不重叠)的子集,用于训练和测试。对于训练集和测试集,样本的分布在上面的表2中给出。

[0129] 生成变形的面部图像的过程涉及对图像进行一系列的预处理和后处理操作,与真实图像相比,它们会产生不同的质量。由于分类器可以检测质量/压缩差异而不是变形的影响,因此这种不同质量度量可能会使变形检测性能产生偏差。因此,发明人格外注意通过在真实面部图像和变形面部图像这二者上应用等同的预处理和后处理操作序列,以具有相同质量的变形面部图像和真实面部图像。

[0130] 使用检测错误权衡(DET)图来报告变形面部检测算法的性能,该图描述了攻击呈现分类错误率(APCER)与真实呈现分类错误率(BPCER)之间的相关性,它们根据ISO/IEC 30107-3被定义为:

[0131] ●APCER:在特定情况下被错误分类为真实呈现的攻击呈现(即,变形的面部样本)的比例;

[0132] ●BPCER:在特定情况下被错误分类为呈现攻击(即,变形的面部样本)的真实呈现的比例。

[0133] 此外,通过将APCER固定设置为5%和10%并报告相应的BPCER,可以显示结果。通过变形面部检测算法获得的较低的APCER和BPCER值将代表最佳的检测性能。

[0134] 进行了两个不同的实验,以评估发明人的方法与现有技术的变形面部检测算法相比的健壮性:

[0135] 实验1:在该实验中,在数字和打印-扫描的变形面部图像数据库上独立评估了所提出方法和现有技术的方法的检测性能。该实验分析了检测方案的性能,使得训练数据和测试数据来自同一来源(数字的或打印-扫描的(HP)或打印扫描的(RICOH))。因此,检测方法将具有关于打印机/扫描仪类型(用于生成待检测的打印-扫描的变形面部图像)的先备知识。

[0136] 实验2:该实验被设计为,测试变形面部检测算法(包括该实施例的算法)的健壮

性。在不同的数据源上对变形检测方法进行了训练和测试。例如，变形面部检测算法在数字版本的数据库上进行了训练，并在打印扫描版(HP或RICOH)上进行了测试。该实验表明了变形检测算法的推广能力。此外，实验2还解决了检测变形的面部图像的现实情况，因为在全球不同的护照发行机构使用了不同的扫描仪。

[0137] 下面的表4显示了在数字子集、HP打印-扫描子集和RICOH打印-扫描子集上独立评估的，本文提出的方法与四种不同的现有技术算法相比的定量结果。

[0138]

数据集	算法	D-EER (%)	BPCER (%) @	
			APCER=10%	APCER=5%
数字的	LBP-SVM [10]	29.28	44.52	56.84
	LPQ-SVM [10]	26.12	42.46	52.73
	IG-SVM [10]	21.88	30.13	41.78
	BSIF-SVM [10]	22.70	38.25	49.31
	提出的方案	8.23	7.53	14.38
HP 打印-扫描	LBP-SVM [10]	34.35	61.64	69.17
	LPQ-SVM [10]	49.82	92.46	94.52
	IG-SVM [10]	38.35	73.28	80.13
	BSIF-SVM [10]	26.12	45.89	55.47
	提出的方案	17.64	32.87	41.78
RICON 打印-扫描	LBP-SVM [10]	22.70	58.21	42.46
	LPQ-SVM [10]	39.18	74.65	82.87
	IG-SVM [10]	34.35	52.05	63.69
	BSIF-SVM [10]	23.29	43.83	54.79
	提出的方案	12.47	16.43	28.76

[0139] 表4

[0140] 图9示出了DET曲线，该曲线显示了在所有三种变形面部数据类型上所提出的方法和基于图像梯度(IG-SVM)的现有技术方法的性能。为简单起见，仅包含两条DET曲线，而所有的结果在上面的表4中详细列出。以下是实验1的主要观察结果：

[0141] ●第二实施例的系统在变形面部数据库的数字版本和打印-扫描的版本上的所有三个独立评估中均表现出最佳的检测性能。它在数字图像数据库上显示最小的检测等错率(D-EER(%))为8.23％，打印-扫描(HP)上的D-EER为17.64％，打印-扫描(RICOH)上的D-EER为12.47％。注意到第二佳的检测性能是BSIF-SVM方案。

[0142] ●总体上，基于微纹理方案的现有技术方案的检测性能在所有三个数据库版本上均显示性能下降。

[0143] ●包括第二实施例的系统的变形检测算法的检测性能，对于打印-扫描版数据库的性能要比数字版数据库有所降低。仍然有趣地注意到，该实施例的性能，特别是在使用

RICOH生成的打印-扫描数据库上，已经表明了合理的性能，当APCER＝5％时，D-EER为12.47％，BPCER＝16.43％。这显示了所提出的方案在检测数字的和打印-扫描的变形面部图像方面的健壮性。

[0144] ●扫描仪的类型也会影响变形面部检测算法的性能。实验表明，将高质量的打印照片与照片扫描仪一起使用会使变形检测更具挑战性。在实验中证明了这一事实，因为包括该实施例的变形面部检测算法的性能表明：与使用来自RICOH的办公室扫描仪创建的打印-扫描数据库相比，使用来自HP的照片扫描仪生成的打印-扫描数据库具有较高的错误率。

[0145] 下面的表5示出了，根据实验2，所提出的方案和现有技术方案对跨数据库评估的性能：

[0146]

训练集	测试集	算法	D-EER (%)	BPCER (%) @	
				APCER=10%	APCER=5%
数字的	打印-扫描（HP）	BSIF-SVM	26.70	48.63	56.16
		提出的方案	15.05	24.65	39.72
数字的	打印-扫描（RICOH）	BSIF-SVM	27.53	57.53	65.75
		提出的方案	15.05	17.80	28.08

[0147]

打印-扫描（HP）	数字的	BSIF-SVM	21.29	40.41	52.05
		提出的方案	19.88	34.24	45.89
打印-扫描（HP）	打印-扫描（RICOH）	BSIF-SVM	28.11	58.90	67.80
		提出的方案	13.06	21.23	28.76
打印-扫描（RICOH）	打印-扫描（HP）	BSIF-SVM	27.53	60.95	65.75
		提出的方案	19.05	28.08	40.41
打印-扫描（RICOH）	数字的	BSIF-SVM	23.29	34.93	52.05
		提出的方案	20.71	30.13	37.67

[0148] 表5

[0149] 为简单起见，展示了所提出的方案的结果以及基于BSIF-SVM的第二佳的方法的结果。以下是从实验2得出的主要观察结果：

[0150] ●与实验1中获得的结果相比，变形检测算法的性能下降。

[0151] ●在该实施例中注意到了最佳的检测性能，显著表明了与现有技术方案相比性能的提高。

[0152] ●当训练数据对应于变形面部数据库的数字版本并测试数据库的打印扫描版本时，观察到实施例的最佳检测性能。当与现有技术方案的健壮性相比时，这表明实施例的健壮性。因此，基于广泛的实验，第二实施例的系统已经成为检测从不同类型的扫描仪生成的

数字和打印扫描版本中的变形面部图像的最佳方法。它还显示了遵循实验2中指示的协议的跨数据库评估的最佳性能，表明了与其他的现有方法相比，所提出的方法的适用性。

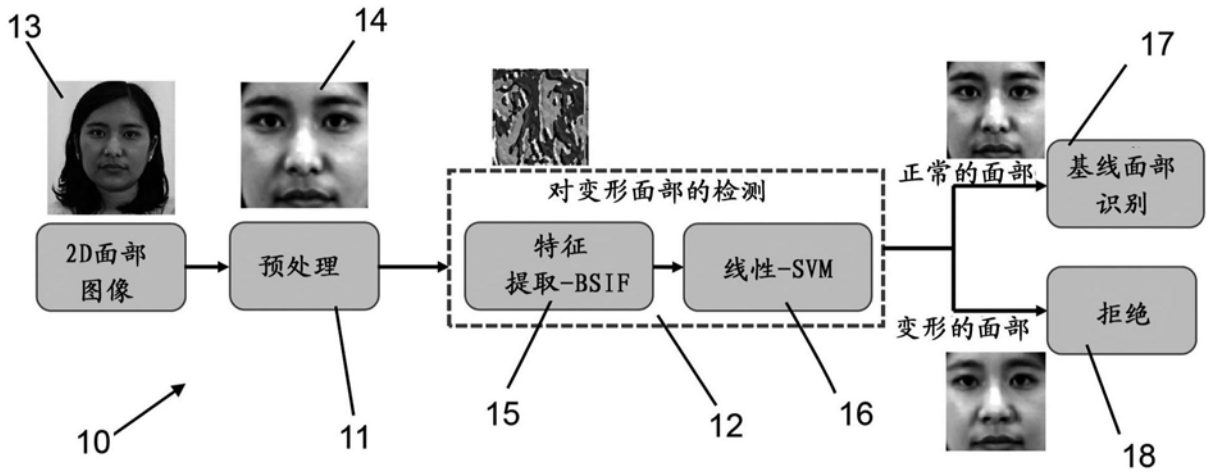

13
14
17
2D面部
图像
预处理
对变形面部的检测
特征
提取-BSIF
线性-SVM
正常的面部
基线面部
识别
变形的面部
拒绝
10
11
15
12
16
18

图1

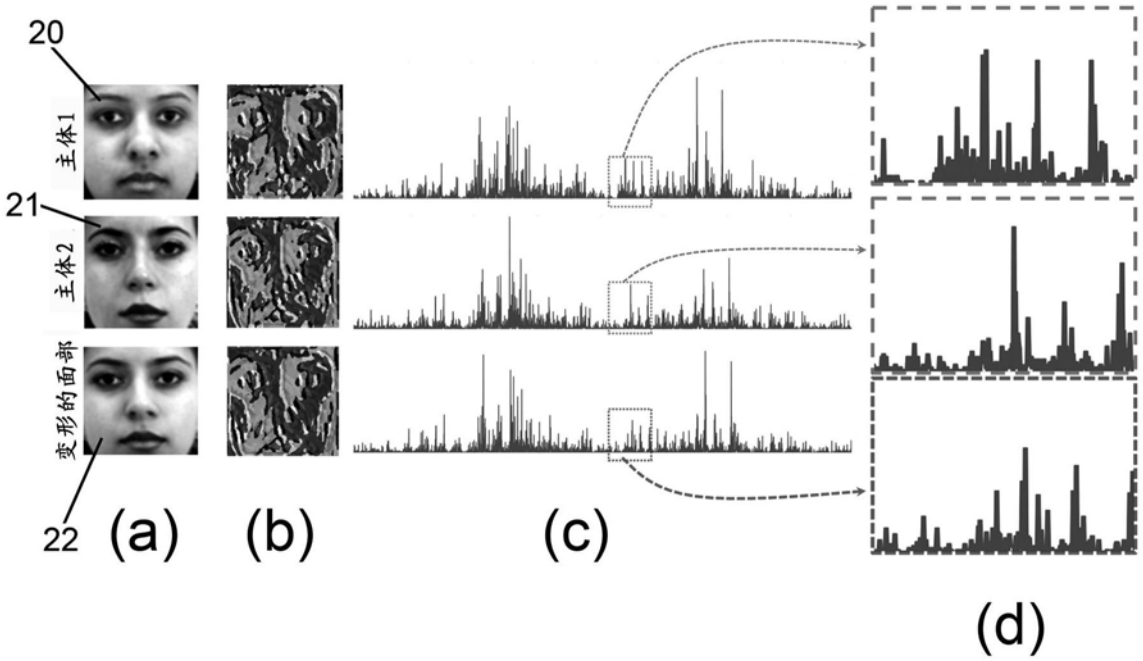

20
主体1
21
主体2
变形的面部
22
(a)
(b)
(c)
(d)

图2

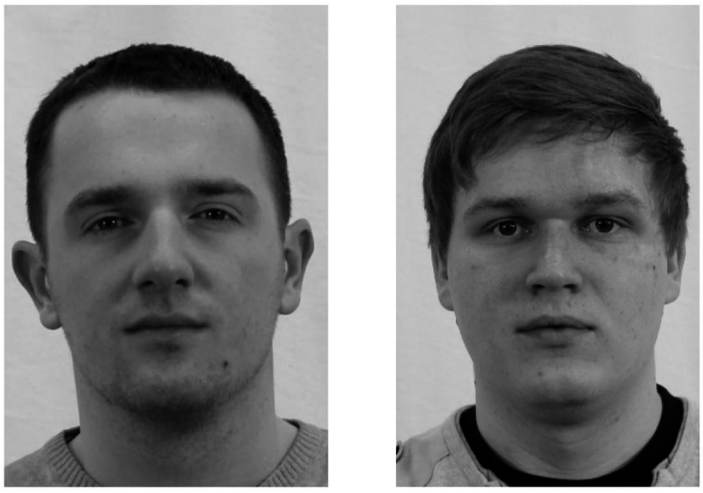


图3

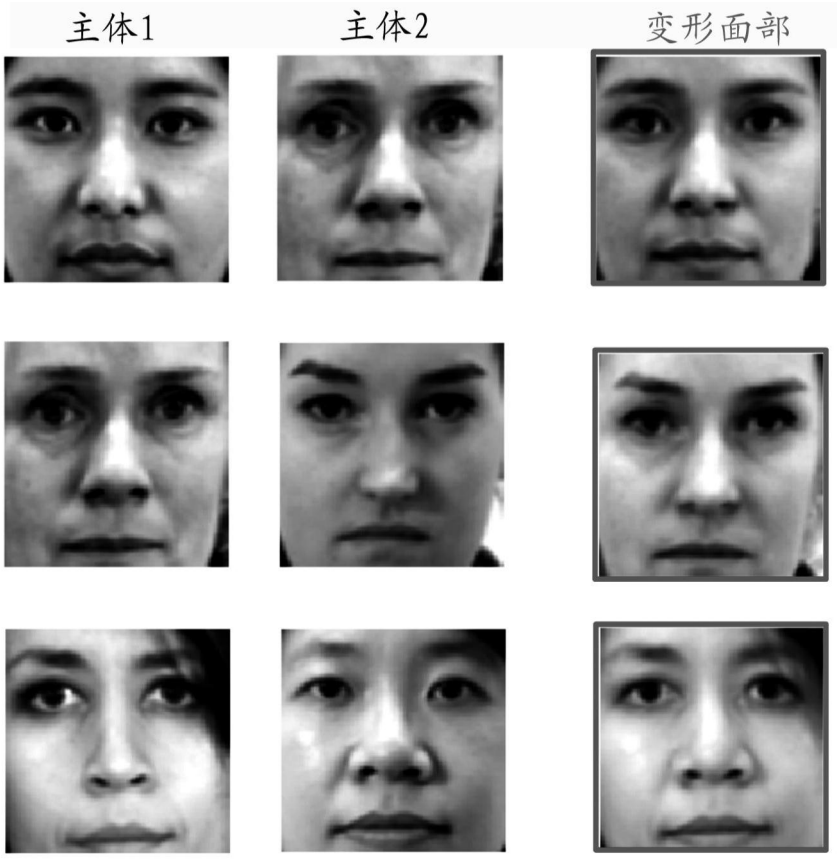


图4

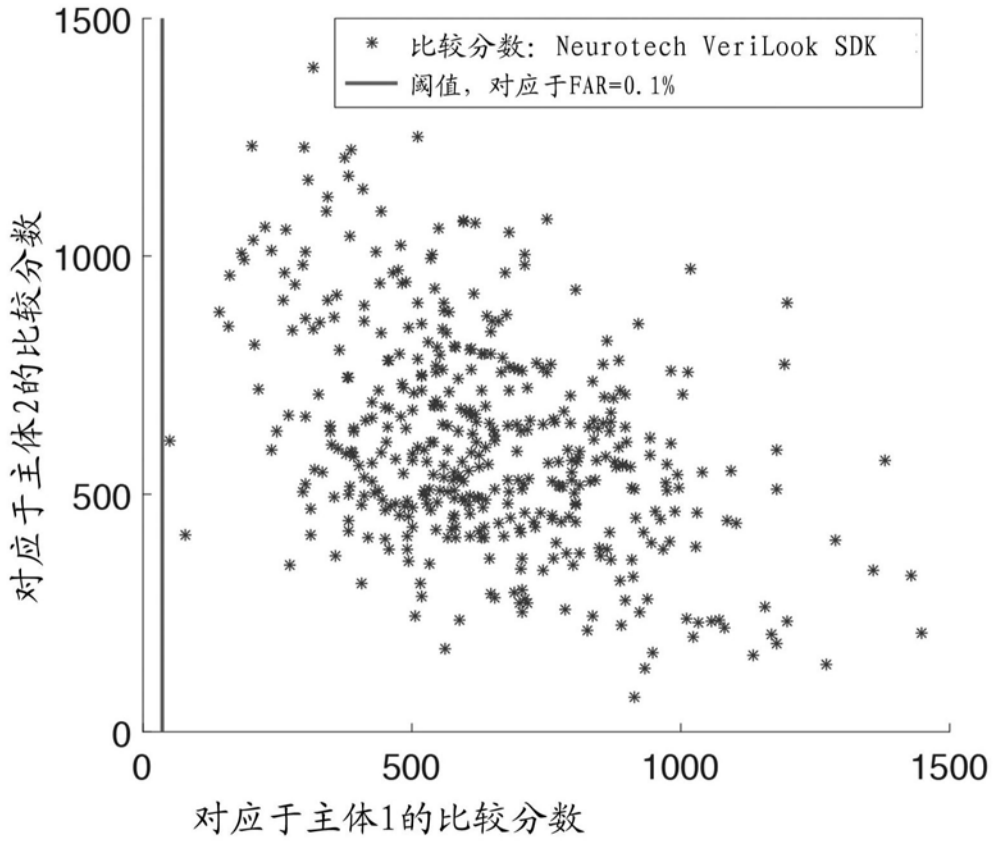

比较分数：Neurotech VeriLook SDK
阈值，对应于FAR=0.1%
对应于主体2的比较分数
对应于主体1的比较分数
0
500
1000
1500

图5

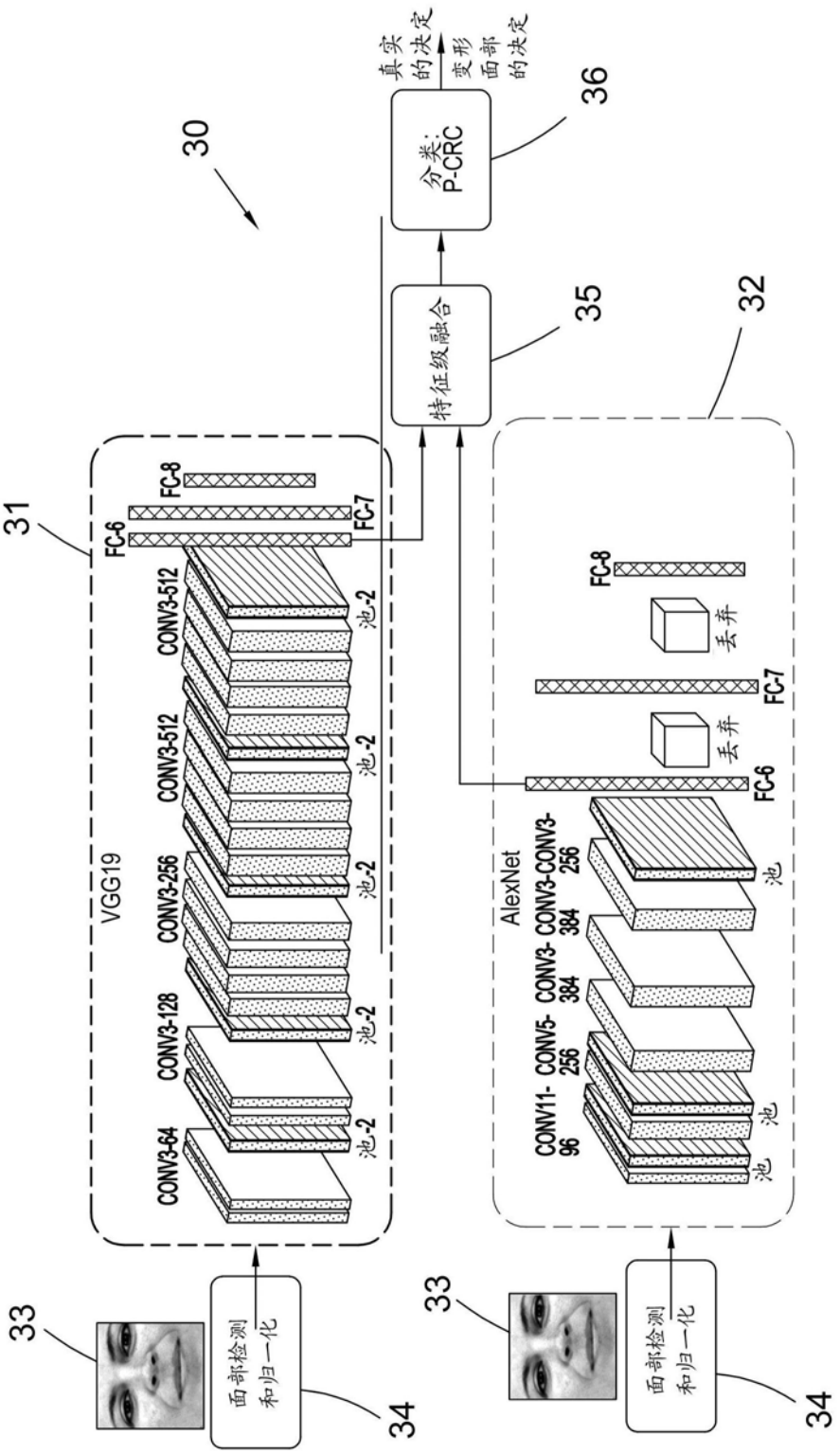

30
真实的决定
变形面部的决定
分类;
P-CRC
36
特征级融合
35
32
31
VGG19
FC-8
FC-6
FC-7
CONV3-512
池-2
CONV3-512
池-2
CONV3-256
池-2
CONV3-128
池-2
CONV3-64
池-2
AlexNet
FC-8
丢弃
FC-7
丢弃
FC-6
CONV3-CONV3-256
池
CONV3-384
CONV5-256
池
CONV11-96
池
33
面部检测和归一化
34
33
面部检测和归一化
34

图6

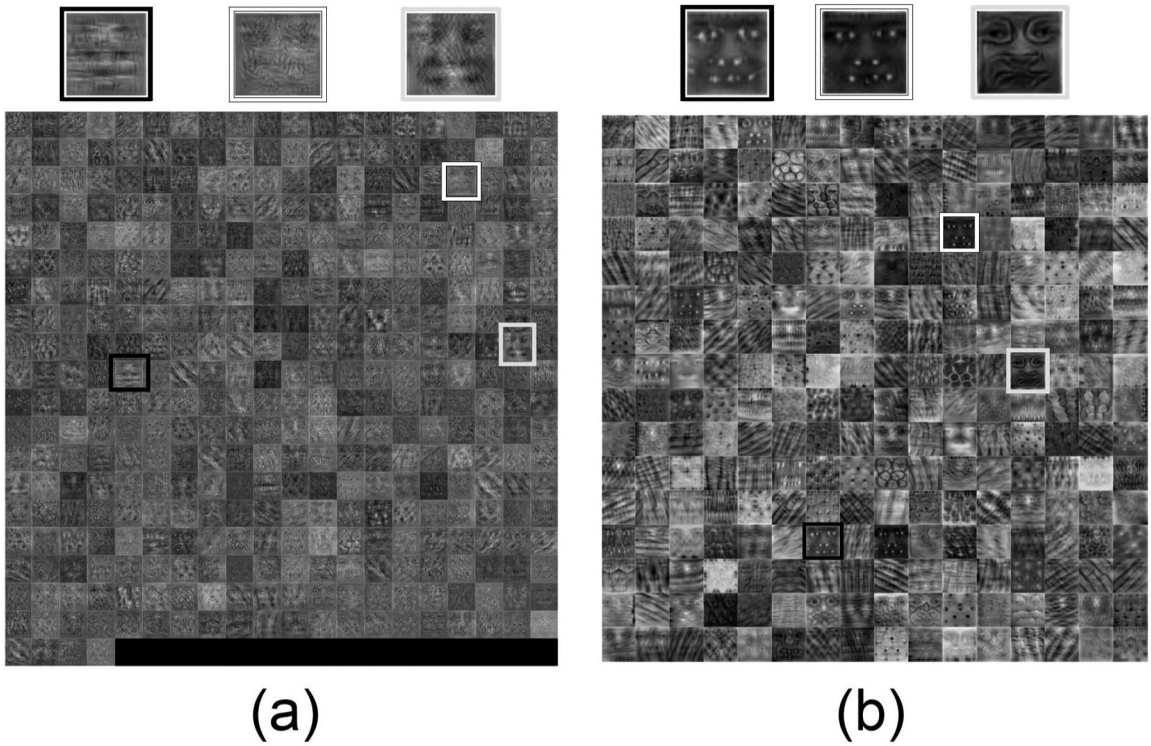


图7

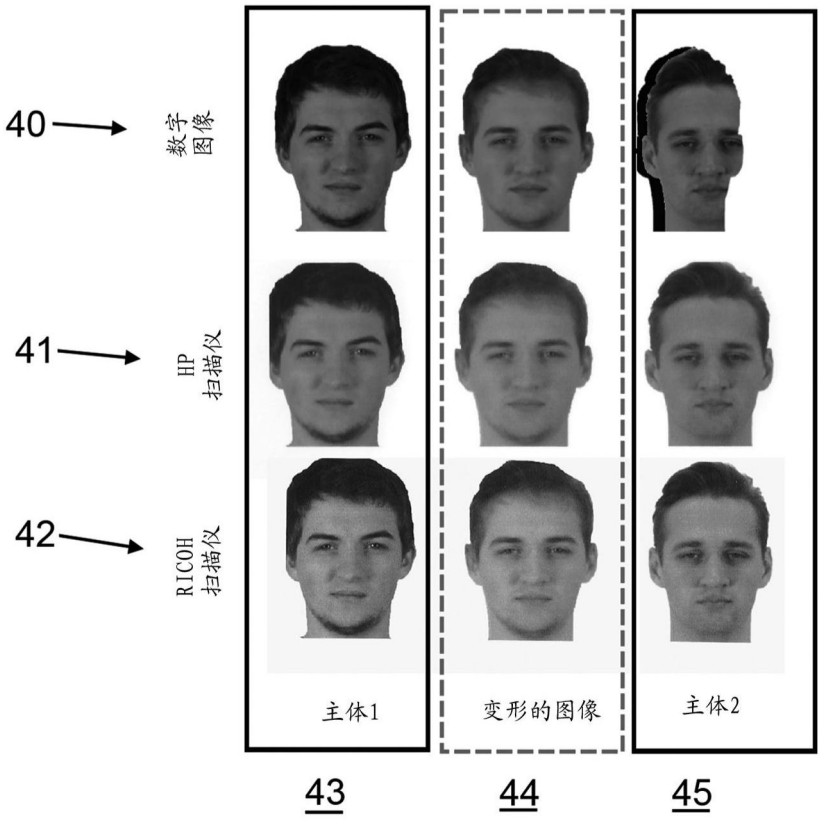


图8

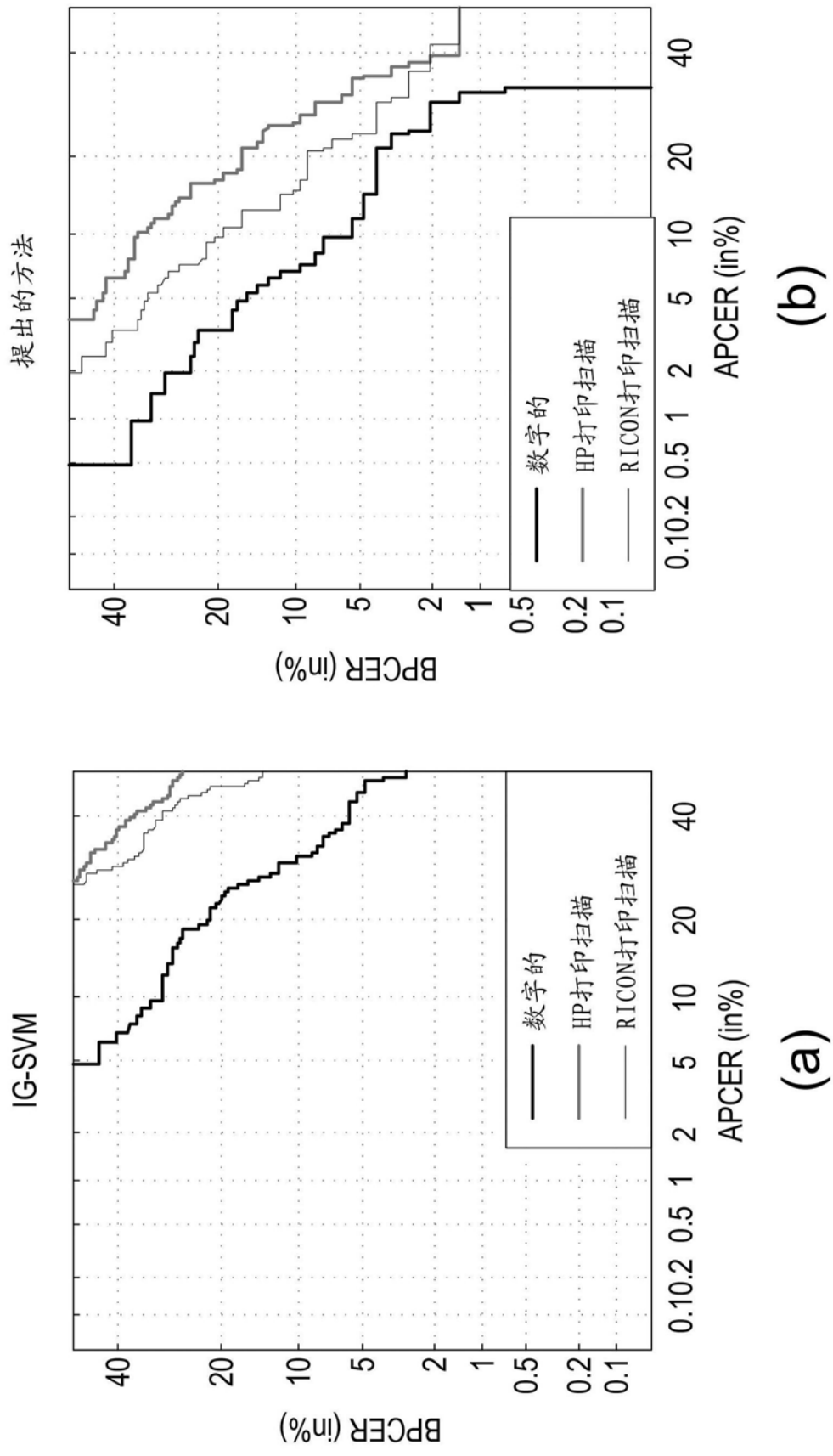


图9