



## Certificate of Invention Patent

Title: METHOD FOR CORRECTION OF THE EYES IMAGE USING MACHINE LEARNING AND METHOD FOR MACHINE LEARNING

Patentee: AUTONOMOUS NON-PROFIT ORGANIZATION FOR HIGHER EDUCATION "SKOLKOVO INSTITUTE OF SCIENCE AND TECHNOLOGY"

Address: Moscow, Russian Federation

Inventor(s): KONONENKO, Daniil Sergeyeovich;  
LEMPITSKY, Victor Sergeyeovich

Patent No.: ZL 202210276561.9      Patent publication No: CN 114863544 B

Patent Filing Date: March 03, 2016      Issue date: January 14, 2025

Applicant(s) recorded on filling date: AUTONOMOUS NON-PROFIT ORGANIZATION FOR HIGHER EDUCATION "SKOLKOVO INSTITUTE OF SCIENCE AND TECHNOLOGY"

Inventor(s) recorded on filling date: KONONENKO, Daniil Sergeyeovich;  
LEMPITSKY, Victor Sergeyeovich

After examining the present invention in accordance with the Patent Law of The People's Republic of China, the office makes a decision to grant the patent right for the invention and announce it.

The patent right for design shall take effect as of the issue date. Legal information such as the validity of the patent right and the change of the patentee shall be subject to the patent register.

Commissioner: SHEN, CHANGYU

National Intellectual Property Administration,  
PRC (Chop)  
January 14, 2025

*Our Ref.: 14091-ID-RUSI*

证书号第7665811号



专利公告信息

# 发明专利证书

发明名称：使用机器学习对眼睛图像进行校正的方法以及机器学习的方法

专利权人：高等教育自主非营利组织 斯科尔科沃科学和技术研究所

地址：俄罗斯联邦莫斯科

发明人：D·S·柯诺南科;V·S·兰皮特斯基

专利号：ZL 2022 1 0276561.9

授权公告号：CN 114863544 B

专利申请日：2016年03月03日

授权公告日：2025年01月14日

申请日时申请人：高等教育自主非营利组织 斯科尔科沃科学和技术研究所

申请日时发明人：D·S·柯诺南科;V·S·兰皮特斯基

国家知识产权局依照中华人民共和国专利法进行审查，决定授予专利权，并予以公告。  
专利权自授权公告之日起生效。专利权有效性及专利权人变更等法律信息以专利登记簿记载为准。

局长  
申长雨

申长雨





## (12) Patent for Invention

(10) patent publication No: CN 114863544 B  
(45) patent publication date: January 14, 2024

(21) Application No. 202210276561.9  
(22) Filing Date March 03, 2016  
(65) Published Document No. of the same Application  
Application publication No. CN 114863544 A  
(43) Publication Date August 05, 2022  
(30) Priority Data  
2015109868 March 20, 2015 RU  
(62) Data of the Parent Application:  
201680028919.9 March 03, 2016  
(73) Patentee AUTONOMOUS NON-PROFIT  
ORGANIZATION FOR HIGHER EDUCATION  
"SKOLKOVO INSTITUTE OF SCIENCE AND  
TECHNOLOGY"  
Address Moscow, Russian Federation  
(72) Inventor(s) KONONENKO, Daniil Sergeyeovich;  
LEMPITSKY, Victor Sergeyeovich  
(74) Agency Beijing Arete Intellectual Property Agency  
11269  
Agent WANG Wei YAN Shen

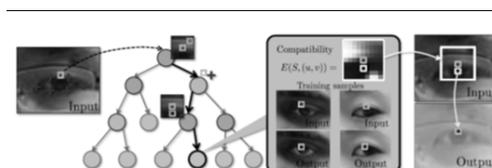
(51) Int. Cl.  
G06V 40/18(2022.01)  
G06V 10/56(2022.01)  
G06V 10/764(2022.01)  
G06V 10/70(2022.01)  
G06N 20/00(2019.01)  
(56) References  
US 2013070046 A1, 2013.03.21  
US 2018035886 A1, 2018.02.08

Examiner LIU Jiangping

Claims 2 pages; Description 6 pages;  
Drawings 3 pages

(54) Title:  
METHOD FOR CORRECTION OF THE EYES  
IMAGE USING MACHINE LEARNING AND  
METHOD FOR MACHINE LEARNING  
(57) Abstract:

The present invention refers to the method for correction of the eyes image using machine learning and method for machine learning. The present invention refers to automatics and computing technology, namely to the field of processing images and video data, namely to correction the eyes image of interlocutors in course of video chats, video conferences with the purpose of gaze redirection. A method of correction of the image of eyes wherein the method obtains, at least, one frame with a face of a person, whereupon determines positions of eyes of the person in the image and forms two rectangular areas closely circumscribing the eyes, and finally replaces color components of each pixel in the eye areas for color components of a pixel shifted according to prediction of the predictor of machine learning. Technical effect of the present invention is rising of correction accuracy of the image of eyes with the purpose of gaze redirection, with decrease of resources required for the process of handling a video image.





# (12) 发明专利

(10) 授权公告号 CN 114863544 B

(45) 授权公告日 2025. 01. 14

(21) 申请号 202210276561.9

(22) 申请日 2016.03.03

(65) 同一申请的已公布的文献号  
申请公布号 CN 114863544 A

(43) 申请公布日 2022.08.05

(30) 优先权数据  
2015109868 2015.03.20 RU

(62) 分案原申请数据  
201680028919.9 2016.03.03

(73) 专利权人 高等教育自主非营利组织 斯科  
尔科沃科学和技术研究所  
地址 俄罗斯联邦莫斯科

(72) 发明人 D·S·柯诺南科  
V·S·兰皮特斯基

(74) 专利代理机构 北京嘉和天工知识产权代理  
事务所(普通合伙) 11269  
专利代理师 王维 严谨

(51) Int. Cl.  
G06V 40/18 (2022.01)  
G06V 10/56 (2022.01)  
G06V 10/764 (2022.01)  
G06V 10/70 (2022.01)  
G06N 20/00 (2019.01)

(56) 对比文件  
US 2013070046 A1, 2013.03.21  
US 2018035886 A1, 2018.02.08  
审查员 刘江平

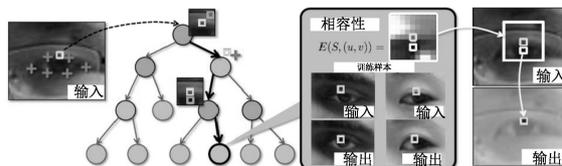
权利要求书2页 说明书6页 附图3页

## (54) 发明名称

使用机器学习对眼睛图像进行校正的方法  
以及机器学习的方法

## (57) 摘要

本发明涉及使用机器学习对眼睛图像进行校正的方法以及机器学习的方法。本发明涉及自动化和计算技术,即,涉及处理图像和视频数据的领域,即,涉及以目光重定向为目的对视频聊天、视频会议中的对话者的眼睛图像进行校正。一种校正眼睛图像的方法,其中该方法对人的脸至少获得一帧,于是确定这个人的眼睛在图像中的位置,并且形成紧密地包围眼睛的两个矩形区域,并且最后将眼睛区域中的每个像素的颜色分量取代成根据机器学习预测器的预测移位的像素的颜色分量。本发明的技术效果是以目光重定向为目的的眼睛图像的校正精度提高,其中处理视频图像的过程所需的资源减少。



1. 一种用于计算装置上的图像增强的方法,所述方法包括:  
接收描绘人脸的数字输入图像;  
经由目光调整机器学习模型,基于将所述数字输入图像中的眼睛中的每个像素的颜色分量取代成所述数字输入图像中的另一像素的颜色分量,通过改变所述人脸上眼睛的目光方向从所述数字输入图像生成经目光调整的图像,其中所述目光调整机器学习模型包括基于多个学习图像对的预测器,所述多个学习图像对不同之处在于目光方向;以及  
输出所述经目光调整的图像。
2. 如权利要求1所述的方法,其中所述预测器是单层或多层神经网络。
3. 如权利要求1所述的方法,其中所述预测器是至少一个决策树或决策树集合。
4. 如权利要求1所述的方法,其中所述学习图像对之间的目光方向相差在5至20度的范围内的角度。
5. 如权利要求1所述的方法,其中所述预测器生成取决于涉及眼睛的多个特性点的像素位置的固定位移矢量。
6. 如权利要求1所述的方法,其中所述预测器基于包含同一人的视图的多个学习图像对。
7. 如权利要求1所述的方法,其中对于所有的学习图像对,学习对中的每个图像中的目光方向的差异是相同的。
8. 如权利要求1所述的方法,其中所述方法还包括从存储所述预测器的信息介质加载所述预测器,所述信息介质包括下列中的一者:硬盘、固态驱动器、闪存、光学盘、混合驱动器、随机存取存储器(RAM)、远程计算机系统和远程数据存储。
9. 如权利要求1所述的方法,其中生成经目光调整的图像的步骤包含识别所述数字输入图像中的多个界标。
10. 如权利要求9所述的方法,其中所述生成经目光调整的图像的步骤还包含基于所述多个界标生成二维位移矢量场,所述二维位移矢量场针对所述数字输入图像中的一个或更多个像素中的每个像素指示所述像素的位移矢量,以及根据所述二维位移矢量场,通过使所述数字输入图像中的所述一个或更多个像素移位来生成所述经目光调整的图像。
11. 如权利要求1所述的方法,其中所述预测器通过下列步骤获得:  
获得多个学习图像对,所述多个学习图像对在每个图像对中包含同一人的视图,所述学习图像对不同之处在于目光方向,并且其中对于所述学习图像对,学习对中的每个图像中的目光方向的差异是相同的;  
确定所述学习图像对中的每个中的眼睛的位置;以及  
根据所述预测器的预测,通过对中的第一图像的每个像素的颜色分量取代成所述对中的所述第一图像的另一像素的颜色分量来确定用于对每个学习图像对产生调整位移矢量场的预测器,从而得到类似于所述对中的第二图像的图像。
12. 如权利要求1所述的方法,其中所述数字输入图像是包含多个帧的视频流中的一个帧。
13. 一种计算装置,所述计算装置包括:  
逻辑机器;以及  
存储机器,所述存储机器保存可由所述逻辑机器执行的指令,以便:

接收描绘人眼的数字输入图像；

经由目光调整机器学习模型,基于将所述数字输入图像中的所述人眼中的每个像素的颜色分量取代成所述数字输入图像中的另一像素的颜色分量,通过改变所述人眼的表观目光方向从所述数字输入图像生成经目光调整的图像,其中所述目光调整机器学习模型包括基于多个学习图像对的预测器,所述多个学习图像对不同之处在于目光方向;以及  
输出所述经目光调整的图像。

14. 如权利要求13所述的计算装置,其中生成经目光调整的图像的步骤包含基于所述数字输入图像中识别的多个界标生成二维位移矢量场,所述二维位移矢量场针对所述数字输入图像中的一个或更多个像素中的每个像素指示所述像素的位移矢量,以及根据所述二维位移矢量场,通过使所述数字输入图像中的所述一个或更多个像素移位来生成所述经目光调整的图像。

15. 一种用于计算装置上的图像增强的方法,所述方法包括:

加载预测器,所述预测器基于包含同一人的视图的多个学习图像对;

接收具有人的脸至少一个帧;

确定所述人的眼睛在所述帧中的位置,并且形成紧密包围所述眼睛的两个矩形眼睛区域;以及

将所述帧中的所述矩形眼睛区域中的每个像素的颜色分量取代成所述帧中的根据所述预测器的基于机器学习的预测的像素的颜色分量。

16. 如权利要求15所述的方法,其中对于所有的学习图像对,学习对中的每个图像中的目光方向的差异是相同的。

17. 如权利要求15所述的方法,其中对于所有的学习图像对,学习对中的每个图像中的目光方向的差异在5至20度的范围内。

18. 如权利要求15所述的方法,其中所述预测器是单层或多层神经网络。

19. 如权利要求15所述的方法,其中所述预测器是至少一个决策树或决策树集合。

20. 如权利要求15所述的方法,其中所述至少一个帧来自包含多个帧的视频流。

## 使用机器学习对眼睛图像进行校正的方法以及机器学习的方法

[0001] 本申请是2016年03月03日递交的PCT国际申请PCT/RU2016/000118于2017年11月17日进入中国国家阶段的中国专利申请号为201680028919.9、发明名称为“使用机器学习对眼睛图像进行校正的方法以及机器学习的方法”的发明专利申请的分案申请。

### 技术领域

[0002] 本组发明涉及自动化和计算技术,一般地涉及处理图像和视频数据的领域,即,涉及校正视频聊天、视频会议的过程中的对话者的眼睛的图像。

### 背景技术

[0003] 目前,商业伙伴之间的视频会议以及在国外给亲戚打视频电话变为日常发生的事。特定软件和硬件的选择被与各方之间缺少视觉接触有关的复杂性所阻碍。该问题长时间吸引了研究者和工程师,并且它是由捕捉视频图像的摄像头的位置和脸在屏幕上的图像之间的不可避免的差异引起的。

[0004] 最成功的目前已知的解决方案需要额外的设备,除了网络摄像头之外,比如:半透明镜/屏幕、立体摄像头或RGB-D摄像头。

[0005] 在现有技术中已知申请人Ramot At Tel-Aviv University Ltd.的于2011年12月1日公布的专利申请W02011148366“用于校正目光偏移的方法和系统(Method and system for correcting gaze offset)”。该方法包括以下步骤:对图像进行处理以便提取至少一个眼睛在该图像上的位置;对该图像进行处理以用取代数据取代与每个眼睛的每个位置相关联的影像数据;并且将所述校正后的图像发送到显示装置。该给定方法的缺点是,首先,在视频会议开始之前必需对视频会议的每个参与者预先记录朝向摄像头的目光的一组影像数据,第二,在视频会议期间目光方向的不自然固定。

[0006] 从现有技术还已知作者D.Giger、J.C.Bazin、a C.Kuster、T.Popa、M.Gross发表于Proceedings of IEEE ICME 2014(中国,成都,2014年7月14-18日)中的文章“用单个网络摄像头进行目光校正(Gaze Correction with a Single Webcam)”。该指定方法包括:确定脸部特征点并且根据确定的脸部特征来匹配人头的几何模型,将图像投影到调整后的模型的纹理,将模型旋转到某个角度并且将转动的模型投影在图像中,匹配模型图像和源图像。该给定方法的限制是:整体脸比例变形,必需预先记录头部的被屏蔽部分(例如,下巴)的纹理,并且需要图形加速器来实现实时系统操作所必需的生产率。

### 发明内容

[0007] 给定的一组发明的目的是在视频会议期间使用仅包括一个视频记录工具(视频摄像头)的装置来校正对话者的眼睛的图像。

[0008] 给定发明的技术效果是,眼睛图像的校正精度提高,同时减少了处理视频帧的过程所需的资源。

[0009] 该技术效果是由于(用于校正图像中的目光方位的)预测器的机器学习方法而实现的,其中该方法首先获得多个图像对,在每对中包含同一个人的图像,但是多个图像对的不同之处仅在于目光方向,然后确定眼睛在每个图像对中的位置,接着学习预测器,生成调整位移矢量场,以使得对于每个图像对,将该对中的第一图像的每个像素中的颜色分量取代成该对中的第一图像的根据预测器的预测移位的一像素的颜色分量,得到尽可能地类似于该对中的第二图像的图像,以及将预测器作为结果保存。

[0010] 本发明提供了一种用于校正图像中的目光方位的预测器的机器学习方法,其特征在于:

[0011] • 获得多个图像对,在每对中包含同一个人的图像,所述多个图像对的不同之处仅在于目光方向;

[0012] • 确定眼睛在每个图像对中的位置;

[0013] • 学习所述预测器,生成调整位移矢量场,以使得对于每个图像对,将所述对中的第一图像的每个像素中的颜色分量取代成所述对中的所述第一图像的根据所述预测器的预测移位的一像素的颜色分量,得到尽可能地类似于所述对中的第二图像的图像;

[0014] • 保存所述预测器。

[0015] 位移场的预测器可以采取以下形式:

[0016] -单层或多层神经网络。

[0017] -至少一个决策树或决策树集合。

[0018] -生成仅取决于涉及眼睛的特征点的像素位置的固定位移矢量的预测器。

[0019] -机器学习中已知的任何其他形式的预测器。

[0020] 使图像对中的眼睛区域为预设像素尺度。

[0021] 预测器被存储在信息介质中,所述信息介质包括硬盘或固态驱动器、或闪存、或光学盘、或混合驱动器、或随机存取存储器(RAM)、或远程计算机系统、或远程数据存储。

[0022] 此外,指定的技术效果是用用于校正眼睛在眼睛图像中的位置的方法来实现的,其中该方法首先加载预测器,然后对人的脸至少获得一个帧,然后确定这个人的眼睛在图像中的位置并且形成紧密包围眼睛的两个矩形区域,并且结果将眼睛区域中的每个像素的颜色分量取代成根据预测器的预测移位的像素的颜色分量。

[0023] 本发明提供了一种用于校正眼睛图像的方法,其特征在于:

[0024] • 加载预测器;

[0025] • 对人的脸至少获得一个帧;

[0026] • 确定所述人的眼睛在所述图像中的位置,并且形成紧密包围所述眼睛的两个矩形区域;

[0027] • 将所述眼睛区域中的每个像素的颜色分量取代成根据所述预测器的基于机器学习的预测移位的像素的颜色分量。

[0028] 预测器是根据上述标准训练的(以对于每个训练对中的第二图像最小化实际图像和获得的图像之间的差异)。

[0029] 预测器从信息介质被加载,所述信息介质可以包括硬盘或固态驱动器、或闪存、或光学盘、或混合驱动器、或随机存取存储器(RAM)、或远程计算机系统、或远程数据存储。

[0030] 预测器是单层或多层神经网络。

- [0031] 预测器是至少一个决策树或决策树集合。
- [0032] 预测器生成仅取决于涉及眼睛的特性点的像素位置的固定位移矢量。
- [0033] 在训练应用于眼睛区域中的每个像素的预测器之后,应满足以下要求:
- [0034] a. 当被应用于像素时,预测器应生成调整位移矢量,该调整位移矢量用于用由该给定位移矢量确定的另一像素的颜色分量取代给定像素的颜色分量;
- [0035] b. 预测器在多个图像对上被训练,其中一对中的一个图像包含人脸在眼睛被调整之前的初始图像,并且另一图像包含眼睛凝视不同方向的人的图像。
- [0036] 应用于所述眼睛区域中的每个像素的预测器满足以下要求:
- [0037] a. 被应用于像素,预测器应生成调整位移矢量,所述调整位移矢量用于用由所述给定位移矢量确定的取代像素的颜色分量取代所述给定像素的颜色分量;
- [0038] b. 所述预测器在多个图像对上被训练,其中一对中的一个图像包含所述人脸在所述眼睛的位置被调整之前的初始图像,并且另一图像包含具有期望眼睛位置的人的图像。
- [0039] 预测器预测的位移矢量根据调整后的图像中的眼睛的大小与训练图像的比率而被缩放。

#### 附图说明

- [0040] 图1示出对话者的眼睛图像的实时校正。左侧是目光朝向摄像头下面十度的输入帧。中间是目光朝向比左侧帧中高出十度的输入帧(在这种情况下,目光朝向摄像头)。右侧帧是本发明的效果的实例,在该实例中,图像中的对话者的目光方向被相对于输入图像(在左侧示出)向上调整十度。
- [0041] 图2示出预测器是决策树的情况下的像素处理。
- [0042] 图3示出用于校正眼睛图像的预测器的机器学习方法的框图。
- [0043] 图4示出用于校正眼睛图像的方法的框图。

#### 具体实施方式

- [0044] 下面将描述本发明的详细公开所需的概念和定义。
- [0045] 预测器是用于与给定的某些对象(例如,类别标签、标量或矢量值)的样本进行比较的程序或硬件实现,该程序或硬件实现取决于借助于针对训练组的机器学习过程而被训练的相当多的参数。决策树、神经网络、决策树集合可以被作为预测器架构的实施例给出。根据预测器,参数可以包括例如神经网络的权重、决策树中的阈值等。
- [0046] 预测器的内部状态是预测器的影响其预测的所有参数值的集合。
- [0047] 决策树是以分层连续结构的表示规则的方法,其中唯一节点(树叶)针对每个对象给予解决方案。
- [0048] 颜色分量是定义用于描述颜色表示的颜色模型的数字元组中的三个或四个值。
- [0049] 存在许多各种类型的颜色模型,但是计算机图形作为规则使用以下颜色空间:CMY、CMYK、CMYK256、RGB、HSB、HLS、L\*a\*b、YIQ、Grayscale(灰色程度)。
- [0050] 例如,在最知名的颜色模型RGB中,元组包含三个颜色分量。例如,(255,64,23)是包含强红色分量、较小量的绿色以及甚至更小量的蓝色的颜色。
- [0051] 给定发明可以在系统形式的计算机上实现,或者在包含用于执行前述方法的指令

的机器可读介质上实现。

[0052] 用于校正图像中的目光方位的预测器的机器学习方法包括以下步骤:

[0053] 得到多个图像对,在每对中包含同一个人的图像,这些图像对的不同之处仅在于目光方向;

[0054] 多个图像对302(图3)来自于视频摄像头301(图3)。为了提供机器学习的可能性,需要满足以下要求的预定图像对集合 $(I_j, O_j)$ :每个图像对表示同一个人的头部位置相同并且可视化条件相等的脸,不同之处仅在于这个人的目光方向(图1)。目光方向的差异在所有学习对中都是相等的。为所述方法的正确操作,目光方向的给定差异应为5度至20度。

[0055] 针对每个图像对限定眼睛的位置;

[0056] 受控梯度下降算法(其描述可以在参考文章[1]中找到)例如可以被用来确定眼睛的位置。基于决策树的连续应用的方法也可以被使用。实现指定方法的各种替换方案在文章[2]和[3]中被描述。

[0057] 此外,使图像为相等大小。对于与右眼匹配每个图像对,聚焦于与图像 $I^j$ 中的这个眼睛匹配的点 $(f_1, g_1)$ 、 $(f_2, g_2) \cdots (f_N, g_N)$ 上。在文章[1]中,点数 $N$ 为6个。显见的是,它不是固定参数。此外,确定每个眼睛的轴对齐矩形 $B'$ 边界点,并且将特性半径 $\Delta$ 定义为 $\sqrt{\text{Area}(B')}$ 。此外,考虑具有在与 $B'$ 相同位置上的中心的矩形 $B$ ,还有与 $\Delta$ 成比例的宽度 $W$ 和高度 $H$ (即,对于一些常数 $\alpha$ 和 $\beta$ , $W = \alpha * \Delta$ , $H = \beta * \Delta$ )。常数 $\alpha$ 和 $\beta$ 是在1至20的范围内任意选择的。在进行的实验中,值 $\alpha$ 和 $\beta$ 相应地为2.0和1.6。由此,矩形 $B$ 对于尺度和眼睛位置是共变的,并且具有边长比 $\alpha:\beta$ 。根据矩形 $B$ 修剪图像 $I^j$ 和 $O^j$ ,并且将这些图像缩放 $R/\Delta$ 倍,以使得图像中的眼睛的特性半径变为常数。在相对于垂直轴规则反射之后,类似地对左眼图像进行处理。结果,获得多个被修剪和缩放的图像对303(图3)。

[0058] 学习预测器,生成调整位移矢量,以使得对于每个图像对,将该对中的第一图像的每个像素中的颜色分量取代成该对中的第一图像的根据预测器的预测位移的另一像素的颜色分量得到尽可能地类似于该对中的第二图像的图像;

[0059] 在每个像素 $(x, y)$ 中,值 $O^j(x, y)$ 借助于运算 $O(x, y) = I(x + u(x, y), y + v(x, y))$ 而被值 $I^j(x, y)$ 取代。由此,边界矩形 $B$ 内的每个像素 $(x, y)$ 指定元组 $S = [(x, y), I, [f_i, g_i], O(x, y)]$ 的学习,该元组包括 $(x, y)$ 像素位置、输入图像 $I$ 、眼睛轮廓上的点 $[f_i, g_i]$ 以及输出图像中的像素的颜色 $O(x, y)$ 。然后,基于教育元组(学习样本)来学习决策树。

[0060] 每个学习样本不包括位移矢量 $(u(x, y), v(x, y))$ ,而是仅包括所需颜色 $O(x, y)$ ,而相同的颜色分量可以借助于各种偏移来获得。

[0061] 眼睛图像通过在眼睛区域中的每个像素中应用机器训练的预测器304(图3)来调整,并且预测器 $P$ 通过以下两个性质来确定:

[0062] (1) 在预测器 $P$ 对像素进行处理时,预测位移矢量,该位移矢量根据校正过程被用来用移位后的像素的强度(颜色分量)取代该像素的强度(颜色分量)。

[0063] (2) 在训练期间,对于所述多个图像对 $(im1, im2)$ ,应用 $P$ 应导致 $im1$ 转换为 $im2$ ,其中 $im1$ 包含输入图像, $im2$ 包含所需图像。

[0064] 保存预测器;

[0065] 根据预测器的类型(决策树、决策树集合、神经网络),给定预测器的参数被存储在

信息介质305(图3)中,信息介质305是硬盘或固态驱动器、或光学驱动器、或闪存。此外,预测器可以被保存在随机存取存储器(RAM)中。

[0066] 例如,如果决策树被用作预测器,则参数可以是:测试的特性、阈值 $\tau$ 、节点之间的连接、叶数、树的深度、节点的值、树的整个拓扑。

[0067] 校正眼睛图像的方法包括:

[0068] 加载预测器;

[0069] 根据预测器404的类型(决策树、决策树集合、神经网络),给定预测器的参数从信息介质403(图4)被加载,信息介质403可以是硬盘或固态驱动器、或光学驱动器、或闪存。此外,预测器可以从随机存取存储器(RAM)被加载。

[0070] 例如,如果决策树被用作预测器,则参数可以是:测试的特性、阈值 $\tau$ 、节点之间的连接、叶数、树的深度、节点的值、树的整个拓扑。

[0071] 获得人的脸的至少一个帧;

[0072] 图像可以以实时地来自视频数据源401(图4)、或来自存储器、来自本地视频服务器或中央服务器的流视频的形式被发送。为了发送流视频,标准协议RTSP(RealTimeStreamingProtocol)、RTMP(RealTimeMessagingProtocol)、HLS(HTTPLiveStreaming)和DASH(DynamicAdaptiveStreamingover HTTP)可以被使用。由此,发送的视频数据的速度和质量可以自动地适应装置通信信道。

[0073] 视频数据可以以压缩形式被发送,例如,借助于编码器H.264、VP8、MJPEG、JPEG、JPEG2000。

[0074] 视频数据可以以单独的文件的形式被发送。由此,标准容器(例如,WebM、OGV、MKV、MP4、TS、JPG及其他)可以被使用。

[0075] 视频数据可以在无线网上被发送,比如GSM(全球移动通信系统)、CDMA(码分多址)、LTE(长期演进)、Wi-Fi(无线保真)。在本发明的一些实现中,获得和/或发送数据是使用上述几种技术或接收/发送技术来执行的。

[0076] 限定图像中的人眼的位置以及紧密地包围眼睛的两个矩形区域;

[0077] 每个眼睛聚焦于眼睛轮廓上分配的匹配该眼睛的点 $(f_1, g_1)$ 、 $(f_2, g_2) \cdots (f_N, g_N)$ 上。此外,确定每个眼睛的轴对齐矩形边界 $B'$ 点,并且将特性半径 $\Delta$ 定义为 $\sqrt{Area(B')}$ 。此外,考虑具有在与 $B'$ 相同位置上的中心的矩形 $B$ ,还有与 $\Delta$ 成比例的宽度 $W$ 和高度 $H$ (即,对于一些常数 $\alpha$ 和 $\beta$ , $W = \alpha * \Delta$ , $H = \beta * \Delta$ )。由此,矩形 $B$ 对于尺度和眼睛位置是共变的,并且具有边长比 $\alpha:\beta$ 。

[0078] 用根据预测器的预测移位的像素的颜色分量取代眼睛区域中的每个像素的颜色分量;

[0079] 在确定围绕眼睛的边界矩形405(图4)之后,改变像素的颜色分量以用于重定向目光。该方法的给定步骤是借助于通过使用预测器404(图4)的机器学习来实现的。作为匹配输入图像和学习集中的图像的矩形中的像素的结果,在 $(x, y)$ 像素坐标处获得二维(2-D)位移矢量 $(u(x, y), v(x, y))$ 。通过公式 $O(x, y) = I(x+u(x, y), y+v(x, y))$ 来计算输出图像中的像素 $O(x, y)$ 的最终值。由此,对于围绕眼睛的边界矩形405(图4)的每个像素获得一组二维偏移矢量406(图4)。

[0080] 更详细地说(图2),预测器(在这种情况下是决策树)将每个像素 $(x, y)$ 与学习数据

进行匹配,由此确定偏移。当像素被传递通过给定树时,它顺序地经过两种类型的测试。在随机化决策树的每个非叶节点中,确定应用于像素的测试的类型。第一类型的测试(外观测试)是由通道 {R;G;B} 中的一个中的偏移 (dx, dy) 以及阈值  $\tau$  确定的,并且在该测试的帧内,该颜色通道中的像素的两个值的差值被如下与阈值进行比较。

$$[0081] \quad I(x+dx, y+dy)[c] - I(x, y)[c] \begin{matrix} > \\ < \end{matrix} \tau$$

[0082] 第二类型的测试(用于像素相对于特征点的位置的测试)是由特征点的数量  $i \in \{1, \dots, N\}$  和阈值  $\tau$  确定的,并且在该测试的帧内,  $x - f_i$  或  $y - g_i$  被如下与阈值  $\tau$  进行比较:

$$[0083] \quad x - f_i \begin{matrix} > \\ < \end{matrix} \tau \text{ 或 } y - g_i \begin{matrix} > \\ < \end{matrix} \tau$$

[0084] 每个叶子包含未规一化的偏移误差分布,该分布被作为相容性图引用。在相容性卡片的每个像素中,总结了输出图像中的像素的真实颜色分量和输入图像中的就 (u, v) 移位的像素的颜色分量之间的差值。如果该差值足够小,则可以用输入图像中的就 (u, v) 移位后的像素来接近输出图像中的像素。

[0085] 所应用的测试的总数取决于决策树的深度,并且当前应用的测试的类型取决于我们此刻所在的节点类型。

[0086] 为了提高校正精度,可以应用一组几个独立被学习的决策树。对于给定像素,每个树预测相容性图(其被记录在树的对应叶子中)。然后将来自不同树的相容性图加在一起,这使得可以更精确地估计对于给定类型的像素典型的相容性图。对于聚合的相容性图,选择对于给定类型的像素最相容的移位(作为图上的最小值的位置)。该值 (u, v) 被用来进行恢复  $O(x, y) = I(x+u(x, y), y+v(x, y))$  时的运算。

[0087] 本详述说明书是通过呈现没有限制性和穷举性特征的各种实施方案来撰写的。同时,对于本领域技术人员,显见的是,在本发明的范围内,这里所公开的(还有部分地公开地)任何实施方案的各种取代、修改或组合可以被再现。由此,意图是且清楚的是,本发明的目前的说明书包括实施方案的其本质在这里未以明确表达的形式陈述的另外的替换方案。参照这里给出的并且没有限制性特征的实施方案,这样的实施方案可以在例如任何动作、部件、装置、性质、方面、性能、约束等的组合、修改或转换的结果中被获得。

[0088] 参考文献

[0089] 1. X. Xiong 和 F. De la Torre. Supervised descent method and its applications to face alignment (受监控下降方法及其对人脸对其的应用). Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference, 第532-539页, 2013年 IEEE。

[0090] 2. B. A. Smith, Q. Yin, S. K. Feiner 和 S. K. Nayar. Gaze locking: passive eye contact detection for human-object interaction (目光锁定: 针对人类对象交互的被动眼睛接触检测). Proceeding of the 26<sup>th</sup> annual ACM symposium on User interface software and technology, 第271-280页, 2013年 ACM。

[0091] 3. S. Ren, X. Cao, Y. Wei 和 J. S. 0001. Face alignment at 3000fps via regressing local binary feature (经由回归本地二进制特征的在3000fps的人脸对齐). CVPR, 第1685-1692页, 2014年。



图1

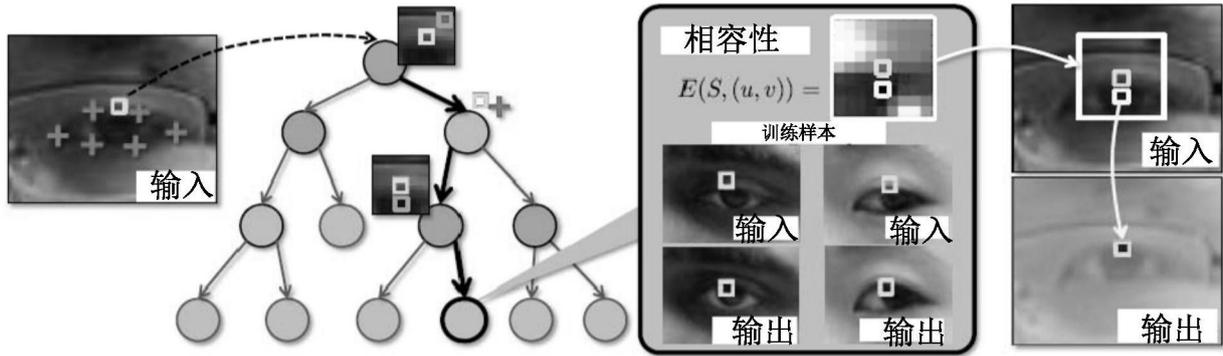


图2

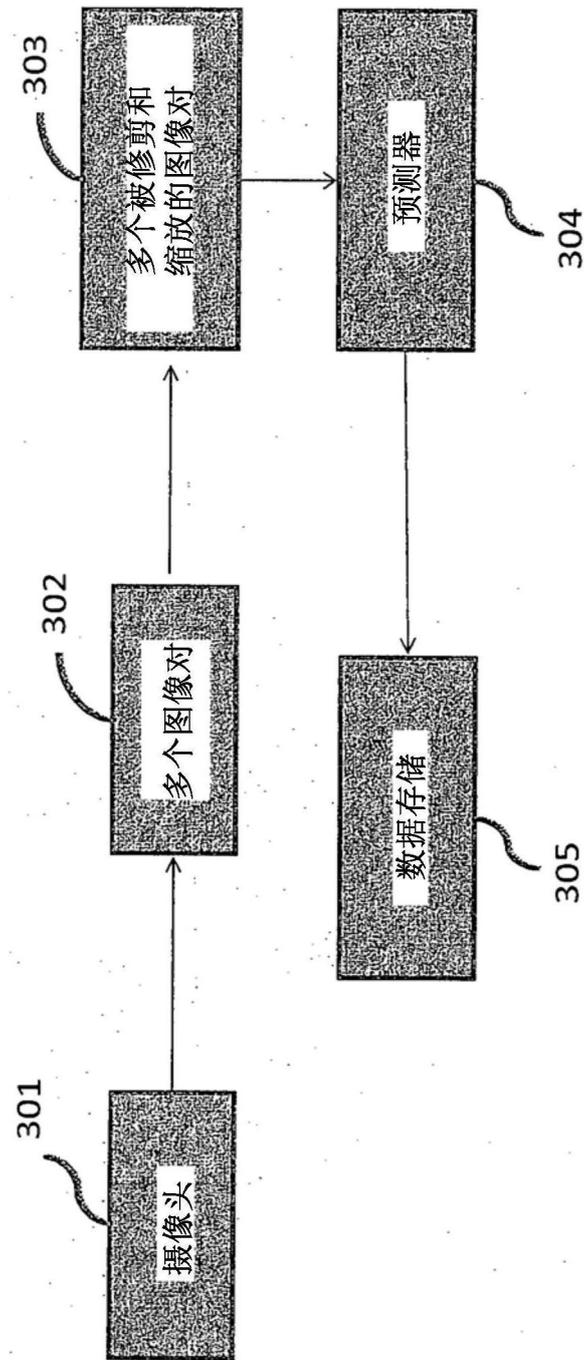


图3

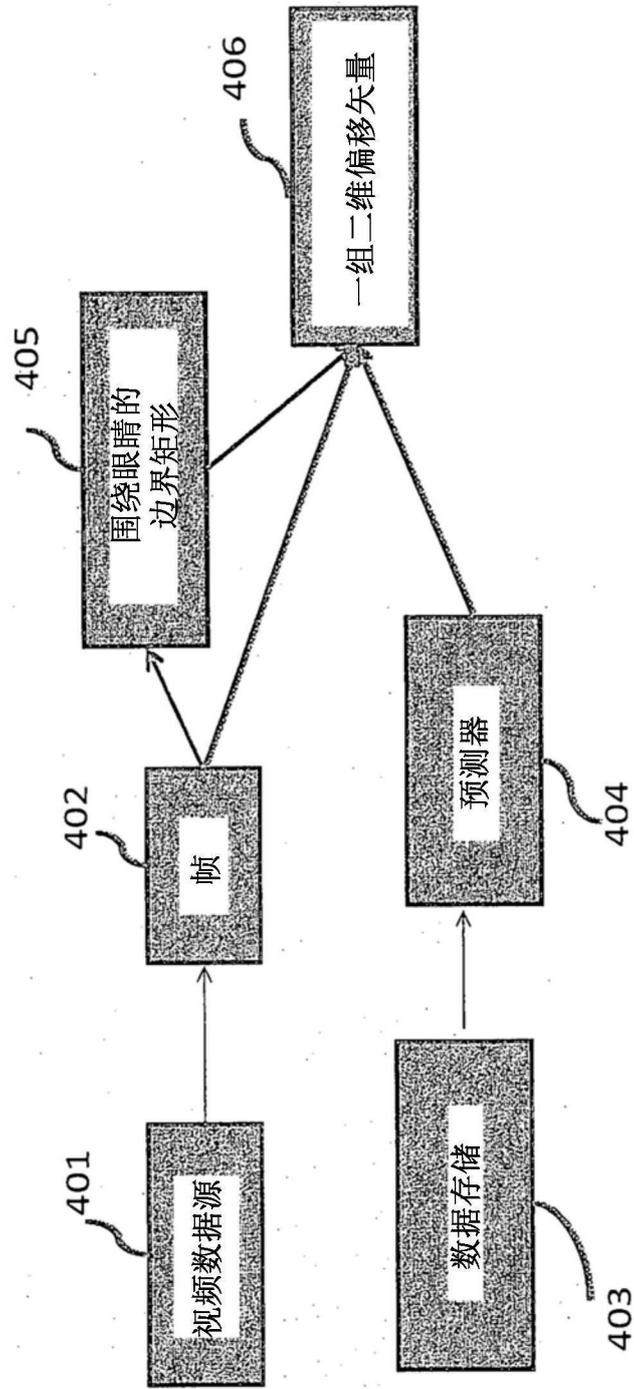


图4