



# (12)发明专利申请

(10)申请公布号 CN 108345869 A

(43)申请公布日 2018.07.31

(21)申请号 201810195342.1

(22)申请日 2018.03.09

(71)申请人 南京理工大学

地址 210094 江苏省南京市孝陵卫200号

(72)发明人 刘柯柯 刘亚洲 孙权森

(74)专利代理机构 南京理工大学专利中心

32203

代理人 王玮

(51)Int.Cl.

G06K 9/00(2006.01)

G06K 9/62(2006.01)

G06T 15/00(2011.01)

G06T 13/40(2011.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

权利要求书3页 说明书6页 附图4页

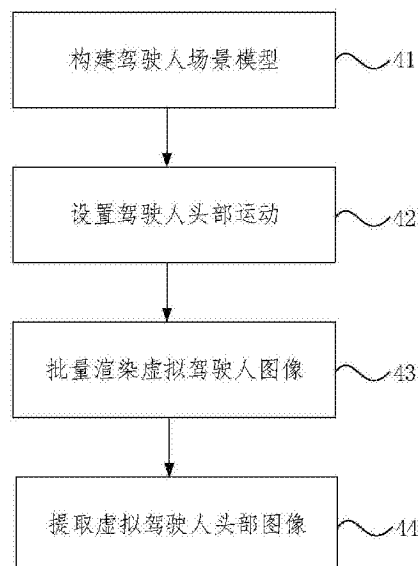
## (54)发明名称

基于深度图像和虚拟数据的驾驶人姿态识别方法

## (57)摘要

本发明公开了一种基于深度图像和虚拟数据的驾驶人姿态识别方法,驾驶人关节检测:将驾驶人深度图像和关节点标签作为深度学习框架Caffe的输入,训练深度卷积神经网络模型;利用深度学习模型检测图像中驾驶人的关节点位置。虚拟驾驶人头部姿态数据集构建:利用三维建模软件的建模功能Modeling构建驾驶人模型;利用动画功能Animation设置驾驶人头部运动;利用渲染功能Rendering批量渲染驾驶人图像;对驾驶人图像进行头部分割处理,构建虚拟驾驶人头部姿态数据集。驾驶人头部姿态分析:利用由虚拟驾驶人头部姿态数据集构建过程获取的虚拟驾驶人头部姿态数据,采用迁移学习的方法,训练深度迁移学习模型;根据深度迁移学习模型,估计目标域图像中驾驶人头部姿态,并检测其面部特征点位置。

CN 108345869 A



1. 一种基于深度图像和虚拟数据的驾驶人姿态识别方法,其特征在于:包括驾驶人关节检测、虚拟驾驶人头部姿态数据集构建和驾驶人头部姿态分析三个过程;

驾驶人关节检测过程,包括以下步骤:

1) 将驾驶人深度图像和关节点标签作为深度学习框架Caffe的输入,训练深度卷积神经网络模型;

2) 利用检测正确率最高的深度学习模型,检测图像中驾驶人的关节点位置;

虚拟驾驶人头部姿态数据集构建过程,包括以下步骤:

3) 利用三维建模软件的建模功能Modeling构建驾驶人模型;

4) 利用三维建模软件的动画功能Animation设置驾驶人头部运动;

5) 利用三维建模软件的渲染功能Rendering批量渲染驾驶人图像;

6) 对驾驶人图像进行头部分割处理,构建虚拟驾驶人头部姿态数据集;

驾驶人头部姿态分析过程,包括以下步骤:

7) 利用由虚拟驾驶人头部姿态数据集构建过程获取的虚拟驾驶人头部姿态数据,采用迁移学习的方法,训练深度迁移学习模型;

8) 根据深度迁移学习模型,估计目标域图像中驾驶人头部姿态,并检测其面部特征点位置。

2. 根据权利要求1所述的驾驶人姿态识别方法,其特征在于:所述步骤1)中驾驶人深度图像通过双目视觉设备或者红外结构光投影设备得到,在图像上的关节点设置标签得到关节点标签。

3. 根据权利要求1或2所述的驾驶人姿态识别方法,其特征在于:所述步骤1)包括以下具体步骤:

11) 将驾驶人深度图像和关节点标签转化为HDF5格式,作为深度学习框架Caffe的输入,调整驾驶人深度图像大小为 $120 \times 96$ ,关节点标签是标定的关节点像素坐标,用(x,y)表示;数据转化工具是Python开源库h5py;

12) 采用多个卷积层、池化层、全连接层构建深度卷积神经网络;卷积神经网络采用端到端的模型训练方式,通过传统的梯度下降方法训练模型中的参数,自动学习图像中具有层次结构的特征;网络基本结构由输入层、多次交替的卷积层和池化层、一个或者多个全连接层以及输出层构成;

13) 采用平方差损失函数微调深度卷积神经网络,训练深度学习模型;在输出层应用平方差损失函数,通过反向传播方式微调深度卷积神经网络,直至收敛。

4. 根据权利要求1或2所述的驾驶人姿态识别方法,其特征在于:所述步骤2)包括以下具体步骤:

21) 从所有深度学习模型中,根据所有测试结果,选择具有最大检测正确率的模型,将检测正确率最高的模型作为驾驶人关节点深度模型;

22) 利用选择的模型,检测图像中驾驶人的关节点位置,包括驾驶人头部、颈部、腰部、右肩关节、右肘关节、右手关节、左肩关节、左肘关节、左手关节;将待测试驾驶人图像转化为HDF5格式,作为驾驶人关节点深度模型的输入,得到与待测驾驶人图像对应的关节点坐标,连接相邻关节点得到驾驶人姿态。

5. 根据权利要求1所述的驾驶人姿态识别方法,其特征在于:所述步骤3)包括以下具体

步骤:

31) 利用三维建模软件的建模功能Modeling构建不同着装、发型、佩戴不同眼镜的人物模型和汽车模型,或从网上模型库中下载人物模型和汽车模型;

32) 调整人物姿态为驾驶人姿态,合并人物模型和汽车模型,模拟驾驶人场景。

6. 根据权利要求1所述的驾驶人姿态识别方法,其特征在于:所述步骤4) 包括以下具体步骤:

41) 利用偏航角、俯仰角、翻滚角三个角度来描述驾驶人头部姿态,称为姿态角;

42) 利用三维建模软件的动画功能Animation设置驾驶人头部不同姿态角的运动;设置好驾驶人模型头部的起始姿态、结束姿态以及关键姿态,动画模块自动平滑处理一系列符合人体运动学的头部运动,模拟驾驶人头部的旋转动作。

7. 根据权利要求1所述的驾驶人姿态识别方法,其特征在于:所述步骤5) 包括以下具体步骤:

51) 在驾驶人面部特征点位置放置不同颜色的球体表示驾驶人面部特征点标签,包括左眼、右眼、鼻尖、左嘴角和右嘴角;

52) 利用三维建模软件的渲染功能Rendering批量渲染驾驶人不同头部旋转角度下的图像,包括驾驶人彩色图像、深度图像、带面部特征点标签的图像,并将头部姿态角和面部特征点位置坐标写入txt文件中;在驾驶室反光镜处放置摄像机模拟现实中的深度摄像机,摄像机镜头对准驾驶人,以摄像机视角进行批量渲染操作。

8. 根据权利要求1所述的驾驶人姿态识别方法,其特征在于:所述步骤6) 包括以下具体步骤:

61) 根据带面部特征点标签的图像和面部特征点坐标信息,确定驾驶人头部位置;

62) 分割驾驶人彩色图像和深度图像,提取驾驶人头部图像,构建虚拟驾驶人头部姿态数据集;提取的驾驶人头部图像大小为 $100 \times 100$ ,同时将头部姿态角度和面部特征点坐标归一化写入txt文件中。

9. 根据权利要求1或8所述的驾驶人姿态识别方法,其特征在于:所述步骤7) 包括以下具体步骤:

71) 将步骤62) 获取的虚拟驾驶人头部姿态数据集作为迁移学习的源域,将大量无标注的真实驾驶人头部姿态数据和少量带有标签的真实数据作为目标域;采用深度迁移学习的方法解决虚拟数据与真实数据之间的特性差异;源域用 $x_s = \{(x_{si}, y_{si}) \mid i = 1, 2, \dots, N_s\}$ 表示,其中 $N_s$ 表示源域样本数量, $x_{si}$ 表示驾驶人头部深度图像, $y_{si}$ 表示 $x_{si}$ 的头部姿态空间角度和面部特征点空间位置坐标;目标域中带标注样本定义为 $x_t = \{(x_{ti}, y_{ti}) \mid i = 1, 2, \dots, N_t\}$ ,采用直推式迁移学习的方法进行特征迁移;

72) 将源域数据作为训练集,将带标签的目标域数据作为测试集,转化为HDF5格式,作为深度学习框架Caffe的输入;

73) 采用多个卷积层、池化层、全连接层构建深度迁移学习网络;利用深度迁移学习网络初始网络权重参数和偏置参数;网络结构包含一个输入层、多个中间层和一个输出层;每一层表示为一个非线性映射的过程,样本 $x$ 在第 $m$ 层的输出计算公式如下:

$$f^{(m)}(x) = h^{(m)} = \varphi(W^{(m)}h^{(m-1)} + b^{(m)}) \quad (1)$$

其中, $W^{(m)} \in R^{p^{(m)} \times p^{(m-1)}}$ 表示网络的 $m$ 层和 $m-1$ 层之间的权重矩阵, $b^{(m)} \in R^{p^{(m)}}$ 表示偏移参

数,  $p^{(m)}$  表示  $m$  层的节点数目,  $\varphi()$  表示非线性激励函数, 使用  $\tanh$ 、 $\text{sigmoid}$  或者  $\text{ReLU}$  函数; 在第一层, 假设  $h^{(0)} = x$ ,  $p^{(0)} = d$ ; 在中间隐含层的特征空间中, 任意两个样本之间的距离度量通过平方欧氏距离来表示, 计算公式如下:

$$d^{(m)}(x_i, x_j) = \left\| f^{(m)}(x_i) - f^{(m)}(x_j) \right\|_2^2 \quad (2)$$

采用最大平均差异衡量源域与目标域在第  $m$  层的分布差异, 计算公式如下:

$$D_{is}^{(m)}(\mathcal{X}_t, \mathcal{X}_s) = \left\| \frac{1}{N_t} \sum_{i=1}^{N_t} f^{(m)}(x_{ti}) - \frac{1}{N_s} \sum_{i=1}^{N_s} f^{(m)}(x_{si}) \right\|_2^2 \quad (3)$$

此时, 深度迁移学习网络的中间层  $m$  的损失函数定义为:

$$\min_{W^{(m)}, b^{(m)}} J = S_c^{(m)} - \alpha S_b^{(m)} + \beta D_{is}^{(m)}(\mathcal{X}_t, \mathcal{X}_s) + \gamma \sum_{i=1}^m (\|W^{(i)}\|_F^2 + \|b^{(i)}\|_2^2) \quad (4)$$

其中,  $\alpha$  ( $\alpha > 0$ )、 $\gamma$  ( $\gamma > 0$ ) 和  $\beta$  表示正则化参数;  $\|Z\|_F$  表示矩阵  $Z$  的 Frobenius 范数;  $S_c^{(m)}$  表示数据的紧致性度量, 用  $S_b^{(m)}$  表示数据的可分性度量; 采用随机梯度下降方法来优化得到参数  $W^{(m)}$  和  $b^{(m)}$ ;

联合优化网络隐含层与最高层的损失函数, 计算整个网络的目标函数, 计算公式如下:

$$\min_{W^{(m)}, b^{(m)}} J = J^{(M)} + \sum_{m=1}^{M-1} \omega^{(m)} h(J^{(m)} - \tau^{(m)}) \quad (5)$$

其中,

$$J^{(m)} = S_c^{(m)} - \alpha S_b^{(m)} + \beta D_{is}^{(m)}(\mathcal{X}_t, \mathcal{X}_s) + \gamma (\|W^{(m)}\|_F^2 + \|b^{(m)}\|_2^2) \quad (6)$$

其中,  $J^{(m)}$  和  $J^{(M)}$  分别表示第  $m$  ( $m=1, 2, \dots, M-1$ ) 层的损失函数和最高层的损失函数; 函数  $h(x) = \max(x, 0)$  用于测量损失;  $\tau^{(m)}$  表示正阈值, 用于控制第  $m$  层损失函数  $J^{(m)}$ , 表示其在学习过程中的重要性;  $\omega^{(m)}$  用于平衡最高层与隐含层的损失重要性; 在学习过程中, 如果第  $m$  层损失函数小于阈值  $\tau^{(m)}$ , 式 (5) 的第二项为 0, 采用梯度下降方法来迭代更新参数  $W^{(m)}$  和  $b^{(m)}$ , 直到网络达到终止条件;

74) 采用平方差损失函数微调深度迁移学习网络, 训练深度迁移学习网络模型;

75) 选择检测准确率最高的模型, 测试真实驾驶人头部图像;

76) 从测试结果中, 如果图像中驾驶人关节点坐标和头部姿态角度与真实值的差值小于设定的阈值, 将该图像选为可信度高的图像, 加入目标域数据中, 扩充数据量;

77) 返回步骤 72), 迭代更新深度迁移学习网络, 直至收敛。

10. 根据权利要求 1 所述的驾驶人姿态识别方法, 其特征在于: 所述步骤 8) 包括以下具体步骤:

81) 从步骤 7) 获得的所有深度迁移学习模型中, 根据所有测试结果, 选择具有最大检测正确率的模型;

82) 利用选择的模型, 估计目标域图像中驾驶人头部姿态角, 并检测其面部特征点位置; 将待测试驾驶人头像图像转化为 HDF5 格式, 作为驾驶人头部深度迁移模型的输入, 得到与待测驾驶人头部图像对应的姿态角和特征点坐标。

## 基于深度图像和虚拟数据的驾驶人姿态识别方法

### 技术领域

[0001] 本发明涉及涉及计算机视觉和车辆辅助驾驶领域,更具体地说,涉及一种基于深度图像和虚拟数据的驾驶人姿态识别方法。

### 背景技术

[0002] 驾驶人姿态识别是车辆辅助驾驶领域中的一个重要的研究课题,是人体姿态识别技术在真实场景的应用。基于图像的驾驶人姿态识别是指在一副给定驾驶人图像中检测驾驶人的上半身关节、估计头部姿态、以及检测面部特征点。计算机可以对驾驶人姿态进行分析,最终实现对于驾驶人危险行为发出警告,达到提醒驾驶人的目的。

[0003] 基于图像的驾驶人姿态识别的输入通常为可见光图像。从图像中提取人体各个部位特征,如颜色、边缘、轮廓、形状等,求解构造的人体模型的各种参数,或者直接学习构造人体特征与人体姿态之间的映射。然而,这种基于可见光图像的人体姿态识别方法存在以下不足:其一,易受光照变化的影响,驾驶室光线明暗变化大,白天较亮,夜晚较暗,人体的外观特征会变得不可辨识;其二,易受到复杂背景的影响,驾驶室环境复杂,极易与人体产生相似的纹理特征,产生误检。其三,人体的体型着装差异变化多,人体关节自由度大,姿态变化多,对识别算法的准确性产生严重的影响。

[0004] 基于图像的驾驶人姿态识别过程中包括两个关键步骤:特征提取和分类器(回归器)训练。特征提取是指提取图像中最能描述人体姿态的信息。Yang Yi和Ramanan Deva提取人体各部位HOG特征,实现精确的关节点检测和人体姿态估计。为了保证算法的精确性,必须提取鲁棒的特征表示,所以需要大量的计算和测试工作。但是这些工作都由人工完成,不仅需要大量的时间,而且非常依赖每个人的经验。分类器训练是指在已确定的特征的基础上,寻找一系列规则,用这些规则来判定当前图像中人体的姿态。确定规则的过程称为(机器)学习的过程,目前较主流的分类器(回归器)包括支持向量机(SVM)、最大熵方法和基于Adaboost的方法。但是这些方法都是“浅层”的学习方法,只能学习单层特征,不能充分学习图像的多层次结构的特征。

[0005] 深度学习方法可以自动学习多层次结构的特征,但是深度模型训练需要海量带标签数据的支持,才能充分挖掘数据更加本质的特征。而带标签数据的获取通常是通过人工完成,这部分工作相当费时且标注误差因人而异。目前,很多研究人员从合成虚拟数据中提取特征应用到真实世界的的数据中,但很难消除虚拟数据特性与真实应用场景的差异。

### 发明内容

[0006] 本发明的目的在于针对上述驾驶人姿态识别技术中,人工提取图像特征费时低效,浅层学习方法不能学习图像多层次结构特征,虚拟数据特性与真实应用场景差异较大的问题,提供一种基于深度图像和虚拟数据的驾驶人姿态识别方法。

[0007] 实现本发明目的的技术解决方案为:基于深度图像的驾驶人姿态识别方法,包括驾驶人关节点检测、虚拟驾驶人头部姿态数据集构建和驾驶人头部姿态分析三个过程:

- [0008] 驾驶人关节检测过程包括以下步骤：
- [0009] 1) 将驾驶人深度图像和关节点标签作为深度学习框架Caffe的输入，训练深度卷积神经网络模型；
- [0010] 2) 利用检测正确率最高的深度学习模型，检测图像中驾驶人的关节点位置。
- [0011] 虚拟驾驶人头部姿态数据集构建过程包括以下步骤：
- [0012] 3) 利用三维建模软件的建模功能 (Modeling) 构建驾驶人模型；
- [0013] 4) 利用三维建模软件的动画功能 (Animation) 设置驾驶人头部运动；
- [0014] 5) 利用三维建模软件的渲染功能 (Rendering) 批量渲染驾驶人图像；
- [0015] 6) 对驾驶人图像进行头部分割处理，构建虚拟驾驶人头部姿态数据集。
- [0016] 驾驶人头部姿态分析过程包括以下步骤：
- [0017] 7) 利用由虚拟驾驶人头部姿态数据集构建过程获取的虚拟驾驶人头部姿态数据，采用迁移学习的方法，训练深度迁移学习模型；
- [0018] 8) 根据深度迁移学习模型，估计目标域图像中驾驶人头部姿态，并检测其面部特征点位置。
- [0019] 上述方法中，所述步骤1) 中的驾驶人深度图像可以通过双目视觉设备或者红外结构光投影设备得到；
- [0020] 上述方法中，所述步骤1) 中的关节点标签通过标记得到的；
- [0021] 上述方法中，所述步骤1) 包括以下具体步骤：
- [0022] 11) 将驾驶人深度图像和关节点标签转化为HDF5格式，作为深度学习框架Caffe的输入；
- [0023] 12) 采用多个卷积层、池化层、全连接层构建深度卷积神经网络；
- [0024] 13) 采用平方差损失函数微调深度卷积神经网络，训练深度学习模型。
- [0025] 上述方法中，所述步骤2) 包括以下具体步骤：
- [0026] 21) 从所有深度学习模型中，根据所有测试结果，选择具有最大检测正确率的模型；
- [0027] 22) 利用选择的模型，检测图像中驾驶人的关节点位置，包括驾驶人头部、颈部、腰部、右肩关节、右肘关节、右手关节、左肩关节、左肘关节、左手关节。
- [0028] 上述方法中，所述步骤3) 包括以下具体步骤：
- [0029] 31) 利用三维建模软件的建模功能 (Modeling) 构建不同着装、发型、佩戴不同眼镜的人物模型和汽车模型；
- [0030] 32) 调整人物姿态为驾驶人姿态，合并人物模型和汽车模型，模拟驾驶人场景。
- [0031] 上述方法中，所述步骤4) 包括以下具体步骤：
- [0032] 41) 利用偏航角、俯仰角、翻滚角三个角度来描述驾驶人头部姿态，称为姿态角；
- [0033] 42) 利用三维建模软件的动画功能 (Animation) 设置驾驶人头部不同姿态角的运动；
- [0034] 上述方法中，所述步骤5) 包括以下具体步骤：
- [0035] 51) 在驾驶人面部特征点位置放置不同颜色的球体表示驾驶人面部特征点标签，包括左眼、右眼、鼻尖、左嘴角和右嘴角；
- [0036] 52) 利用三维建模软件的渲染功能 (Rendering) 批量渲染驾驶人不同头部旋转角

度下的图像,包括驾驶人彩色图像、深度图像、带面部特征点标签的图像,并将头部姿态角和面部特征点位置坐标写入txt文件中;

[0037] 上述方法中,所述步骤6)包括以下具体步骤:

[0038] 61) 根据带面部特征点标签的图像和面部特征点坐标信息,可以确定驾驶人头部位置;

[0039] 62) 分割驾驶人彩色图像和深度图像,提取驾驶人头部图像,构建虚拟驾驶人头部姿态数据集。

[0040] 上述方法中,所述步骤7)包括以下具体步骤:

[0041] 71) 将步骤62)获取的虚拟驾驶人头部姿态数据集作为迁移学习的源域,将大量无标注的真实驾驶人头部姿态数据和少量带有标签的真实数据作为目标域;

[0042] 72) 将源域数据作为训练集,将带标签的目标域数据作为测试集,转化为HDF5格式,作为深度学习框架Caffe的输入;

[0043] 73) 采用多个卷积层、池化层、全连接层构建深度迁移学习网络;

[0044] 74) 采用平方差损失函数微调深度迁移学习网络,训练深度迁移学习网络模型;

[0045] 75) 选择检测准确率最高的模型,测试真实驾驶人头部图像;

[0046] 76) 从测试结果中,筛选可信度高的图像,加入目标域数据中,扩充数据量;

[0047] 77) 跳到步骤72),迭代更新深度迁移学习网络,直至收敛。

[0048] 所述步骤8)包括以下具体步骤:

[0049] 81) 从步骤76)获得的所有深度迁移学习模型中,根据所有测试结果,选择具有最大检测正确率的模型;

[0050] 82) 利用选择的模型,估计目标域图像中驾驶人头部姿态角,并检测其面部特征点位置。

[0051] 本发明与现有技术相比,其显著优点:其一,深度图像具有颜色无关特性,包含物体深度信息,当驾驶室光照变化较大或者光照较弱的环境下,本发明仍然能够对驾驶人姿态辨识;其二,不易受到复杂的背景的影响,即使当背景纹理很丰富时,仍然能够有效的将目标与背景进行区分,从而减小误检数目;其三,不易受到驾驶人复杂的体型着装差异,人体关节自由度,姿态变化的影响,检测准确率高;其四,基于深度学习的特征提取方法可以自动地提取具有结构层次的特征,大大减少了人工设置特征带来的费时和低效性的同时,提高了检测准确率;其五,利用虚拟数据模拟真实场景,可以有效地扩充训练集数据,利用深度迁移学习的思想,可以有效降低虚拟数据特性与真实应用场景存在的差异。

## 附图说明

[0052] 图1是驾驶人深度图像。

[0053] 图2是驾驶人关节点表示。

[0054] 图3是驾驶人关节点检测过程。

[0055] 图4是虚拟驾驶人头部姿态数据集构建过程。

[0056] 图5是驾驶人头部姿态角表示。

[0057] 图6是驾驶人面部特征点表示。

[0058] 图7是驾驶人头部姿态分析过程。

## 具体实施方式

[0059] 下面结合附图对本发明作进一步说明。

[0060] 本发明的驾驶人姿态识别方法包括驾驶人关节检测、虚拟驾驶人头部姿态数据集构建和驾驶人头部姿态分析三个主要过程。

[0061] 本发明的数据源为从深度图像采集设备得到的驾驶人深度图像,如图1所示。深度图像上的每个像素点的值表示该物点到相机投影中心的距离,通过该值可以得到驾驶人的形状信息以及三维位置信息。深度图像可以由双目视觉装置获得或者结构光投影装置获得。

[0062] 驾驶人关节检测过程是指检测图像中驾驶人关节位置,驾驶人关节是指驾驶人头部、颈部、腰部、右肩关节、右肘关节、右手关节、左肩关节、左肘关节、左手关节,如图2所示。检测过程如图3所示,包括以下具体步骤:

[0063] 将驾驶人深度图像与关节点标签转化为HDF5格式数据过程31。调整驾驶人深度图像大小为 $120 \times 96$ ,关节点标签是人工标定的关节点像素坐标,用 $(x, y)$ 表示。利用Python开源库h5py将驾驶人深度图像与对应的关节点标签转化为HDF5格式数据,作为深度学习框架Caffe的输入数据格式。

[0064] 驾驶人关节点深度模型训练过程32。卷积神经网络采用了一种端到端的模型训练方式,通过传统的梯度下降方法训练模型中的参数,自动学习图像中具有层次结构的特征。网络基本结构由输入层、多次交替的卷积层和池化层、一个或者多个全连接层以及输出层构成。在输出层应用平方差损失函数,通过反向传播方式微调深度卷积神经网络,直至收敛。比较每个模型的测试准确率,选择准确率最高的模型作为驾驶人关节点深度模型。

[0065] 驾驶人关节检测过程33。将待测试驾驶人图像转化为HDF5格式,作为驾驶人关节点深度模型的输入,得到与待测驾驶人图像对应的关节点坐标,连接相邻关节点得到驾驶人姿态。

[0066] 虚拟驾驶人头部姿态数据集构建过程是指利用三维建模软件构建虚拟驾驶人头部姿态数据集,如图4所示。包括以下具体步骤:

[0067] 构建驾驶人场景模型过程41。利用三维建模软件(如3DSMAX)的建模功能(Modeling)构建不同着装、发型、佩戴不同眼镜的人物模型和汽车模型,亦可从网上模型库中下载人物模型和汽车模型。调整人物模型大小与汽车模型相符,将其姿态调整为驾驶人开车姿势,合并人物模型与汽车模型,模拟驾驶人驾驶汽车场景。

[0068] 设置驾驶人头部运动过程42。利用三维建模软件的动画功能(Animation)设置好驾驶人模型头部的起始姿态、结束姿态以及关键姿态,动画模块会自动平滑处理一系列的符合人体运动学的头部运动,模拟驾驶人头部的旋转动作。

[0069] 批量渲染虚拟驾驶人图像过程43。在驾驶人模型的左眼、右眼、鼻尖、左嘴角和右嘴角位置放置不同颜色的球体表示驾驶人面部特征点标签。在驾驶室反光镜处放置摄像机模拟现实中的深度摄像机,摄像机镜头对准驾驶人,以摄像机视角批量渲染驾驶人不同头部旋转角度下的图像,包括驾驶人彩色图像、深度图像、带面部特征点标签的图像,并将头部姿态角和面部特征点位置坐标写入txt文件中。

[0070] 提取虚拟驾驶人头部图像过程44。根据带面部特征点标签的图像和面部特征点坐



标信息,确定驾驶人头部位置,分割驾驶人彩色图像和深度图像,提取 $100 \times 100$ 大小的驾驶人头部图像,构建虚拟驾驶人头部姿态数据集。同时将头部姿态角度和面部特征点坐标归一化写入txt文件中。

[0071] 驾驶人头部姿态分析过程是指估计驾驶人头部姿态角和检测驾驶人面部特征点位置,驾驶人头部姿态角指驾驶人头部的偏航角(Yaw)、俯仰角(Pitch)和翻滚角(Roll),如图5所示,驾驶人面部特征点指驾驶人的左眼、右眼、鼻尖、左嘴角和右嘴角,如图6所示。分析过程如图7所示,包括以下具体步骤:

[0072] 初始化源域和目标域数据集过程71。本发明采用深度迁移学习的方法解决虚拟数据的与真实数据之间的特性差异。将过程44构建的大量带标签的虚拟驾驶人头部深度图像作为源域,用 $x_s = \{(x_{si}, y_{si}) \mid i=1, 2, \dots, N_s\}$ 表示,其中 $N_s$ 表示源域样本数量, $x_{si}$ 表示驾驶人头部深度图像, $y_{si}$ 表示 $x_{si}$ 的头部姿态空间角度和面部特征点空间位置坐标;将大量无标签的真实驾驶人头部深度图像(少数部分带标签)作为目标域,其中目标域的带标注样本定义为 $x_t = \{(x_{ti}, y_{ti}) \mid i=1, 2, \dots, N_t\}$ 。采用直推式迁移学习的方法进行特征迁移。

[0073] 初始训练深度迁移学习网络过程72。将大量带标签的源域样本和少量具有标签的目标域样本作为训练集,利用深度迁移学习网络初始网络权重参数和偏置参数。网络结构包含一个输入层、多个中间层和一个输出层。每一层可以表示为一个非线性映射的过程,样本 $x$ 在第 $m$ 层的输出计算公式如下:

$$[0074] \quad f^{(m)}(x) = h^{(m)} = \varphi(W^{(m)}h^{(m-1)} + b^{(m)}) \quad (1)$$

[0075] 其中, $W^{(m)} \in R^{p^{(m)} \times p^{(m-1)}}$ 表示网络的 $m$ 层和 $m-1$ 层之间的权重矩阵, $b^{(m)} \in R^{p^{(m)}}$ 表示偏移参数, $p^{(m)}$ 表示 $m$ 层的节点数目, $\varphi()$ 表示非线性激励函数,通常使用tanh、sigmoid或者ReLU函数。在第一层,假设 $h^{(0)} = x$ , $p^{(0)} = d$ 。在中间隐含层的特征空间中,任意两个样本之间的距离度量可以通过平方欧氏距离来表示,计算公式如下:

$$[0076] \quad d^m(x_i, x_j) = \|f^{(m)}(x_i) - f^{(m)}(x_j)\|_2^2 \quad (2)$$

[0077] 采用最大平均差异衡量源域与目标域在第 $m$ 层的分布差异,计算公式如下:

$$[0078] \quad D_{ts}^{(m)}(\mathcal{X}_t, \mathcal{X}_s) = \left\| \frac{1}{N_t} \sum_{i=1}^{N_t} f^{(m)}(x_{ti}) - \frac{1}{N_s} \sum_{i=1}^{N_s} f^{(m)}(x_{si}) \right\|_2^2 \quad (3)$$

[0079] 此时,深度迁移学习网络的中间层 $m$ 的损失函数定义为:

$$[0080] \quad \min_{W^{(m)}, b^{(m)}} J = S_c^{(m)} - \alpha S_b^{(m)} + \beta D_{ts}^{(m)}(\mathcal{X}_t, \mathcal{X}_s) + \gamma \sum_{i=1}^m (\|W^{(i)}\|_F^2 + \|b^{(i)}\|_2^2) \quad (4)$$

[0081] 其中, $\alpha$  ( $\alpha > 0$ )、 $\gamma$  ( $\gamma > 0$ )和 $\beta$ 表示正则化参数; $\|Z\|_F$ 表示矩阵 $Z$ 的Frobenius范数; $S_c^{(m)}$ 表示数据的紧致性度量,用 $S_b^{(m)}$ 表示数据的可分性度量。采用随机梯度下降方法来优化得到参数 $W^{(m)}$ 和 $b^{(m)}$ 。

[0082] 为了最大化利用隐含层的判别信息,联合优化网络隐含层与最高层的损失函数,整个网络的目标函数计算公式如下:

$$[0083] \quad \min_{W^{(M)}, b^{(M)}} J = J^{(M)} + \sum_{m=1}^{M-1} \omega^{(m)} h(J^{(m)} - \tau^{(m)}) \quad (5)$$

[0084] 其中,

$$[0085] \quad J^{(m)} = S_c^{(m)} - \alpha S_b^{(m)} + \beta D_{ts}^{(m)}(\chi_t, \chi_s) + \gamma (\|W^{(m)}\|_F^2 + \|b^{(m)}\|_2^2) \quad (6)$$

[0086] 其中,  $J^{(m)}$  和  $J^{(M)}$  分别表示第  $m$  ( $m=1, 2, \dots, M-1$ ) 层的损失函数和最高层的损失函数; 函数  $h(x) = \max(x, 0)$  用于测量损失;  $\tau^{(m)}$  表示正阈值, 用于控制第  $m$  层损失函数  $J^{(m)}$ , 表示其在学习过程中的重要性;  $\omega^{(m)}$  用于平衡最高层与隐含层的损失重要性; 在学习过程中, 如果第  $m$  层损失函数小于阈值  $\tau^{(m)}$ , 式 (5) 的第二项为 0。采用梯度下降方法来迭代更新参数  $W^{(m)}$  和  $b^{(m)}$ , 直到网络达到终止条件。

[0087] 迭代更新网络过程 73。将带标签的目标域加入训练集, 更新网络权重参数和偏置参数。将得到的深度迁移模型测试无标签的真实驾驶人图像, 筛选出可信度高的样本, 加入训练集, 迭代更新网络权重参数和偏置参数, 直至收敛, 得到驾驶人头部深度迁移模型。

[0088] 驾驶人头部分析过程 74。将待测试驾驶人头像图像转化为 HDF5 格式, 作为驾驶人头部深度迁移模型的输入, 得到与待测驾驶人头部图像对应的姿态角和特征点坐标。

[0089] 相比于目前最较先进的深度网络 AlexNet 和 GoogleNet, 本发明设计的深度卷积神经网络, 取得了至少 12% 的检测正确率提升。本发明采用计算机动画软件批量渲染虚拟数据的方法, 能够很好地解决缺少大量带标签数据的问题。通过深度迁移学习方法, 本发明能够很好地在虚拟数据中学习模型应用到真实数据。



图1

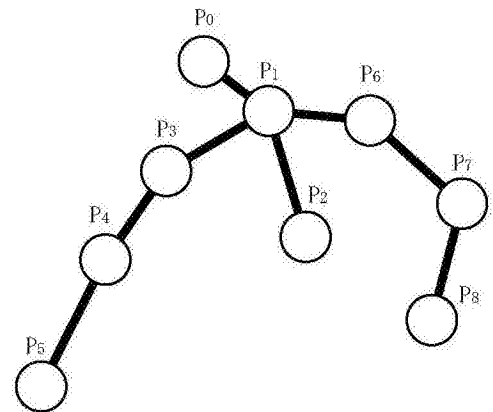


图2

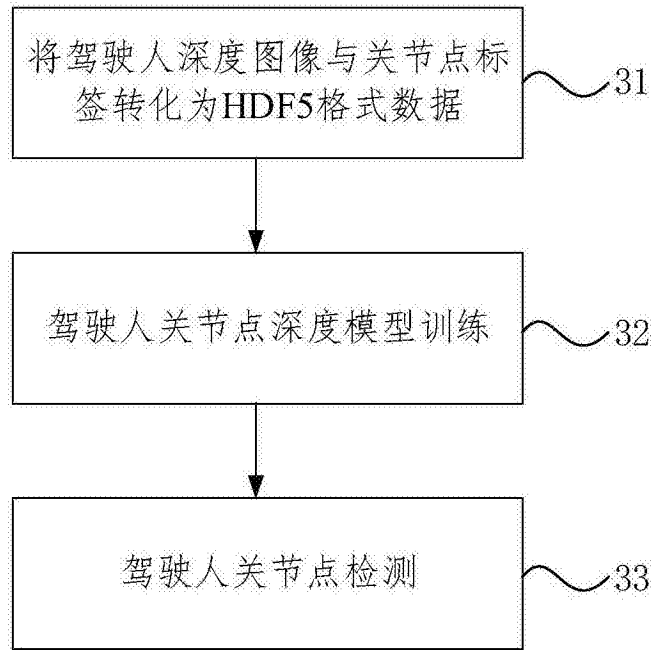


图3

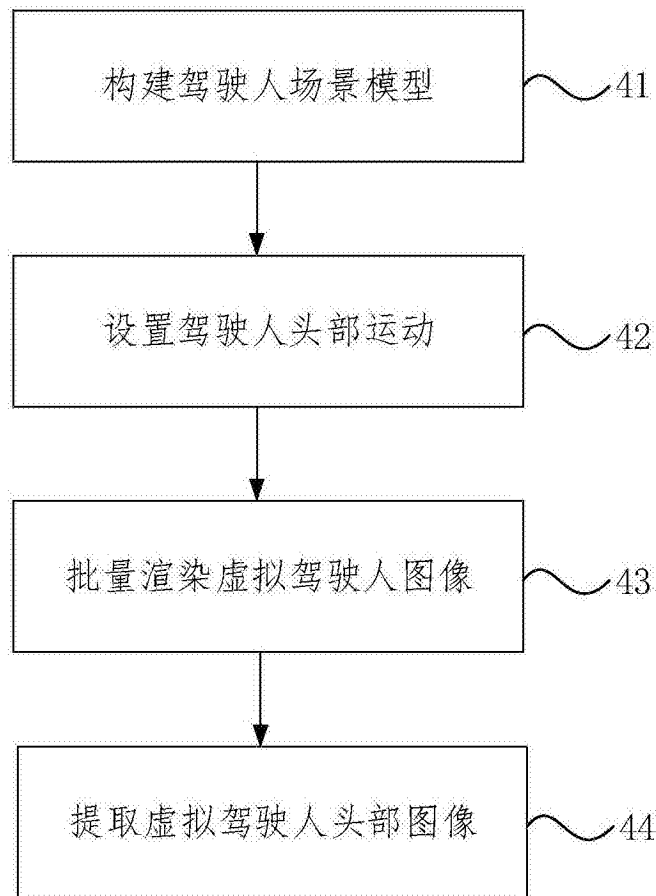


图4

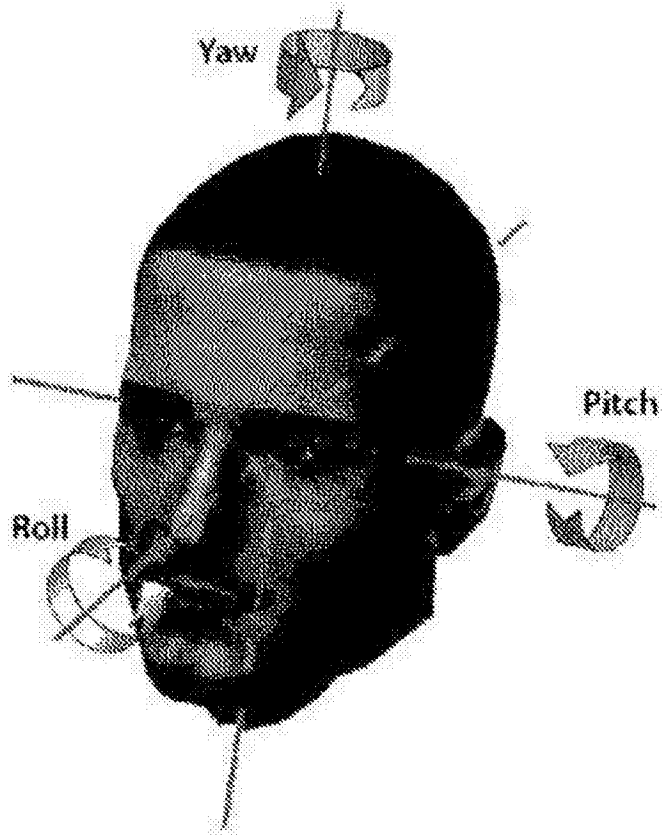


图5

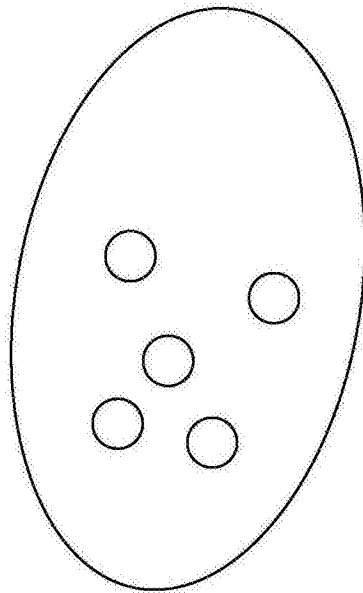


图6

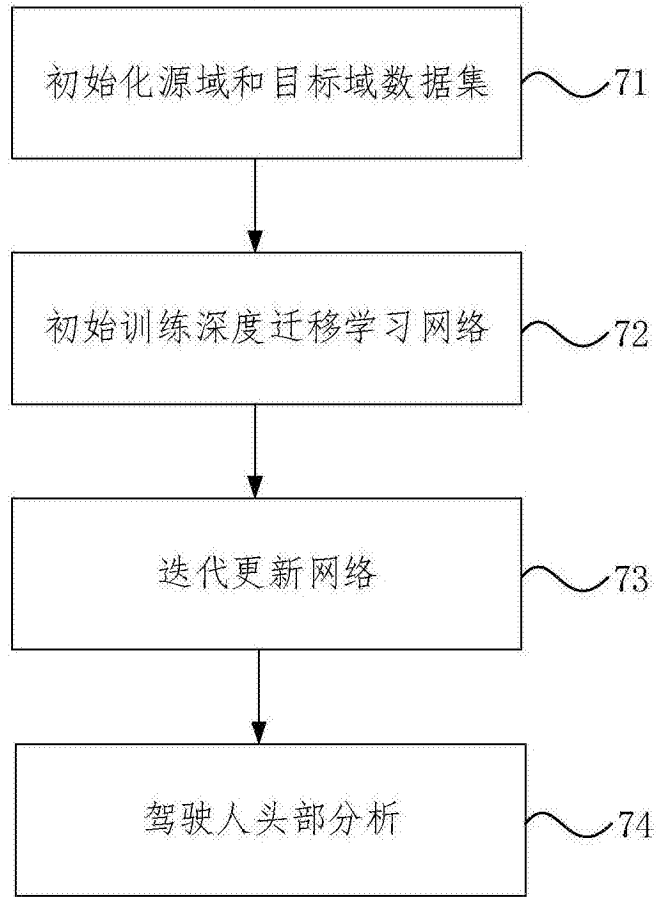


图7