

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2018 年 7 月 5 日 (05.07.2018)



(10) 国际公布号
WO 2018/120964 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2017/104990
- (22) 国际申请日: 2017 年 9 月 30 日 (30.09.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201611251820.3 2016 年 12 月 30 日 (30.12.2016) CN
- (71) 申请人: 山东大学 (SHANDONG UNIVERSITY)
[CN/CN]; 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。
- (72) 发明人: 许宏吉 (XU, Hongji); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。

孙君凤 (SUN, Junfeng); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。 房海腾 (FANG, Haiteng); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。 周英明 (ZHOU, Yingming); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。 潘玲玲 (PAN, Lingling); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。 党娟 (DANG, Juan); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。 李菲菲 (LI, Feifei); 中国山东省济南市历城区山大南路 27 号, Shandong 250199 (CN)。

(74) 代理人: 济南金迪知识产权代理有限公司 (JINAN JINDI INTELLECTUAL PROPERTY AGENCY CO.,

(54) Title: POSTURE CORRECTION METHOD BASED ON DEPTH INFORMATION AND SKELETON INFORMATION

(54) 发明名称: 一种基于深度信息和骨骼信息的姿势矫正方法

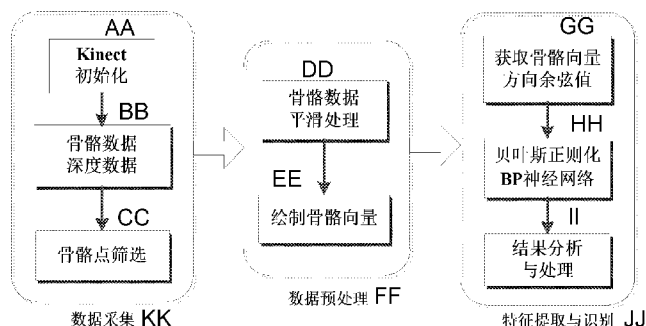


图 1

- AA Kinect initialization
BB Skeleton data and depth data
CC Skeleton point screening
DD Skeleton data smoothing processing
EE Draw skeleton vectors
FF Data preprocessing
GG Obtain direction cosine values of skeleton vectors
HH Bayesian Regularization BP neural network
II Result analysis and processing
JJ Feature extraction and recognition
KK Data acquisition

(57) Abstract: A posture correction method based on depth information and skeleton information, comprising: (1) performing effective skeleton point screening based on a user ID and a skeleton space coordinate depth value; (2) smoothing skeleton information, and standardizing coordinates of skeleton points in a skeleton space coordinate system; (3) drawing skeleton vectors, and calculating a direction cosine value of each skeleton vector; (4) obtaining a training data set and a test data set; (5) using the data sets as the input of a Bayesian Regularization BP neural network, and recognizing 25 skeleton points and 24 skeleton segments via the Bayesian



LTD); 中国山东省济南市历城区山大南路
27号, Shandong 250199 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

根据细则4.17的声明:

- 关于发明人身份(细则4.17(i))

本国际公布:

- 包括国际检索报告(条约第21条(3))。

Regularization BP neural network; and (6) performing result analysis and processing. The method overcomes the defect that a color camera is liable to be affected by external light, can accurately track a human body in a visual field, and also improves the naturalness of man-machine interaction.

(57) 摘要: 一种基于深度信息和骨骼信息的姿势矫正方法, 包括: (1)基于用户ID和骨骼空间坐标深度值的有效骨骼点筛选; (2)对骨骼信息进行平滑处理, 将骨骼点在骨骼空间坐标系中的坐标标准化; (3)绘制骨骼向量, 计算每个骨骼向量的方向余弦值; (4)获取训练数据集和测试数据集; (5)数据集作为贝叶斯正则化BP神经网络的输入, 通过贝叶斯正则化BP神经网络对25个骨骼点和24个骨骼段进行识别; (6)结果分析和处理。本方法克服了彩色摄像头易受光线等外界干扰的缺点, 可以准确追踪到视野范围内的人体, 同时也提高了人机交互的自然性。

一种基于深度信息和骨骼信息的姿势矫正方法

技术领域

本发明涉及一种基于深度信息和骨骼信息的姿势矫正方法，属于智能感知与智能计算技术领域。

背景技术

日常生活中，姿势矫正被应用于诸多领域，例如，医学方面，可用于帕金森综合症、肌肉痉挛等运动障碍疾病的康复治疗；教育方面，可用于篮球、艺术体操、舞蹈等项目的教学；健身娱乐方面，可用于瑜伽、普拉提等项目的体态纠正，帮助健身者达到预期的健身效果。据报道，在 40 岁以上的亚洲人中，每 3 人中就有 1 人可能患有运动障碍综合征(Locomotive Syndrome)，合理的运动治疗是这类病症必不可少的康复手段，而对于体操、瑜伽、普拉提等项目，动作姿势完成的质量会影响后期训练的效果，错误的动作姿势还可能会造成骨骼错位或肌肉拉伤。所以为了实现更好的医疗康复效果和减少训练中不必要的身体伤害，研究一种智能、高效的姿势监督矫正方法显得尤为重要。

医疗方面，传统的姿势矫正方法都需要专业人员的陪同与指导，浪费诸多人力物力的同时也令参与者感到枯燥乏味，往往不能达到预期的康复疗效。此外，已有一些研究机构将索尼体感设备 EyeToy 和任天堂 Wii 应用于肢体运动康复训练中，但其二维图像处理的局限性制约了此项技术在康复领域的发展。

微软发布的 Kinect 体感设备有即时 3D 运动捕捉、麦克输入和影音识别等功能，特别是第二代 Kinect 体感器提供了优化版骨骼追踪功能，增强保真度的深度相机结合改进版软件，带来了一系列骨骼追踪功能提升，除了现在可以追踪 6 套完整骨骼（第一代 Kinect 最多追踪 2 套）以及每个目标 25 个骨骼点（第一代 Kinect 为 20 个）之外，自动追踪定位也比上一代更加精确稳定，且追踪范围也更大。

发明内容

针对现有方法的不足，本发明提出了一种基于深度信息和骨骼信息的姿势矫正方法。

本发明利用 Kinect 摄像头获取用户深度数据和骨骼数据的优势，旨在开发一种抗干扰能力强、便捷实用有效的姿势矫正方法。该方法包括：1) 数据采集：使用 Kinect2.0 采集传感器视野范围内

的深度数据和骨骼数据（可采集 6 个人体身上的 25 个骨骼点），基于用户 ID（Kinect 唯一的骨骼跟踪 ID，用于分配给视野中的每个用户，用以区分现在这个骨骼数据是哪个用户的）和骨骼空间坐标深度值进行骨骼点筛选；2）数据预处理：对骨骼数据平滑处理将骨骼点的坐标标准化，根据人体结构学原理绘制骨骼向量；3）特征提取与识别：计算骨骼向量的方向余弦值，并将每个骨骼向量的 3 个方向余弦值作为特征进行提取，并输入到贝叶斯正则化 BP 神经网络进一步识别，最后对识别结果进行分析与处理，处理结果显示在用户界面上，标准的骨骼段和骨骼点用亮绿显示，非标准的骨骼段和骨骼点用亮红显示同时使用语音进行提示。

本发明提高了姿势矫正的实用性、准确性及鲁棒性。

本发明的技术方案为：

术语解释：

1、BP（Back Propagation）神经网络，是 1986 年由 Rumelhart 和 McClland 为首的科学家小组提出，是一种按误差逆传播算法训练的多层前馈网络，是目前应用最广泛的神经网络模型之一。BP 神经网络能学习和存储大量的输入—输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法，通过反向传播来不断调整网络的权值和阈值，使网络的误差平方和最小。BP 神经网络模型拓扑结构包括输入层（input layer）、隐含层(hidden layer)和输出层(output layer)。

2、深度数据流，指 kinect 实时采集的一系列深度数据；

3、彩色数据流，指 kinect 实时采集的一系列彩色数据；

一种基于深度信息和骨骼信息的姿势矫正方法，具体步骤包括：

（1）基于用户 ID 和骨骼空间坐标深度值的有效骨骼点筛选：利用 Kinect2.0 摄像头获取用户骨骼数据和深度数据，选择目标用户的所有骨骼点；用户 ID 是指 Kinect2.0 传感设备唯一的用户跟踪标识，用于分配给有效视距内的每个用户，用以区分骨骼信息是哪个用户的；

（2）对步骤（1）选定的骨骼数据进行平滑处理，将骨骼点在骨骼空间坐标系中的坐标标准化；减少骨骼帧之间的骨骼点位置差异，kinect 开发工具 kinect for windows sdk 的数据对象类型以骨骼帧的形式提供，每一帧包括 25 个骨骼点。

（3）根据步骤（2）中已经实现骨骼空间坐标标准化的骨骼点在骨骼空间坐标系中的坐标，绘制骨骼向量，计算每个骨骼向量的方向余弦值，将这些方向余弦值作为特征进行提取；

（4）获取训练数据集和测试数据集；这些数据集均是用户骨骼向量的方向余弦值，每一帧骨骼帧包含 24 个骨骼向量，每一个骨骼向量又包含三个方向余弦值；

（5）步骤（4）获取的训练数据集和测试数据集作为贝叶斯正则化 BP 神经网络的输入，通过

贝叶斯正则化 BP 神经网络对 25 个骨骼点和 24 个骨骼段进行识别；

(6) 结果分析和处理，将每个骨骼段和骨骼点的分析与处理结果实时显示在用户界面上。Kinect2.0 摄像头捕捉到的有效视距内目标用户的骨骼数据帧会实时呈现在用户界面上，与标准姿势相匹配的骨骼向量所对应的骨骼段和骨骼点会呈现出亮绿色，而与姿势错误的骨骼向量相对应的骨骼段和骨骼点会呈现出亮红色，同时会伴随语音提示。

根据本发明优选的，所述步骤(1)，包括以下步骤：

A、利用 Kinect2.0 摄像头获取有效视距内所有物体的骨骼数据和深度数据，所述骨骼数据是指骨骼点在骨骼空间坐标系中的坐标；骨骼空间坐标系是指：以 Kinect2.0 摄像头为原点，Z 轴与 Kinect2.0 摄像头的朝向一致，Y 轴正半轴向上延伸，X 轴正半轴向 Kinect 感应器的视角延伸；所述有效视距指 Kinect2.0 摄像头能正确采集信息的视野范围，所述有效视距的取值范围为 0.8-3.0m；所述骨骼点包括：头部、颈部、右手食指、右手拇指、右手掌心、右手腕、右肘部、右肩、肩部中心、左肩、左肘部、左手腕、左手掌心、左手拇指、左手食指、脊柱、髋部中心、右髋、左髋、右膝盖、左膝盖、右脚踝、左脚踝、右脚、左脚共 25 个骨骼点；

B、确定目标用户的骨骼点，对其他用户的骨骼点进行滤除：利用 Kinect2.0 摄像头获取每一个用户的骨骼点的深度值，并进行累加，计算每一个用户的所有骨骼点的平均深度值，对比这些平均深度值，平均深度值最小的用户即目标用户，保存该目标用户的骨骼点，滤除其他用户的骨骼点。在有效视距内仍有多个用户，就会出现许多无效骨骼点影响后期特征值提取的精度，因此，我们需要确定目标人体的骨骼点，对其他无效骨骼点进行滤除。

Kinect2.0 可以同时追踪 6 个目标用户身上的 25 个骨骼点以及视野范围内物体的深度数据流和彩色数据流，Kinect for windows sdk 中的骨骼 API 可以提供位于 Kinect 前方至多 5 个人的位置信息，包括详细的姿势和骨骼点的三维坐标信息还有用户 ID 信息，数据对象类型以骨骼帧的形式提供，每一帧可以保存 25 个骨骼点。

根据本发明，所述步骤(2)，包括步骤如下：

C、设置平滑值(Smoothing)属性，所述平滑值的取值范围为 0 到 1 之间的浮点型数据，平滑值越大，平滑的越多，平滑值为 0 表示不进行平滑；

设置修正值(Correction)属性，所述修正值的取值范围为 0 到 1 之间的浮点型数据，修正值越小，骨骼信息越平滑；

设置抖动半径(JitterRadius)属性，所述抖动半径的取值范围为 0 到 1 之间的浮点型数据，当骨骼点抖动超过了设置的抖动半径，将会被纠正到这个抖动半径之内；

设置抖动半径的最大边界(MaxDeviationRadius)属性，取值范围为 0 到 1 之间的浮点型数据，

任何超过抖动半径的最大边界的点都不会认为是抖动产生的，而被认定为是一个新的点；

设置预测帧大小(Prediction)属性，取值范围为 0 到 1 之间的浮点型数据，默认值为 0；

D、调用 kinect 骨骼数据平滑算法即向 SkeletonStream.Enable()中传递步骤 C 中的各个平滑参数对骨骼数据进行平滑处理。

对骨骼数据进行平滑处理会产生性能开销。平滑处理的越多，性能消耗越大。设置各个平滑参数没有经验可以遵循。需要不断的测试和调试已达到最好的性能和效果。在程序运行的不同阶段，可能需要设置不同的平滑参数。

在骨骼点跟踪过程中，有些情况会导致骨骼运动呈现出跳跃式的变化。例如，目标用户的动作不够连贯，Kinect 硬件性能差等等。骨骼点的相对位置可能在帧与帧之间变动很大，这会对应用程序产生一些负面的影响，例如，会影响用户体验和给控制造成意外等。通过对骨骼数据进行平滑，将骨骼点的坐标标准化，来减少帧与帧之间的骨骼点位置差异。

根据本发明优选的，所述步骤（3），包括步骤如下：

E、根据人体结构学原理，针对步骤（2）提取到的 25 个骨骼点，依次将相邻的两个骨骼点连接组成一个骨骼段，并将这一个骨骼段定义为一个骨骼向量，共得到 24 个骨骼向量，即：

$$\{\vec{d}_1, \vec{d}_2, \vec{d}_3, \dots, \vec{d}_{24}\};$$

F、设任一骨骼向量 $\vec{d}_i$ 为在骨骼空间坐标系中的坐标为 (x_1, y_1, z_1) 的骨骼点与在骨骼空间坐标系中的坐标为 (x_2, y_2, z_2) 的骨骼点连接所得， $\vec{d}_i \in \{\vec{d}_1, \vec{d}_2, \vec{d}_3, \dots, \vec{d}_{24}\}$ ；骨骼向量 $\vec{d}_i$ 表示如式（I）所示：

$$\vec{d}_i = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \quad (\text{I})$$

G、设骨骼向量 $\vec{d}_i$ 与骨骼空间坐标系 X 轴、Y 轴和 Z 轴的夹角分别为 α 、 β 和 γ ，由此，可得到骨骼向量 $\vec{d}_i$ 的三个方向余弦值，计算公式如式（II）、式（III）和式（IV）所示：

$$\cos \alpha = \frac{x_2 - x_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{II})$$

$$\cos \beta = \frac{y_2 - y_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{III})$$

$$\cos \lambda = \frac{z_2 - z_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{IV})$$

通过上述算法，依次计算 24 个骨骼向量的三个方向余弦值，并将这些余弦值作为特征进行提

取。

当用户做出不同姿势时，对于人体每一个骨骼段而言，都具有不同的位置和角度信息，因此，可以利用定义的 24 个骨骼向量的方向余弦值来表征某一类动作。

根据本发明优选的，所述步骤（4），包括步骤如下：

H、获取训练数据集：目标用户在专业人员的指导下完成一系列的标准动作，通过步骤（1）至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即训练数据集；

I、获取测试数据集：目标用户在 Kinect2.0 摄像头有效视距内独自完成一系列动作，通过步骤（1）至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即测试数据集。

姿势矫正前期，用户会在专业人员的指导下完成一系列的标准动作，Kinect2.0 摄像头会以骨骼帧的形式记录下这些动作，然后经过骨骼点筛选、骨骼数据平滑处理、骨骼向量的方向余弦特征值提取，最后生成我们所需要的训练数据集，训练集总共有若干数据帧（帧数越多精度越高），每一帧骨骼数据包含一个姿势的 24 个骨骼向量，每个骨骼向量有三个方向余弦值；姿势矫正阶段，用户会在 Kinect 视野范围内独自完成一系列动作，Kinect 同样会以骨骼帧的形式记录下这些可能不规范的动作，然后经过相应的骨骼点筛选、骨骼数据平滑处理、骨骼向量的方向余弦特征值提取，最后生成我们所需要的测试数据集。

根据本发明优选的，所述步骤（5），包括步骤如下：

G、确定 BP 神经网络层数为 3，确定输入层的节点数为 3，确定输出层的节点数为 1，并通过式（V）确定隐含层的节点数 m ，式（V）如下所示：

$$m = \sqrt{n+l} + \alpha \quad (V)$$

式（V）中， n 为输入层的节点数； l 为输出层的节点数； α 为 1-10 间的整数；

通过对 BP 神经网络的学习，对于简单的动作识别问题通过单隐层网络即可实现，所以在此选择三层 BP 神经网络。输入层和输出层的节点数，一般是根据实际问题，由输入变量和输出变量的维数确定。BP 神经网络输入变量为三维方向余弦特征值，所以输入层的节点数为 3，输出变量每个骨骼向量与标准姿势的骨骼向量匹配与否即 1 或 0，所以输出层节点数为 1。隐含层节点数对网络性能影响很大，节点数不同，最终结果将会差异很大，节点数太少，网络的迭代速率很快，但网络的建模不充分，导致网络性能较差；节点数太多，会导致网络结构复杂，计算量增大，训练时间较长，所以要选择合适的节点数。

K、将训练数据集作为训练集输入到贝叶斯正则化 BP 神经网络，对每一个骨骼向量训练学习，在姿势矫正阶段，将测试数据集输入到已训练完毕的贝叶斯正则化 BP 神经网络中，对每一个骨骼向量的正确性进行识别，即与训练集的标准骨骼姿势的相应骨骼向量匹配与否。

本发明的有益效果为：

1、与传统彩色摄像头不同，Kinect 传感器能够提供第三维深度数据。它能克服彩色摄像头易受光线等外界干扰的缺点，准确追踪到视野范围内的人体，同时也提高了人机交互的自然性。

2、Kinect 能够提取人体骨骼信息，当用户做出不同动作时，相应的骨骼点与骨骼段具有不同的位置和角度信息，这些信息将提供非常可靠和直接的方法来实现人体姿态识别。

3、Kinect2.0 是对第一代 Kinect 传感器的全面升级，骨骼追踪方面可追踪 6 套完整骨骼（第一代 Kinect 最多追踪 2 套）以及每人身上的 25 个骨骼点（第一代 Kinect 为 20 个），并且自动追踪定位功能也比上一代更加精确稳定，且追踪范围也更大，使用 Kinect2.0 可以大大提高姿态识别的准确度。

4、传统的特征提取方法需要采用复杂的数学算法，实现困难并且运行效率不高，本发明直接研究人体骨骼的角度信息则更加便捷直观，同时也提高了程序的运行速度。

5、BP 神经网络存在很多缺陷，尤其是过拟合现象引起的泛化能力下降异常明显，贝叶斯正则化算法可以有效抑制过拟合现象，具有较高的泛化能力。

6、利用用户 ID 和骨骼空间坐标深度值对目标用户进行定位，提高了骨骼追踪的准确性。

7、本发明对人体的 25 个骨骼点、24 个骨骼段的动作姿势进行识别矫正，提高了姿势矫正的正确性和鲁棒性。

附图说明

图 1 为本发明基于深度信息和骨骼信息的姿势矫正方法的流程框图；

图 2 为本发明骨骼点筛选流程示意图；

图 3 为本发明定义的 24 个骨骼向量示意图；

图 4 为实施例 1 定义的骨骼向量 $\vec{h}_4$ 方向余弦示意图；

图 5 为本发明方向余弦特征值提取流程示意图。

具体实施方式

下面结合说明书附图和实施例对本发明作进一步限定，但不限于此。

实施例 1

一种基于深度信息和骨骼信息的姿势矫正方法，如图 1 所示，具体步骤包括：

(1) 基于用户 ID 和骨骼空间坐标深度值的有效骨骼点筛选：利用 Kinect2.0 摄像头获取用户骨骼数据和深度数据，选择目标用户的所有骨骼点；用户 ID 是指 Kinect2.0 传感设备唯一的用户跟踪标识，用于分配给有效视距内的每个用户，用以区分骨骼信息是哪个用户的；如图 2 所示，包括以下步骤：

A、利用 Kinect2.0 摄像头获取有效视距内所有物体的骨骼数据和深度数据，所述骨骼数据是指骨骼点在骨骼空间坐标系中的坐标；骨骼空间坐标系是指：以 Kinect2.0 摄像头为原点，Z 轴与 Kinect2.0 摄像头的朝向一致，Y 轴正半轴向上延伸，X 轴正半轴向 Kinect 感应器的视角延伸；所述有效视距指 Kinect2.0 摄像头能正确采集信息的视野范围，所述有效视距的取值范围为 0.8-3.0m；所述骨骼点包括：头部、颈部、右手食指、右手拇指、右手掌心、右手腕、右肘部、右肩、肩部中心、左肩、左肘部、左手腕、左手掌心、左手拇指、左手食指、脊柱、髋部中心、右髋、左髋、右膝盖、左膝盖、右脚踝、左脚踝、右脚、左脚等 25 个骨骼点；

B、确定目标用户的骨骼点，对其他用户的骨骼点进行滤除：利用 Kinect2.0 摄像头获取每一个用户骨骼点的深度值，并进行累加，计算每一个用户的所有骨骼点的平均深度值，对比这些平均深度值，平均深度值最小的用户即目标用户，保存该目标用户的骨骼点，滤除其他用户的骨骼点。在有效视距内仍有多个用户，就会出现许多无效骨骼点影响后期特征值提取的精度，因此，我们需要确定目标人体的骨骼点，对其他无效骨骼点进行滤除。

Kinect2.0 可以同时追踪 6 个目标用户身上的 25 个骨骼点以及视野范围内物体的深度数据流和彩色数据流，Kinect for windows sdk 中的骨骼 API 可以提供位于 Kinect 前方至多 5 个人的位置信息，包括详细的姿势和骨骼点的三维坐标信息还有用户 ID 信息，数据对象类型以骨骼帧的形式提供，每一帧可以保存 25 个骨骼点。

(2) 对步骤(1)选定的骨骼数据进行平滑处理，将骨骼点在骨骼空间坐标系中的坐标标准化；减少骨骼帧之间的骨骼点位置差异，kinect 开发工具 kinect for windows sdk 的数据对象类型以骨骼帧的形式提供，每一帧包括 25 个骨骼点。包括步骤如下：

C、设置平滑值(Smoothing)属性，所述平滑值的取值范围为 0 到 1 之间的浮点型数据，平滑值越大，平滑的越多，平滑值为 0 表示不进行平滑；

设置修正值(Correction)属性，所述修正值的取值范围为 0 到 1 之间的浮点型数据，修正值越小，骨骼信息越平滑；

设置抖动半径(JitterRadius)属性，所述抖动半径的取值范围为 0 到 1 之间的浮点型数据，当骨骼点抖动超过了设置的抖动半径，将会被纠正到这个抖动半径之内；

设置抖动半径的最大边界(MaxDeviationRadius)属性，取值范围为 0 到 1 之间的浮点型数据，

任何超过抖动半径的最大边界的点都不会认为是抖动产生的，而被认定为是一个新的点；

设置预测帧大小(Prediction)属性，取值范围为 0 到 1 之间的浮点型数据，默认值为 0；

D、调用 kinect 骨骼数据平滑算法即向 SkeletonStream.Enable()中传递步骤 C 中各个平滑参数对骨骼数据进行平滑处理。

对骨骼数据进行平滑处理会产生性能开销。平滑处理的越多，性能消耗越大。设置各个平滑参数没有经验可以遵循。需要不断的测试和调试已达到最好的性能和效果。在程序运行的不同阶段，可能需要设置不同的平滑参数。

在骨骼点跟踪过程中，有些情况会导致骨骼运动呈现出跳跃式的变化。例如，目标用户的动作不够连贯，Kinect 硬件性能差等等。骨骼点的相对位置可能在帧与帧之间变动很大，这会对应用程序产生一些负面的影响，例如，会影响用户体验和给控制造成意外等。通过对骨骼数据进行平滑，将骨骼点的坐标标准化，来减少帧与帧之间的骨骼点位置差异。

(3) 根据步骤(2)中已经实现骨骼空间坐标标准化的骨骼点在骨骼空间坐标系中的坐标，绘制骨骼向量，计算每个骨骼向量的方向余弦值，将这些方向余弦值作为特征进行提取；包括步骤如下：

E、根据人体结构学原理，针对步骤(2)提取到的 25 个骨骼点，依次将相邻的两个骨骼点连接组成一个骨骼段，并将这一个骨骼段定义为一个骨骼向量，共得到 24 个骨骼向量，即：

$\{b_1, b_2, b_3, \dots, b_{24}\}$ ；如图 3 所示；

F、设右手的肩关节到肘关节这个骨骼向量为 b_4 ，如图 4 所示，肩关节的三维坐标为 (x_1, y_1, z_1) ，肘关节的三维坐标为 (x_2, y_2, z_2) ，骨骼向量 b_4 表示如式 (I) 所示：

$$b_4 = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \quad (I)$$

G、设骨骼向量 b_4 与骨骼空间坐标系 X 轴、Y 轴和 Z 轴的夹角分别为 α 、 β 和 γ ，由此，可得到骨骼向量 b_4 的三个方向余弦值，计算公式如式 (II)、式 (III) 和式 (IV) 所示：

$$\cos \alpha = \frac{x_2 - x_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (II)$$

$$\cos \beta = \frac{y_2 - y_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (III)$$

$$\cos \lambda = \frac{z_2 - z_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (IV)$$

通过上述算法，依次计算 24 个骨骼向量的三个方向余弦值，并将这些余弦值作为特征进行提取，如图 5 所示。

当用户做出不同姿势时，对于人体每一个骨骼段而言，都具有不同的位置和角度信息，因此，可以利用定义的 24 个骨骼向量的方向余弦值来表征某一类动作。

(4) 获取训练数据集和测试数据集：这些数据集均是用户骨骼向量的方向余弦值，每一帧骨骼帧包含 24 组骨骼向量，每一个骨骼向量又包含三个方向余弦值；包括步骤如下：

H、获取训练数据集：目标用户在专业人员的指导下完成一系列的标准动作，通过步骤（1）至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即训练数据集；

I、获取测试数据集：目标用户在 Kinect2.0 摄像头有效视距内独自完成一系列动作，通过步骤（1）至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即测试数据集。

姿势矫正前期，用户会在专业人员的指导下完成一系列的标准动作，Kinect2.0 摄像头会以骨骼帧的形式记录下这些动作，然后经过骨骼点筛选、骨骼数据平滑处理、骨骼向量的方向余弦特征值提取，最后生成我们所需要的训练数据集，训练集总共有若干数据帧（帧数越多精度越高），每一帧骨骼数据包含一个姿势的 24 个骨骼向量，每个骨骼向量有三个方向余弦值；姿势矫正阶段，用户会在 Kinect 视野范围内独自完成一系列动作，Kinect 同样会以骨骼帧的形式记录下这些可能不规范的动作，然后经过相应的骨骼点筛选、骨骼数据平滑处理、骨骼向量的方向余弦特征值提取，最后生成我们所需要的测试数据集。

(5) 步骤（4）获取的训练数据集和测试数据集作为贝叶斯正则化 BP 神经网络的输入，通过贝叶斯正则化 BP 神经网络对 25 个骨骼点和 24 个骨骼段进行识别；包括步骤如下：

G、确定 BP 神经网络层数为 3，确定输入层的节点数为 3，确定输出层的节点数为 1，并通过式（V）确定隐含层的节点数 m ，式（V）如下所示：

$$m = \sqrt{n+l} + \alpha \quad (V)$$

式（V）中， n 为输入层的节点数； l 为输出层的节点数； α 为 1-10 间的整数；

通过对 BP 神经网络的学习，对于简单的动作识别问题通过单隐层网络即可实现，所以在此选择三层 BP 神经网络。输入层和输出层的节点数，一般是根据实际问题，由输入变量和输出变量的维数确定。BP 神经网络输入变量为三维方向余弦特征值，所以输入层的节点数为 3，输出变量每个骨骼向量与标准姿势的骨骼向量匹配与否即 1 或 0，所以输出层节点数为 1。隐含层节点数对网

络性能影响很大，节点数不同，最终结果将会差异很大，节点数太少，网络的迭代速率很快，但网络的建模不充分，导致网络性能较差；节点数太多，会导致网络结构复杂，计算量增大，训练时间较长，所以要选择合适的节点数。

K、将训练数据集作为训练集输入到贝叶斯正则化 BP 神经网络，对每一个骨骼向量训练学习，在姿势矫正阶段，将测试数据集输入到已训练完毕的贝叶斯正则化 BP 神经网络中，对每一个骨骼向量的正确性进行识别，即与训练集的标准骨骼姿势的相应骨骼向量匹配与否。

（6）结果分析和处理，将每个骨骼段和骨骼点的分析与处理结果实时显示在用户界面上。Kinect2.0 摄像头捕捉到的有效视距内目标用户的骨骼数据帧会实时呈现在用户界面上，与标准姿势相匹配的骨骼向量所对应的骨骼段和骨骼点会呈现出亮绿色，而与姿势错误的骨骼向量相对应的骨骼段和骨骼点会呈现出亮红色，同时会伴随语音提示。

权 利 要 求 书

1、一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，具体步骤包括：

(1) 基于用户 ID 和骨骼空间坐标深度值的有效骨骼点筛选：利用 Kinect2.0 摄像头获取用户骨骼数据和深度数据，选择目标用户的所有骨骼点；用户 ID 是指 Kinect2.0 传感设备唯一的用户跟踪标识，用于分配给有效视距内的每个用户，用以区分骨骼信息是哪个用户的；

(2) 对步骤(1)选定的骨骼数据进行平滑处理，将骨骼点在骨骼空间坐标系中的坐标标准化；

(3) 根据步骤(2)中已经实现骨骼空间坐标标准化的骨骼点在骨骼空间坐标系中的坐标，绘制骨骼向量，计算每个骨骼向量的方向余弦值，将这些方向余弦值作为特征进行提取；

(4) 获取训练数据集和测试数据集；

(5) 步骤(4)获取的训练数据集和测试数据集作为贝叶斯正则化 BP 神经网络的输入，通过贝叶斯正则化 BP 神经网络对 25 个骨骼点和 24 个骨骼段进行识别；

(6) 结果分析和处理，将每个骨骼段和骨骼点的分析与处理结果实时显示在用户界面上。

2、根据权利要求 1 所述的一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，所述步骤(1)，包括以下步骤：

A、利用 Kinect2.0 摄像头获取有效视距内所有物体的骨骼数据和深度数据，所述骨骼数据是指骨骼点在骨骼空间坐标系中的坐标；所述有效视距的取值范围为 0.8-3.0m；所述骨骼点包括：头部、颈部、右手食指、右手拇指、右手掌心、右手腕、右肘部、右肩、肩部中心、左肩、左肘部、左手腕、左手掌心、左手拇指、左手食指、脊柱、髋部中心、右髋、左髋、右膝盖、左膝盖、右脚踝、左脚踝、右脚、左脚 25 个骨骼点；

B、确定目标用户的骨骼点，对其他用户的骨骼点进行滤除：利用 Kinect2.0 摄像头获取每一个用户骨骼点的深度值，并进行累加，计算每一个用户的所有骨骼点的平均深度值，对比这些平均深度值，平均深度值最小的用户即目标用户，保存该目标用户的骨骼点，滤除其他用户的骨骼点。

3、根据权利要求 2 所述的一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，所述步骤(2)，包括步骤如下：

C、设置平滑值属性，所述平滑值的取值范围为 0 到 1 之间的浮点型数据，平滑值越大，平滑的越多，平滑值为 0 表示不进行平滑；

设置修正值属性，所述修正值的取值范围为 0 到 1 之间的浮点型数据，修正值越小，骨骼信息越平滑；

设置抖动半径属性，所述抖动半径的取值范围为 0 到 1 之间的浮点型数据，当骨骼点抖动超过

了设置的抖动半径，将会被纠正到这个抖动半径之内；

设置抖动半径的最大边界属性，取值范围为 0 到 1 之间的浮点型数据，任何超过抖动半径的最大边界的点都不会认为是抖动产生的，而被认定为是一个新的点；

设置预测帧大小属性，取值范围为 0 到 1 之间的浮点型数据，默认值为 0；

D、调用 kinect 骨骼数据平滑算法即向 SkeletonStream.Enable()中传递步骤 C 中的各个平滑参数对骨骼数据进行平滑处理。

4、根据权利要求 2 所述的一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，所述步骤（3），包括步骤如下：

E、根据人体结构学原理，针对步骤（2）提取到的 25 个骨骼点，依次将相邻的两个骨骼点连接组成一个骨骼段，并将这一个骨骼段定义为一个骨骼向量，共得到 24 个骨骼向量，即：

$$\{\vec{d}_1, \vec{d}_2, \vec{d}_3, \dots, \vec{d}_{24}\};$$

F、设任一骨骼向量 $\vec{d}_i$ 为在骨骼空间坐标系中的坐标为 (x_1, y_1, z_1) 的骨骼点与在骨骼空间坐标系中的坐标为 (x_2, y_2, z_2) 的骨骼点连接所得， $\vec{d}_i \in \{\vec{d}_1, \vec{d}_2, \vec{d}_3, \dots, \vec{d}_{24}\}$ ；骨骼向量 $\vec{d}_i$ 表示如式（I）所示：

$$\vec{d}_i = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \quad (\text{I})$$

G、设骨骼向量 $\vec{d}_i$ 与骨骼空间坐标系 X 轴、Y 轴和 Z 轴的夹角分别为 α 、 β 和 γ ，由此，可得到骨骼向量 $\vec{d}_i$ 的三个方向余弦值，计算公式如式（II）、式（III）和式（IV）所示：

$$\cos \alpha = \frac{x_2 - x_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{II})$$

$$\cos \beta = \frac{y_2 - y_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{III})$$

$$\cos \lambda = \frac{z_2 - z_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}} \quad (\text{IV})$$

通过上述算法，依次计算 24 个骨骼向量的三个方向余弦值，并将这些余弦值作为特征进行提取。

5、根据权利要求 4 所述的一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，所述步骤（4），包括步骤如下：

H、获取训练数据集：目标用户在专业人员的指导下完成一系列的标准动作，通过步骤（1）

至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即训练数据集；

I、获取测试数据集：目标用户在 Kinect2.0 摄像头有效视距内独自完成一系列动作，通过步骤（1）至步骤（3），Kinect2.0 摄像头获取每一个动作对应的 24 个骨骼向量的三个方向余弦值，即测试数据集。

6、根据权利要求 5 所述的一种基于深度信息和骨骼信息的姿势矫正方法，其特征在于，所述步骤（5），包括步骤如下：

G、确定 BP 神经网络层数为 3，确定输入层的节点数为 3，确定输出层的节点数为 1，并通过式（V）确定隐含层的节点数 m ，式（V）如下所示：

$$m = \sqrt{n+l} + \alpha \quad (V)$$

式（V）中， n 为输入层的节点数； l 为输出层的节点数； α 为 1-10 间的整数；

K、将训练数据集作为训练集输入到贝叶斯正则化 BP 神经网络，对每一个骨骼向量训练学习，在姿势矫正阶段，将测试数据集输入到已训练完毕的贝叶斯正则化 BP 神经网络中，对每一个骨骼向量的正确性进行识别，即与训练集的标准骨骼姿势的相应骨骼向量匹配与否。

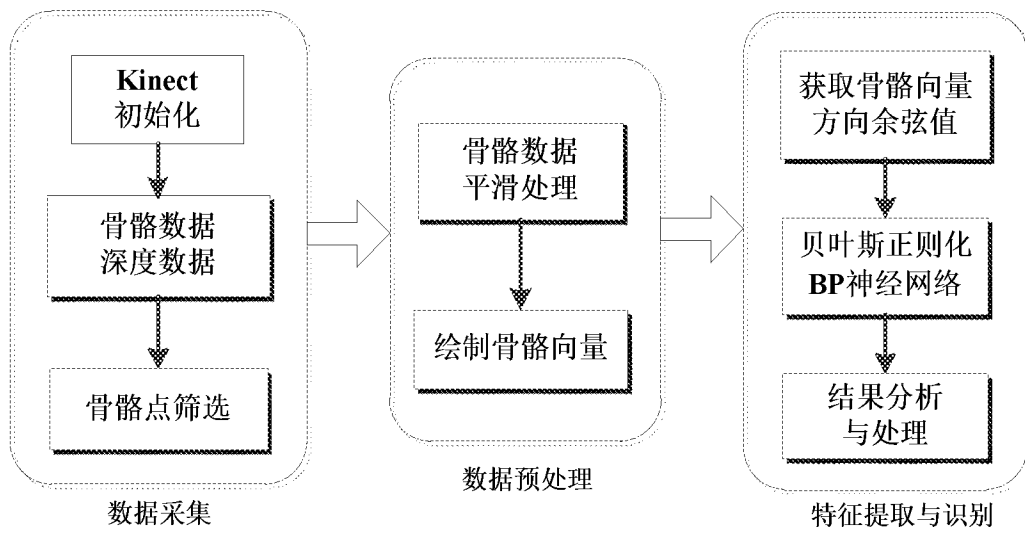


图 1

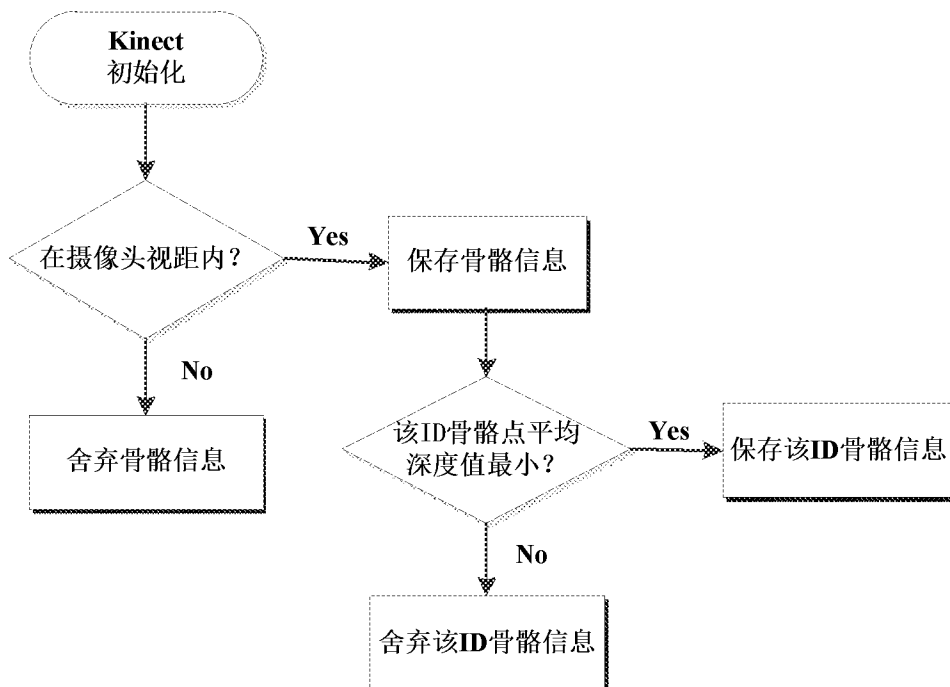


图 2

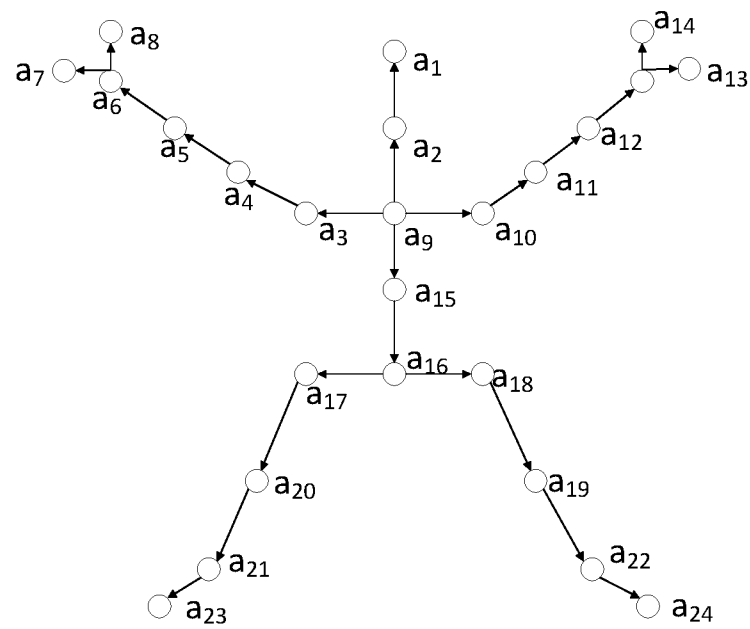


图 3

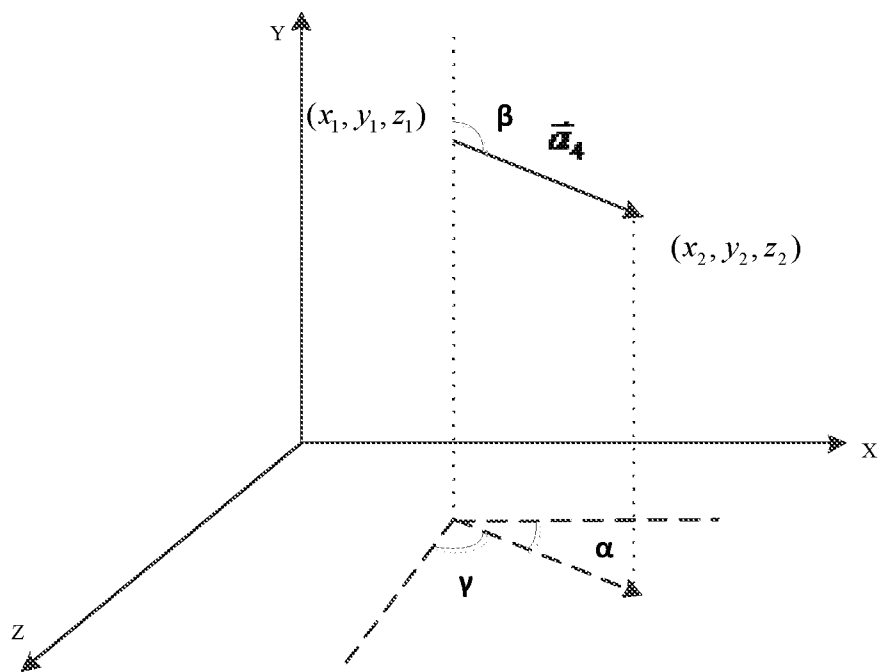


图 4

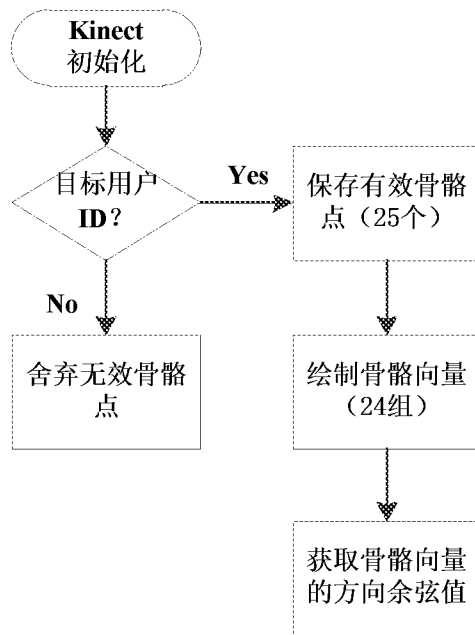


图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/104990

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K, G06F, H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE: 深度信息, 骨骼, 姿势, 姿态, 矫正, kinect, 坐标, BP, 神经网络, depth, information, skeleton, gesture, correction, coordinate, neural, network

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106650687 A (SHANDONG UNIVERSITY), 10 May 2017 (10.05.2017), claims 1-6	1-6
A	CN 105307017 A (TCL CORPORATION), 03 February 2016 (03.02.2016), description, paragraphs [0004] and [0028]-[0038], and figure 4	1-6
A	CN 104484574 A (DONGHUA UNIVERSITY), 01 April 2015 (01.04.2015), entire document	1-6
A	CN 105005769 A (SHANDONG UNIVERSITY), 28 October 2015 (28.10.2015), entire document	1-6
A	US 6788809 B1 (INTEL CORPORATION), 07 September 2004 (07.09.2004), entire document	1-6

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 19 December 2017	Date of mailing of the international search report 05 January 2018
---	---

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer
JU, Bo
Telephone No. (86-10) 62413661

International application No.
PCT/CN2017/104990

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2017/104990

A. 主题的分类

G06K 9/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06K G06F H04N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, EPODOC, CNPAT, CNKI, IEEE: 深度信息, 骨骼, 姿势, 姿态, 矫正, kinect, 坐标, BP, 神经网络, depth, information, skeleton, gesture, correction, coordinate, neural, network

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 106650687 A (山东大学) 2017年 5月 10日 (2017 - 05 - 10) 权利要求1-6	1-6
A	CN 105307017 A (TCL集团股份有限公司) 2016年 2月 3日 (2016 - 02 - 03) 说明书第[0004]、[0028]-[0038]段、附图4	1-6
A	CN 104484574 A (东华大学) 2015年 4月 1日 (2015 - 04 - 01) 全文	1-6
A	CN 105005769 A (山东大学) 2015年 10月 28日 (2015 - 10 - 28) 全文	1-6
A	US 6788809 B1 (INTEL CORPORATION) 2004年 9月 7日 (2004 - 09 - 07) 全文	1-6

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2017年 12月 19日

国际检索报告邮寄日期

2018年 1月 5日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

鞠博

传真号 (86-10)62019451

电话号码 (86-10)62413661

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/104990

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106650687	A	2017年 5月 10日	无	
CN	105307017	A	2016年 2月 3日	无	
CN	104484574	A	2015年 4月 1日	无	
CN	105005769	A	2015年 10月 28日	无	
US	6788809	B1	2004年 9月 7日	无	

表 PCT/ISA/210 (同族专利附件) (2009年7月)