

(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
2023 年 7 月 27 日 (27.07.2023)



(10) 国际公布号
WO 2023/138445 A1

(51) 国际专利分类号:
G06V 40/20 (2022.01)

(21) 国际申请号: PCT/CN2023/071657

(22) 国际申请日: 2023 年 1 月 10 日 (10.01.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210078571.1 2022 年 1 月 24 日 (24.01.2022) CN

(71) 申请人: 索尼半导体解决方案公司 (SONY SEMICONDUCTOR SOLUTIONS CORPORATION)
[JP/JP]; 日本神奈川县厚木市旭町 4-14-1,
Kanagawa 243-0014 (JP)。

(72) 发明人: 及

(71) 申请人 (仅对 MG): 萨里尼约瑟夫 (SALINI, Joseph)
[FR/CN]; 中国上海市黄浦区湖滨路 222 号领展企
业广场一座 8 楼, Shanghai 200021 (CN)。

(74) 代理人: 北京信慧永光知识产权代理有限
责任公司 (BEIJING SUNHOPE INTELLECTUAL

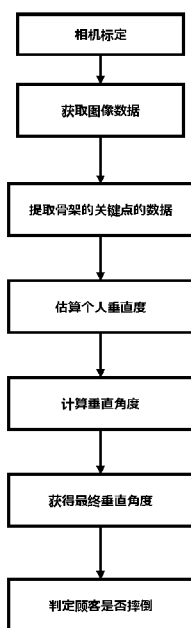
PROPERTY LTD.); 中国北京市海淀区知春路 9 号
坤讯大厦 1106 室, Beijing 100191 (CN)。

(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,
CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ,
IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ,
LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, 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, CV, 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, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,

(54) **Title:** DETECTION METHODS AND DEVICES FOR DETECTING IF PERSON HAS FALLEN AND PICK-UP OR
PUT-BACK BEHAVIOR OF PERSON

(54) 发明名称: 用于检测人摔倒、人的拾取或放回行为的检测方法和设备



AA: Camera calibration
BB: Obtain image data
CC: Extract key point data of skeleton
DD: Estimate personal verticality
EE: Calculate vertical angles
FF: Obtain final vertical angle
GG: Determine whether customer has fallen

(57) **Abstract:** The present invention discloses a detection method and device for detecting if a person has fallen on premises and a detection method and device for detecting a pick-up or put-back behavior of a person in front of a shelf. The method for detecting if a person has fallen comprises the steps: calibrating cameras, so as to have a suitable vertical vector in a field of view of each camera; obtaining image data comprising a person on premises by means of least a portion of the cameras, and extracting from the image data key point data of the skeleton of the person; estimating a personal verticality of the person by means of using the key point data; calculating vertical angles of the person on the basis of the personal verticality and the vertical vectors in the field of view of at least a portion of the cameras; aggregating all the vertical angles of the person in the fields of view of at least a portion of the cameras at some moment to obtain a final vertical angle of the person; and determining whether the person has fallen down on the basis of a fall score calculated from the final vertical angle. The right to privacy of a customer can be protected, and detection can be accurately and quickly performed.

(57) **摘要:** 本发明公开了用于检测场所中的人的摔倒的检测方法和设备以及用于检测人在货架前的拾取或放回行为的检测方法的设备。用于检测人的摔倒的方法包括步骤: 对相机进行标定, 以使在各相机的视场中均具有恰当的垂直向量; 通过至少一部分相机获得包含场所中的人的图像数据, 并且从图像数据中提取出人的骨架的关键点的数据; 通过使用关键点的数据估算人的个人垂直度; 基于针对至少一部分相机的视场中的垂直向量和个人垂直度计算人的垂直角度; 通过对人在某一时刻在至少一部分相机的各视场中的所有垂直角度进行聚合, 获得人的最终垂直角度; 基于由最终垂直角度求出的摔倒分值判定人是否摔倒。根据本发明, 能够保护顾客隐私权并且准确快速地进行检测。

图 1

CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布：

- 包括国际检索报告(条约第21条(3))。

用于检测人摔倒、人的拾取或放回行为的检测方法和设备

相关申请的引用

本申请要求于 2022 年 01 月 24 日向中华人民共和国国家知识产权局提交的第 202210078571.1 号中国专利申请的权益，在此将其全部内容以援引的方式整体并入本文中。

技术领域

本发明涉及计算机视觉中的人类行为检测和识别。具体地，本发明涉及能够准确地检测和识别人的摔倒、人的拾取或放回行为的检测方法和设备。

背景技术

作为一门研究如何使机器通过“看”去理解世界的学科，计算机视觉技术是目前人工智能领域最热门的研究领域之一。具体来说，通过相机与计算单元的结合，在一定场景下机器视觉系统能够代替人眼对目标进行识别、跟踪和测量等工作。在计算机视觉技术的诸多应用领域之中，基于计算机视觉技术的人类行为的检测和识别是非常重要的一个方面，并且被应用于例如视频监控、群体互动、消费者行为识别等等场景。

尤其是近年来，随着无人超市、自助零售商店等智能零售场景模式的普及，大量的研究都关注于如何检测并识别无人零售环境中的人（即，顾客）的行为。在当前典型的无人零售环境（例如，小型无人超市）中，顾客在通过身份识别入场后，往往随意地在场所中移动并随机地在货架前拾取或放回货品，然后在无购物或经确认已自助结账的情况下离场。因此，通过计算机视觉检测并识别出顾客在购物场所中的行为，尤其是货架前对于商品的拾取和放回行为，不仅对于追踪并确认顾客的最终购买行为，更加快速地确定顾客的消费金额，提升顾客的消费体验有着非常直接的帮助，而且能够对无人零售环境中的货架布置、货品摆放等工作提供非常有价值的参考信息。另外，由于无人超市等购物场所往往为人流量较少的封闭空间，因此基于最基本的安全需求，还要求能够通过计算机视觉及时地发现在其中的顾客的例如突然倒地等具有潜在安全风险的行为。

目前，用于检测并识别顾客上述行为的最佳方法是创建一个自动系统，该系统通常具有 RGB 摄像机和智能信息处理设备。该系统首先通过为垂直安装的 RGB 摄像机检测到的每名顾客分配一个唯一的 ID，跟踪他们在商店内的活动，并且检测他们与货架的互动。接着进行分析和分类交互，记录顾客在商店内的活动，指示商品是否已被顾客从货架取出或者取出后的商品是否又被放回原处。由此，系统能够识别货架前顾客的行为活动。

发明内容

要解决的技术问题

然而，当前应用于无人超市等智能零售场景的现有的顾客行为检测和识别的技术方案通常都需要 RGB 摄像机拍摄大量的图像视频文件，并涉及进行基于这些图像视频文件的大量的文件传输和处理。因此，导致整个系统的数据处理速度较为缓慢。另外，如上所述，这些方法往往需要为每名顾客分配 ID，而这样的处理通常是基于面部识别完成的。在未获得顾客授权的情况下，这样的信息收集和处理会有侵犯顾客隐私权的风险。

鉴于上述问题，本发明期望提供一种能够在保护顾客隐私权的前提下准确、快速地检测和识别顾客的摔倒、顾客的拾取或放回行为的检测方法和设备。

技术方案

本发明所要解决的技术问题是通过如下技术方案实现的。

根据本发明的第一实施例，提供了一种用于检测分布有相机的场所中的人的摔倒的检测方法，包括如下步骤：对所有所述相机进行标定，以使在各所述相机的视场中均具有恰当的垂直向量；通过至少一部分所述相机获得包含所述场所中的人的图像数据，并且从所述图像数据中提取出所述人的骨架的关键点的数据；通过使用所述关键点数据估算所述人的个人垂直度；针对所述至少一部分相机中的各者，基于相应的相机的所述视场中的所述垂直向量和所述个人垂直度计算所述人的垂直角度；通过对所述人在某一时刻在所述至少一部分相机的各所述视场中的所有所述垂直角度进行聚合，获得所述人的最终垂直角度；基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。

根据本发明的第一实施例，还提供了一种用于检测场所内的人的摔倒的检测设备，包括：多个相

机,所述多个相机分布在所述场所内并且具有不同的视场,所述多个相机能够获得包含所述场所中的人的图像数据;处理单元,所述处理单元对所述多个相机获得的所述图像数据进行处理,以判定所述场所内的所述人是否摔倒。所述处理单元包括:标定模块,对所有所述相机进行标定,以使在各所述相机的视场中均具有恰当的垂直向量;数据处理模块,对从所述多个相机传输来的所述图像数据进行处理,从而获得所述人在所述多个相机之中的至少一部分相机的所述视场中的个人垂直度;计算模块,基于从所述标定模块发送来的所述垂直向量和从所述数据处理模块发送来的所述个人垂直度计算所述人的最终垂直角度;以及判定模块,基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。

根据本发明的第一实施例,还提供了一种存储介质,其上存储有计算机可读程序,所述程序在处理器上执行时实施前述的根据本发明的第一实施例的检测方法。

根据本发明的第二实施例,提供了一种用于检测人在货架前的拾取或放回行为的检测方法,包括如下步骤:从图像数据中获取所述人的骨架的包含手部关键点的多个关键点的数据,并且从所述图像数据提取所述货架的外轮廓线,其中,所述外轮廓线包括所述货架的外部多边形以及与所述货架的真实外轮廓相对应的内部多边形,所述外部多边形在所述内部多边形的外部的接近区域中;在所述人的至少一只手的所述手部关键点被检测到进入所述外部多边形的情况下,针对所述人的进入所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的进入物品检测;在所述人的至少一只手的所述手部关键点被检测到退出所述外部多边形的情况下,针对所述人的退出所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的退出物品检测;基于所述进入物品检测的结果和所述退出物品检测的结果来判定所述人在所述货架前的所述拾取或放回行为。优选地,在执行所述进入物品检测与进行所述退出物品检测之间还包括步骤:针对所述人的进入所述外部多边形的每只手,分别记录所述手部关键点附近的所述物品在所述外部多边形与所述内部多边形之间的轨迹。

根据本发明的第二实施例,还提供了一种用于检测人在货架前的拾取或放回行为的检测设备,包括:至少一台相机或图像传感器,用于获取图像数据;处理单元,所述处理单元根据前述的用于检测人在货架前的拾取或放回行为的检测方法对所述图像数据进行处理。

根据本发明的第二实施例,还提供了一种存储介质,其上存储有计算机可读程序,所述程序在处理器上执行时实施前述的根据本发明第二实施例的检测方法。

技术效果

根据本发明,由于在获取视频或图片数据后就立即对这些数据进行了简化提取处理,在后续的所有步骤中均只需要传输和处理提取出的图片中的关键点的数据,因此大大减少了数据处理量,能够快速准确地实现顾客的行为的检测和识别。此外,由于不会对顾客进行面部识别,并且在经过最初的数据简化提取处理后,存储和传输的数据中将只包含顾客骨架关键点的相关信息,因此保护了顾客隐私权。

应当理解,本发明的有益效果不限于上述效果,而可以是本文中说明的任何有益效果。

附图说明

构成本申请的一部分的说明书附图用来提供对本发明的进一步理解,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

图1是图示了根据本发明第一实施例的用于检测和识别顾客的摔倒的方法的主要步骤的框图;

图2是图示了根据本发明第一实施例的相机的标定的示意图;

图3是图示了根据本发明第一实施例的从图像数据中提取的顾客的骨架的关键点的示意图;

图4是图示了根据本发明第一实施例的在相机的拍摄视场中的顾客的个人垂直度的示意图;

图5的a和b分别图示了根据本发明第一实施例的在不同的拍摄视场中的顾客的个人垂直度以及垂直角度的示意图;

图6的a至e分别图示了根据本发明第一实施例的在相机的拍摄视场中当顾客站立时的垂直角度的示例;

图7的a至d分别图示了根据本发明第一实施例的当同一顾客摔倒时在不同相机的拍摄视场中的不同垂直角度的示意图;

图8图示了根据本发明第一实施例的最终垂直角度与摔倒分值的转关系的示意图;

图 9 的 a 和 b 图示了根据本发明第一实施例的顾客的最终垂直角度和摔倒分值的转换以及顾客是否摔倒的判定的示例图；

图 10 是图示了根据本发明第一实施例的用于检测和识别顾客的摔倒的设备的示意性框图；

图 11 是图示了根据本发明第二实施例的用于检测和识别顾客的拾取或放回行为的方法的主要步骤的示意性框图；

图 12 是图示了顾客在货架前的主要拾取或放回行为的分类示意图；

图 13 的 a 和 b 图示了在根据本发明的第二实施例的检测方法中的物品存在的判定条件的示例。

图 14 图示了在根据本发明的第二实施例的检测方法中的物品的轨迹比较的示意图；

图 15 图示了在根据本发明的第二实施例的检测方法中的第一个 FSM 运行时的状态判定的示例性流程图；

图 16 图示了在根据本发明的第二实施例的检测方法中的第二个 FSM 运行时的状态判定的示例性流程图；

图 17 图示了在根据本发明的第二实施例的检测方法中的仅有一个 FSM 运行的情况下获得的 FSM 的状态表；

图 18 中的 a 和 b 图示了在根据本发明的第二实施例的检测方法中的两个 FSM 运行的情况下获得的 FSM 的状态表；

图 19 图示了在根据本发明的第二实施例的检测方法中的将图 18 中所示的两个 FSM 状态表相结合后的最终状态表；

图 20 图示了在根据本发明的第二实施例的检测方法中的考虑基于多组图像数据的检测结果时的物品存在的判定条件的示例。

具体实施方式

下面，将参照附图详细说明根据本发明的各具体实施例。需要强调的是，附图中的所有尺寸、形状、位置关系等仅是示意性的，并且为了图示方便而不一定是按照真实比例图示的，因而不具有限定性。此外，在以下所述的实施方案中，相同的部件、配置和步骤等由相同的附图标记表示，并且将省略对它们的重复说明。

此外，将根据以下项的顺序来说明本发明。

- 1、第一实施例（用于检测和识别顾客的摔倒的方法和设备）
- 2、第二实施例（用于检测和识别顾客在货架前的拾取或放回行为的方法和设备）
- 1、第一实施例

1. 1、用于检测和识别顾客的摔倒的方法和设备的概述

首先，将说明根据本发明第一实施例的用于检测和识别顾客的摔倒的方法和设备的概况。图 1 示出了根据本发明第一实施例的用于检测和识别顾客的摔倒的方法的主要步骤。

在诸如无人超市、自助服务商店等无人智能售卖环境中，由于没有服务人员巡场和值守，所以当顾客在这样的环境中例如因突发疾病或意外事故而摔倒时，就存在着因无法及时发现而耽误救治的隐患。为了解决上述问题，根据本发明第一实施例，在这样的无人售卖环境中分布有能够拍摄获取环境的图像数据的相机。为了能够准确地识别出环境中顾客的姿态，首先需要对所有的相机进行标定处理，以使每一台相机都能够真实地反映出各自的拍摄视场中的垂直向量。经过标定后的相机获取无人售卖环境的环境图像数据，并对环境的图像数据执行数据提取处理。当提取到的环境图像数据中存在顾客图像时，则能够通过数据提取处理将图像数据中的顾客图像简化为骨架（skeleton）图像，进而提取出的顾客的关键点数据信息。后续的处理都是基于这些不会透露顾客隐私信息的关键点数据信息进行的。随后，通过使用这些关键点信息估算出顾客在各个相机的拍摄视场中的个人垂直度。容易理解地，这样的个人垂直度例如可以通过关键点数据中从代表顾客足部的关键点到代表顾客的头部是关键点的向量来表示。然后，针对各个相机的拍摄视场，基于在前面步骤中求出的垂直向量和个人垂直度计算出各拍摄视场中的顾客的垂直角。容易理解地，由于各个相机的拍摄视角不同，因此需要从不同拍摄视场内的垂直角中选择出最能够反映顾客真实身体姿态的垂直角。换言之，需要对在某时刻由各个相机拍摄的垂直角进行数据聚合，并从中获得最终的能够反映顾客真实身体姿态的垂直角。显然，当人站立时和摔倒在地时，垂直角显然存在巨大差异。因此，随后就能够基于由垂直角得出的摔倒分值来确定被拍摄到的顾客是否摔倒。

上述的根据本发明第一实施例的用于检测和识别顾客的摔倒的方法中的除了获取环境的图像数据的步骤是由各相机执行之外，其余的各步骤的处理均可通过与各相机通信连接的中央处理器等处理器或数据处理芯片执行。或者，优选地，这些步骤中的处理也可以由集成在各相机中的具有 AI 处理功能的传感器执行。这样的传感器兼具数据处理能力和数据存储能力，无需额外的硬件即可执行机器学习驱动的计算机视觉处理任务。

在下文中，将详细说明根据本发明第一实施例的用于检测和识别顾客的摔倒的方法中的各个步骤中的处理。

1. 2、相机的标定

图 2 示出了根据本发明第一实施例的用于检测和识别顾客的摔倒的方法中的相机的标定处理的示意图。

由于相机往往以不同的角度分布在无人超市等场所中的不同位置处，因此，如图 2 中的短实线所示，在各个相机的拍摄视场中，垂直向量的分布状态取决于拍摄相机的拍摄角度和图像中的物体所处的位置。换言之，在不同的相机的拍摄视场中，垂直向量的分布状态通常是不相同的。因而，需要对各相机进行垂直向量的标定处理，以使各相机能够真实地反映出所拍摄的视场中不同位置处的垂直向量方向，从而能够在后续的步骤中为判断拍摄到的顾客是否站立提供基准。

相机的标定(calibration)简单来说就是从世界坐标系转换为相机坐标系，再由相机坐标系转换为图像坐标系的过程。换言之，也就是求出相机的最终的投影矩阵的过程。具体而言，世界坐标系(world coordinate system)是指用户定义的三维世界的坐标系，为了描述目标物在真实世界里的位置而被引入；相机坐标系(camera coordinate system)是指在相机上建立的坐标系，为了从相机的角度描述物体位置而定义，作为沟通世界坐标系和图像/像素坐标系的中间一环；图像坐标系(image coordinate system)是指为了描述成像过程中物体从相机坐标系到图像坐标系的投影透射关系而引入，方便进一步得到像素坐标系下的坐标。从世界坐标系到相机坐标系的转换是从三维点到三维点的转换，需要使用旋转矩阵 R 和平移矢量 t 等相机外参(camera extrinsic information)。从相机坐标系到图像坐标系的转换是从三维点到二维点的转换，需要使用焦距、像主点坐标、畸变参数等相机内参。例如，可以通过使用棋盘格图片等作为标定图片，并利用 OpenGL、OpenGL 等软件工具中的相机标定功能实现上述的相机的标定。

1. 3、顾客的骨架及其关键点数据的提取

在根据本发明第一实施例的用于检测和识别顾客的摔倒的方法中，在经过标定的相机拍摄到图像数据之后，在进行后续的处理之前，会对图像数据进行数据提取处理。在拍摄到顾客的情况下，这样的数据提取处理会将原始的图像数据转换为仅包含图像中所需的诸如顾客的骨架信息等简化图像数据，从而极大地减少了后续数据传输和处理的数据量，并且避免了侵犯顾客隐私权的风险。需要说明的是，现有技术中已经存在大量的用于从多种类型的包含人物图像的图像数据中提取人体骨架特征点的技术，本文中为了突出说明重点，在此就不再详述。可以采用诸如 Hrnet 等任何合适的已知的专用软件和算法来进行本步骤中的人体骨架及其关键点的数据提取。例如，如图 3 所示，可以提取出包括人的眼部、耳朵、鼻子、肩部、肘部、腕部、胯部、膝部、脚踝等部位在内的 17 个部位的图像数据作为人体骨架的关键点信息。

需要特别说明的是，在用于图示本实施例的附图中，在示出提取出的顾客骨架关键点信息之外，为了便于理解和说明，也示出了完整的图片信息。但在实际的使用过程中，在提取步骤之后的各步骤中，均不需要传输、存储和处理完整的图片或视频数据。

1. 4、顾客的个人垂直度的估算

图 4 示出了基于提取出的顾客的人体骨骼关键点信息来获得顾客在相应的相机拍摄视场中的个人垂直度的示意图。

如图 4 中的箭头所示，可以通过计算从表示人的脚部的关键点到表示人的头部的关键点的身体向量来获得个人垂直度信息。具体地，如上文提到的，提取出的与人的头部相关的关键点例如可以包括人的眼部、耳朵、鼻子等部位的关键点；提取出的与人的脚部相关的关键点例如可以包括人的踝部等部位的关键点。另外，由于不同的相机具有不同的拍摄视场，可以想象的是，处于某一特定位置和姿态下的人对于不同的相机而言，其可拍摄到的关键点可能是不同的。因此，可以使用对于相应的相机而言是可见的顾客的眼部、耳朵、鼻子的关键点信息的平均值来代表在该相机拍摄的图片中的顾客的头部位置信息，使用对于相应的相机而言是可见的顾客的踝部的关键点信息的平均值来代表在该相机拍摄的图片中的顾客的脚部位置信息。例如，图 5 中的 a 示出了在某相机的拍摄视场中，顾客的一个踝部是不可见的，因此仅使用可见的那一个踝部的关键点信息来计算在该拍摄视场中在该拍摄时刻的顾客的个人垂直度。图 5 的 b 示出了在另一相机的拍摄视场中，顾客的一个踝部不可见，且眼部和鼻子也不可见。因此，使用双耳的关键点信息的平均值以及可见的那一个踝部的关键点信息来计算该拍摄视场中在该拍摄时刻的顾客的个人垂直度。

此外, 还可以定义置信因子 $c=N_v/N_t$, 其中, N_v 表示从某相机拍摄的图片数据中 (例如, 特定帧中) 提取的顾客的可见的关键点数量, N_t 表示定义的关键点总数。例如, 可以设定当 $c<0.5$ (也即是, 当有一半以上的关键点不可见时) 时, 我们认为从该相机拍摄的该图片数据中提取的顾客的关键点信息不足 (例如, 在此时顾客并没有完全走入该相机的拍摄视场中), 该相机在此情况下是无效相机。那么, 将不使用这样的无效相机所拍摄的图片中的关键点信息来计算顾客的个人垂直度。置信因子的置信阈值不限于 0.5, 可以根据需要而更改和设定。应当理解的是, 随着顾客在环境中的移动, 在不同的拍摄时刻 (例如, 不同帧), 某 RGB 相机的有效/无效状态可能会发生转变。

1. 5、顾客的垂直角度的确定

如图 5 的 a 和 b 中的箭头所示, 通过将在上一步骤中求出的表示个人垂直度的关键点向量与所在拍摄视场的对应的垂直向量进行比较, 就可以获得这两个向量之间的角度 α_v , 作为表示顾客的姿态的垂直角度。关于顾客在不同姿态下的垂直角度, 容易想象的是, 如图 6 的 a 至 e 所示, 当顾客站立时, 对于所有相机而言, 垂直角度 α_v 都接近于 0。

然而, 在顾客摔倒在地 (例如, 如图 7 的 a 至 d 所示), 由于不同相机的拍摄视场不同, 基于各相机所拍摄的图片信息所求出的反映同一顾客相同姿态的垂直角度却可能存在差异。换言之, 在顾客摔倒的情况下, 往往每个相机所拍摄到的图片信息只能反映出顾客姿态的真实信息的一部分, 因此需要对从所有有效相机拍摄的图片信息中提取出的数据进行数据聚合 (data aggregation), 以获得反映真实信息的全局结果。例如, 这里的数据的聚合可以定义为:

$$\alpha_v = \max_c(\alpha_{v_c}) \quad \text{式(1)}$$

其中, c 表示有效相机。

由上式 (1) 可知, 可以使用基于各有效相机拍摄的图片信息而求出的垂直角度中的最大值作为反映拍摄时刻顾客的姿态的最终垂直角度。当然, 可以采用其他的数据聚合方式获得最终垂直角度。例如, 也可以采用基于各有效相机拍摄的图片信息而求出的垂直角度中的第二大值作为反映拍摄时刻顾客的姿态的最终垂直角度。或者, 也可以采用基于各有效相机拍摄的图片信息而求出的垂直角度中的去掉最大值和最小值之后的其余垂直角度的平均值作为反映拍摄时刻顾客的姿态的最终垂直角度。

1. 6、基于摔倒分值确定顾客是否摔倒

由于在最终的判定结果中, 顾客仅存在 “站立” 和 “摔倒” 两种姿态判定结果, 因此, 可以将最终垂直角度转换为摔倒分值作为最终判定顾客是否摔倒的基准。

例如, 可以如下定义摔倒分值。

若 $\alpha_v < T_l$, 则 $s_f = 0$;

若 $T_l < \alpha_v < T_h$, 则 $s_f = \frac{\alpha_v - T_l}{T_h - T_l}$;

若 $\alpha_v > T_h$, 则 $s_f = 1$

式 (2)

其中, T_l 为设定的垂直角度的下限, T_h 为设定的垂直角度的上限, s_f 为摔倒分值。 T_l 和 T_h 可以根据需要而进行设定和调整, 例如 T_l 可以设定为 40 度, T_h 可以设定为 80 度。

根据上面的式 (2) 可知, 摔倒分值可以被看成是一个模糊逻辑值。其如图 8 所示, 具有在 0 与 1 (包括端点) 之间浮动变化的值, 以反映 “站立” 和 “摔倒” 状态。可以根据需要, 设定摔倒分值的判定阈值 s_T , 从而能够基于摔倒分值与判定阈值的比较, 准确地对顾客的 “站立” 和 “摔倒” 状态进行判定。例如, 若 $s_f > s_T$, 则判定顾客处于 “摔倒” 状态; 否则, 则判定顾客处于 “站立” 状态。 s_T 的值可以根据安全需求等因素而进行设定。优选地, s_T 的值可以被设定为 0.5 与 0.8 之间 (包含端点) 的值。例如, s_T 的值可以设定为 0.5。

图 9 的 a 和 b 示出了在 T_l 设定为 40 度, T_h 设定为 80 度, 且 s_T 设定为 0.5 的情况下, 基于某顾客的所有有效相机拍摄的图片信息获得的最终垂直角度 α_v 与摔倒分值 s_f 的转换关系以及最终的状态判定结果。

1. 7、用于检测和识别顾客的摔倒的设备的示例

图 10 图示了根据本发明第一实施例的用于检测和识别顾客的摔倒的设备的构成的示意性框图。例如, 根据本发明第一实施例的用于检测和识别顾客的摔倒的检测设备 1 可以包括多个相机 101 和处理单元 102。

在诸如无人超市等无人零售场所内可以布置有一个或多个相机 101, 其具有不同的视野并且能够获得包括该场所内的顾客的图像数据。相机 101 可以是 RGB 相机, 也可以是 RGB-D 相机等其它类型的相机, 以获取诸如 IR 图像、RGB 图像或激光图像等类型的图像信息。

处理单元 102 接收由多个相机拍摄获得的图像信息, 并包含多个数据处理模块。处理单元 102 能够通

过多个组成模块执行作为软件或固件而被存储在其中的存储元件中或与其互连的存储器或数据存储单元中的应用程序或例程，从而执行上文中所述的根据本发明第一实施例的用于检测和识别顾客的摔倒的方法。这里的处理单元 102 例如由中央处理单元 (CPU) 和存储元件构成。例如，处理单元 102 可以包括一个或多个通用处理器、控制器、现场可编程门阵列 (FPGA)、图形处理单元 (GPU)、专用集成电路 (ASIC) 或它们的组合等，作为与各个相机 101 数据通信的专用数据处理器或数据处理芯片。或者，处理单元 102 也可以是集成在各相机 101 中的彼此能够互相进行数据交互的具有 AI 处理功能的传感器。这样的传感器兼具数据处理能力和数据存储能力，无需额外的硬件即可与相机 101 配合执行根据本发明第一实施例的用于检测和识别顾客的摔倒的方法。

例如，处理单元 102 包括：标定模块 1021，其用于对多个相机 101 进行标定，以使各相机 101 的视野中具有正确的垂直向量；数据处理模块 1022，其用于处理从所述多个相机 101 发送来的图像数据，以获得在多个相机 101 的至少一部分相机的拍摄视场中的顾客的个人垂直度；计算模块 1023，用于基于从标定模块 1021 发送的垂直向量和从数据处理模块 1022 发送的个人垂直度来计算顾客在至少一部分相机的拍摄视场中的垂直角度，并对各个垂直角度进行聚合以获得顾客的最终垂直角度；和判定模块 1024，用于将最终垂直角度转换为摔倒分数，并根据摔倒分数判定顾客是否摔倒。其中，数据处理模块 1022 可以包括：提取模块 10221，其从图像数据中提取出顾客的骨架的关键点的数据；和估算模块 10222，其通过使用关键点的数据估算出顾客的个人垂直度。其中，计算模块 1023 可以包括角度计算模块 10231 和聚合模块 10232。

估算模块，通过使用来自所述提取模块的所述关键点的数据估算所述人的所述个人垂直度。

根据本发明的第一实施例，由于在获取视频或图片数据后就立即对这些数据进行了提取关键点信息的数据简化处理，在后续的所有步骤中均只需要传输和处理提取出的人骨架的关键点数据，因此大大减少了数据处理量，能够快速、准确地识别出无人购物场所中顾客的摔倒状态和站立状态。此外，由于不需要对顾客进行面部识别，并且在经过最初的数据简化处理后，存储和传输的数据中将只包含顾客的骨架关键点数据，因此保护了顾客隐私权。

2、第二实施例

2.1、用于检测和识别顾客的拾取和放回行为的方法和设备的概述

首先，将说明根据本发明第二实施例的用于检测和识别顾客的拾取和放回物品行为的方法和设备的概况。图 11 示出了根据本发明第二实施例的用于检测和识别顾客的拾取和放回物品行为的方法的主要步骤。

在诸如无人超市、自助服务商店等无人智能售卖环境中，由于没有服务人员巡场和值守，所以需要借助计算机视觉手段确定顾客对货架上陈列的商品的取放行为，这样不仅能够辅助自助结算系统更快速地进行商品结算，还可以更有效地防止偷盗行为。因此，根据本发明第二实施例，提供了用于检测和识别顾客的拾取和放回行为的方法和设备。

首先，将概述根据本发明第二实施例的用于检测和识别顾客的拾取和放回行为的方法的主要步骤。需要说明的是，在开始执行根据本发明第二实施例的用于检测和识别顾客的拾取和放回行为的方法之前，需要获得布置有货架的购物场所的图像数据。这样的图像数据可以由布置在无人购物场所中的至少一台相机拍摄的，也可以是已经被传输至并存储在用于执行根据本发明第二实施例的用于检测和识别顾客的拾取和放回行为的方法的诸如 AI 传感器等处理器中。因此，这样的图像数据获取步骤既可以被看做是根据本发明第二实施例的用于检测和识别顾客的拾取和放回行为的方法中的一个步骤，也可以被看做是根据本发明第二实施例的用于检测和识别顾客的拾取和放回行为的方法的准备步骤。

在获得图像数据后，执行从图像数据中提取后续处理所需的数据的数据提取处理。在本实施例中，这样的数据提取处理包括两个部分：货架的轮廓数据的提取和顾客的包括手部、腕部等骨架的关键点数据的提取。顾客的包括手部、腕部等骨架的关键点数据的提取与第一实施例中的信息提取步骤类似。因此，同样地，极大地减少了后续数据传输和处理的数据量，并且避免了侵犯顾客隐私权的风险。需要说明的是，现有技术中已经存在大量的用于从多种类型的包含人物图像的图像数据中提取人体骨架特征点的技术，本文中为了突出说明重点，在此就不再详述。可以采用诸如 Hrnet 等任何合适的已知的专用软件来进行本步骤中的人体骨架及其关键点的数据提取。此外，货架的轮廓数据的提取也可以采用诸如 LabelMe 等任何已知的适合的软件来进行。考虑到货架的实际形状，例如可以将货架图像定义并提取为包括至少两个多边形轮廓。其中，内部多边形是基于货架的真实外部轮廓提取出的多边形轮廓线，其表示货架的实际边界；处于内部多边形之外的外部多边形是用于定义内部多边形的接近区域的范围的轮廓线。

此后，在检测到顾客的手的手部关键点进入到外部多边形的情况下，针对顾客的每只手分别执行手部关键点附近是否存在物品的进入物品检测，从而确定所拍摄到的货架前的顾客的手在进入货架前是否持握有物品以及在持有物品的情况下物品的种类、数量等。需要说明的是，对于手部关键点附近的物品的检测可以通过借助诸如 YoloX 等工具采用任何已知的适合的方法来进行。随后，在检测到顾客的手的手部关键

点退出外部多边形的情况下，针对顾客的每只手再分别执行手部关键点附近是否存在物品的退出物品检测，从而确定所拍摄到的货架前的顾客的手在离开货架时是否持握有物品以及持有物品的情况下物品的种类、数量等。

可选地，在检测到顾客的两只手都分别进入外部多边形的情况下，还可以分别记录手部关键点附近的各物品在外部多边形与内部多边形之间轨迹。通过对不同物品的轨迹进行比对，例如能够有助于分辨出某物品是否是被顾客的双手持握。最后，根据进入物品检测结果和退出物品检测结果，优选地，再结合手部关键点附近的物品的轨迹，能够判定顾客在货架前的拾取或放回行为。例如，这样的判定过程可以通过基于各前序步骤中的检测结果，对顾客的双手分别使用有限状态机（FSM）来实现。

根据本发明的第二实施例，发明人定义了位于代表货架的实际轮廓线的内部多边形外部的接近区域的外部多边形，从而使用在内部多边形与外部多边形之间的区域内获得的图像数据作为进行检测和识别的基础，并且以顾客的手进出外部多边形的时间点来作为检测和识别的触发点。通过这样的方法，能够有效地消除顾客的手在货架附近的多余或复杂动作对拾取或放回动作识别的干扰，从而更加准确有效地进行检测和识别。另外，容易理解的是，在真实情况下，顾客的手是在货架内（即，货架的内部多边形内）进行的物品的拾取或放回行为。然而，由于货架的遮挡，我们通常无法基于手部关键点及其附近的区域在内部多边形内的图像数据来直接判定出顾客在货架前可能出现的拾取或放回行为。因此，通过这样的方法，还能够有效消除货架的遮挡对于拾取或放回动作识别的影响。

优选地，根据本发明的检测方法在在所述步骤 S2 之后，在顾客的手进入货架的所述外部多边形并针对顾客的每只手分别执行进入物品检测之后，还包括判定所述人的所述手部关键点是否进入所述内部多边形变形的步骤。在此情况下，只有在确定所述人的所述手部关键点进入所述内部多边形的情况下，根据本发明的检测方法才继续进行后续步骤。例如，只有在顾客的手部关键点在获取的图像数据的至少 3 个连续帧中都处于货架的内部多边形内，则该手部关键点才被判定为进入了货架的内部多边形。

在上述的根据本发明第二实施例的用于检测和识别顾客的拾取或放回行为的方法中，除了获取环境的图像数据的步骤是由各相机执行之外，其余的各步骤的处理均可通过与各相机通信连接的诸如电脑、CPU、TPU、GPU、FPGA 等处理器或专用数据处理芯片执行。或者，优选地，这些步骤中的处理也可以由具有 AI 处理功能的传感器执行。这样的传感器兼具数据处理能力和数据存储能力，无需额外的硬件即可执行机器学习驱动的计算机视觉处理任务。这样的传感器例如可以集成设置在相机中。

下面，将详细地说明如何利用有限状态机基于进入物品检测结果和物品检测结果等来判定顾客在货架前的拾取或放回行为。

与第一实施例中类似地，在用于图示本实施例的附图中，在示出提取出的顾客骨架关键点信息、物品信息以及货架轮廓信息之外，为了便于理解和说明，也示出了完整的图片信息。但在实际的使用过程中，在提取步骤之后的各步骤中，均不需要传输、存储和处理完整的图片或视频数据。

2. 2、顾客在货柜前的拾取或放回行为的简要分析

为了便于理解，这里先将顾客在货架前可能出现的拾取或放回行为进行简要说明和分类。

图 12 示意性地示出了顾客在货架期的主要四种拾取或放回行为。

(1) 如图 12 右上部分所示，只有远离相机或图形传感器那侧的一只手进入货架的外部多边形，并且继而进入内部多边形放回和/或拾取一件物品。在此情况下，顾客两只手的关键点及其附近的区域均没有被遮挡，因此能够容易地检测到手部的关键点的动作以及手部的关键点附近的区域的物品。需要注意的是，在此情况下，两只手对于手部的关键点附近区域的物品（即，被手持握的物品）的遮挡情况是不一样的，因此在使用软件检测物品时，针对两只手不同遮挡情况设定不同的检测阈值是优选的。

(2) 如图 12 左上部分所示，只有靠近相机或图形传感器那侧的一只手进入货架的外部多边形，并且继而进入内部多边形放回和/或拾取一件物品。在此情况下，顾客的远侧的另一手的关键点及其附近的区域很可能被遮挡。因此，在根据本发明第二实施例的方法中，为了便于进行运算中，将这种情况看做是下文中将要说明的第四种情况进行处理。

(3) 如图 12 左下部分所示，顾客的两只手均进入货架的外部多边形，在此情况下，如果检测到两只手的关键点附近的区域中的物品是不同的物品，则认为两只手分别放回和/或拾取了一件物品。也即是，该顾客放回和/或拾取了两件物品。

(4) 如图 12 右下部分所示，顾客的两只手均进入货架的外部多边形，在此情况下，如果检测到两只手的关键点附近的区域中的物品是相同的物品，则认为两只手共同放回和/或拾取一件物品。也即是，该顾客一共放回和/或拾取了一件物品。

需要注意的是，在顾客的两只手均进入货架的外部多边形的上述第（3）和第（4）种情况下，存在着

两只手先后在货架上放回和/或拾取物品以及换手持握物品等复杂情况。但是，无论在放回和/或拾取物品的过程中的动作如何复杂，只要当双手退出货架的外部多边形时分别针对双手的手部关键点附近的物品进行检测并基于下文中说明的记录的物品轨迹判定是否为同一物体，即可简化放回和/或拾取物品行为的判定。

根据以上分析，更加具体地，上述 4 种主要拾取或放回行为可以进一步扩展为包括如下 12 种情况。

只有一只手进入货架的情况下：

- (1) 一只手放回一件物品；
- (2) 一只手拾取一件物品；
- (3) 一只手放回一件物品并且拾取一件物品；

两只手均进入货架的情况下：

- (4) 两只手放回一件物品；
- (5) 两只手放回两件物品；
- (6) 两只手拾取一件物品；
- (7) 两只手拾取两件物品；
- (8) 两只手放回一件物品，一只手拾取一件物品；
- (9) 两只手放回一件物品，两只手拾取一件物品；
- (10) 两只手放回一件物品，两只手拾取两件物品；
- (11) 两只手放回两件物品，一只手拾取一件物品；
- (12) 两只手放回两件物品，两只手拾取两件物品。

因此，在下文的关于在本实施例中使用的有限状态机的说明中，将根据此处分析的顾客在货架前的可能存在的行为状态类别对有限状态机的参数进行定义和使用。

2. 3、物品存在的判定条件的设定

在上文中已经提及，在根据本发明的第二实施例的用于检测和识别顾客的拾取或放回行为的方法中，需要采用诸如 YoloX 等任何已知的适合的软件对手部关键点附近的物品进行进入物品检测和退出物品检测等物品检测。

需要说明的是，为了使物品检测更加的准确，需要设定判定条件。例如，只有在所拍摄或获取的图像数据中的具有检测到的物品的帧数等于或大于预定的最小帧数，并且具有检测到的所述物品的帧数与图像数据的总帧数的比值等于或大于预定的最小比值的情况下，才确定为检测出进入物品。这里，进入物品是指顾客的手在进入货架时就持有的物品。类似地，只有在所拍摄或获取的图像数据中的具有检测到的物品的帧数等于或大于预定的最小帧数，并且具有检测到的所述物品的帧数与图像数据的总帧数的比值等于或大于预定的最小比值的情况下，才确定为检测出退出物品。这里，退出物品是指顾客的手在退出货架时持有的物品。这里的最小帧数和最小比值可以根据需要设定。例如，最小帧数可以为 2 帧，最小比值可以为 0.03。或者，最小帧数可以为 5 帧，最小比值可以为 0.06。

例如，图 13 的 a 和 b 是图示了根据本发明的第二实施例的物品存在的判定条件的示例的列表。如图 13 的 a 所示，在针对顾客的左手进行的进入物品检测过程中，在 10 个帧（时间戳）中，检测到物品的帧数为 5。因此，判定左手在进入时持有物品（存在）。如图 13 的 b 所示，在针对顾客的左手进行的退出物品检测过程中，在 10 个帧（时间戳）中，检测到物品的帧数为 1。因此，判定左手在退出时未持有物品（不存在）。

2. 4、本实施例中使用的有限状态机的说明

有限状态机（FSM）是一种用于进行对象行为建模的工具，其作用主要是描述对象在它的生命周期内所经历的状态序列，以及如何响应来自外界的各种事件。在计算机科学中，有限状态机已经被广泛用于建模应用行为、硬件电路系统设计、软件工程，编译器、网络协议、和计算与语言的研究。因此，有限状态机非常适合用来协助顾客在货架前的拾取或放回行为。

简而言之，有限状态机可归纳为 4 个要素，即现态、条件、动作、次态。“现态”和“条件”是因，“动作”和“次态”是果。具体定义如下：

现态：是指对象当前所处的状态。例如，在本实施例中，顾客的手部进入外部多边形时是否持握有物品的状态可以作为 FSM 的现态。

条件：又称为“事件”。当一个条件被满足，将会触发一个动作，或者执行一次状态的迁移。

动作：条件满足后执行的动作。动作执行完毕后，可以迁移到新的状态，也可以仍旧保持原状态。动作不是必需的，当条件满足后，也可以不执行任何动作，直接迁移到新状态。

例如，在本实施例中，顾客的手部在货柜内的拾取或放回动作就可以作为 FSM 的“事件”或者“条件满足后执行的动作”，其导致迁移至新状态。

次态：条件满足后要迁往的新状态。“次态”是相对于“现态”而言的，“次态”一旦被激活，就转变成新的“现态”了。例如，在本实施例中，顾客的手部退出外部多边形时是否持握有物品的状态可以作为 FSM 的次态。

在本实施例中，当用于一只手的一个 FSM 结束时，能够询问该 FSM 的当前估计，从而判断与这只手相关的物品的拾取或放回行为。例如，为了实施根据本发明的第二实施例的上述检测方法，可以为各 FSM 定义并获取用于顾客的各只手的内部变量 `accumulate_entry`、`accumulate_exit` 和 `accumulate_inside`。其中，`accumulate_entry` 是当手部关键点进入货架的外部多边形时的帧的列表，该内部变量在上述进入物品检测时用于表达是否存在检测到的物品；`accumulate_exit` 是当手部关键点退出货架的外部多边形时的帧的列表，该内部变量在上述退出物品检测时用于表达是否存在检测到的物品；`accumulate_inside` 是当手部关键点位于货架内部（即，位于货架的内部多边形）时的帧的列表，该内部变量用于确认顾客的手部是否实际进行了放回或拾取动作，并且如前文所述，例如以至少连续的 3 帧作为判断基础。各 FSM 能够通过调用前文所述的在进入状态期间的进入物品检测和在退出状态期间的退出物品检测的检测结果，来设置和修改这些内部参数。FSM 根据进入物品检测的检测结果（即，FSM 的现态）和退出物品检测的检测结果（即，FSM 的次态）就能判断出顾客在货架前的拾取或放回动作（即，FSM 的事件）。

另外，在本实施例的检测方法中，在通过利用 Java 等工具使用 FSM 来完成状态判定的过程中，例如可以定义以下多种可能的状态：

`unknown`：顾客进入了货架的外部多边形，但没有进入内部多边形；

`did_not_enter_shelf`：顾客进入了货架的外部多边形但没有进入内部多边形，并随后退出了外部多边形；

`other_hand`：顾客用两只手拾取/放回物品，一只手退出了外部多边形，但另一只手还没有退出外部多边形；

`both_hand_entry`：顾客使用两只手在持有物品的情况下进入货架，仅仅是提醒；

`both_hand_exit`：顾客使用两只手在持有物品的情况下退出货架，仅仅是提醒；

`pick`：顾客拾取了一件物品（在此状态下，无进入物品但存在退出物品）；

`release`：顾客放回了一件物品（在此状态下，存在进入物品但不存在退出物品）

`no_change`：顾客在退出时与进入时具有相同的状态（认为不存在进入物品也不存在退出物品）。

需要说明的是，以上的定义仅仅是示例。FSM 作为一种用于进行对象行为建模的工具，其具体实现方法在本领域内已经比较成熟，在本文中就不再赘述。

2. 5、基于物品轨迹的相同物品判定

由前文中的说明可知，顾客在货架前的拾取/放回行为的最终判定是通过使用 FSM 基于进入货架时的进入物品的数量、退出货架时的退出物品的数量的比较来确定的。这又取决于进入的手的数量以及双手是否持握着相同的物品。

如前文所述，根据本发明的第二实施例的检测方法还包括记录顾客的手部关键点附近的物品在外部多边形与内部多边形之间的轨迹。因此，通过比较顾客的双手的手部关键点附近的进入物品的轨迹，就能够确定左手的手部关键点附近的进入物品与右手的手部关键点附近的进入物品是否为同一物品，例如如图 12 左下和右下两部分所示的情况。例如，这样的轨迹比较是通过核查拍摄或获取的图像数据的各帧中的左手的手部关键点附近的进入物品与右手的手部关键点附近的进入物品之间的距离而进行的。例如，仅在各帧中的距离均低于预定的距离阈值的情况下，确定为轨迹类似，进而确定为两只手的手部关键点附近的进入物品为同一物品。或者，在各帧中的距离的平均值低于预定的距离阈值的情况下，确定为轨迹类似，进而确定为两只手的手部关键点附近的进入物品为同一物品。上述距离阈值可以根据不同应用场景中的物品的最小尺寸或者平均尺寸等因素而设定。例如，可以将距离阈值设定为 25 个像素。在这里，如果确定为是同一物品，则说明顾客将用两只手放回一件物品；如果确定为是不同的物品，则说明顾客将用两只手放回两件物品。

类似地，通过比较顾客的双手的手部关键点附近的退出物品的轨迹，就能够确定左手的手部关键点附近的退出物品与右手的手部关键点附近的进入物品是否为同一物品。例如，这样的轨迹比较是通过核查拍

摄或获取的图像数据的各帧中的左手的手部关键点附近的退出物品与右手的手部关键点附近的退出物品之间的距离而进行的。例如，仅在各帧中的距离均低于预定的距离阈值的情况下，确定为轨迹类似，进而确定为两只手的手部关键点附近的退出物品为同一物品。或者，在各帧中的距离的平均值低于预定的距离阈值的情况下，确定为轨迹类似，进而确定为两只手的手部关键点附近的退出物品为同一物品。类似地，例如可以将距离阈值设定为 25 个像素。在这里，如果确定为是同一物品，则说明顾客用两只手从货架上拾取了一件物品；如果确定为是不同的物品，则说明顾客用两只手从货架上拾取了两件物品。

图 14 示出了上述轨迹比较的示例。如图 14 中所示，在图像数据的 10 个帧（时间戳）中，针对左手的进入物品检测共在 6 帧中检测到了进入物品，针对右手的进入物品检测共在 7 帧中检测到了进入物品。通过比较各帧中物品的坐标之间的距离，可知所有的距离均低于 25 个像素。因此，可以认为左手和右手在进入时是持握的相同的物品。

另外，在进行上述的距离比较时，可以设想，如果通过比较判定左手附近的进入物品与右手附近的退出物品具有相同的轨迹，或者右手附近的进入物品与左手附近的退出物品具有相同的轨迹，则我们将难以判断这样的物品到底是被那一只手持握的。因此，当发生这样的情况时，需要去除这样的干扰点。例如，当在某帧图像数据中，左手的退出物品与右手的进入物品的距离均小于 25 个像素时，将这样的干扰点的数据直接去除。

2. 6、基于 FSM 的拾取或放回行为的行为判定

下面，将结合上文中的说明，论述根据本发明第二实施例的检测方法中的基于 FSM 的拾取或放回行为的行为判定的情况。在下面的说明中，可以将 FSM 返回的状态判定信息作为顾客在货架前的拾取或放回行为的行为判定结果。

当至少一只手离开货架的外部多边形时，触发 FSM 的状态判定并且针对各帧进行核查。此时，FSM 根据前文中说明的进入物品的判定结果、退出物品的判定结果和物品轨迹等数据中的至少一部分来进行拾取或放回行为的分析和判定。应当理解的是，由于 FSM 是针对两只手分别运行的，所以可以先针对其中的一只手（例如，先离开的那只手）进行状态判定，然后再针对另一只手（例如，后离开的那只手）进行状态判定。

图 15 示出了第一个 FSM 运行时的状态判定的示例性流程图。如图 15 中所示，当 FSM 的编号为 1 时，当对应的手部离开货架的外部多边形时，基于前文说明的进入物品检测和退出物品检测进行判定。当没有进入物品（~entry）且具有退出物品（exit）时，返回状态“拾取一件物品（拾取+1）”。当具有进入物品（entry）且没有退出物品（~exit）时，返回状态“放回一件物品（放回-1）”。当没有进入物品（~entry）且没有退出物品（~exit）时，返回状态“没有变化”。当具有进入物品（entry）且具有退出物品（exit）时，返回状态“放回，拾取”。

接着，如果另一只手也退出了货架的外部多边形，则进行第二个 FSM 的状态判定，如图 16 的示例性流程图所示。当 FSM 的编号为 2 时，首先检查第二只手在进入时的进入状态。当这只手在进入不持有物品（~object）时，则继续检查这只手在退出时的状态。如果退出时未持有物品（~object），则返回状态“没有变化”；如果退出时持有物品（object），则比较两只手的物品的退出轨迹。如果退出轨迹相同，则返回状态“双手退出”；如果退出轨迹不同，则返回状态“拾取一件物品（拾取+1）”。

此外，在第二只手在进入时持有物品（object）的情况下，则比较两只手的物品的进入轨迹，并分别继续检查退出状态。在两只手的物品的进入轨迹相同的情况下，若第二只手的退出状态为未持有物品（~object），则返回状态“双手进入”；若第二只手的退出状态为持有物品（object），则比较两只手的物品的退出轨迹，如果退出轨迹相同，则返回状态“放回，拾取”，如果退出轨迹不同，则返回状态“拾取一件物品（拾取+1）”。在两只手的物品的进入轨迹不同的情况下，若第二只手的退出状态为未持有物品（~object），则返回状态“放回一件物品（放回-1）”；若第二只手的退出状态为持有物品（object），则比较两只手的物品的退出轨迹，如果退出轨迹不同，则返回状态“放回，拾取”，如果退出轨迹不同，则返回状态“放回一件物品（放回-1）”。

图 17 图示了在仅有针对一只手的 FSM 运行的情况下获得的 FSM 的状态表。可以理解的是，当仅有一只手进入货架的情况下，图 17 的状态表中反映出的拾取或放回行为状态即是顾客的拾取或放回行为检测最终状态分类。图 18 图示了在针对顾客的两只手的 FSM 运行的情况下获得的 FSM 的状态表。图 19 示出了通过将图 18 中的两个 FSM 的状态表相结合之后获得的最终状态表。可以理解的是，当顾客的两只手均进入货架的情况下，图 19 的状态表中反映出的拾取或放回行为状态是顾客的拾取或放回行为检测的最终判定情况。

2. 7、变型例

在上面的根据本发明的第二实施例的说明中，图像数据是来自于一台相机或图像传感器从一个视角拍

摄的一组图像数据。然而，本发明不限于此，例如，可以使用来自多个相机或图像传感器的从不同的视角拍摄的多组图像数据。在此情况下，能够避免因拍摄视角的盲区或物体的遮挡而导致的检测错误。例如，在此情况下，只要在来自多个相机或图像传感器的多组图像数据中的具有检测到的物品的帧数之和等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，进入物品检测/退出物品检测就能够确认检测出进入物品/退出物品。例如，如图 20 所示，在第一台相机获取的图像数据中，仅检测出在右手关键点附近的区域中存在退出物品。在第二台相机获取的图像数据中，仅检测出在左手关键点附近的区域中存在退出物品。当结合考虑基于多组图像数据的检测结果时，可以判定为在左手关键点附近的区域中和在右手关键点附近的区域中分别都检测到了退出物品。

另外，在上面的根据本发明的第二实施例的说明中，已经说明了使用 FSM 作为工具基于手部进入外部多边形时的进入物品检测的检测结果和退出外部多边形时的退出物品检测的检测结果以及物品的轨迹比较的比较结果的至少一部分来实现顾客在货柜前的拾取或放回动作的检测和识别。但是，本发明不限于此，而是可以使用任何本领域中适合的用于描述对象的状态序列以及如何响应来自外界的各种事件的工具。例如，通过上文中的说明，可以知道顾客在货柜前的拾取或放回动作是有限的，其与手部进出货架的外部多边形时的物品检测结果以及物品的轨迹比较结果的关联性也是相对固定的。因此，例如还可以预先将完全地记载有拾取或放回动作的各种情况以及上述各种检查结果与各种情况的关联性和对应关系的查找表存储在处理器的存储元件中或处理器能够访问的存储元件中。这样，能够使用查找表作为工具，基于手部进入外部多边形时的进入物品检测的检测结果和退出外部多边形时的退出物品检测的检测结果以及物品的轨迹比较的比较结果的至少一部分，能够查表的方式来实现根据本发明的第二实施例的检测方法。

2. 8、用于检测和识别顾客的拾取或放回行为的设备的示例

根据本发明第二实施例的用于检测和识别顾客的拾取或放回行为的检测设备例如可以包括：至少一台相机或图像传感器，以用于获取图像数据；处理单元，所述处理单元接收由所述至少一台相机或图像传感器获取的这些图像数据，并且能够采用前文中所述的根据本发明第二实施例的检测方法对所述图像数据进行处理，以检测和识别顾客在货架前的拾取或放回行为。处理单元例如能够通过多个组成模块执行作为软件或固件而被存储在其中的存储元件中或与其互连的存储器或数据存储单元中的应用程序或例程，从而执行上文中所述的根据本发明第二实施例的用于检测和识别顾客在货架前的拾取或放回行为的方法。这里的处理单元例如由中央处理单元（CPU）和存储元件构成。例如，处理单元可以包括一个或多个通用处理器、控制器、现场可编程门阵列（FPGA）、图形处理单元（GPU）、专用集成电路（ASIC）或它们的组合等，作为与各个相机或图像传感器数据通信的专用数据处理器或数据处理芯片。或者，处理单元也可以是集成在各相机或图像传感器中的彼此能够互相进行数据交互的具有 AI 处理功能的传感器元件。这样的传感器元件兼具数据处理能力和数据存储能力，无需额外的硬件即可执行根据本发明第二实施例的用于检测和识别顾客在货架前的拾取或放回行为的方法。

根据本发明的第二实施例，由于在获取视频或图片数据后就立即对这些数据进行了提取关键点信息的数据简化处理，在后续的所有步骤中均只需要传输和处理提取出的人骨架的关键点数据，因此大大减少了数据处理量，能够快速、准确地识别出无人购物场所中顾客在货架前的拾取或放回行为。此外，由于不需要对顾客进行面部识别，并且在经过最初的数据简化处理后，存储和传输的数据中将只包含顾客的骨架的包含手部在内的关键点的数据，因此保护了顾客隐私权。

本发明例如可以如下进行实施、构造或配置。

（1）一种用于检测分布有相机的场所中的人的摔倒的检测方法，其特征在于，包括如下步骤：

步骤 S1：对所有所述相机进行标定，以使在各所述相机的视场中均具有恰当的垂直向量；

步骤 S2：通过至少一部分所述相机获得包含所述场所中的人的图像数据，并且从所述图像数据中提取出所述人的骨架的关键点的数据；

步骤 S3：通过使用所述关键点的数据估算所述人的个人垂直度；

步骤 S4：针对所述至少一部分相机中的各者，基于相应的相机的所述视场中的所述垂直向量和所述个人垂直度计算所述人的垂直角度；

步骤 S5：通过对所述人在某一时刻在所述至少一部分相机的各所述视场中的所有所述垂直角度进行聚合，获得所述人的最终垂直角度；

步骤 S6：基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。

（2）

根据（1）中所述的方法，其特征在于，在所述步骤 S1 中，所述标定时基于各所述相机的内部参数和外部参数进行的。

(3)

根据(1)中所述的方法,其特征在于,在所述步骤S2中,所述人的所述关键点的数据的提取是在不进行所述人的面部识别的情况下进行的。

(4)

根据(1)中所述的方法,其特征在于,在所述步骤S2中,提取所述人的17个所述关键点的数据。

(5)

根据(1)中所述的方法,其特征在于,在所述步骤S3中,只在有效视场中估算所述人的所述个人垂直度,所述有效视场是指其中的置信因子 c 等于或大于预定的置信阈值的视场,所述置信因子 c 如下定义:

$$c = N_v / N_t$$

其中, N_v 是在相应的视场中的所述人的能够观察到的所述关键点的数量, N_t 是所述人的所述关键点的总数。

(6)

根据(5)中所述的方法,其特征在于,在所述步骤S3中,通过计算在所述有效视场中的所述人的从代表足部的所述关键点到代表头部的所述关键点的身体向量而获得所述人的个人垂直度。

(7)

根据(6)中所述的方法,其特征在于,基于所述人的能够观察到的代表所述头部的所述关键点的平均坐标和能够观察到的代表所述足部的所述关键点的平均坐标计算所述身体向量。

(8)

根据(1)中所述的方法,其特征在于,在所述步骤S5中,所述人的所述最终垂直角度等于所述人在所述至少一部分相机的各所述视场中的所述垂直角度的最大值。

(9)

根据(1)中所述的方法,其特征在于,在所述步骤S6中,所述最终垂直角度 α_v 与所述摔倒分值 s_f 之间的关系满足如下等式:

若 $\alpha_v < T_l$, 则 $s_f = 0$;

若 $T_l < \alpha_v < T_h$, 则 $s_f = \frac{\alpha_v - T_l}{T_h - T_l}$;

若 $\alpha_v > T_h$, 则 $s_f = 1$;

其中, T_l 为设定的所述最终垂直角度的下限, T_h 为设定的所述最终垂直角度的上限,并且

仅在所述摔倒分值 s_f 大于判定阈值的情况下,才判定并检测出所述人的摔倒。

(10)

根据(9)中所述的方法,其特征在于,所述最终垂直角度的所述下限为40度,所述最终垂直角度的所述上限为80度,并且所述判定阈值为0.5。

(11)

一种用于检测场所内的人的摔倒的检测设备,其特征在于,包括:

至少一个相机,所述至少一个相机分布在所述场所内并且具有不同的视场,所述至少一个相机能够获得包含所述场所中的人的图像数据;

处理单元,所述处理单元对所述多个相机获得的所述图像数据进行处理,以判定所述场所内的所述人是否摔倒,其中,所述处理单元包括:

标定模块,对所有所述相机进行标定,以使在各所述相机的视场中均具有恰当的垂直向量;

数据处理模块,对从所述多个相机传输来的所述图像数据进行处理,从而获得所述人在所述多个相机之中的至少一部分相机的所述视场中的个人垂直度;

计算模块,基于从所述标定模块发送来的所述垂直向量和从所述数据处理模块发送来的所述个人垂直度计算所述人的最终垂直角度;以及

判定模块,基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。

(12)

根据(11)中所述的检测设备,其特征在于,所述数据处理模块包括:

提取模块，从所述图像数据中提取出所述人的骨架的关键点的数据；和
估算模块，通过使用来自所述提取模块的所述关键点的数据估算所述人的所述个人垂直度。

(13)

根据(12)中所述的检测设备，其特征在于，所述计算模块包括：

角度计算模块，针对所述至少一部分相机的各所述视场分别基于所述垂直向量和所述个人垂直度计算所述人的垂直角度；以及

聚合模块，通过对所述人在某一时刻在所述至少一部分相机的各所述视场中的所有所述垂直角度进行聚合，获得所述人的所述最终垂直角度。

(14)

根据(11)至(13)中任一项所述的检测设备，其特征在于，所述标定模块基于所述多个相机的各者的内部参数和外部参数进行所述标定。

(15)

根据(12)中所述的检测设备，其特征在于，所述提取模块仅将所述人的所述关键点的数据发送给所述估算模块。

(16)

根据(12)中所述的检测设备，其特征在于，所述提取模块提取所述人的 17 个所述关键点的数据。

(17)

根据(12)中所述的检测设备，其特征在于，所述估算模块只估算所述多个相机的各所述视场之中的有效视场中的所述人的所述个人垂直度，所述有效视场是指其中的置信因子 c 等于或大于预定的置信阈值的视场，所述置信因子 c 如下定义：

$$c = N_v / N_t$$

其中， N_v 是在相应的视场中的所述人的能够观察到的所述关键点的数量， N_t 是所述人的所述关键点的总数。

(18)

根据(17)中所述的检测设备，其特征在于，所述估算模块通过计算在所述有效视场中的所述人的从代表足部的所述关键点到代表头部的所述关键点的身体向量而获得所述人的个人垂直度。

(19)

根据(18)中所述的检测设备，其特征在于，所述估算模块基于所述人的能够观察到的代表所述头部的所述关键点的平均坐标和能够观察到的代表所述足部的所述关键点的平均坐标计算所述身体向量。

(20)

根据(13)中所述的检测设备，其特征在于，所述聚合模块将所述人在所述至少一部分相机的各所述视场中的所述垂直角度的最大值设定为所述人的所述最终垂直角度。

(21)

根据(11)至(13)中任一项所述的检测设备，其特征在于，所述判定模块仅在摔倒分值 s_f 大于判定阈值的情况下，才判定并检测出所述人的摔倒，

其中，所述最终垂直角度 α_v 与所述摔倒分值 s_f 之间的关系满足如下等式：

若 $\alpha_v < T_l$ ，则 $s_f = 0$ ；

若 $T_l < \alpha_v < T_h$ ，则 $s_f = \frac{\alpha_v - T_l}{T_h - T_l}$ ；

若 $\alpha_v > T_h$ ，则 $s_f = 1$ ，

其中， T_l 为设定的所述最终垂直角度的下限， T_h 为设定的所述最终垂直角度的上限。

(22)

根据(21)中所述的检测设备，其特征在于，所述最终垂直角度的所述下限为 40 度，所述最终垂直角度的所述上限为 80 度，并且所述判定阈值为 0.5。

(23)

一种存储介质，其上存储有计算机可读程序，所述程序在处理器上执行时实施如前述(1) - (10)中

任一项所述的方法。

(24)

一种用于检测人在货架前的拾取或放回行为的检测方法，其特征在于，包括如下步骤：

步骤 S1：从图像数据中获取所述人的骨架的包含手部关键点的多个关键点的数据，并且从所述图像数据提取所述货架的外轮廓线，其中，所述外轮廓线包括所述货架的外部多边形以及与所述货架的真实外轮廓相对应的内部多边形，所述外部多边形在所述内部多边形的外部的接近区域中；

步骤 S2：在所述人的至少一只手的所述手部关键点被检测到进入所述外部多边形的情况下，针对所述人的进入所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的进入物品检测；

步骤 S4：在所述人的至少一只手的所述手部关键点被检测到退出所述外部多边形的情况下，针对所述人的退出所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的退出物品检测；

步骤 S5：基于所述进入物品检测的结果和所述退出物品检测的结果来判定所述人在所述货架前的所述拾取或放回行为。

(25)

根据(24)中所述的方法，其特征在于，在所述步骤 S1 中，从图像数据中获取所述人的所述多个关键点的数据包括：

通过至少一部相机拍摄包含所述货架和所述货架前的所述人的所述图像数据；

从所述图像数据中提取所述人的包含所述手部关键点的所述多个关键点的数据。

(26)

根据(24)中所述的方法，其特征在于，在所述步骤 S2 之后，还包括判定所述人的所述手部关键点是否进入所述内部多边形的步骤，其中

在确定所述人的所述手部关键点进入所述内部多变形的情况下，执行所述步骤 S2 的后续步骤。

(27)

根据(26)中所述的方法，其特征在于，

只有在所述人的所述手部关键点在所述图像数据的至少 3 个连续帧中都处于所述内部多边形内，所述手部关键点才被确定为进入所述内部多边形。

(28)

根据(24)中所述的方法，其特征在于，在所述步骤 S5 中，通过基于所述进入物品检测的结果和所述退出物品检测的结果针对所述人的每只手分别使用一个有限状态机，来判定所述人在所述货架前的所述拾取或放回行为。

(29)

根据(24)中所述的方法，其特征在于，在所述步骤 S2 与所述步骤 S4 之间还包括步骤 S3：

针对所述人的进入所述外部多边形的每只手，分别记录所述手部关键点附近的所述物品在所述外部多边形与所述内部多边形之间的轨迹。

(30)

根据(29)中所述的方法，其特征在于，在所述步骤 S5 中，通过基于所述进入物品检测的结果、所述退出物品检测的结果和所述物品的所述轨迹，针对所述人的每只手分别使用一个有限状态机，来判定所述人在所述货架前的所述拾取或放回行为。

(31)

根据(30)中所述的方法，其特征在于，若在所述步骤 S2 中针对所述人的两只手的所述进入物品检测均确认检测出进入物品，则在所述步骤 S5 中，通过比较所述人的左手的所述手部关键点附近的所述进入物品的所述轨迹与右手的所述手部关键点附近的所述进入物品的所述轨迹，确定左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品是否为同一物品。

(32)

根据(31)中所述的方法，其特征在于，所述人的左手的所述手部关键点附近的所述进入物品的所述轨迹与右手的所述手部关键点附近的所述进入物品的所述轨迹的所述比较是通过核查所述图像数据的各帧中的左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品之间的距离而进行的，并且

仅在各帧中的所述距离均低于预定的距离阈值的情况下，确定左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品为同一物品。

(33)

根据(30)至(32)中任一项所述的方法，其特征在于，若在所述步骤S4中针对所述人的两只手的所述退出物品检测均确认检测出退出物品，则在所述步骤S5中，通过比较所述人的左手的所述手部关键点附近的所述退出物品的所述轨迹与右手的所述手部关键点附近的所述退出物品的所述轨迹，确定左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品是否为同一物品。

(34)

根据(33)中所述的方法，其特征在于，所述人的左手的所述手部关键点附近的所述退出物品的所述轨迹与右手的所述手部关键点附近的所述退出物品的所述轨迹的所述比较是通过核查所述图像数据的各帧中的左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品之间的距离而进行的，并且

仅在各帧中的所述距离均低于预定的距离阈值的情况下，确定左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品为同一物品。

(35)

根据(24)至(30)中任一项所述的方法，其特征在于，在所述步骤S2中，只有在所述图像数据中的具有检测到的物品的帧数等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述进入物品检测确认检测出进入物品。

(36)

根据(24)至(30)中任一项所述的方法，其特征在于，在所述步骤S4中，只有在所述图像数据的具有检测到的物品的帧数等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述退出物品检测确认检测出退出物品。

(37)

根据(35)中所述的方法，其特征在于，所述图像数据包括来自多个相机或图像传感器的多组图像数据，并且

在所述步骤S2中，只有在来自所述多个相机或图像传感器的所述多组图像数据中的具有检测到的物品的帧数之和等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述进入物品检测确认检测出进入物品。

(38)

根据(28)或(30)所述的方法，其特征在于，所述有限状态机包括下列拾取或放回行为：无变化、拾取一个物品、拾取两个物品、放回一个物品和放回两个物品。

(39)

一种用于检测人在货架前的拾取或放回行为的检测设备，其特征在于，所述检测设备包括：

至少一台相机或图像传感器，获取图像数据；

处理单元，所述处理单元根据如(24)至(38)中任一项所述的用于检测人在货架前的拾取或放回行为的检测方法对所述图像数据进行处理。

(40)

一种存储介质，其上存储有计算机可读程序，所述程序在处理器上执行时实施如(24)至(38)中任一项所述的方法。

尽管在上面已经参照附图说明了根据本发明的补偿方法、成像装置和存储介质，但是本发明不限于此，且本领域技术人员应理解，在不偏离本发明随附权利要求书限定的实质或范围的情况下，可以做出各种改变、组合、次组合以及变型。

权利要求书

1. 一种用于检测分布有相机的场所中的人的摔倒的检测方法，其特征在于，包括如下步骤：
步骤 S1：对所有所述相机进行标定，以使在各所述相机的视场中均具有恰当的垂直向量；
步骤 S2：通过至少一部分所述相机获得包含所述场所中的人的图像数据，并且从所述图像数据中提取出所述人的骨架的关键点的数据；
步骤 S3：通过使用所述关键点的数据估算所述人的个人垂直度；
步骤 S4：针对所述至少一部分相机中的各者，基于相应的相机的所述视场中的所述垂直向量和所述个人垂直度计算所述人的垂直角度；
步骤 S5：通过对所述人在某一时刻在所述至少一部分相机的各所述视场中的所有所述垂直角度进行聚合，获得所述人的最终垂直角度；
步骤 S6：基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。
2. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S1 中，所述标定时基于各所述相机的内部参数和外部参数进行的。
3. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S2 中，所述人的所述关键点的数据的提取是在不进行所述人的面部识别的情况下进行的。
4. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S2 中，提取所述人的 17 个所述关键点的数据。
5. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S3 中，只在有效视场中估算所述人的所述个人垂直度，所述有效视场是指其中的置信因子 c 等于或大于预定的置信阈值的视场，所述置信因子 c 如下定义：
$$c = N_v / N_t$$
其中， N_v 是在相应的视场中的所述人的能够观察到的所述关键点的数量， N_t 是所述人的所述关键点的总数。
6. 根据权利要求 5 所述的方法，其特征在于，在所述步骤 S3 中，通过计算在所述有效视场中的所述人的从代表足部的所述关键点到代表头部的所述关键点的身体向量而获得所述人的个人垂直度。
7. 根据权利要求 6 所述的方法，其特征在于，基于所述人的能够观察到的代表所述头部的所述关键点的平均坐标和能够观察到的代表所述足部的所述关键点的平均坐标计算所述身体向量。
8. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S5 中，所述人的所述最终垂直角度等于所述人在所述至少一部分相机的各所述视场中的所述垂直角度的最大值。
9. 根据权利要求 1 所述的方法，其特征在于，在所述步骤 S6 中，所述最终垂直角度 α_v 与所述摔倒分值 s_f 之间的关系满足如下等式：
若 $\alpha_v < T_l$ ，则 $s_f = 0$ ；
若 $T_l < \alpha_v < T_h$ ，则 $s_f = \frac{\alpha_v - T_l}{T_h - T_l}$ ；
若 $\alpha_v > T_h$ ，则 $s_f = 1$ ，
其中， T_l 为设定的所述最终垂直角度的下限， T_h 为设定的所述最终垂直角度的上限，并且仅在所述摔倒分值 s_f 大于判定阈值的情况下，才判定并检测出所述人的摔倒。

10. 根据权利要求 9 所述的方法，其特征在于，所述最终垂直角度的所述下限为 40 度，所述最终垂直角度的所述上限为 80 度，并且所述判定阈值为 0.5。

11. 一种用于检测场所内的人的摔倒的检测设备，其特征在于，包括：

至少一个相机，所述至少一个相机分布在所述场所内并且具有不同的视场，所述至少一个相机能够获得包含所述场所中的人的图像数据；

处理单元，所述处理单元对所述多个相机获得的所述图像数据进行处理，以判定所述场所内的所述人是否摔倒，其中，所述处理单元包括：

标定模块，对所有所述相机进行标定，以使在各所述相机的视场中均具有恰当的垂直向量；

数据处理模块，对从所述多个相机传输来的所述图像数据进行处理，从而获得所述人在所述多个相机之中的至少一部分相机的所述视场中的个人垂直度；

计算模块，基于从所述标定模块发送来的所述垂直向量和从所述数据处理模块发送来的所述个人垂直度计算所述人的最终垂直角度；以及

判定模块，基于由所述最终垂直角度求出的摔倒分值判定所述人是否摔倒。

12. 根据权利要求 11 所述的检测设备，其特征在于，所述数据处理模块包括：

提取模块，从所述图像数据中提取出所述人的骨架的关键点的数据；和

估算模块，通过使用来自所述提取模块的所述关键点的数据估算所述人的所述个人垂直度。

13. 根据权利要求 12 所述的检测设备，其特征在于，所述计算模块包括：

角度计算模块，针对所述至少一部分相机的各所述视场分别基于所述垂直向量和所述个人垂直度计算所述人的垂直角度；以及

聚合模块，通过对所述人在某一时刻在所述至少一部分相机的各所述视场中的所有所述垂直角度进行聚合，获得所述人的所述最终垂直角度。

14. 根据权利要求 11 至 13 中任一项所述的检测设备，其特征在于，所述标定模块基于所述多个相机的各者的内部参数和外部参数进行所述标定。

15. 根据权利要求 12 所述的检测设备，其特征在于，所述提取模块仅将所述人的所述关键点的数据发送给所述估算模块。

16. 根据权利要求 12 所述的检测设备，其特征在于，所述提取模块提取所述人的 17 个所述关键点的数据。

17. 根据权利要求 12 所述的检测设备，其特征在于，所述估算模块只估算所述多个相机的各所述视场之中的有效视场中的所述人的所述个人垂直度，所述有效视场是指其中的置信因子 c 等于或大于预定的置信阈值的视场，所述置信因子 c 如下定义：

$$c = N_v / N_t$$

其中， N_v 是在相应的视场中的所述人的能够观察到的所述关键点的数量， N_t 是所述人的所述关键点的总数。

18. 根据权利要求 17 所述的检测设备，其特征在于，所述估算模块通过计算在所述有效视场中的所述人的从代表足部的所述关键点到代表头部的所述关键点的身体向量而获得所述人的个人垂直度。

19. 根据权利要求 18 所述的检测设备，其特征在于，所述估算模块基于所述人的能够观察到的代表所述头部的所述关键点的平均坐标和能够观察到的代表所述足部的所述关键点的平均坐标计算所述身体向量。

20. 根据权利要求 13 所述的检测设备, 其特征在于, 所述聚合模块将所述人在所述至少一部分相机的各所述视场中的所述垂直角度的最大值设定为所述人的所述最终垂直角度。

21. 根据权利要求 11 至 13 中任一项所述的检测设备, 其特征在于, 所述判定模块仅在摔倒分值 s_f 大于判定阈值的情况下, 才判定并检测出所述人的摔倒,

其中, 所述最终垂直角度 α_v 与所述摔倒分值 s_f 之间的关系满足如下等式:

若 $\alpha_v < T_l$, 则 $s_f = 0$;

若 $T_l < \alpha_v < T_h$, 则 $s_f = \frac{\alpha_v - T_l}{T_h - T_l}$;

若 $\alpha_v > T_h$, 则 $s_f = 1$,

其中, T_l 为设定的所述最终垂直角度的下限, T_h 为设定的所述最终垂直角度的上限。

22. 根据权利要求 21 所述的检测设备, 其特征在于, 所述最终垂直角度的所述下限为 40 度, 所述最终垂直角度的所述上限为 80 度, 并且所述判定阈值为 0.5。

23. 一种存储介质, 其上存储有计算机可读程序, 所述程序在处理器上执行时实施如权利要求 1-10 中任一项所述的方法。

24. 一种用于检测人在货架前的拾取或放回行为的检测方法, 其特征在于, 包括如下步骤:

步骤 S1: 从图像数据中获取所述人的骨架的包含手部关键点的多个关键点的数据, 并且从所述图像数据提取所述货架的外轮廓线, 其中, 所述外轮廓线包括所述货架的外部多边形以及与所述货架的真实外轮廓相对应的内部多边形, 所述外部多边形在所述内部多边形的外部的接近区域中;

步骤 S2: 在所述人的至少一只手的所述手部关键点被检测到进入所述外部多边形的情况下, 针对所述人的进入所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的进入物品检测;

步骤 S4: 在所述人的至少一只手的所述手部关键点被检测到退出所述外部多边形的情况下, 针对所述人的退出所述外部多边形的每只手分别执行用于检测所述手部关键点的附近的物品的退出物品检测;

步骤 S5: 基于所述进入物品检测的结果和所述退出物品检测的结果来判定所述人在所述货架前的所述拾取或放回行为。

25. 根据权利要求 24 所述的方法, 其特征在于, 在所述步骤 S1 中, 从图像数据中获取所述人的所述多个关键点的数据包括:

通过至少一部相机拍摄包含所述货架和所述货架前的所述人的所述图像数据;

从所述图像数据中提取所述人的包含所述手部关键点的所述多个关键点的数据。

26. 根据权利要求 24 所述的方法, 其特征在于, 在所述步骤 S2 之后, 还包括判定所述人的所述手部关键点是否进入所述内部多边形的步骤, 其中

在确定所述人的所述手部关键点进入所述内部多变形的情况下, 执行所述步骤 S2 的后续步骤。

27. 根据权利要求 26 所述的方法, 其特征在于,

只有在所述人的所述手部关键点在所述图像数据的至少 3 个连续帧中都处于所述内部多边形内, 所述手部关键点才被确定为进入所述内部多边形。

28. 根据权利要求 24 所述的方法, 其特征在于, 在所述步骤 S5 中, 通过基于所述进入物品检测的结果和所述退出物品检测的结果针对所述人的每只手分别使用一个有限状态机, 来判定所述人在所述货架前的所述拾取或放回行为。

29. 根据权利要求 24 所述的方法，其特征在于，在所述步骤 S2 与所述步骤 S4 之间还包括步骤 S3：

针对所述人的进入所述外部多边形的每只手，分别记录所述手部关键点附近的所述物品在所述外部多边形与所述内部多边形之间的轨迹。

30. 根据权利要求 29 所述的方法，其特征在于，在所述步骤 S5 中，通过基于所述进入物品检测的结果、所述退出物品检测的结果和所述物品的所述轨迹，针对所述人的每只手分别使用一个有限状态机，来判定所述人在所述货架前的所述拾取或放回行为。

31. 根据权利要求 30 所述的方法，其特征在于，若在所述步骤 S2 中针对所述人的两只手的所述进入物品检测均确认检测出进入物品，则在所述步骤 S5 中，通过比较所述人的左手的所述手部关键点附近的所述进入物品的所述轨迹与右手的所述手部关键点附近的所述进入物品的所述轨迹，确定左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品是否为同一物品。

32. 根据权利要求 31 所述的方法，其特征在于，所述人的左手的所述手部关键点附近的所述进入物品的所述轨迹与右手的所述手部关键点附近的所述进入物品的所述轨迹的所述比较是通过核查所述图像数据的各帧中的左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品之间的距离而进行的，并且

仅在各帧中的所述距离均低于预定的距离阈值的情况下，确定左手的所述手部关键点附近的所述进入物品与右手的所述手部关键点附近的所述进入物品为同一物品。

33. 根据权利要求 30 至 32 中任一项所述的方法，其特征在于，若在所述步骤 S4 中针对所述人的两只手的所述退出物品检测均确认检测出退出物品，则在所述步骤 S5 中，通过比较所述人的左手的所述手部关键点附近的所述退出物品的所述轨迹与右手的所述手部关键点附近的所述退出物品的所述轨迹，确定左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品是否为同一物品。

34. 根据权利要求 33 所述的方法，其特征在于，所述人的左手的所述手部关键点附近的所述退出物品的所述轨迹与右手的所述手部关键点附近的所述退出物品的所述轨迹的所述比较是通过核查所述图像数据的各帧中的左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品之间的距离而进行的，并且

仅在各帧中的所述距离均低于预定的距离阈值的情况下，确定左手的所述手部关键点附近的所述退出物品与右手的所述手部关键点附近的所述退出物品为同一物品。

35. 根据权利要求 24 至 30 中任一项所述的方法，其特征在于，在所述步骤 S2 中，只有在所述图像数据中的具有检测到的物品的帧数等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述进入物品检测确认检测出进入物品。

36. 根据权利要求 24 至 30 中任一项所述的方法，其特征在于，在所述步骤 S4 中，只有在所述图像数据中的具有检测到的物品的帧数等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述退出物品检测确认检测出退出物品。

37. 根据权利要求 35 所述的方法，其特征在于，所述图像数据包括来自多个相机或图像传感器的多组图像数据，并且

在所述步骤 S2 中，只有在来自所述多个相机或图像传感器的所述多组图像数据中的具有检测到的物品的帧数之和等于或大于预定的最小帧数并且具有检测到的所述物品的帧数与总帧数的比值等于或大于预定的最小比值的情况下，所述进入物品检测确认检测出进入物品。

38. 根据权利要求 28 或 30 所述的方法，其特征在于，所述有限状态机包括下列拾取或放回行为：无变化、拾取一个物品、拾取两个物品、放回一个物品和放回两个物品。

39. 一种用于检测人在货架前的拾取或放回行为的检测设备，其特征在于，所述检测设备包括：
至少一台相机或图像传感器，获取图像数据；
处理单元，所述处理单元按照根据权利要求 24 至 38 中任一项所述的用于检测人在货架前的拾取或放回行为的检测方法对所述图像数据进行处理。
40. 一种存储介质，其上存储有计算机可读程序，所述程序在处理器上执行时实施根据权利要求 24 至 38 中任一项所述的方法。

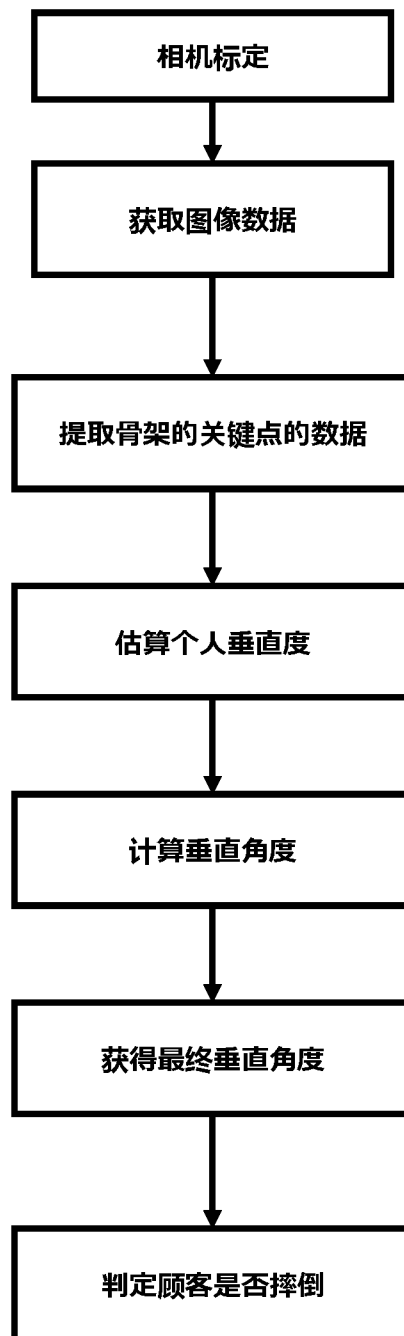


图 1



图 2

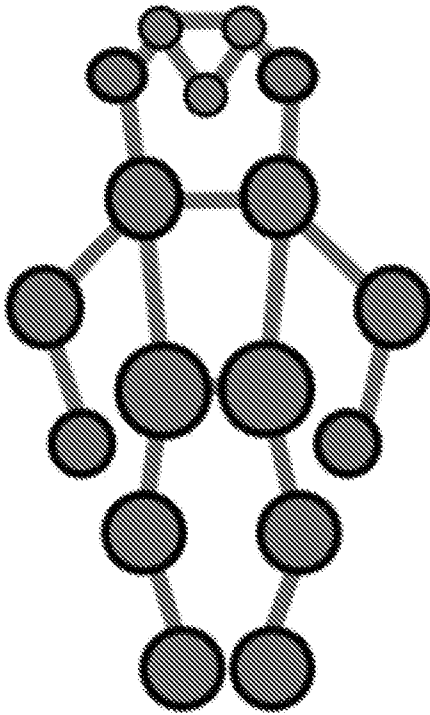


图 3

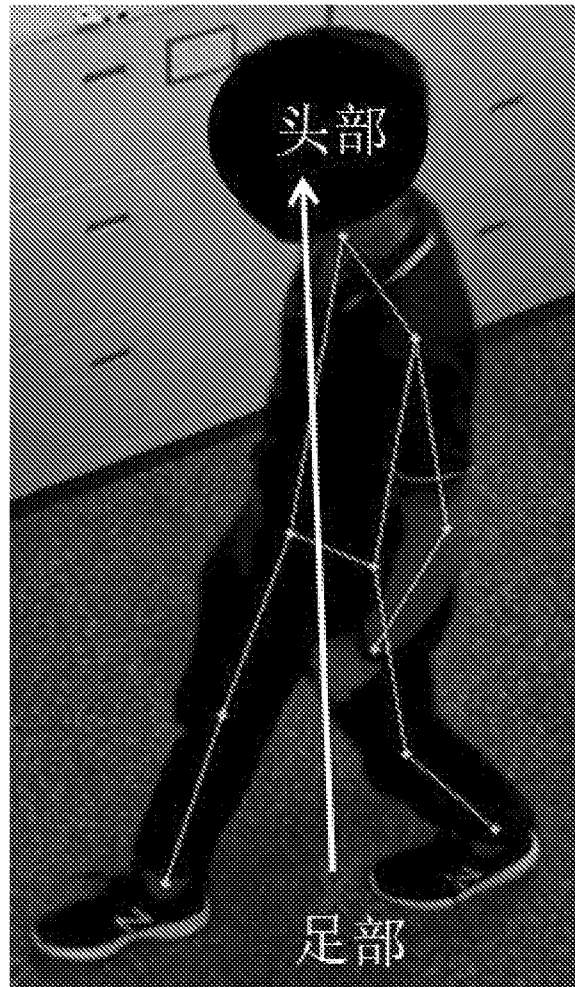
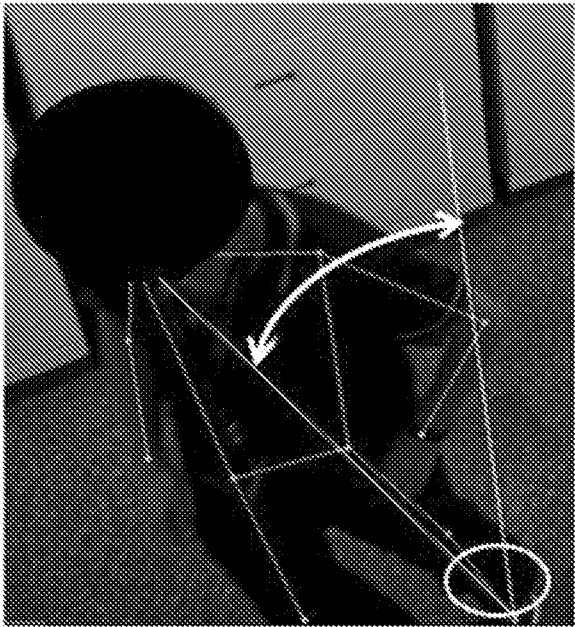
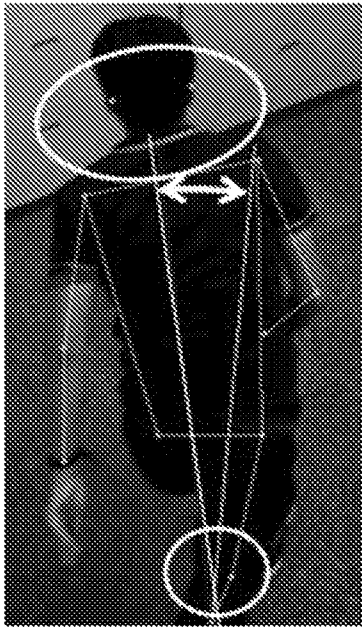


图 4

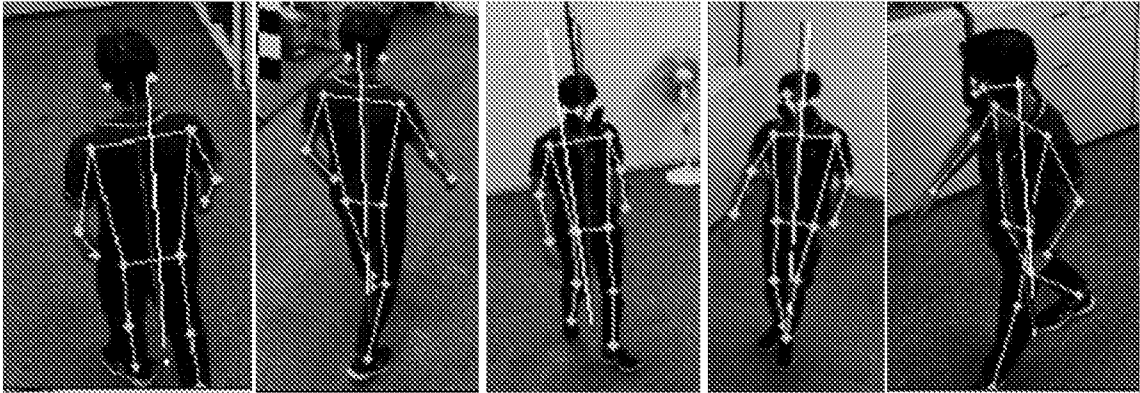


a



b

图 5



a

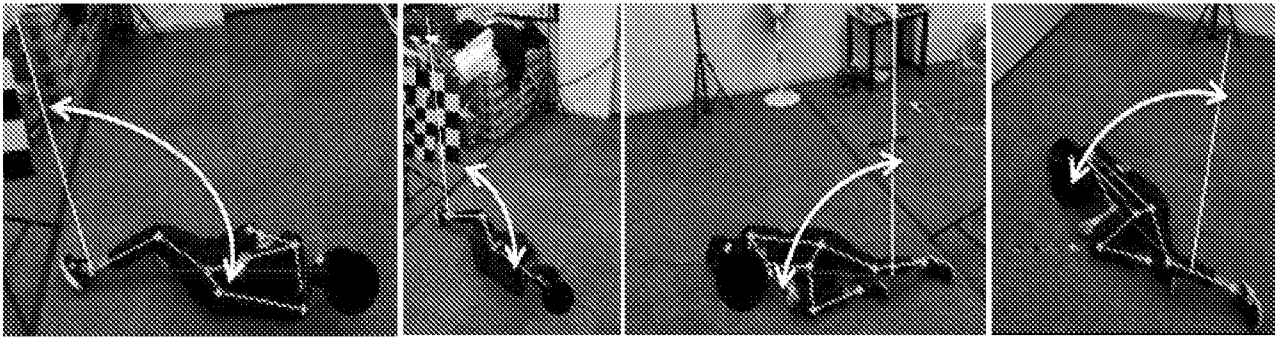
b

c

d

e

图 6



a b c d

图 7

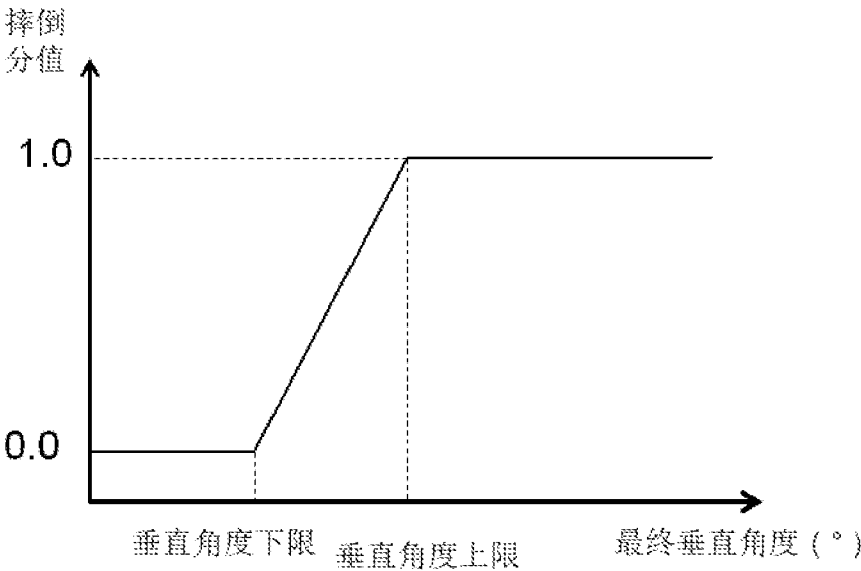
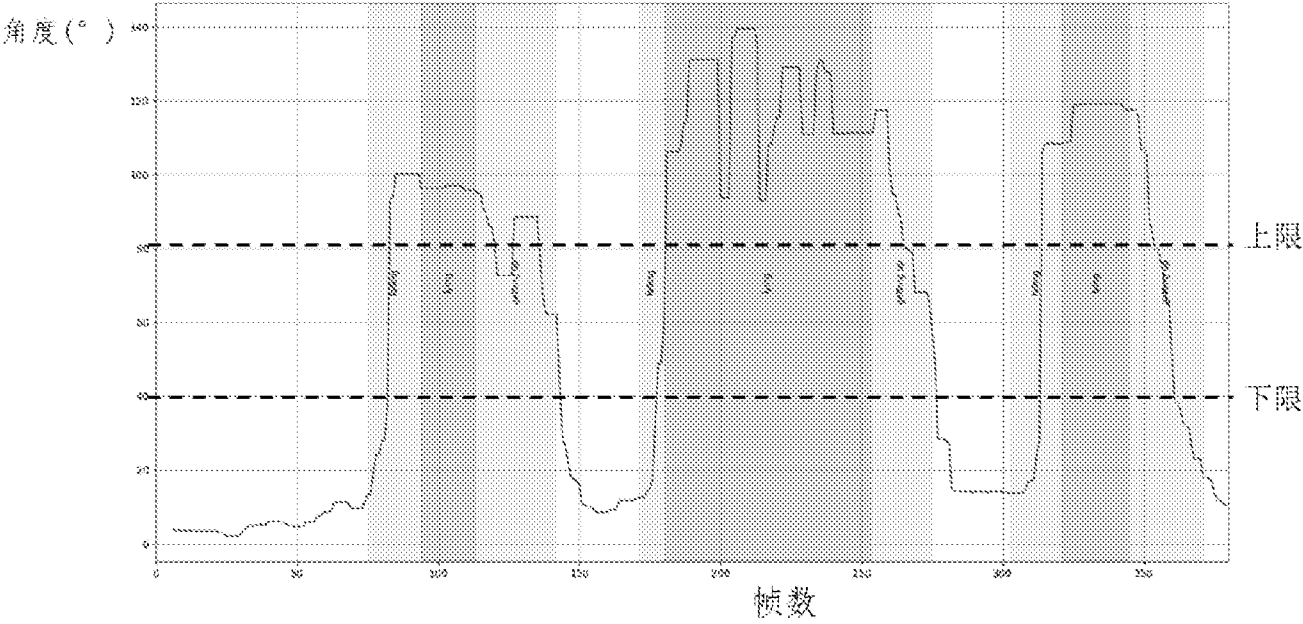
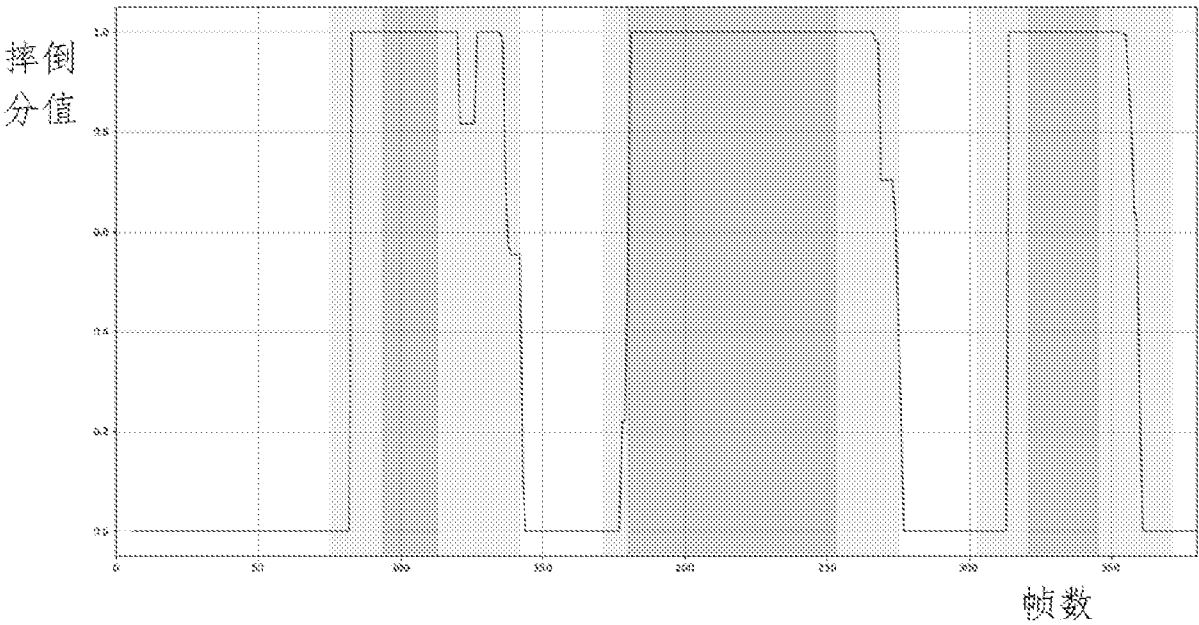


图 8



a



b

图 9

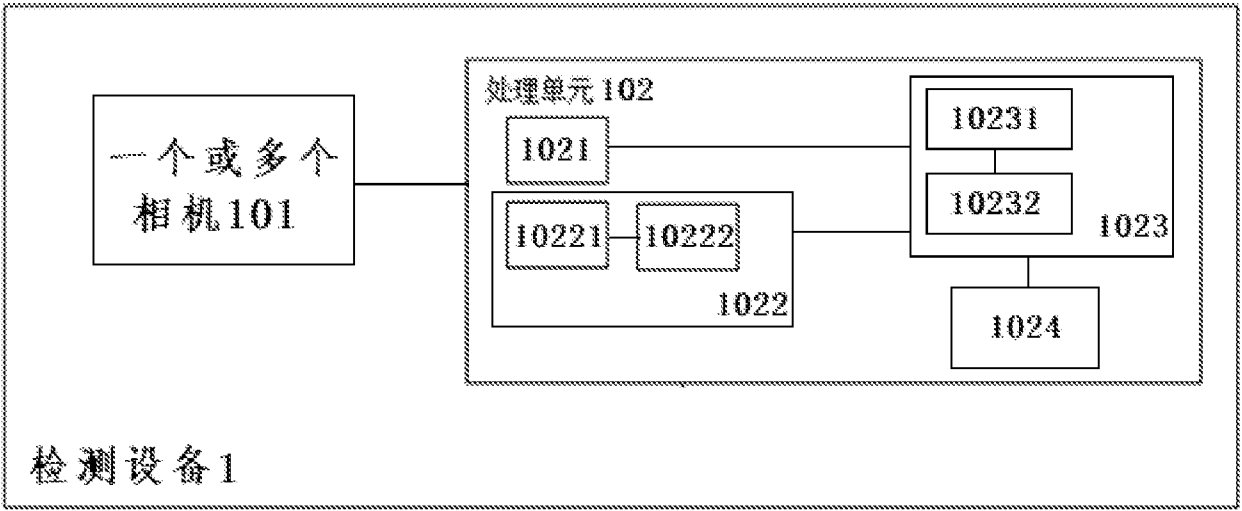


图 10

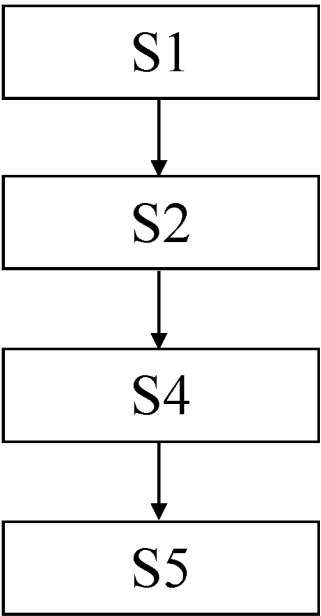


图 11

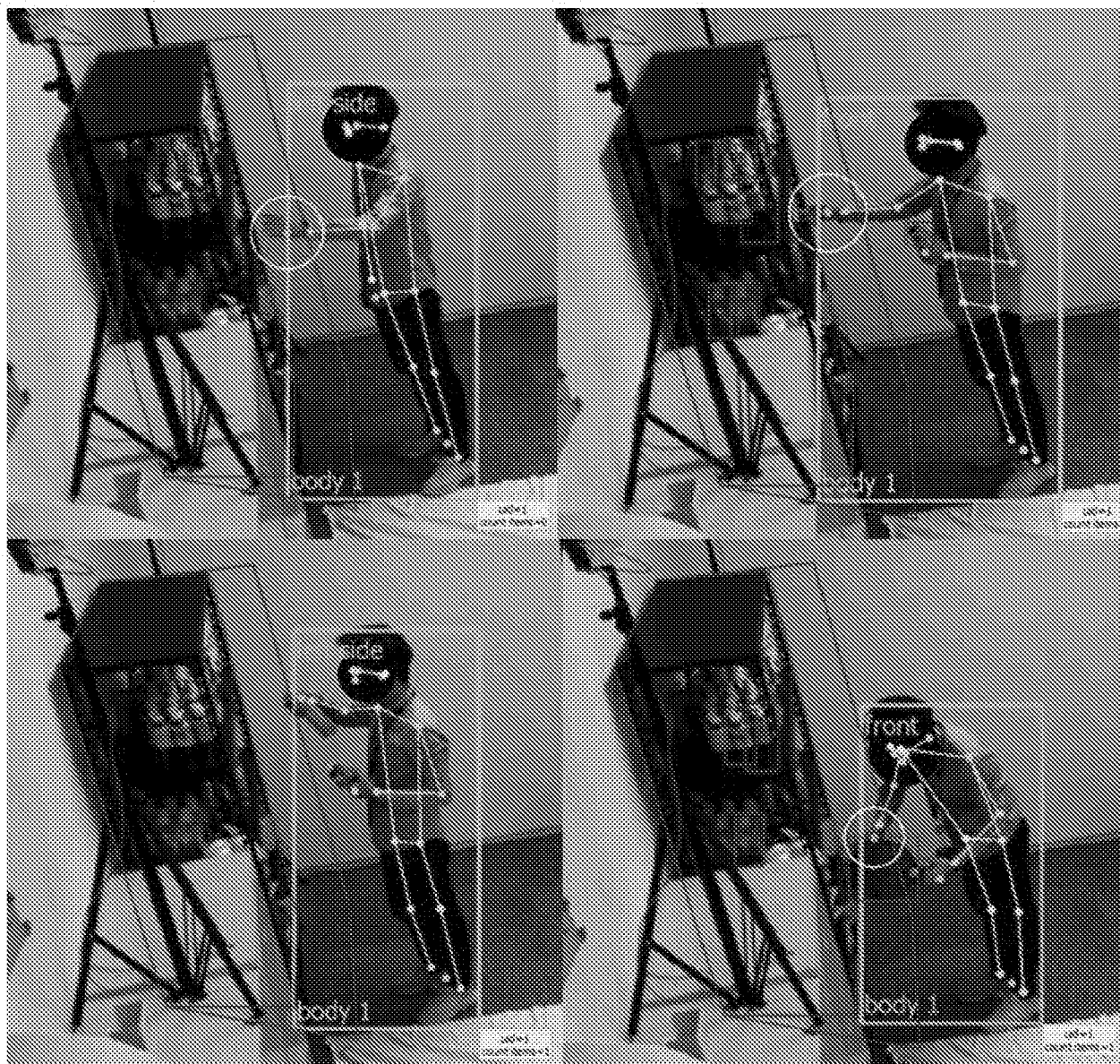


图 12

时间戳 (帧)	1	2	3	4	5	6	7	8	9	10
左手 进入		○		○				○	○	○

a

时间戳 (帧)	1	2	3	4	5	6	7	8	9	10
左手 退出		○								

b

图 13

时间戳 (帧)	1	2	3	4	5	6	7	8	9	10
左手 进入		○		(100, 20)○	○	(250, 30)○	(270, 380)○		(320, 40)○	
右手 进入			○	(110, 20)○		(260, 30)○	(270, 380)○	○	(310, 40)○	○
像素中的 区别				10		-14	0		-14	

图 14

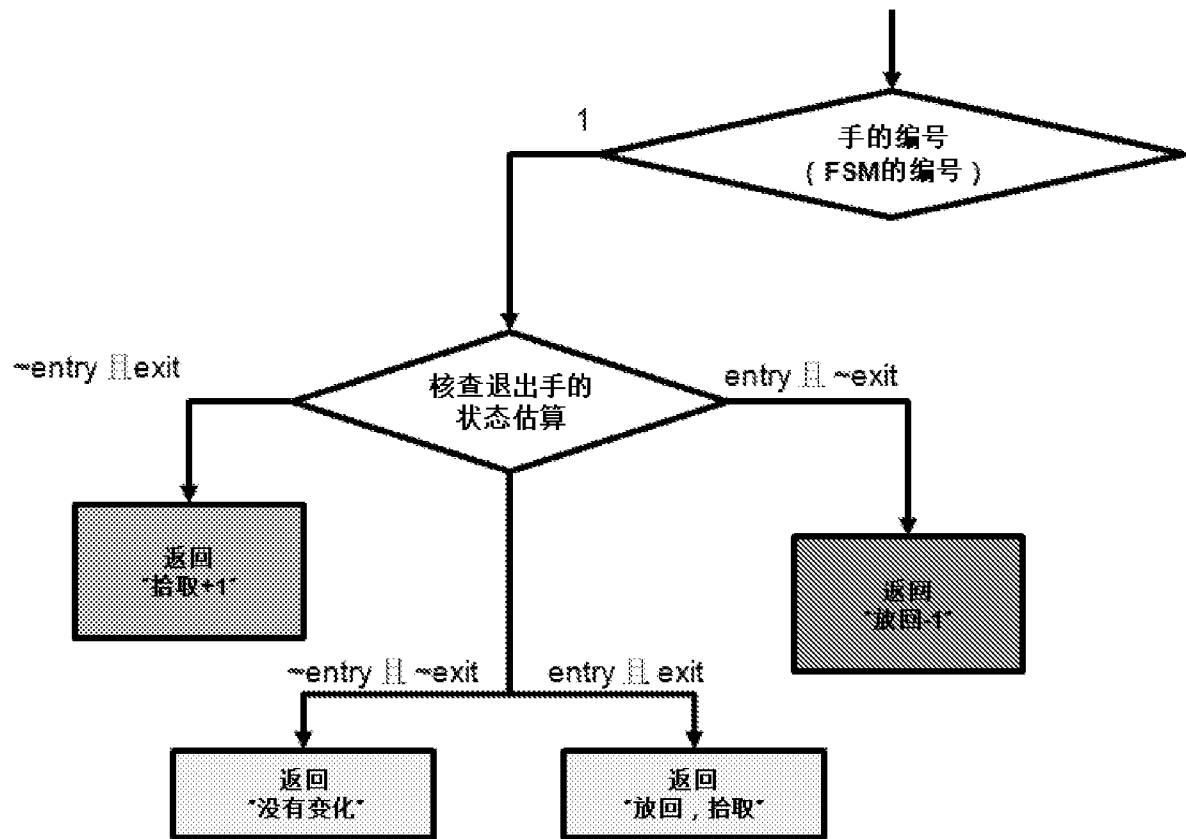
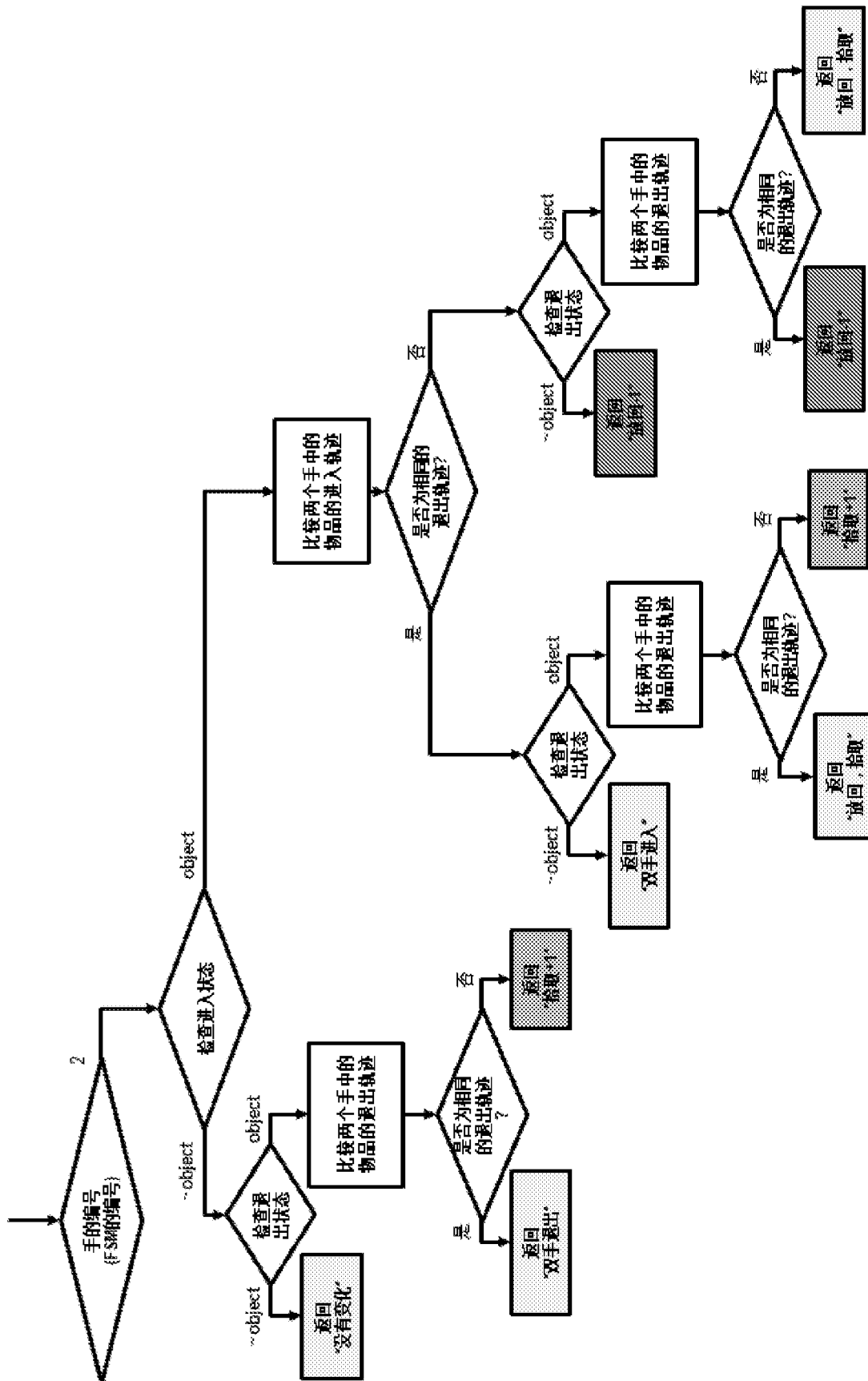


图 15



进入 \ 退出	没有物品	一个物品
没有物品	没有变化	拾取+1
一个物品	放回-1	放回, 拾取

图 17

退出 进入	没有物品	2只手相同的物品	2只手2个不同物品
没有物品	没有变化	双手退出	拾取+1
2只手相同的物品	双手进入	放回, 拾取	拾取+1
2只手2个不同物品	放回-1	放回-1	放回, 拾取

a

进入 \ 退出	没有物品	一个物品
没有物品	没有变化	拾取+1
一个物品	放回-1	放回, 拾取

b

图 18

退出 进入	没有物品	1只手拾取1个物品	2只手拾取1个相同的物品	2只手拾取两个物品
没有物品	没有变化	拾取+1	拾取+1	拾取，拾取 +2
1只手放回1个物品	放回-1	放回，拾取	放回，拾取	放回，拾取，拾取 +1
2只手放回1个相同的物品	放回-1	放回，拾取	放回，拾取	放回，拾取，拾取 +1
2只手放回两个物品	放回，放回 -2	放回，放回，拾取 -1	放回，放回，拾取 -1	放回，放回，拾取，拾取

图 19

相机1:

时间戳 (帧)	1	2	3	4	5
左手退出	○				
右手退出	○	○	○	○	○

仅右手附近区域检测到一个物品

相机2:

时间戳 (帧)	1	2	3	4	5	6
左手退出		○			○	○
右手退出						○

仅左手附近区域检测到一个物品

相机1+2 :

时间戳 (帧)	1	2	3	4	5	6
左手退出	○	○			○	○
右手退出	○	○	○	○	○	○

两只手附近区域分别检测到一个物品

图 20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/071657

A. CLASSIFICATION OF SUBJECT MATTER

G06V40/20(2022.01)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)

G06V,G06K

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)

VEN, CNTXT, ENTXT, DWPI, CNKI, IEEE: 监控, 相机, 摄像头, 视频, 图像, 跌倒, 摔倒, 骨骼, 关键点, 角度, 货架, 货物, 物品, 取, 放, 同一, 相同, monitor, camera, videos, image, fall, bone, key, point, angles, shelve, good, item, fetch, put, same

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 113392681 A (SHENZHEN INTELLIFUSION TECHNOLOGIES CO., LTD.) 14 September 2021 (2021-09-14) description, paragraphs [0018] and [0039]-[0092], and figure 3	1-23
Y	CN 112907892 A (SHANGHAI DIANJI UNIVERSITY) 04 June 2021 (2021-06-04) description, paragraphs [0024]-[0039]	1-23
X	CN 111507315 A (HANGZHOU HIKVISION DIGITAL TECHNOLOGY CO., LTD.) 07 August 2020 (2020-08-07) description, paragraphs [0099]-[0150] and [0217]-[0218], and figures 2-3	24-40
A	CN 112287759 A (ZHEJIANG HANRUI INTELLIGENT TECHNOLOGY CO., LTD.) 29 January 2021 (2021-01-29) entire document	1-40
A	US 2021304007 A1 (VENTECH SOLUTIONS, INC.) 30 September 2021 (2021-09-30) entire document	1-40



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

28 March 2023

Date of mailing of the international search report

04 April 2023

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/071657

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113392681	A	14 September 2021	None			
CN	112907892	A	04 June 2021	None			
CN	111507315	A	07 August 2020	None			
CN	112287759	A	29 January 2021	None			
US	2021304007	A1	30 September 2021	US	2023048309	A1	16 February 2023

国际检索报告

国际申请号

PCT/CN2023/071657

A. 主题的分类 G06V40/20 (2022.01) 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) G06V, G06K 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)) VEN, CNTXT, ENTXT, DWPI, CNKI, IEEE: 监控, 相机, 摄像头, 视频, 图像, 跌倒, 摔倒, 骨骼, 关键点, 角度, 货架, 货物, 物品, 取, 放, 同一, 相同, monitor, camera, videos, image, fall, bone, key, point, angles, shelve, good, item, fetch, put, same		
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
Y	CN 113392681 A (深圳云天励飞技术有限公司) 2021年9月14日 (2021 - 09 - 14) 说明书第[0018], [0039]-[0092]段; 附图3	1-23
Y	CN 112907892 A (上海电机学院) 2021年6月4日 (2021 - 06 - 04) 说明书第[0024]-[0039]段	1-23
X	CN 111507315 A (杭州海康威视数字技术股份有限公司) 2020年8月7日 (2020 - 08 - 07) 说明书第[0099]-[0150], [0217]-[0218]段; 附图2-3	24-40
A	CN 112287759 A (浙江汉德瑞智能科技有限公司) 2021年1月29日 (2021 - 01 - 29) 全文	1-40
A	US 2021304007 A1 (VENTECH SOLUTIONS, INC.) 2021年9月30日 (2021 - 09 - 30) 全文	1-40
☐ 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 “D” 申请人在国际申请中引证的文件 “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 2023年3月28日		国际检索报告邮寄日期 2023年4月4日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 孔昕 电话号码 (+86) 010-53961371

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/071657

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	113392681	A	2021年9月14日	无	
CN	112907892	A	2021年6月4日	无	
CN	111507315	A	2020年8月7日	无	
CN	112287759	A	2021年1月29日	无	
US	2021304007	A1	2021年9月30日	US 2023048309 A1	2023年2月16日