

基于影像监测的货品感知系统及货品感知方法

技术领域

[0001] 本发明涉及一种用于零售业的货品感知技术，具体地说，涉及一种基于影像监测的货品感知系统及感知方法。

背景技术

[0002] 传统零售业的购物方式，每一家超市货便利店需要有专门的销售人员和收款人员，人力成本较高。随着电子支付技术、身份感知技术及云计算技术的发展，无人超市项目在技术上具备很高的可行性。在无人超市项目中，急需解决的一个基本问题就是用户选购货品的判断和记录问题，具体地说，服务器需要准确判定用户从货架上取走货品的种类、取走货品的数量及单价等，以便自动为用户实现结算。

[0003] 现有技术中，有人采用 RFID 技术感知用户选购的货品，这种方式需要在每一个货品上设置 RFID 标签，在门禁处设置 RFID 读写器，其不足之处在于，首先，硬件成本较高，每一 RFID 标签的价格大约在 0.5-1 元左右，标签会提升每一种货品的成本，降低超市的竞争力，对于成本为 5 元的货品来说，加装 RFID 标签会提升其成本的 10-20%；其次，货品感知存在被屏蔽、被去除的可能性，从而出现用户蒙蔽 RFID 阅读器的现象，导致货品丢失；再次，该方案中只有在超市门禁处才能实现结算，如果有用户在离店前将可食用的货品吃掉，将包装留在超市里，RFID 阅读器也无法感知和确定用户的真实消费金额。也就是说，这一方案高度依靠市民用户的自律性和道德水平，而不是用技术手段加以约束，这样的无人超市营业过程中的风险较大。

发明内容

[0004] 本发明的目的在于，提供一种基于影像监测的货品感知技术，可以解决现有技术存在的货品取放状态的感知精准度差、监控误差较大、成本较高、货品容易丢失的技术问题。

[0005] 为实现上述目的，本发明提供一种基于影像监测的货品感知系统，包括：样本采集单元，用以采集多组图片样本，每一组图片样本包括一种货品在多角度下的多张样本图片；同一种类货品的一组图片样本被设有相同的组别标识，该组别标识即为该组图片样本对应的货品的种类；模型训练单元，用以根据多组图片样本中的每一样本图片及其组别标识训练卷积神经网络模型，获取货品识别模型；实时图片采集单元，用以连续采集货架前方空间的至少一实时图片，每一实时图片包括货品图片的部分或全部；以及货品种类获取单元，用以根据所述实时图片及所述货品识别模型获取所述实时图片中显示的货品的种类及数量。

[0006] 进一步地，在不同实施例中，所述基于影像监测的货品感知系统，还包括第一摄像头，连接至所述样本采集单元，用以拍摄每一货品多个角度多个距离的图片；第二摄像头，连接至所述实时图片采集单元，用以拍摄一货架前方空间的实时图片；其中，所述第二摄像头的镜头的视野范围覆盖所述货架前方空间。

[0007] 进一步地，在不同实施例中，所述第二摄像头的数量为两个或四个；所述第二摄像头的镜头朝向所述货架前方的空间区域的中部；所述第二摄像头的镜头的中轴线与水平面夹角为 30-60 度；和/或，所述第二摄像头的镜头与所述货架上端或下端的距离为 0.8-1.2 米；和/或，所述第二摄像头的镜头与所述货架一侧边的距离为 0.8-1.2 米。

[0008] 进一步地，在不同实施例中，所述模型训练单元包括分组模型训练单元，将多组训练样本的每一样本图片及其组别标识输入至卷积神经网络模型，经训练后获取分组模型；其中，所述训练样本为所述多组图片样本的全部或部分，所述分组模型即为货品识别模型。

[0009] 进一步地，在不同实施例中，所述模型训练单元包括样本分类单元，用以将多组所述图片样本随机分成两类，分别为训练样本和测试样本；分组模型训练单元，用以将多组训练样本的每一样本图片及其组别标识输入至卷积神经网络模型，经训练后获取分组模型；以及交叉验证单元，用以根据多组测试样本的每一图片及每一组测试样本的组别标识对所述分组模型进行验证，计算模型准确率；当所述分组模型的准确率小于预设阈值时，返回所述样本分类单元；当所述分组模型的准确率大于或等于所述预设阈值时，所述分组模型为货品识别模型。

[0010] 进一步地，在不同实施例中，所述分组模型训练单元包括特征提取单元，用以将每一训练样本的图片输入卷积神经网络，进行特征提取，获取特征图像；候选区域生成单元，用以将每一训练样本的图片输入候选区域网络，生成多个候选区域；候选区域映射单元，用以把每一训练样本的图片候选区域映射到卷积神经网络最后一层卷积层的特征图像上；分类器生成单元，用以收集多个训练样本的图片特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络，生成分类器。

[0011] 进一步地，在不同实施例中，所述交叉验证单元包括：测试样本输入单元，用以将多个测试样本的每一图片输入至所述分组模型，获取多个测试样本的测试组别标识；以及准确率计算单元，用以将多个测试样本的测试组别标识与所述测试样本的组别标识对比，计算相同标识的数量与所述测试样本的数量的比值，该比值即为所述初级数据模型的准确率。

[0012] 进一步地，在不同实施例中，所述货品种类获取单元还包括组别标识获取单元，用以将多个实时图片输入至所述货品识别模型，获取多个实时图片所对应的多个组别标识，作为可能性结论；以及标识可信度计算单元，用以所述可能性结论中每一种组别标识的数量与所述可能性结论中全部组别标识总数的比值，该比值即为每一种组别标识的可信度，可信度最大的组别标识所对应的货品的种类和/或数量即为所述实时图片上显示的货品的种类和/或数量。

[0013] 进一步地，在不同实施例中，所述组别标识获取单元包括：第二特征提取单元，用以将一实时图片输入卷积神经网络，进行特征提取，获取特征图像；第二候选区域生成单元，用以将所述实时图片输入候选区域网络，生成多个候选区域；第二候选区域映射单元，用以把所述实时图片的候选区域映射到卷积神经网络最后一层的特征图像上；以及组别获取单元，用以收集多个图片的特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络，获取该实时图片所对应的组别标识。

[0014] 为实现上述目的，本发明还提供一种基于影像监测的货品感知方法，包括如下步骤：样本采集步骤，用以采集多组图片样本，每一组图片样本包括一种货品在多角度下的多张样本图片；同一种类货品的一组图片样本被设有相同的组别标识，该组别标识即为该组图片样本对应的货品的种类；模型训练步骤，用以根据多组图片样本中的每一样本图片及其组别标

识训练卷积神经网络模型，获取货品识别模型；实时图片采集步骤，用以连续采集所述货架前方空间的至少一实时图片，每一实时图片包括货品图片的一部分或全部；以及货品种类获取步骤，用以根据所述实时图片及所述货品识别模型获取所述实时图片中显示的货品的种类及数量。

[0015] 进一步地，在不同实施例中，所述模型训练步骤包括如下步骤：分组模型训练步骤，将多组训练样本的每一样本图片及其组别标识输入至卷积神经网络模型，经训练后获取分组模型；其中，所述训练样本为所述多组图片样本的全部或部分，所述分组模型即为货品识别模型。

[0016] 进一步地，在不同实施例中，所述模型训练步骤包括如下步骤：样本分类步骤，用以将多组所述图片样本随机分成两类，分别为训练样本和测试样本；分组模型训练步骤，用以将多组训练样本的每一样本图片及其组别标识输入至卷积神经网络模型，经训练后获取分组模型；以及交叉验证步骤，用以根据多组测试样本的每一图片及每一组测试样本的组别标识对所述分组模型进行验证，计算模型准确率；当所述分组模型的准确率小于预设阈值时，返回所述样本分类步骤；当所述分组模型的准确率大于或等于所述预设阈值时，所述分组模型为货品识别模型。

[0017] 进一步地，在不同实施例中，所述分组模型训练步骤包括如下步骤：特征提取步骤，用以将每一训练样本的图片输入卷积神经网络，进行特征提取，获取特征图像；候选区域生成步骤，用以将每一训练样本的图片输入候选区域网络，生成多个候选区域；候选区域映射步骤，用以把每一训练样本的图片候选区域映射到卷积神经网络最后一层卷积层的特征图像上；分类器生成步骤，用以收集多个训练样本的图片特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络，生成分类器。

[0018] 进一步地，在不同实施例中，所述交叉验证步骤，包括如下步骤：测试样本输入步骤，用以将多个测试样本的每一图片输入至所述分组模型，获取多个测试样本的测试组别标识；以及准确率计算步骤，用以将多个测试样本的测试组别标识与所述测试样本的组别标识对比，计算相同标识的数量与所述测试样本的数量的比值，该比值即为所述初级数据模型的准确率。

[0019] 进一步地，在不同实施例中，所述货品种类获取步骤还包括如下步骤：组别标识获取步骤，用以将多个实时图片输入至所述货品识别模型，获取多个实时图片所对应的多个组别标识，作为可能性结论；以及标识可信度计算步骤，用以所述可能性结论中每一种组别标识的数量与所述可能性结论中全部组别标识总数的比值，该比值即为每一种组别标识的可信度，可信度最大的组别标识所对应的货品的种类和/或数量即为所述实时图片上显示的货品的种类和/或数量。

[0020] 进一步地，在不同实施例中，所述组别标识获取步骤包括如下步骤：第二特征提取步骤，用以将一实时图片输入卷积神经网络，进行特征提取，获取特征图像；第二候选区域生成步骤，用以将所述实时图片输入候选区域网络，生成多个候选区域；第二候选区域映射步骤，用以把所述实时图片的候选区域映射到卷积神经网络最后一层的特征图像上；以及组别获取步骤，用以收集多个图片的特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络，获取该实时图片所对应的组别标识。

[0021] 本发明的有益效果在于，提供一种基于影像监测的货品感知系统及货品感知方法，可以判断货架上货品的取放状态，可以根据货架前方空间的实时图片判断出被取走或被放回的货品的具体种类。若结合基于重量监测的货品感知系统，还可以准确判断货品数量，若结合用户身份识别技术和用户定位跟踪技术，还可以准确判断被取走或被放回的货品的用户身份，从而在该用户购物数据库中准确添加或删除购物记录，以使用户购物结束后自动结算。

附图说明

[0022] 图1为本发明实施例所述的无人超市俯视图；

图2为本发明实施例所述货架的整体结构示意图；

图3为本发明实施例所述用户身份识别系统的结构框图；

图4为本发明实施例所述用户定位系统的结构框图；

图5为本发明实施例所述影像传感器在封闭空间内的分布图；

图6为本发明实施例所述基于影像监测的货品感知系统的结构框图；

图7为本发明实施例所述第二摄像头与货架的位置关系图；

图8为本发明实施例所述购物用户判断系统的结构框图；

图 9 为本发明实施例所述购物数据库系统的结构框图；

图 10 为本发明实施例所述结算系统的结构框图。

[0023] 图 11 为本发明实施例所述基于影像监测的货品感知方法的流程图；

图 12 为本发明实施例所述模型训练步骤的流程图；

图 13 为本发明实施例所述分组模型训练步骤的流程图；

图 14 为本发明实施例所述交叉验证步骤的流程图；

图 15 为本发明实施例所述种类判断步骤的流程图；

图 16 为本发明实施例所述组别标识获取步骤的流程图。

[0024] 图中各个部件标号如下：

1 封闭空间，2 货架，3 支架，4 托盘，5 架板，7 数据处理设备；

100 用户身份识别系统，101 门禁装置，102 身份识别装置；1021 扫码装置，1022 身份获取单元，103 用户入口，104 用户出口；

200 目标物定位系统，201 三维影像采集装置，202 目标物坐标获取单元；

2011 影像传感器，2012 深度图像传感器，2013 RGB 图像传感器，2014 三维影像整合单元；

2021 坐标系建立单元，2022 参数获取单元，2023 背景去除单元，2024 目标物坐标计算单元；

20221 传感器坐标获取单元，20222 相对坐标获取单元，20223 坐标修正单元；

400 基于影像监测的货品感知系统，401 样本采集单元，402 模型训练单元，403 实时图片采集单元，404 货品种类获取单元，405 第一摄像头，406 第二摄像头；407 被放货品确认单元；408 取走货品确认单元；

4021 样本分类单元，4022 分组模型训练单元，4023 交叉验证单元；

40221 特征提取单元，40222 候选区域生成单元，40223 候选区域映射单元，40224 分类器生成单元；40231 测试样本输入单元，40232 准确率计算单元；

4041 组别标识获取单元，4042 标识可信度计算单元；

40411 第二特征提取单元，40412 第二候选区域生成单元，40413 第二候选区域映射单元，40414 组别获取单元；

500 购物用户判断系统, 501 货品信息存储单元, 502 架板坐标存储单元, 503 架板与用户匹配判断单元, 504 货品与用户匹配判断单元;

600 购物数据库系统, 601 购物数据库生成单元, 602 购物数据库更新单元;

700 结算系统, 701 总金额计算单元, 702 支付单元。

具体实施方式

[0025] 以下参考说明书附图完整介绍本发明的优选实施例, 使其技术内容更加清楚和便于理解。本发明可以通过许多不同形式的实施例来得以体现, 其保护范围并非仅限于文中提到的实施例。

[0026] 在附图中, 结构相同的部件以相同数字标号表示, 各处结构或功能相似的部件以相似数字标号表示。本发明所提到的方向用语, 例如, 上、下、前、后、左、右、内、外、上表面、下表面、侧面、顶部、底部、前端、后端、末端等, 仅是附图中的方向, 只是用来解释和说明本发明, 而不是用来限定本发明的保护范围。

[0027] 当某些部件被描述为“在”另一部件“上”时, 所述部件可以直接置于所述另一部件上; 也可以存在一中间部件, 所述部件置于所述中间部件上, 且所述中间部件置于另一部件上。当一个部件被描述为“安装至”或“连接至”另一部件时, 二者可以理解为直接“安装”或“连接”, 或者一个部件通过一中间部件间接“安装至”或“连接至”另一个部件。

[0028] 本实施例涉及一种基于影像监测的货品感知系统, 是一种用于无人超市的无人售货系统中的一部分。如图 1~2 所示, 所述无人售货系统包括一封闭空间 1, 其内设有多组货架 2, 每一货架 2 包括支架 3 及可拆卸式安装在支架 3 上的多个托盘 4, 多个托盘 4 在不同高度彼此平行或者在同一高度彼此平齐。每一托盘 4 上设有多个并列设置的架板 5, 每一架板 5 上放置有至少一种货品。本实施例架板 5 上放置的货品需要便于用户取走或放回, 因此, 以架板 5 朝向用户的一端作为架板 5 的前端。本实施例中, 每一架板 5 皆为一个敞口的箱体, 可以被放置有一种或多种货品, 所述货品为标准货品, 同一种类的货品的外观和重量都相同或近似。

[0029] 本实施例还包括数据处理设备 7，如服务器或计算机等，数据处理设备 7 内部设有多个数据处理软件，具有多个功能模块，可以通过数据线连接至多个硬件，以软硬件结合方式实现多种功能。

[0030] 如图 1、图 3 所示，本实施例还包括用户身份识别系统 100，用以识别每一用户的身份信息。用户身份识别系统 100 包括门禁装置 101 及身份识别装置 102，本实施例所述封闭空间 1 不是绝对密封空间，而是相对封闭的空间，封闭空间 1 设有一个或多个出入口，优选一个用户入口 103 及一个用户出口 104，所有用户由用户入口 103 进入此封闭空间 1，由用户出口 104 离开此封闭空间 1。

[0031] 封闭空间 1 的每一出入口处都设置有门禁装置 101，优选一自动闸机。身份识别装置 102 用以获取用户的身份信息，包括连接至数据处理设备 7 的扫码装置 1021 以及数据处理设备 7 内的身份获取单元 1022。扫码装置 1021 设置于用户入口 103 处的门禁装置 101 的内部或外部，优选设于自动闸机的外表面，用以扫描身份识别码，优选二维码；身份获取单元 1022 为数据处理设备 7 中的一个功能模块，可以根据所述身份识别码获取用户的身份信息。用户出口 104 处的门禁装置 101 无需设置身份识别装置 102。

[0032] 本实施例中，每个用户将与无人超市配套使用的专用应用软件（APP）下载到移动通信终端（手机、平板电脑等）中，在应用软件（APP）中注册账号，关联至支付软件；或者，每个用户将支付软件（如微信/支付宝）下载到移动通信终端中，在支付软件中嵌入与无人超市配套使用的小程序，在支付软件中注册账号，专用应用软件（APP）或支付软件内设有用户注册信息及电子支付信息，包括用户身份信息、银行账号信息、支付密码等。注册完成后，用户身份信息会存储于数据处理设备 7 的用户数据库中。

[0033] 移动通信终端中的应用软件（APP）可以生成一二维码，该二维码内存储有用户的身份信息，当某一用户需要从外部进入封闭空间 1 时，将应用软件生成的二维码正对扫码装置 1021 的扫描端，扫码后扫码装置 1021 对此二维码进行解码处理，将解码结果传送至数据处理设备 7，如果二维码是可识别的，且识别出的身份信息与预先存储在用户数据库的身份信息相匹配，确定该用户身份合法，门禁装置 101 打开允许该用户进入封闭空间 1，用户入口 103 处的门禁装置 101 设有感应装置，如红外传感器等，当该用户进入封闭空间 1 之后，

门禁装置 101 感知到有人走过门禁, 然后自动关闭。当用户购物结束需要离开封闭空间 1 时, 用户出口 104 处的门禁装置 101 感应到有人从封闭空间 1 内部靠近门禁装置 101 时, 该门禁装置会自动打开, 待用户离开封闭空间 1 后, 门禁装置 101 感知到有人走过门禁, 然后自动关闭。

[0034] 身份获取单元 1022 根据所述身份识别码获取用户的身份信息后, 数据处理设备 7 可以生成该用户的购物数据库, 在用户购物过程中根据用户每一次购物行为获取购物信息更新该购物数据库。由于用户随身携带的移动通信终端通过应用软件 (APP) 与数据处理设备 7 进行实时数据交换, 用户的购物数据库也可以显示在移动通信终端中的应用软件 (APP) 中, 形成购物车界面, 以使用户了解自己的购物记录及后续结算。

[0035] 如图 4 所示, 本实施例还包括目标物定位系统 200, 用以获取每一目标物在封闭空间 1 的实时位置, 所述目标物为用户及其延伸部的全部或局部, 用以获取用户整体或局部 (如头部、手部等) 的坐标集。目标物定位系统 200 包括连接至数据处理设备 7 的三维影像采集装置 201 以及设于数据处理设备 7 内的目标物坐标获取单元 202。

[0036] 如图 5 所示, 三维影像采集装置 201 包括至少一影像传感器 2011, 用以实时采集至少一帧三维影像, 影像传感器 2011 平均分布于封闭空间 1 顶部, 其镜头朝向下方, 镜头中轴线可以与水平面垂直也可以有一定的倾角, 影像传感器 2011 的视野范围覆盖封闭空间 1 的全部底面。用户在无人超市内行动或购物时, 始终处于影像传感器的监视之下, 此时影像传感器采集到的三维影像中包括用户影像, 所述用户影像是指该用户身体及其延伸部的全部或局部。如果该封闭空间内空无一人, 每一时刻下的三维影像都与前一时刻下相同, 可以判断该时刻下的三维影像都为背景, 不包括任何用户影像。

[0037] 每一影像传感器 2011 包括并列设置的深度图像传感器 2012 及 RGB 图像传感器 2013 以及三维影像整合单元 2014, 深度图像传感器 2012 连续采集多帧深度图像, RGB 图像传感器 2013 连续采集多帧 RGB 图像, 三维影像整合单元 2014 将同一时刻采集到的一帧深度图像及一帧 RGB 图像结合为一帧三维影像。

[0038] 上述两个传感器同步采集 (同时采集且采集频率相同), 三维影像整合单元 2014 可以连续获得多帧三维影像并传送至数据处理设备 7 的目标物坐标获取单元 202。

[0039] 目标物坐标获取单元 202 为数据处理设备 7 中的一个功能模块, 在所述封闭空间内建立三维坐标系, 根据连续多帧包括用户影像的三维影像实时获取所述用户在所述三维坐标系下的坐标集或坐标。目标物坐标获取单元 202 包括坐标系建立单元 2021、参数获取单元 2022、背景去除单元 2023 以及目标物坐标计算单元 2024。坐标系建立单元 2021 在所述封闭空间内建立三维坐标系, 优选地, 选择封闭空间底面(无人超市的地面)的中心点作为坐标系原点, 在水平方向设置 X 轴、Y 轴, 在竖直方向设置 Z 轴。

[0040] 由于用户身体的全部或局部在三维坐标系下占用较大的空间, 因此可以用一个坐标集代表用户的位置, 如果考虑到位置精确控制和计算方便, 也可以用该坐标集中某一个特定的点的坐标代表用户的位置, 例如可以采用该用户坐标集中最高的一个点(Z 轴数值最大的点)的坐标来表示用户位置。

[0041] 参数获取单元 2022 对连续多帧包括用户影像的三维影像进行处理, 获取每一帧三维影像的每一像素点的位置参数和色彩参数; 所述位置参数为 x 、 y 、 z , 代表该像素点在所述三维坐标系下的位置坐标; 所述色彩参数为 r 、 g 、 b , 分别代表该像素点的三原色强度。当某一用户行进至任一影像传感器的视野内之后, 数据处理设备 7 可以获取多帧三维影像, 每一帧三维影像中都包括用户影像和背景影像, 每一像素点可能是用户的一部分也可以是背景的一部分。

[0042] 在不同的影像传感器采集到三维影像中, 表示用户身体及其延伸部的同样位置的像素点, 其色彩参数 r 、 g 、 b 都是相同的。由于不同位置的影像传感器与用户的距离不同, 每一影像传感器直接采集的初级位置参数都是用户身体及其延伸部上的一点相对于该影像传感器的位置坐标, 因此要进行坐标变换, 将不同位置的影像传感器采集的初级位置参数都转换为在所述封闭空间内建立的三维坐标系下的位置参数。

[0043] 参数获取单元 2022 包括传感器坐标获取单元 20221、相对坐标获取单元 20222 以及坐标修正单元 20223, 传感器坐标获取单元 20221 获取采集该帧三维影像的影像传感器的中心点(即并列设置的深度图像传感器 2012 及 RGB 图像传感器 2013 的镜头中心点连线的中分点)在所述封闭空间内建立的所述三维坐标系下的坐标; 相对坐标获取单元 20222 以所述影像传感器的中心点为第二原点建立第二三维坐标系, 其 X 轴、Y 轴及 Z 轴的方向与所述三

维坐标系相同,从所述三维影像获取每一像素点在所述第二三维坐标系下的坐标;坐标修正单元 20223 用以根据所述影像传感器中心点在所述三维坐标系下的坐标及所述三维影像中每一像素点在所述第二三维坐标系下的坐标,计算并修正所述三维影像的每一像素点在所述三维坐标系下的坐标,从而获得用户及其延伸部的每一像素点的位置参数。

[0044] 在连续采集的 M 帧三维影像中,每一帧三维影像包括且仅包括一个用户的影像,若分属于不同三维影像、位置参数相同的 N 个像素点的色彩参数相同时,且 N 大于 $0.9 \times M$ 且小于或等于 M,背景去除单元 2023 判定该 N 个像素点为背景像素点,从所述 M 帧三维影像中去除 N 个所述背景像素点,获得 M 帧无背景三维影像,即为该用户的影像。在连续获取的三维影像中,如果分属于不同三维影像的、位置参数相同的像素点的色彩参数相同,或者大部分(如 90%)相同,即可认定像素点的位置为背景,从而可以将该像素点从相应的三维影像中去除。

[0045] 在目标物坐标计算单元 2024 中,若目标物为用户及其延伸部全部,所述 M 帧无背景三维影像中所有像素点的位置参数的集合即为所述用户及其延伸部的全部的坐标集;在所述坐标集中,参数 z 最大的像素点的位置参数被定义为用户的坐标。在连续获取的三维影像中,去除背景像素点后,剩下的像素点即可代表该用户整体的行进轨迹。若连续采集的 M 帧三维影像中,每一帧三维影像包括多个用户的影像,需要先在每一 M 帧三维影像截取只含一个用户全部或部分的三维影像。

[0046] 若目标物为用户及其延伸部的局部,可以获取所述用户局部的坐标集,如头部、肩部、肘部、腕部、手部等。深度图像传感器 2012 及 RGB 图像传感器 2013 分别设有一个镜头,如果将两个镜头的中轴线垂直水平面设置,两个镜头就会俯视封闭空间内的货品及用户。正常情况下,两个镜头可以捕捉到用户头部及肩部的位置坐标集,当用户伸出手时,也可以捕捉到用户臂部、肘部、腕部、手部的位置坐标集。如果将某一时刻下该用户的头部、肩部、肘部、腕部、手部都连成一条折线或曲线,即可将用户的手部与头部位置建立对应关系,也就是说,可以实时获取到某一手部的的位置,同时可以判断出该手部属于哪一用户。

[0047] 进一步地,影像传感器 2011 的视野范围也可以覆盖出入口外部的部分空间,当用户在出入口外部时,该用户的影像就可以被影像传感器 2011 获取到。用户使用所述无人售

货系统的全部过程，包括出入口处身份识别过程、进入封闭空间 1 的过程、在封闭空间 1 内行走或驻留过程、离开封闭空间 1 过程，全部处于影像传感器 2011 的监控下，可以实时监控已知身份的某一用户及其身体的一部分在封闭空间 1 内的实时位置。扫码装置 1021 读取用户的二维码时，数据处理设备 7 即可获取其身份信息，影像传感器 2011 从扫码装置 1021 读码时开始定位及实时跟踪该用户位置，监控该用户是否与某一货架匹配。当影像传感器 2011 无法获取该用户的实时三维影像时，可以认定该用户购物结束，从而对其进行结算。

[0048] 如图 6 所示，本实施例涉及一种基于影像监测的货品感知系统 400，其包括样本采集单元 401、模型训练单元 402、实时图片采集单元 403 以及货品种类获取单元 404，上述四个单元为数据处理设备 7 中的功能模块，基于影像监测的货品感知系统 400 可以监控货架前方空间的实时影像，判断被取走或被放回的货品的种类。

[0049] 基于影像监测的货品感知系统 400 还包括第一摄像头 405 及第二摄像头 406，第一摄像头 405 连接至数据处理设备 7 中的样本采集单元 401，用以拍摄每一货品多个角度多个距离的图片。第二摄像头 406 连接至数据处理设备 7 中的实时图片采集单元 403，用以拍摄一货架前方空间的实时图片。

[0050] 如图 7 所示，优选地，第二摄像头 406 的数量为两个或四个，设置于货架 2 的外部，每一第二摄像头 406 朝向货架 2 的一个角落处。货架 2 的多个架板 5 的最前端位于同一平面上，该平面称之为货架平面，第二摄像头 406 设有镜头，该镜头的视野范围覆盖所述货架前方空间；当货品被从所述货架上被取下或者被放置在货架上时，所述货品被取下过程或被放回过程的影像被所述第二摄像头拍摄到。所述货架前方空间是指货架前方对应货架平面的空间区域，所述货架前方空间一般是指货架前方 30~50 厘米宽度的区域范围，每一个第二摄像头 406 的镜头朝向所述货架前方空间的中心区域。

[0051] 优选地，第二摄像头 406 的镜头的中轴线与水平面夹角为 30-60 度；和/或，第二摄像头 406 的镜头与货架 2 上端或下端的距离为 0.8-1.2 米；和/或，第二摄像头 406 的镜头与货架 2 一侧边的距离为 0.8-1.2 米，确保第二摄像头 406 的视野范围可以完全覆盖货架前方空间，当货品被从货架 2 上被取下或者被放置在货架 2 上时，取下过程或放置过程的影像被第二摄像头 406 拍摄到。

[0052] 如图 6 所示, 样本采集单元 401 用以采集至少一组图片样本, 每一组图片样本包括一种货品在多角度下的多张样本图片; 同一种类货品的一组图片样本被设有相同的组别标识, 该组别标识代表该组图片样本对应的货品的种类。优选地, 第一摄像头 405 对货架 2 上每种货品需要拍摄不同角度不同距离的 3000~5000 张图片, 并传送至数据处理设备 7 的样本采集单元 401。这些图片中, 有些是独立拍摄货品的图片, 有些是有背景的货品图片, 有些是货品被某人拿在手中的图片, 有些是多个同类产品叠放在一起的图片。由于本实施例涉及的无人超市中所销售的货品为标准货品, 同一种类货品的外观六面图都是相同或相近似的, 因此在同一种类货品中只要选择一个或几个产品, 对其进行多次拍照处理即可完成该类货品训练样本的采样。

[0053] 模型训练单元 402 用以根据多组图片样本中的每一样本图片及每一样本图片的组别标识训练卷积神经网络 (CNN) 模型, 获取货品识别模型。卷积神经网络 (Convolutional Neural Network, 简称 CNN) 是一种前馈神经网络, 对于大型图像处理有出色表现。优选地, 本实施例中的卷积神经网络模型是目前运算量最小、响应速度最快的 Faster RCNN 网络模型, 该模型的最快响应速度只要 0.2 秒左右, 可以在极短的时间准确识别出图片的物品的种类及数量。如果模型训练的样本较少, 或者, 样本的清晰度比较低, 只训练一次获得的分组模型在判断图片中显示的货品的种类时可能误差较大, 因此最好能有交叉验证的过程, 可以获得更加精确的模型。模型训练单元 402 包括样本分类单元 4021、分组模型训练单元 4022 以及交叉验证单元 4023。

[0054] 样本分类单元 4021 将多组所述图片样本随机分成两类, 分别为训练样本和测试样本; 分组模型训练单元 4022 将多组训练样本的每一样本图片及其组别标识输入至卷积神经网络模型, 经训练后获取分组模型; 交叉验证单元 4023 用以根据多组测试样本的每一图片及每一组测试样本的组别标识对所述分组模型进行验证, 计算模型准确率; 当所述分组模型的准确率小于预设阈值 (如 90%) 时, 返回所述样本分类单元; 当所述分组模型的准确率大于或等于所述预设阈值 (如 90%) 时, 所述分组模型为货品识别模型。如果模型训练的样本较少, 或者, 样本的清晰度比较低, 只训练一次获得的分组模型, 在判断图片中显示的货品的种类时可能误差较大, 因此最好能有交叉验证单元。

[0055] 如果模型训练的样本够多，样本的清晰度比较高，直接训练一次，就可以利用 Faster RCNN 网络模型训练出一个分组模型，该分组模型可以有效判断图片中显示的货品的种类。在其他实施例中，模型训练单元 402 可以只包括分组模型训练单元 4022，将样本采集单元 401 采集的多组图片样本的全部或部分作为训练样本，将每一样本图片及其组别标识输入至卷积神经网络模型，经训练后获取分组模型，所述分组模型即为货品识别模型。

[0056] 分组模型训练单元 4022 包括特征提取单元 40221、候选区域生成单元 40222、候选区域映射单元 40223 以及分类器生成单元 40224。特征提取单元 40221 将每一训练样本的图片输入卷积神经网络，进行特征提取，获取特征图像（feature map），该特征图像对应训练图片中显示全部或部分货品的区域；候选区域生成单元 40222 将每一训练样本的图片输入候选区域网络（RPN），生成多个候选区域（region proposals）；候选区域映射单元 40223 把每一训练样本的图片候选区域映射到卷积神经网络最后一层卷积层的特征图像上；分类器生成单元 40224 收集多个训练样本的图片特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络（classifier 网络），生成分类器。在该分类器网络中，将对应同一货品的所有训练样本图片的特征图像与该货品的组别标识形成对应关系，该组别标识即为该组图片样本对应的货品的种类。

[0057] 交叉验证单元 4023 包括测试样本输入单元 40231，准确率计算单元 40232，测试样本输入单元 40231 将多个测试样本的每一图片输入至所述分组模型，获取多个测试样本的测试组别标识；准确率计算单元 40232 将多个测试样本的测试组别标识与所述测试样本的组别标识对比，计算相同标识的数量与所述测试样本的数量的比值，该比值即为所述初级数据模型的准确率。优选地，将预设阈值设置为 90%，当分组模型的准确率小于 90% 时，样本分类单元 4021 将样本重新分组、重新训练；当所述分组模型的准确率大于或等于 90% 时，可以认为所述分组模型即为货品识别模型。

[0058] 实时图片采集单元 403 用以连续采集货架前方空间的至少一实时图片，每一实时图片包括一个或多个货品图片的部分或全部。实时图片采集单元 403 连接至多个第二摄像头 406，设置于所述货架的边缘处，实时拍摄一货架前方区域的图片。当有用户从货架的某一架板上取走货品时，或者，当有用户放置货品或物品至货架的某一架板时，第二摄像头 406

可以拍摄到货品在货架前的图片，该图片中包括此货品的全部或局部的照片，显示该货品的形状、图案及色彩。第二摄像头获取该货品的多张实时图片，传送至实时图片采集单元 403。

[0059] 在连续获取的多帧图片中，每一帧图片中货品与货架边缘处的距离都是不同的，数据处理设备 7 可以计算所述距离的变化判断货品与货架距离变化，如果所述距离变大，可以认为货品被从货架上取走，如果所述距离变小，可以认为货品被放置在货架上。

[0060] 假设用户手持货品在货架前方停留 1 秒钟，第二摄像头 406 在 1 秒钟内获取不同角度的 60 张实时图片，每一图片的显示内容可以包括用户的手部及货品的局部。如果实时图片中不包含任何货品内容，例如某些货品体型较小，某些图片中只能显示用户的手部，或者，某些图片只显示背景影像，这样的图片可以筛除。本实施例不能适用于尺寸极小的货品，如果货品完全被用户手部包住，就不能识别出货品种类。

[0061] 第二摄像头 406 可以为长时间持续运行的设备，也可以在货架板上设置一个红外开关，连接至第二摄像头 406，平时为断电状态，当有用户行进至货架附近时，红外开关感应到热量并连通，第二摄像头 406 通电启动；当用户离开时，红外开关感应到不到热量从而断开，第二摄像头 406 断电关闭。

[0062] 货品种类获取单元 404 用以根据所述实时图片及所述货品识别模型获取所述实时图片中显示的货品的种类及数量。货品种类获取单元 404 包括组别标识获取单元 4041 以及标识可信度计算单元 4042。

[0063] 组别标识获取单元 4041 获取多幅包含某一货品内容的实时图片，将其输入至所述货品识别模型，获取该模型输出的、多个实时图片所对应的多个组别标识，作为可能性结论。组别标识获取单元 4041 包括第二特征提取单元 40411、第二候选区域生成单元 40412、第二候选区域映射单元 40413 以及第二组别获取单元 40414。第二特征提取单元 40411 将一实时图片输入卷积神经网络，进行特征提取，获取特征图像；第二候选区域生成单元 40412 将所述实时图片输入候选区域网络，生成多个候选区域；第二候选区域映射单元 40413 把所述实时图片的候选区域映射到卷积神经网络最后一层的特征图像上；组别获取单元 40414 收集多个图片的特征图像和候选区域，计算出候选区域的特征图像，并送入分类器网络，获取该实时图片所对应的组别标识。每一实时图片输入至所述货品识别模型之后，都可以判断出该图

片对应的组别标识，如果第二摄像头采集到 60 张图片，且每一图片上都有显示某货品，将其输入货品识别模型，就可以获取 60 个组别标识。

[0064] 标识可信度计算单元 4042 计算所述可能性结论中所述可能性结论中每一种组别标识的数量与所述可能性结论中全部组别标识总数的比值，该比值即为每一种组别标识的可信度，可信度最大的组别标识所对应的货品的种类和/或数量即为所述实时图片上显示的货品的种类和/或数量。对比上述的 60 个组别标识，如果与 60 个组别标识相应的 60 个货品种类中，出现货品 A 的次数为 36，出现货品 B 的次数为 18，出现货品 C 的次数为 6，则三者的可信度分别为 60%、30%及 10%，可以认定所述实时图片上所显示的货品的种类为可信度最高的货品 A。如果样本采集单元 401 采集的图片样本中包括多个同类货品叠放在一起的照片，货品种类获取单元 404 还可以进一步判断货品的数量。

[0065] 基于影像监控的货品感知系统 400 的有益效果在于，可以实时监测货架前方空间的影像，判断是否有货品被从货架上取走或被放回至货架，利用机器学习中的卷积算法推断出货品的种类及数量的可能性结论，并选择其中可信度最高的结果作为最后结论。

[0066] 如图 8 所示，本实施例还包括购物用户判断系统 500，其为数据处理设备 7 中的功能模块，当任一种类货品被取走或被放回时，根据所述用户的身份信息及所述用户的实时位置获取取走或放回货品的用户身份。购物用户判断系统 500 包括货品信息存储单元 501、架板坐标存储单元 502、架板与用户匹配判断单元 503 以及货品与用户匹配判断单元 504。

[0067] 数据处理设备 7 内设有货品数据库，存储于货品信息存储单元 501 中，所述货品数据库包括每一货品信息；所述货品信息包括每一货品的货品名、型号、净含量及单价等，还包括放置该货品的货架编号、放置该货品的架板编号及货品编号。

[0068] 目标物定位系统 200 在所述封闭空间内建立三维坐标系，由于货架 2 及架板 5 的位置确定，因此建立坐标系后即可获取各个货架 2 及各个架板 5 的坐标，将货架坐标集及架板坐标集存储于架板坐标存储单元 502 中，设置架板上方的用以放置货品的架板空间的高度（如 30CM），即可获取所述架板空间的坐标集。

[0069] 用户坐标获取单元 202 可以获取每一已知身份用户的手部的实时坐标集,当一架板上方的架板空间的坐标集与一用户手部坐标集有交集时,架板与用户匹配判断单元 503 判定该架板与该用户匹配,可以认为该用户将手部伸入至该架板上方的架板空间。

[0070] 目标物定位系统 200 在所述封闭空间内建立三维坐标系,由于货架 2 及架板 5 的位置确定,因此建立坐标系后即可获取各个货架 2 及各个架板 5 的坐标,将货架坐标集及架板坐标集存储于架板坐标存储单元 502 中,设置架板上方的用以放置货品的架板空间的高度(如 30CM),即可获取所述架板空间的坐标集。

[0071] 用户坐标获取单元 202 可以获取每一已知身份用户的手部的实时坐标集,当一架板上方的架板空间的坐标集与一用户手部坐标集有交集时,架板与用户匹配判断单元 503 判定该架板与该用户匹配,可以认为该用户将手部伸入至该架板上方的架板空间。

[0072] 基于影像监测的货品感知系统 400 通过第二摄像头 406 监控货架前方空间的实时影像,可以根据连续获取的多帧图片中货品与货架的距离变化来判断货品的取放状态,并判断被取走或被放回的货品的种类及数量。所述取放状态包括货品静置状态、被取走状态及被放回状态。

[0073] 当有货品从一架板上被取走或被放置到一架板上时,且同一时刻下有一用户与该架板匹配,货品与用户匹配判断单元 504 判定该货品与该用户匹配,该货品在这一时刻被该用户从该架板上取走或放置到该架板上,从而确定取走货品或放回货品的该用户的身份。

[0074] 如图 9 所示,本实施例还包括购物信息记录单元 600,其为数据处理设备 7 中的功能模块,根据每一用户的身份信息生成至少一购物数据库,用以记录每一用户取走至少一货品的种类及数量。购物信息记录单元 600 包括购物数据库生成单元 601 以及购物数据库更新单元 602。

[0075] 当一用户的身份被用户身份识别系统 100 识别时,身份获取单元 1022 获取用户的身份信息,购物数据库生成单元 601 根据所述用户的身份信息在数据处理设备 7 中生成该用户的购物数据库,初始状态下的购物数据库无任何购物信息。

[0076] 购物数据库更新单元 602 根据被取走货品的种类及数量以及取走货品的用户的身份信息生成一组购物信息,并存储至该用户的购物数据库,该购物信息中包括此刻被取走货

品的种类及数量，以及该货品的货品信息，如货品名、型号、净含量及单价，等等。用户在封闭空间 1 内多次取走货品之后，其购物数据库内包括多组购物信息，由于用户随身携带的移动通信终端与数据处理设备 7 以无线通信方式连接并进行数据交换，因此，购物数据库中的购物信息也可以显示在用户的移动通信终端的 APP 界面上，形成用户的电子购物车。

[0077] 当货品与一用户匹配时，若基于影像监测的货品感知系统 400 监控到某一物品被放置于该架板上，可以判断出该物品的种类及数量。数据处理设备 7 在该用户的购物数据库中查询每一购物信息，判断是否有已购货品的种类值与该物品种类相匹配，也即判断是否用户的购物数据库中是否有一个或多个已购货品与被放置到架板上的货物相同。

[0078] 如果基于影像监测的货品感知系统 400 还可以判断被放回的货品的种类与该架板上原有货品种类是否一致，如果不一致，可选择地生成一个报警信号，提醒管理人员或用户发生误放现象。如果基于影像监测的货品感知系统 400 无法判断被放回货品种类，即可确认被放回架板的物品并非该无人超市中的既有货品，有可能是用户自带的物品，如雨伞、手机等，此时可选择地生成一个报警信号，必要时，可以将该架板的架板编号显示在某一显示器上，以便提醒管理人员或用户。

[0079] 在其他实施例中，所述无人售货系统还可以包括基于重量监测的货品感知系统，在每一架板上只放置一种货品，在架板内设置重量传感器，实时感知每一种架板的重量变化，与本实施例中基于影像监测的货品感知系统配合工作，可以更精准的判断货品取放状态、被取放货品的种类及数量。

[0080] 如图 10 所示，本实施例还包括结算系统 700，其为数据处理设备 7 中的功能模块，用以根据所述用户的购物数据库中所有货品的种类及数量结算费用。用户购物过程结束后，可以自行从出入口的门禁装置处离开离开封闭空间 1。当用户定位系统 200 的影像传感器 2011 无法获取该用户的实时三维影像时，可以认定该用户购物结束，结算系统 700 为该用户结算费用。

[0081] 结算系统 700 包括总金额计算单元 701 及支付单元 702。当所述用户离开所述封闭空间时，总金额计算单元 701 根据所述用户的购物数据库中全部货品的种类及数量计算总金额，由于每一种类货品的单价作为货品信息预存在数据处理设备 7 中，因此多种货品单价与

数量的乘积的总和的金额即为该用户需要支付的总金额。进一步地，在其他实施例中，用户可以享受到货品折扣或使用优惠券、抵用券等，用户需要支付的总金额为多种货品单价与数量的乘积的总和的金额内减去优惠券和/或抵用券金额和/或折扣金额。支付单元 702 为结算系统 700 自带的支付软件或第三方支付软件，可以从所述用户的银行账户或电子账户上扣款，扣除的款项金额与该用户需要支付的总金额相同。

[0082] 如图 11 所示，本实施例还提供一种基于影像监测的货品感知方法，也就是前述基于影像监测的货品感知系统 400 的实现方法，包括如下步骤：步骤 S201) 样本采集步骤、步骤 S202) 模型训练步骤、步骤 S203) 实时图片采集步骤以及步骤 S204) 货品种类获取步骤。由于本实施例涉及的无人超市中所销售的货品为标准货品，同一种类货品的外观六面图都是相同的，因此在同一种类货品中只要选择一个或几个产品，对其进行多次拍照处理即可完成该类货品训练样本的采样。

[0083] 步骤 S201) 样本采集步骤，用以采集多组图片样本，每一组图片样本包括一种货品在多角度下的多张样本图片；对应同一种类货品的一组图片样本被设有相同的组别标识，该组别标识即为该组图片样本对应的货品的种类。在所述样本采集步骤中，拍摄每一货品多个角度多个距离的图片，优选地，拍摄次数为 3000~5000 次，过多则成本太高，过少则模型误差较大。每种货品拍摄不同角度不同距离的 5000 张图片，有些是独立拍摄货品的图片，有些是有背景的货品图片，有些是货品被某人拿在手中的图片，甚至可以是多个同类货品被叠放在一起后的图片。

[0084] 步骤 S202) 模型训练步骤，用以根据多组图片样本中的每一样本图片及其组别标识训练卷积神经网络模型，获取货品识别模型。

[0085] 如果模型训练的样本较少，或者，样本的清晰度比较低，只训练一次获得的分组模型，在判断图片中显示的货品的种类时可能误差较大，因此最好能有交叉验证步骤。如图 12 所示，步骤 S202) 模型训练步骤包括步骤 S2021) 样本分类步骤、步骤 S2022) 分组模型训练步骤以及步骤 S2023) 交叉验证步骤。

[0086] 步骤 S2021) 样本分类步骤，用以将多组所述图片样本随机分成两类，分别为训练样本和测试样本；在前述的步骤 S201) 样本采集步骤，拍摄每一货品拍摄不同角度不同距

离的 4000 张图片。将每一货品对应的 4000 张图片随机分成两部分，训练样本和测试样本各有 2000 张左右的图片。

[0087] 步骤 S2022) 分组模型训练步骤，用以将多组训练样本的每一样本图片及每一样本图片的组别标识输入至卷积神经网络模型，经训练后获取分组模型。如图 13 所示，步骤 S2022) 分组模型训练步骤包括步骤 S20221) 特征提取步骤、步骤 S20222) 候选区域生成步骤、步骤 S20223) 候选区域映射步骤以及步骤 S20224) 分类器生成步骤。步骤 S20221) 特征提取步骤，用以将每一训练样本的图片输入卷积神经网络 (CNN)，进行特征提取，获取特征图像 (feature map)，对应训练图片中显示全部或部分货品的区域。例如，某一货品的 2000 张图片，在每一张图片上找出与货品整体或局部相关的特征图像 (feature map)。对于彩色图片，每一像素点的 RGB 三原色分别对应一个二维矩阵，每个矩阵经过 3*3 或者 5*5 的卷积核进行卷积运算之后，产生三个新的二维矩阵，也就是特征图像 (feature map)。步骤 S20222) 候选区域生成步骤，用以将每一训练样本的图片输入候选区域网络 (RPN)，生成多个候选区域 (region proposals)，优选地，每张图片生成 300 个候选区域。候选区域网络 (Region Proposal Networks, 简称 RPN) 是 Faster RCNN 模型中的 proposal 生成网络，由于目标检测的目标尺度可能相差很大，因此我们需要尽可能产生不同尺寸的候选区域 region proposals。步骤 S20223) 候选区域映射步骤，先将不同尺寸的候选区域转化为同一尺寸的图片，然后把每一训练样本的图片的候选区域 (region proposals) 映射到卷积神经网络最后一层卷积层的特征图像 (feature map) 上。步骤 S20224) 分类器生成步骤用以收集多个训练样本的图片特征图像 (feature maps) 和候选区域 (region proposals)，计算出候选区域的特征图像 (proposal feature maps)，并送入分类器网络 (classifier 网络)，生成分类器。在该分类器网络中，将对应同一货品的所有训练样本图片的特征图像与该货品的组别标识形成对应关系，该组别标识即为该组图片样本对应的货品的种类。

[0088] 步骤 S2023) 交叉验证步骤，用以根据多组测试样本的每一图片及每一组测试样本的组别标识对所述分组模型进行验证，计算模型准确率。如图 14 所示，步骤 S2023) 交叉验证步骤包括步骤 S20231) 测试样本输入步骤以及步骤 S20232) 准确率计算步骤。步骤 S20231) 测试样本输入步骤用以将多个测试样本的每一图片输入至所述分组模型，获取多个

测试样本的测试组别标识。步骤 S20232) 准确率计算步骤, 用以将多个测试样本的测试组别标识与所述测试样本的组别标识对比, 计算相同标识的数量与所述测试样本的数量的比值, 该比值即为所述初级数据模型的准确率。当所述分组模型的准确率小于一个预先设置的预设阈值时, 返回所述样本分类步骤; 当所述分组模型的准确率大于或等于所述预设阈值时, 所述分组模型为货品识别模型。优选地, 将预设阈值设置为 90%, 当分组模型的准确率小于 90% 时, 返回所述样本分类步骤, 将样本重新分组、重新训练; 当所述分组模型的准确率大于或等于 90% 时, 可以认为所述分组模型即为货品识别模型。

[0089] 如果模型训练的样本够多, 样本的清晰度比较高, 直接训练一次, 就可以利用 Faster RCNN 网络模型训练出一个分组模型, 该分组模型可以有效判断图片中显示的货品的种类。步骤 S202) 模型训练步骤包括步骤 S2022) 分组模型训练步骤即可, 将步骤 S201) 样本采集步骤中采集的多组图片样本的全部或部分作为训练样本, 将每一训练图片及其组别标识输入至卷积神经网络模型, 经训练后获取分组模型, 所述分组模型即为货品识别模型。

[0090] 步骤 S203) 实时图片采集步骤, 用以连续采集至少一实时图片, 每一实时图片包括一货品影像的一部分或全部。步骤 S203) 实时图片采集步骤包括第二图片采集步骤, 拍摄每一货品多个图片, 拍摄次数为 10~200 次。所述货架的四个角落处分别设置有一个第二摄像头 206, 每一第二摄像头 206 的视野范围覆盖所述货架前方的空间区域, 每一个第二摄像头 206 的镜头朝向所述货架平面的中心区域。当用户伸手将某一货品从货架上取下或者将某一货品放置会货架上时, 四个第二摄像头 206 会从不同角度拍到该货品整体或局部图片。假设用户手持货品在货架前方停留 1 秒钟, 四个第二摄像头 206 可以在 1 秒钟内获取不同角度的 120 张实时图片, 每一图片的显示内容可以包括用户的手部及货品的局部。第二摄像头 206 可以为长时间通电或启动的设备, 也可以在货架板上设置一个红外开关, 连接至第二摄像头 206, 平时为断电状态, 当有用户行进至货架附近时, 红外开关感应到热量并连通, 第二摄像头 206 通电启动; 当用户离开时, 红外开关感应到不到热量从而断开, 第二摄像头 206 断电关闭。

[0091] 步骤 S204) 货品种类获取步骤, 用以根据所述实时图片及所述货品识别模型判断所述实时图片中显示的货品的种类。如图 15 所示, 步骤 S204) 种类判断步骤还包括步骤 S2041)

组别标识获取步骤以及步骤 S2042) 标识可信度计算步骤。步骤 S2041) 组别标识获取步骤, 用以将多个实时图片输入至所述货品识别模型, 获取多个实时图片所对应的多个组别标识; 步骤 S2042) 标识可信度计算步骤, 计算每一种组别标识的数量与全部组别标识总数的比值, 该比值即为每一种组别标识的可信度, 可信度最大的组别标识所对应的货品的种类即为所述实时图片上显示的货品的种类。步骤 S204) 执行完毕后, 返回步骤 S203) 实时图片采集步骤, 以待下一次有货品被取走或被放回时, 及时拍摄到货品图片。

[0092] 如图 16 所示, 步骤 S2041) 组别标识获取步骤包括: 步骤 S20411) 第二特征提取步骤、步骤 S20412) 第二候选区域生成步骤、步骤 S20413) 第二候选区域映射步骤以及步骤 S20414) 组别获取步骤。步骤 S20411) 第二特征提取步骤用以将一实时图片输入卷积神经网络 (CNN), 进行特征提取, 获取特征图像 feature map。步骤 S20412) 第二候选区域生成步骤用以将所述实时图片输入候选区域网络 (RPN), 生成多个候选区域 (region proposals), 每张实时图片生成 300 个候选区域。步骤 S20413) 第二候选区域映射步骤用以把所述实时图片的候选区域 (region proposals) 映射到卷积神经网络最后一层的特征图像 (feature map) 上。步骤 S20414) 组别获取步骤用以收集多个图片的特征图像 (feature maps) 和候选区域 (region proposals), 计算出候选区域的特征图像 (proposal feature maps), 并送入分类器网络 (classifier 网络), 获取该实时图片所对应的组别标识。

[0093] 步骤 S2042) 标识可信度计算步骤, 用以计算所述可能性结论中所述可能性结论中每一种组别标识的数量与所述可能性结论中全部组别标识总数的比值, 该比值即为每一种组别标识的可信度, 可信度最大的组别标识所对应的货品的种类和/或数量即为所述实时图片上显示的货品的种类和/或数量。对比上述的 60 个组别标识, 如果与 60 个组别标识相应的 60 个货品种类中, 出现货品 A 的次数为 36, 出现货品 B 的次数为 18, 出现货品 C 的次数为 6, 则三者的可信度分别为 60%、30% 及 10%, 可以认定所述实时图片上所显示的货品的种类为可信度最高的货品 A。如果步骤 S201) 样本采集步骤中采集的图片样本中包括多个同类货品叠放在一起的照片, 步骤 S204) 货品种类获取步骤还可以进一步判断货品的数量。

[0094] 前文所述的基于影像监测的货品感知系统及货品感知方法, 可以根据货架前方空间的实时图片判断出货架上货品的取放状态、被取走或被放回的货品的具体种类, 若结合基于

重量监测的货品感知系统可以准确判断货品数量，若结合用户身份识别技术和用户定位跟踪技术，还可以准确判断被取走或被放回的货品的用户身份，从而在该用户购物数据库中准确添加或删除购物记录，以使用户购物结束后自动结算。

[0095] 以上所述仅是本发明的优选实施方式，使本领域的技术人员更清楚地理解如何实践本发明，这些实施方案并不是限制本发明的范围。对于本技术领域的普通技术人员，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本发明的保护范围。