arXiv:2601.00928v1 [cs.CV] 2 Jan 2026

分析购物旅程：计算实体零售店中的货架浏览访问

Luis Yoichi Morales[1],Francesco Zanlungo[2],和 David M. Woollard[1]

[1]Standard AI，旧金山，加利福尼亚州，
美国[2]意大利巴勒莫大学

2026年1月6日

摘要

受近期将机器人部署到零售领域面向顾客岗位所面临挑战的启发，本文对实体店中的顾客活动展开研究，以此迈向对购物者意图的自主理解。我们介绍了一种算法，用于计算购物者的“货架访问”，从而捕捉他们在店内的浏览行为。货架访问是根据通过基于机器视觉的三维跟踪和俯视摄像机获取的轨迹提取出来的。我们使用不同的轨迹集（分别包含 8138 条和 15129 条轨迹），对货架访问算法进行了两次独立校准；这些轨迹采集自不同的店铺，并由人工审核员标注。随后，我们使用未参与校准过程的轨迹评估校准后的模型，这些轨迹既来自实施校准的同一家店铺，也来自另一家店铺。结果分析表明，在不同于校准所用环境的环境中进行评估时，该算法仍能识别顾客的浏览活动。接着，我们使用该模型分析大量轨迹中的顾客“浏览模式”，以及这些模式与店内实际购买之间的关系。最后，我们讨论如何将货架浏览信息用于零售规划以及人机交互场景。

1 引言

尽管机器人有望提升顾客满意度、改善门店绩效并增加销售额，但在面向顾客的岗位引入机器人时，主要限制因素之一是难以赋予其自主行为能力 [1]。尽管人工智能有望彻底改变机器人实现开放世界理解的能力，但普遍共识是，这场变革还很遥远，零售商应采取谨慎态度[2]。软银从 Pepper 转向 Whiz，便体现了当前零售领域面向顾客的服务机器人的局限：Pepper 是一款旨在直接与顾客互动的客户服务机器人，但成效有限；而 Whiz 则是专为清洁等非面向顾客的岗位设计的机器人 [1, 3]。

功能缺失的一个关键方面，是机器人能否迅速判断购物者意图，从而在无需大量协助的情况下直接满足顾客的紧迫需求。虽然购物者意图可以被视为一个复杂的开放世界问题，但我们可以借鉴电子商务场景中购物者行为（及其测量方式）的相关研究，探索如何更好地理解购物者意图。

关于“浏览活动”的数据对于电子商务的分析和建模至关重要，但在实体店的绩效分析中却一直被忽视。不过，实体店的顾客确实会进行与线上浏览类似的行为。例如，购物者在货架前停留的时间可长可短，会查看不同的产品、拿起产品，之后只购买其中一部分（如果有购买的话）。显然，这种“浏览活动”与顾客的购买行为相关，因此对于门店管理层和包装消费品（CPG）制造商而言都是极具价值的信息，他们

不仅关注自身产品的销售表现，也关心购物者在店内如何看待并与产品互动。

大量实体店优化研究聚焦于更易获取的数据，包括分析销售额 [4], ，以及构建模型来优化店铺布局 [5, 6] 和产品摆放 [7, 8]，从而提高收入。

尽管做出了这些努力，实体零售分析中仍有一个明显的缺口，那就是对销售转化漏斗的衡量（即购物者进入商店到完成购买之间发生了什么）。我们认为，了解购物者在实体店中的浏览活动，有助于弥合机器人自主性方面的差距，为购物者提供更有针对性、更有意义的支持，从而改善结果，既能提升顾客支持，也能增加销售额。

本文介绍了一种算法，用于分析购物者在实体店中的移动轨迹，并提取其“货架访问”记录——即顾客在货架前停留的时间、地点及持续时长。我们利用该算法分析顾客的“浏览模式”及其与“购买模式”的关系，并探讨可能的应用、局限性和未来工作。

2 相关工作

关于实体零售店购物者行为分析的相关工作可分为两大类：给定购物清单估计购物者路径，以及零售货架陈列优化。

先前的研究提出了多种方法，用于根据购物清单估计购物者在店内的行进路径及该路径对购买行为的影响。例如， [9] 提出通过计算货架端架上的平均冲动购买量来估计利润，而 [10] 则根据货架的位置估算其客流量，以计算冲动消费。冲动消费尤其值得研究，因为计划外购买会对店铺收入产生重大影响 [11], ，因此，优化促销商品陈列 [12] 是整体门店绩效的关键因素。

近年来，零售业已认识到基于现有零售数据进行货架布局优化的技术所带来的益处 [13], ，并且多项研究着重探讨了在实践中应用货架空间分配模型 [8]。此外，采用货架空间优化的数值研究表明，交叉空间弹性对零售利润的影响有限 [14]。文献 [15]对零售货架规划问题进行了前沿综述。先前研究提出了能够提升店铺收入的布局设计 [16] ，并在文献 [17]中探讨了货架布局与营销效果及销售额之间的关系。机器人技术领域也提出了一些面向零售的应用。文献 [18] 介绍了用于零售物流的数字孪生模型研究，文献 [19]则介绍了用于店铺建模的机器人数据采集。

据我们所知，以往的研究既未提出模型，也未分析货架浏览活动，而这正是本文的重点。

3 数据收集与处理

3.1 目标门店

本研究使用了两家便利店的数据（见图 1）。门店 $s1$的区域面积为 87.39 m^2 ，设有 $n_{s,1} = 19$ 个货架和 $n_{e,1} = 2$ 个出入口。门店 $s2$ 的区域面积为 109.16 m^2 ，设有 $n_{s,2} = 50$个货架和 $n_{e,2} = 2$ 个出入口。

两家门店的所有入口也都可用作出口，且我们的轨迹分析系统会记录顾客是否以及何时将这些入口用作入口或出口（即，在轨迹跟踪开始时，若顾客经过同一入口/出口几何区域，则记录为经过“入口”；在轨迹跟踪结束时，则记录为经过“出口”）。

3.2 商店地图

我们构建并使用商店的高密度彩色三维点云地图。该地图用于布置摄像机、计算其覆盖范围 [20] 并提取商店布局。商店布局包含周界、入口、出口和货架。货架表示为一组三维顶点和一个法向量

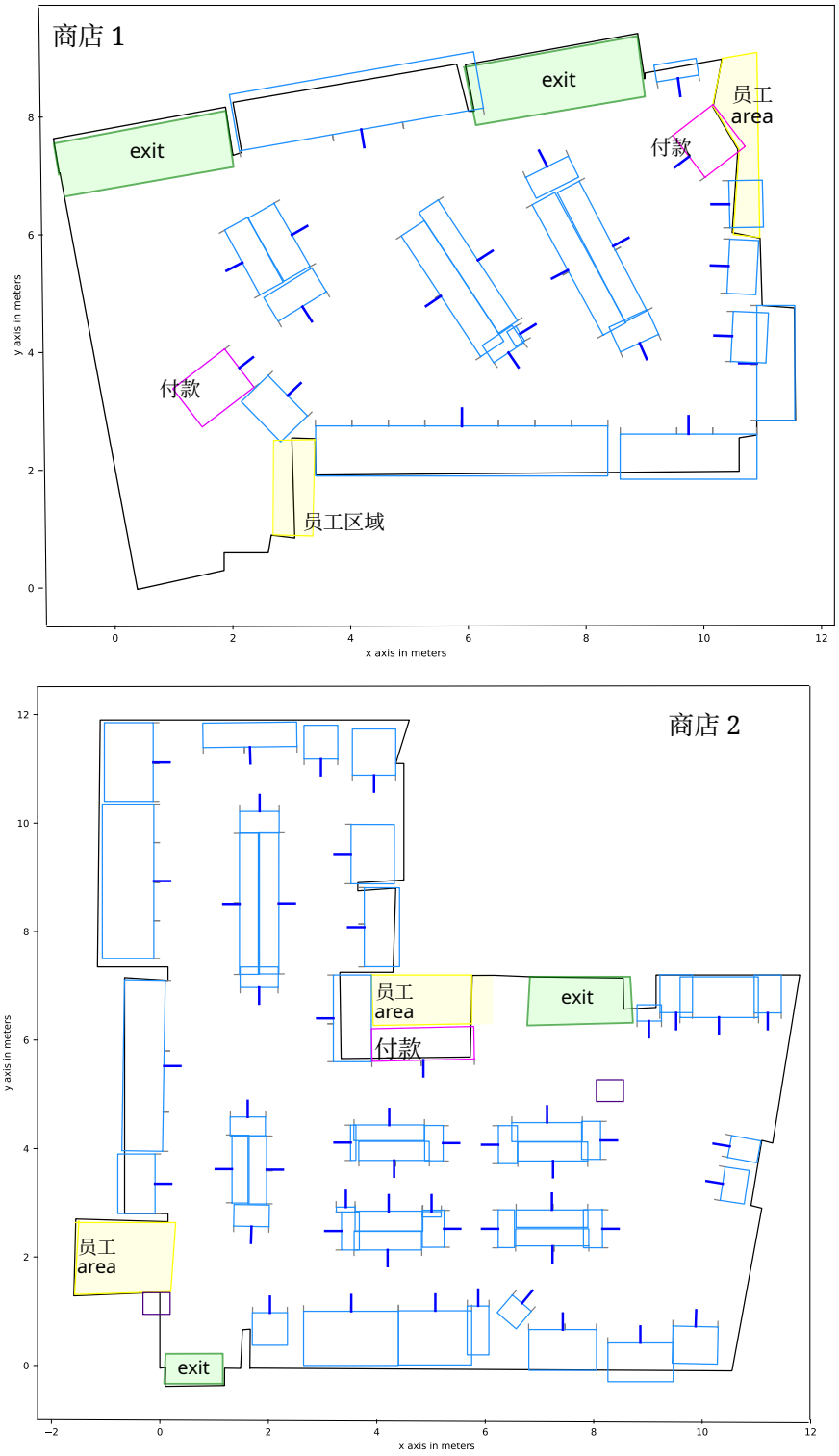

图 1：上方为商店 $s1$ 的布局，下方为商店 $s2$ 的布局。蓝色多边形表示货架，其可交互面以蓝线标示，蓝线代表面法线。出口以绿色标示，仅限员工进入区域以黄色标示。

表示可交互式货架面（即陈列产品的一面）。每个货架只有一个可交互面。本文提出的算法仅考虑货架可交互面的二维地面投影，以及可能遮挡顾客视线的障碍物的二维地面投影，例如货架的非交互面。我们将这些二维投影视为线段，并将对应于第 j 个货架的线段记为 $\mathbf{s}j$，其中 $j \in 1, \ldots, n_s$；将对应于第 k 个障碍物的线段记为 $\mathbf{s}^{n_s+k}$。

3.3 人员追踪系统

我们使用 Standard AI 的视觉机器学习平台追踪购物者移动。每个二维姿态检测模型都以 10 帧/秒的速度运行于一组经过校准的摄像机上，这些摄像机安装在商店天花板上（类似于典型的监控摄像机设置）；随后，根据来自各个摄像机的二维姿态进行三角测量，合成为三维姿态。我们使用颈部关键点追踪人群的质心。

3.4 数据处理

我们将轨迹定义为一组按时间排序的数组，其中包含一个人从商店入口到出口的位置和身体朝向。我们的追踪系统的采样率为 $10Hz$。在计算速度之前，我们会按照下文所述对轨迹点进行低通滤波。

尽管我们的追踪系统提供了更多数据（例如加速度、手部关键点等），这些数据会被存储并用于其他目的和未来的开发，但所提出的算法仅使用以下信息。

二维向量

$$\mathbf{x}_i(t_k), \tag{1}$$

表示 顾客在第 k 个时间戳时的（3D）质心在商店地面上的二维投影 i ，

$$t_k = k\Delta_t, \tag{2}$$

$\Delta_t = 0.1$ 表示追踪时间步长，而角度

$$\theta_i(t_k) \tag{3}$$

通过单位法向量表示顾客的 i 身体朝向（连接肩部的直线在二维平面上的投影的前向法向量；肩部的位置和身体的前向方向由前述三维跟踪系统识别）

$$\mathbf{n}_i(t_k) = (\cos(\theta_i(t_k), \sin(\theta_i(t_k)). \tag{4}$$

速度定义为

$$\mathbf{v}_i(t_k) = \frac{\mathbf{x}_i(t_{k+1}) - \mathbf{x}_i(t_{k-1})}{2\Delta_t}. \tag{5}$$

尽管 尽管这是一个二维向量，但在下文的算法中，我们只使用它的范数（速度） $v_i(t_k)$。

4 货架停留算法

本节将介绍我们用于判断一条轨迹中是否发生货架停留的启发式算法。该算法相对简单，因此只需使用一组相对较小的轨迹进行校准，并且也能适用于较简单的跟踪系统（任何能够提供式（1、3）中的二维向量 $\mathbf{x}$ 和角度 θ 的跟踪系统均可）。

核心思想是，货架前的“停留” j 是顾客轨迹的一部分，并满足以下条件，这些条件由 3 个参数定义：T_B （最小浏览时间）、 Δ_B （到货架的最大距离）和 v_B （最大浏览速度）。

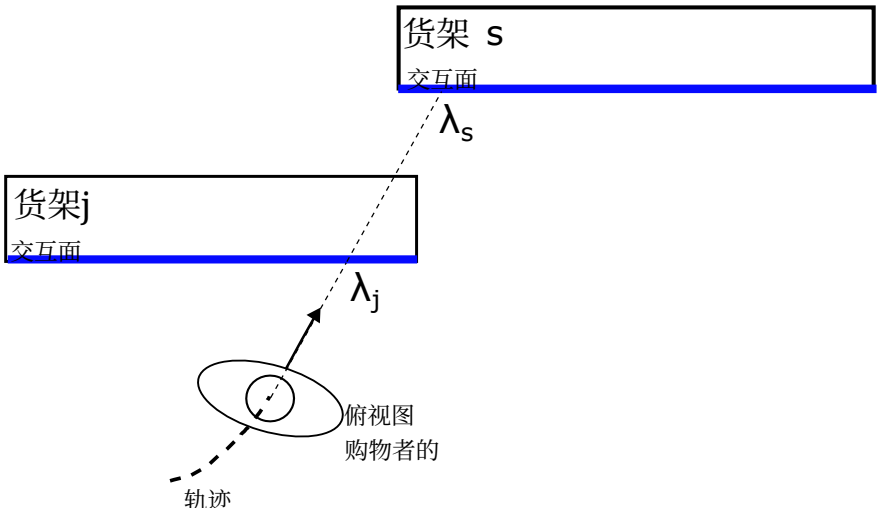

图 2：可能被识别为正在浏览货架 j 的购物者的俯视图。 λ_j 是到最近货架的距离，而 λ_s 是朝向更远货架 s 的距离。

更具体地说，对于每位顾客 i，我们按照以下方式确定一个候选货架（或不确定候选货架）。首先，我们找出货架面或障碍物 $\mathbf{s}_k$ 与由顾客方向向量的正倍数定义的半直线之间的所有交点，

$$\lambda \mathbf{n}_i. \tag{6}$$

然后，我们根据 λ 在交点处所取的值，将每个此类交点与货架或障碍物 l 进行对应，

$$\lambda_l > 0, \tag{7}$$

并查找对应于最小距离的线段，

$$j = \text{argmin}_l \lambda_l. \tag{8}$$

如果最小距离对应于交互式货架面，

$$j \in 1, \ldots, n_s, \tag{9}$$

则将该货架识别为候选货架（否则，不存在候选货架）。

T此计算还为我们提供了到货架的距离，其定义很简单： λ $_j$.

停留定义为一个时间区间

$$t_k \in [t_s, t_f],\ t_f - t_s \geq T_B \tag{10}$$

在此期间，对所有 k

- 候选货架 j 均保持不变，
- 到货架的距离满足

$$\lambda_j(t_k) \leq \Delta_B, \tag{11}$$

- 且顾客速度小于某个阈值

$$v_i(t_k) \leq v_B. \tag{12}$$

对于每位顾客 i、货架 j 和时间 k，算法都会生成一个布尔输出

$$S_j^i(t_k) \tag{13}$$

假设上述条件满足时取值为 1，否则取值为 0。

5 算法参数校准

这 3 个参数（T_B、 Δ_B 和 v_B）通过基于人工标注的校准过程进行优化。

我们构建了两个校准集，第一个包含来自 $s1$ 的 $n_1 = 279$ 条轨迹，第二个包含来自 $s2$ 的 $n_2 = 270$ 条轨迹。对于每条轨迹，我们生成了一段 $2D$ 视频，其中包含位置、速度和方向信息，随后由人工审核员据此识别并编码货架浏览行为。

$n_l = 4$ 人工审核员分析了完整的轨迹集，并根据他们对浏览行为的理解，识别出顾客 i 在货架 j 前停留的时间。根据这些标注，顾客 i 的每个时间戳 k 都会通过投票系统获得一个指向货架 j 的“访问布尔值”。

具体而言，将 $n_j^i(t_k)$ 定义为在时间 k 识别出 i 在 j 前停留的审核员人数，则访问布尔值为

$$V_j^i(t_k) = \begin{cases} 1, \text{ if } n_j^i(t_k) > n_l/2, \\ 0, \text{ if } n_j^i(t_k) \leq n_l/2. \end{cases} \tag{14}$$

随后，我们可以将真正例的数量定义为

$$TP = \#\left(k : S_j^i(t_k) = V_j^i(t_k) = 1\right), \tag{15}$$

其中， $S_j^i(t_k)$ 是算法输出中的值（公式 13）；假正例的数量定义为

$$FP = \#\left(k : S_j^i(t_k) \neq V_j^i(t_k) = 0\right), \tag{16}$$

以及假负例的数量为

$$FN = \#\left(k : S_j^i(t_k) \neq V_j^i(t_k) = 1\right). \tag{17}$$

我们将精确率定义为

$$P = \frac{TP}{TP + FP}, \tag{18}$$

将召回率定义为

$$R = \frac{TP}{TP + FN}. \tag{19}$$

校准 过程基于寻找使 F_1 得分最大化的参数值，其定义为 s

$$F_1 = \frac{2PR}{P + R}. \tag{20}$$

我们进行了两次独立校准。基于 $s1$ 校准集的校准过程得出了以下参数值，对应的最大值为 $F_1 \approx 0.86$: $T_I = 2$s、$\Delta_I \approx 1.2$ m、 $v_I \approx 0.55$ m/s。

另一方面，在 $s2$ 校准测试中获得了 $F_1 \approx 0.89$ ，对应的参数值为: $T_I = 1.7$ s、 $\Delta_I \approx 1.55$ m、 $v_I \approx 0.48$ m/s。

Δ_I 在 $s2$上数值更大，这很可能是由于店铺几何结构不同。

6 算法评估

我们通过随机选取比例为 p 的轨迹进行同店评估，并用这些轨迹校准模型；随后使用优化后的参数在剩余轨迹上运行算法，并计算相应的 F_1 分数。关于 $s2$ 的结果见图 3。

跨店评估再次通过随机选取比例为 p 的轨迹并用其校准模型来进行。随后，在另一家店的完整轨迹集上评估校准后的模型。

在 $s1$ 上校准并在 $s2$ 上评估所得的 F_1 分数为 0。84；在 $s2$ 上校准并在 $s1$ 上评估所得的分数为 0。87。

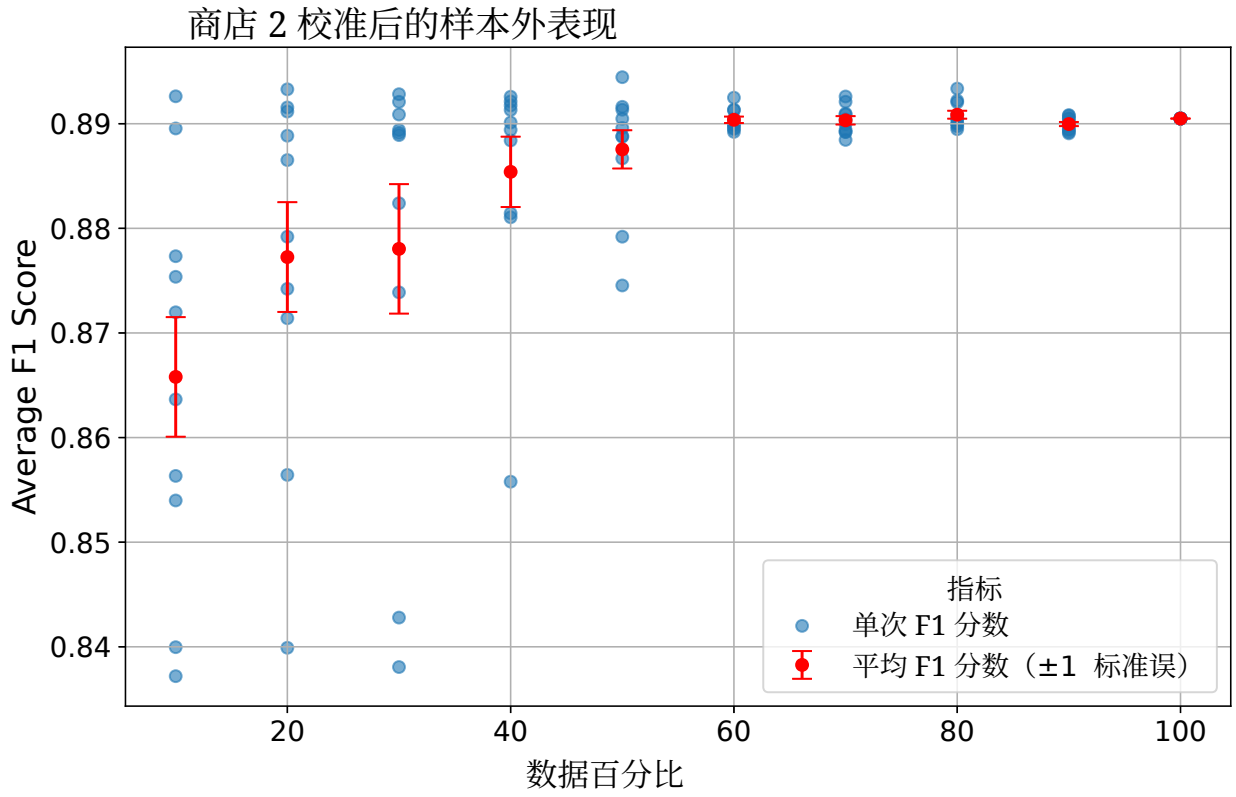

图 3：同一商店的 F_1 分数模型评估。 x 轴表示用于校准的数据百分比， y 轴表示在其余轨迹上的评估分数。蓝点对应于校准集的 10 次随机选择，红点表示平均值，红色误差线表示标准误。

7 浏览行为分析

我们对浏览行为进行了两种分析：一种使用 $s1$ 和 $s2$ 可获得的最大轨迹集，另一种使用规模较小的 $s1$ 轨迹集，这些轨迹集提供了购买活动数据。

7.1 大型数据集分析， $s1$

我们分析了 $s1$ 的 $N_1 = 8138$ 条轨迹，其中有 $n_{s1} = 19$ 个货架可供使用。对于每条轨迹，我们构建一个二进制 n_{s1} 维向量；如果购物者访问了相应的货架，则对应条目为 1，否则为 0。对所有轨迹取平均后，我们得到每次行程的平均访问次数为 2.54。货架访问次数见图 4。

7.2 大型数据集分析 $s2$

我们采用相同的方法分析了 $s2$ 的 $N_2 = 15129$ 条轨迹，其中有 $n_{s2} = 50$ 个货架可供使用，得到的每次行程平均访问次数为 2.82。货架访问次数见图 5。

7.3 与 $s1$ 上购买活动的比较

对于 $s1$ 的 473 条轨迹子集，我们已将这些轨迹与购物者的购买活动进行匹配（数据由零售商通过其销售点系统提供的交易日志）。因此，我们可以将根据这些轨迹构建的平均访问向量与相应的每个货架的平均购买次数进行比较。将每个货架的平均购买量除以每个货架的平均访问量，即可得到一个向量

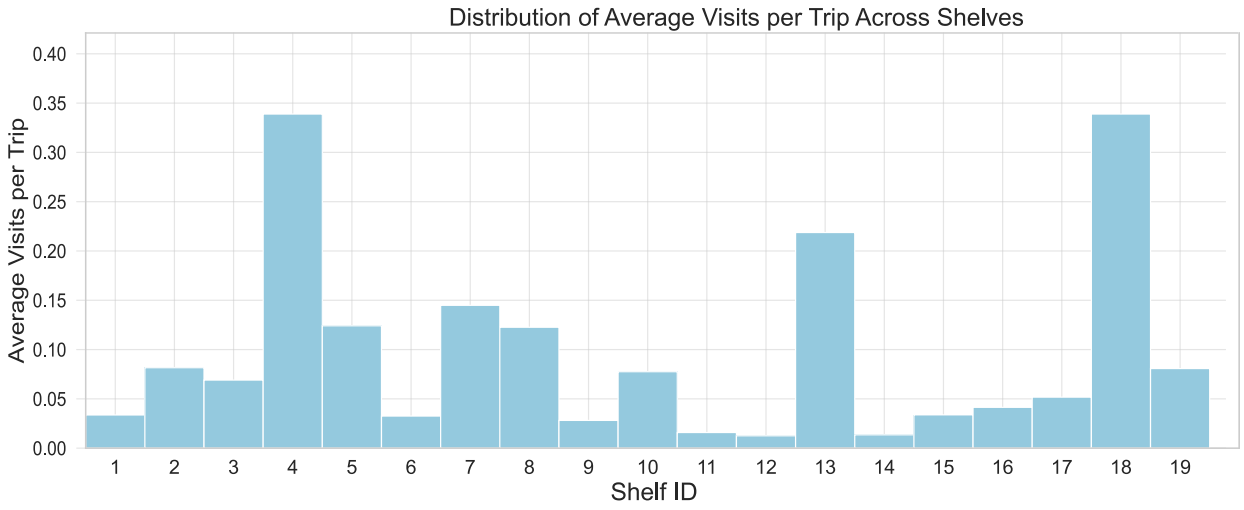

图4：门店 $s1$ 中各货架每次行程平均访问次数的分布。

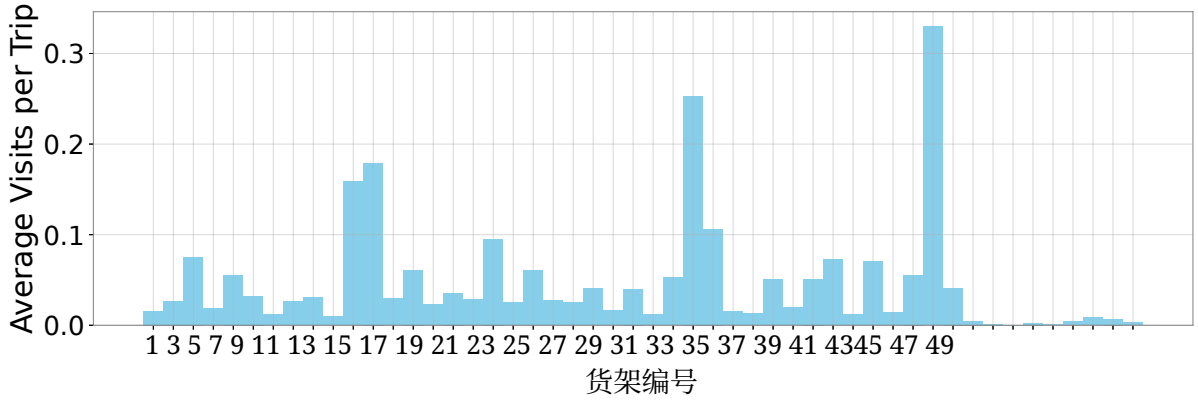

图5：门店 $s2$ 中各货架每次行程平均访问次数的分布。

其中包含访问/购买转化率，如图8所示。尽管这只是非常初步的分析，但它展现了我们的方法在揭示对不同货架的访问如何影响购买，以及最终为了解整体门店绩效提供洞见方面的潜力。

8 结论与未来工作

本文介绍了一种根据实体零售店中购物者的轨迹计算货架访问次数的方法。文中讨论了如何使用人工标注数据校准模型，并展示了使用两组不同轨迹对该方法进行评估的结果。校准后的模型用于分析浏览模式，以及这些模式与购物者购买行为之间的关系。尽管本研究是在便利店中开展的，但同样的方法也适用于其他业态。

未来的研究将进一步完善模型，以检测更复杂的购物者行为；引入机器学习和其他技术，提高浏览检测的准确性；并通过分析注视方向和头部朝向，更有针对性地了解购物者在浏览货架时关注了哪些产品。通过深入了解购物者的意图，我们可以让各种干预措施更有针对性、更有意义，包括与服务机器人互动；这些机器人的任务是为顾客提供有用且

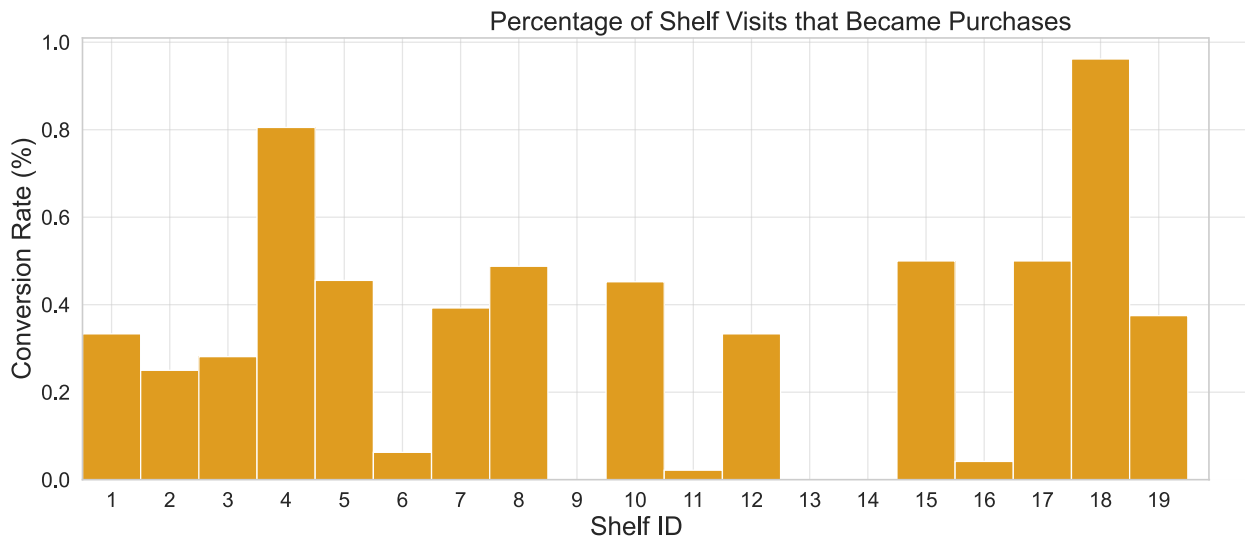

图 6：包含购买信息的店铺 s_1 轨迹子集中转化为购买的访问比例。

实用的帮助 [21]。

此外，通过了解店内购物模式，机器人系统还可以根据购物者的轨迹和先前的货架浏览历史，向购物者提供实时推荐并推广不同的产品类别。与电子商务中的点击历史类似，了解购物者在实体店中的意图，是实现产品推荐个性化的重要因素。尽管这项工作预设了服务机器人的使用场景，但也可以通过移动应用，甚至店内媒体显示屏，向购物者提供这些相同的功能（购物者协助和个性化推荐）。

最后，我们可以深入了解不同布局可能如何影响购物者与产品的互动。这一能力将使我们能够以数据为依据优化店铺布局，或将彻底改变零售商对产品摆放和货架组织的看法，从而最大限度地促进顾客互动，并最终增加收入。

参考文献

[1] A. Guha 和 D. Grewal，《机器人将如何影响零售业的未来》，*AMS Review*，第 12 卷，第 3 期，第 245–252 页，2022 年。[2] A. Guha、D. Grewal、P. K. Kopalle、M. Haenlein、M. J. Schneider、H. Jung、R. Moustafa、D. R. Hegde 和 G. Hawkins，《人工智能将如何影响零售业的未来》，*Journal of Retailing*，第 97 卷，第 1 期，第 28–41 页，2021 年。[3] A. Rindfleisch、N. Fukawa 和 N. Onzo，《零售业中的机器人：推出 Whiz》，*AMS Review*，第 12 卷，第 3 期，第 238–244 页，2022 年。[4] M. Pesendorfer，《零售销售额：超市定价行为研究》，*The Journal of Business*，第 75 卷，第 1 期，第 33–66 页，2002 年。[5] W. Ke，《商品摆放优化》，载于*Retail Space Analytics*，丛书 International Series in Operations Research & Management Science，A. Ghoniem 和 B. Maddah 编。Springer，2023年12月，第 129–159 页。

[6] E. Danisman 和 A. E. Smith，《数据驱动的杂货店分析设计》，载于*Retail Space Analytics*，丛书 International Series in Operations Research & Management Science，A. Ghoniem 和 B. Maddah 编。Springer，2023年12月，第 75–101 页。

[7] T. D¨usterh¨oft 和 A. H¨ubner，“零售空间规划中应用优化模型的问题与机遇”，载于*Retail Space Analytics*, International Series in Operations Research & Management Science 系列，A. Ghoniem 和 B. Maddah 编。Springer，2023年12月，第161–181页。

[8] O. Pak、M. Ferguson、O. Perdikaki 和 S.-M. Wu，“杂货零售商促销展示空间的库存单位选择优化”，载于*Retail Space Analytics*，International Series in Operations Research & Management Science 系列，A. Ghoniem 和 B. Maddah 编。Springer，2023年12月，第103–128页。

[9] A. R. Botsalı、G.-A. Klutke 和 B. A. Peters，“顾客行走行为对零售设施中网格布局与货架空间分配的影响”，载于*Retail Space Analytics*，International Series in Operations Research & Management Science 系列，A. Ghoniem 和 B. Maddah 编。Springer，2023年12月，第1–20页。

[10] T. Flamand、A.Ghoniem 和 B. Maddah，《考虑客流连锁效应的全店货架空间分配》，运筹学，第 71 卷，2023 年 3 月。

[11]R. Abratt 和 S. D. Goodey，《超市中的非计划性购买与店内刺激》，管理与决策经济学，第 11 卷，第 111–121 页，1990 年。 [在线]。链接：https: //api.semanticscholar.org/CorpusID:154448527[12] O. Pak、M. Ferguson、O. Perdikaki 和 S. Wu，《优化杂货零售商促销陈列空间的库存单位选择》，《运营管理杂志》，第 66 卷，2019 年 12 月。[13] T. Bianchi-Aguiar、A. H¨ubner、M. A. Carravilla 和 J. F. Oliveira，《零售货架空间规划问题：全面的综述与分类框架》，《欧洲运筹学杂志》，第 289 卷，第 1 期，第 1–16 页，2021 年。[14]K. Schaal 和 A. H¨ubner，《跨空间弹性何时会影响货架空间规划？一种决策分析方法》，《*Omega*》，第 80 卷，第 135–152 页，2018 年。[在线]。链接：https://www.sciencedirect.com/science/article/pii/S030504831630785X[15] T. Bianchi-Aguiar、A. H¨ubner、M. A. Carravilla 和 J. F. Oliveira，《零售货架空间规划问题：全面的综述与分类框架》，《欧洲运筹学杂志》，第 289 卷，第 1 期，第 1–16 页，2021 年。[16]E. Ozgormus 和 A. E. Smith，《杂货店区块布局的数据驱动方法》，*Computers I&Industrial Engineering*，第 139 卷，文章编号 105562，2020 年。 [在线]。链接：https: //www.sciencedirect.com/science/article/pii/S0360835218306144[17] E. Van Nierop、D. Fok 和 P. Franses，《货架布局与营销效果之间的相互作用及其对优化货架排列的影响》，《营销科学》，第 27 卷，第 1065–1082 页，2008 年 11 月。[18] M. K¨umpel、C. A. Mueller 和 M. Beetz，《零售物流的语义数字孪生》，载于物流动态：不来梅（德国）跨学科物流研究二十五年。施普林格国际出版公司，查姆，2021 年，第 129–153 页。[19] M. Beetz、S. Stelter、D. Beßler、K. Dhanabalachandran、M. Neumann、P. Mania 和 A. Haidu，《style id='36'>机器人采集数据：商店建模。查姆：施普林格国际出版公司，2022 年，第 41–64 页。[20] A. Malhotra、D. Singh、T. Dadlani 和 L. Y. Morales，《通过三维摄像机投影优化重叠覆盖的摄像机布置》，载于*2022*年国际机器人与自动化会议 *(ICRA)*，2022 年，第 5002–5009 页。

[21]M. Niemel¨a、P. Heikkil¨a 和 H. Lammi，《购物中心中的社会服务机器人：管理层、零售商和消费者的期望》，*2017* 年 *ACM/IEEE* 国际人机交互会议附属论文集，2017 年。 [在线]。网址：https://api.semanticscholar.org/CorpusID:802313