# *CPA*: Cyber-Physical Augmentation for Vibration Sensing in Autonomous Retails

Yue Zhang
yzhang58@ucmerced.edu
University of California, Merced
Merced, CA, USA

Carlos Ruiz
carlos@aifi.com
AiFi Inc.
Burlingame, CA, USA

Shubham Rohal
srohal@ucmerced.edu
University of California, Merced
Merced, CA, USA

Shijia Pan
span24@ucmerced.edu
University of California, Merced
Merced, CA, USA

## ABSTRACT

Customer-product interaction (pickup or put down) detection is essential for autonomous retail. Previous studies have explored many sensing technologies, including vision, load, RFID, and piezo. These approaches often require strict deployment conditions, such as Line-of-Sight (LoS), dense deployment, and/or high cost.

In this paper, we present a vibration-based customer-product interaction detection system for autonomous retail. We propose to use one vibration sensor placed on the back panel of the gondola to achieve single-point sensing to monitor all the shelves on the gondola. The key challenge is to detect pickup of small items (e.g., candy bar), which induces extremely low SNR vibration signals. We introduce *CPA*, a cyber-physical augmentation scheme to enhance the vibration sensing signal via a physical arc structure and achieves high accuracy labeling-free event detection. We evaluate *CPA* in a real-world scenario by deploying a vibration sensor on a retail gondola to monitor interactions across multiple shelves. Our approach demonstrates up to 2.9× improvement for light-weight product pickup event detection.

## CCS CONCEPTS

• **Computer systems organization** → **Sensors and actuators**.

## KEYWORDS

Sensing Augmentation; Autonomous Retail; Vibration Sensing; Event Detection; Low Signal-to-Noise Ratio

**ACM Reference Format:**
Yue Zhang, Carlos Ruiz, Shubham Rohal, and Shijia Pan. 2023. *CPA*: Cyber-Physical Augmentation for Vibration Sensing in Autonomous Retails. In *The 24th International Workshop on Mobile Computing Systems and Applications (HotMobile '23), February 22–23, 2023, Newport Beach, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3572864.3580328

## 1 INTRODUCTION

Customer-product interaction detection is essential for autonomous retail [16]. Various sensing methods have been proposed, including vision [16, 18], RFID [13, 15], load [16], and pressure (piezo) [5] based approaches. They have various limitations in real deployments. Vision-based methods require line-of-sight (LoS) and might not work if the human body or shelf blocks the view. RFID-based methods often require assigning a tag for each product item, which causes a continuous cost of labor and tags. Load-based and pressure-based methods usually require a dense deployment, which causes a cost increase. To address these limitations, we present a vibration-based single-point sensing solution to allow low-cost monitoring without the requirement of gondola retrofit.

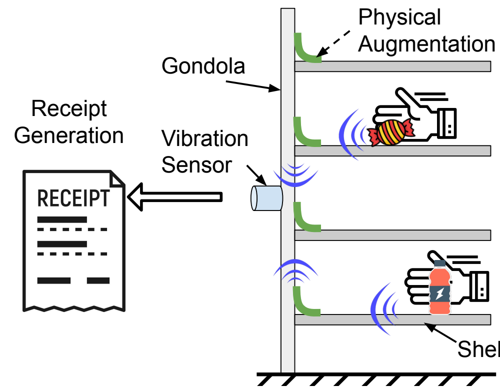


**Figure 1: Illustration of the vibration-based single-point sensing for customer-product interaction detection. The pickup and/or put down events generate vibrations that propagate through the gondola structure and are captured by the vibration sensor on the back panel of the gondola.**

Structural vibration-based sensing methods are emerging and allow various smart building applications with sparse deployments [11, 21]. They can be easily applied to existing infrastructures, such as the retail gondola[1], as shown in Figure 1. First, we deploy the vibration sensor on the back panel of the gondola to monitor the vibration of all shelves. The customer-product interaction-induced vibration propagates through the gondola structure and is captured by the sensor. Then, we can infer the interaction information via the vibration signal and enable autonomous retail and/or inventory monitoring. However, this is challenging for traditional vibration-based human sensing systems because of the **extremely low Signal-to-Noise Ratio (SNR) signals** and/**or expensive cost of labeling**. For example, when we use the traditional vibration-based sensing system to detect events of picking up an energy bar (60 grams), we could only achieve a 5% detection rate (more details in Section 4.2). On the other hand, manually labeling the event from

[1]A gondola is a freestanding fixture used by retailers, and its shelves are mostly used to display merchandise [1].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
*HotMobile '23, February 22–23, 2023, Newport Beach, CA, USA*
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0017-0/23/02.
https://doi.org/10.1145/3572864.3580328

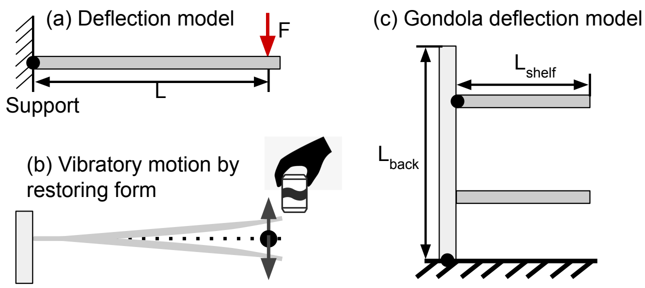


**Figure 2: Deflection models. (a) Cantilever end-load deflection model described with the parameter of force $F$ and force-fixed-end distance $L$. (b) Pick-up event induces vibration by restoring deflection caused by product load. (c) Deflection models applied on the gondola – individual shelf and back wall.**

the video is impractical for a retail store that contains hundreds and thousands types of items [2].

In this paper, we present *CPA*, a cyber-physical augmentation scheme that enhances the traditional vibration sensing system to allow low-cost single-point sensing on each gondola. With one vibration sensor on the gondola's back panel, we are able to monitor customer-product interaction on multiple shelves. *CPA* handles the aforementioned challenge via 1) physical augmentation with a physical arc between each shelf and the back panel to enhance the vibration signal on the back panel; and 2) cyber augmentation with a combination of pseudo labeling and binary classifier to eliminate the need for manual labeling. The **contributions** of this paper can be summarized as follows:

- We introduce a vibration-based single-point sensing system for customer-product interaction detection.
- We present *CPA*, a cyber-physical augmentation scheme to enable accurate interaction detection with extremely low SNR signals and without manual labeling.
- We evaluate *CPA* through real-world experiments on a retail gondola.

## 2 BACKGROUND: DEFLECTION AND VIBRATION

Deflection describes the degree of displacement for the structural element under a load [14]. Figure 2(a) shows a cantilever beam end-load deflection model, and the displacement $\delta$ at $L$ distance from the support calculates as:

$$\delta = \frac{FL^3}{3EI}, \quad I = mL^2, \tag{1}$$

Where $I$ is the moment of inertia [4]. Here, $F$ is the force applied, $E$ is Young's modulus that describes how easily the material would deform and stretch [14], and $L$ is the distance of the beam.

As shown in Figure 2(b), when we remove the force $F$, the cantilever beam will return to its original shape, and the inertia of the beam will cause vibration around that initial location [17]. This vibration is described as a damped oscillation [3], and its amplitude is proportional to the deflection displacement $\delta$. Therefore, the lighter the item is, the less deflection displacement its load will induce, and the vibration induced by interacting with it is of lower amplitude.

Figure 2(c) depicts the deflection models to describe the gondola structure: 1) the shelf ($L_{shelf}$), with the support that connects it to the back panel, and 2) the back panel ($L_{back}$), with the support that connects it to the ground. The customer-product interaction (pickup and put down) can cause a displacement of the shelf and hence generate vibration waves. When a person puts an item on the shelf, this impact deforms the shelf and causes displacement. For a product placed on the shelf, its load induces the deformation of the shelf. When a person picks it up, the restoring force of the shelf induces vibration.

## 3 SYSTEM DESIGN

We present a vibration-based customer-product interaction detection system for autonomous retails. Figure 3 depicts the system overview. When the customer pickup or puts down a product on the shelf, these interactions induce the shelf to vibrate. The vibration propagates through the shelf surface and shelf-gondola connection and is then captured by the sensor placed on the back panel of the gondola shown in Figure 1.

However, this interaction-induced vibration signal significantly dissipates at the structure connector between the shelf and the back panel of the gondola, as shown in Figure 4(a). Therefore, it is challenging to **capture the interaction-induced vibration on shelves via single-point vibration sensing at the back panel of the gondola**.

In addition, as mentioned in Section 2, the pickup interaction induces an extremely low amplitude of vibration signal, which makes event detection with a traditional energy-based approach difficult. On the other hand, the machine learning approach can effectively classify events from noise with time and frequency features. However, it would require intensive labeling, which makes it inapplicable for each shelf. Therefore, it is challenging to **classify these interaction events from noise without label**.

We introduce *CPA*, a cyber-physical augmentation scheme, to enhance its detection performance by augmenting both sensing (Section 3.1) and learning (Section 3.2) perspectives.

### 3.1 Physical Augmentation to Enhance SNR

Vibration sensing signal quality is determined by both the data acquisition hardware and the signal propagation media [20, 22, 23]. Here we focus on the physical augmentation of the signal propagation media. To monitor multiple shelves via single-point vibration sensing, the sensor is placed on the back panel of the gondola.

Figure 4(a) shows the vibration wave propagation from the shelf to the back panel of the gondola. The horizontal grey bar depicts the shelf surface, and the vertical light grey bar depicts the back panel of the gondola. We can consider it as a simplified model of a flat rectangular plate with one edge fixed. The interaction-induced vibration propagates along the plate surface and reaches the fixed edge. Instead of deflecting the plate, the vibration dissipates at the connecting edge, since the vibration particle motion is aligned with (instead of perpendicular to) the surface of the back panel – the

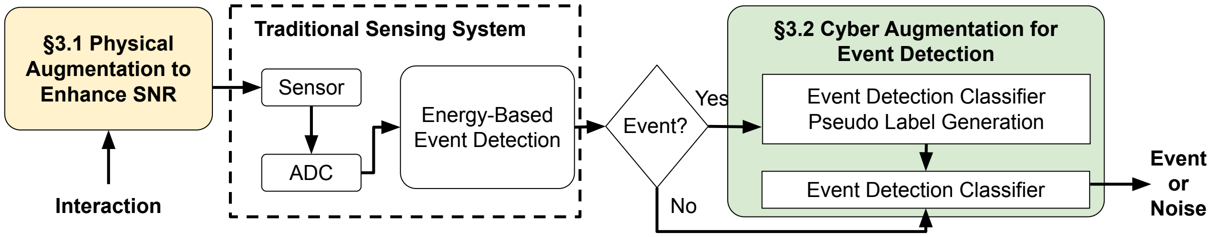


**Figure 3: System architecture. *CPA* contains physical augmentation to enhance SNR and cyber augmentation for event detection.**

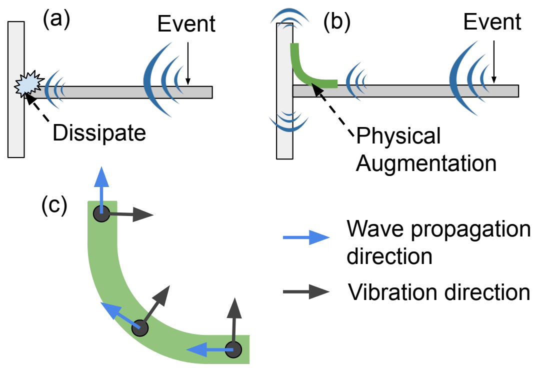


**Figure 4: Physical augmentation. (a) depicts the wave propagation on the original gondola structure, where the shelf vibration dissipates at the connector to the back panel. (b) shows the wave propagation on the gondola with a physical augmentation structure, where the wave is guided to propagate on the back panel. (c) shows the physical augmentation structure and the wave propagation direction changing over the arc.**

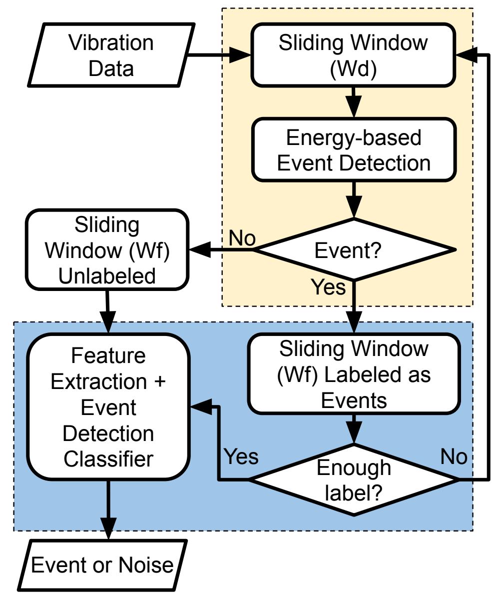


**Figure 5: Flowchart of the cyber augmentation for event detection. The cyber augmentation contains two main parts: event pseudo label generation (yellow box) and event detection classifier (blue box).**

deflection along the surface of the back panel of the gondola is almost zero.

On the other hand, the back panel of the gondola can be considered as a plate with a fixed edge on the floor. If the vibration is perpendicular to the plate, it will propagate along the plate and be captured by the vibration sensor. We present a thin arc structure that can align the vibration direction on the shelf and the back panel. We adopt a 90-degree arc architecture with a radius of 25mm and a thickness of 2mm. Figure 4(c) depicts the wave propagation along the arc structure, where the blue arrows represent the wave propagation direction, and the black arrows represent the particle motion direction. When the vibration wave propagates through the arc architecture, the vibration direction changes with the arc direction [8, 19]. Therefore, our 90° arc architecture can change the vibration direction for 90°, allowing the back panel of the gondola to be used as an efficient vibration signal propagation media for sensing.

### 3.2 Cyber Augmentation for Event Detection

*CPA* leverages the target data's heterogeneity and similarity to achieve labeling free accurate event detection. As discussed in Section 2, items of different weights induce vibration signals of different amplitude; hence these events' SNR varies. As a result, the interaction with heavier items can be detected with a higher detection rate using signal energy-based approaches [9]. These detected events, which are mostly of heavy items, are then used as pseudo labels of events to train an **event detection classifier**. Figure 5 depicts the flowchart of our cyber augmentation components, which contains two main parts: event pseudo label generation (yellow box) and event detection classifier (blue box).

*Event Detection Classifier Pseudo Label Generation.* We apply the traditional energy-based event detection [11] to generate pesudo-positive labels for the event detection classifier. First, a sliding window (with a window size $W_d$) is applied to the vibration signals, and the windowed signal energy is calculated. A Gaussian model ($\mu$,$\sigma$) is then established based on the windowed noise signal energy. Next, for an incoming sliding window, its energy is compared to the Gaussian model, and if the signal energy is higher than a threshold,

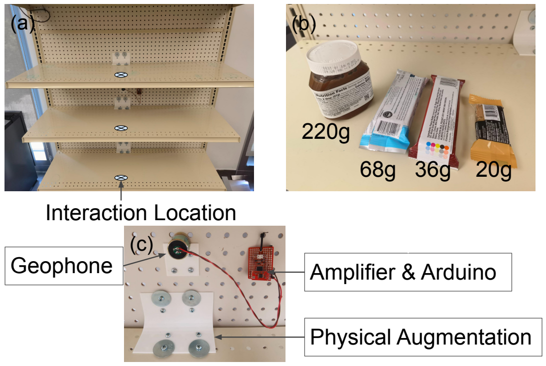


**Figure 6: Hardware and experiment setup.**

e.g., $\mu + 3\sigma$, the window is considered as part of an event. When there are consecutive windows detected as events, we output them as an interaction event *IE*. These detected *IE*, which has a low false positive rate and high false negative rate, are labeled as the positive pseudo label. On the other hand, the ambient noise signals acquired when the store is idle or closed, i.e. no customer in the area is labeled as the negative pseudo label.

*Event Detection Classifier.* When enough pseudo labels are collected, we use them and the corresponding signal segments to train a binary classifier for event detection on each gondola. Because part of the target signals (pickup of small items) often have extremely low SNR values and short duration – therefore are often failed to be detected through the energy-based method. We apply a sliding window with a small window size $W_f$ to effectively capture their features. For each windowed signal, we apply the continuous wavelet transform (CWT) to calculate the component in multiple scales and take the energy of each scale as a feature. Then, the Min-Max normalized [12] energy array of all scales compose the features of this window data. Extracted features and pseudo labels are taken as inputs to train a binary classifier to determine if the unlabeled windows are events or noises.

Multiple classifiers can be applied here, such as neural networks, decision trees, and support vector machine (SVM). Exploring these algorithms, however, is not the focus of this paper. Here, we adopt SVM due to its generally good performance over a small amount of training data. In addition, we choose the per-gondola design to allow robustness and flexibility over different configurations (e.g., materials, total load, and items) of gondolas in stores.

## 4 EVALUATION

We conduct real-world experiments to evaluate *CPA* with multiple product pickup and put down events.

### 4.1 Data Collection and Experiment Setup

*4.1.1 Implementation.* We 3D print our physical augmentation structure with Acrylonitrile Butadiene Styrene (ABS) material. The width, radius, and thickness of the physical augmentation are 120mm, 25mm, and 2mm, respectively. We use screws to mount the physical augmentation between each shelf and gondola, as shown in Figure 6(c). The geophone sensor is horizontally installed on the back panel of the gondola with a 3D-printed mount. To enhance the connection between the sensor and the back panel, we apply resin glue between them. We estimate the cost to retrofit one gondola with our prototype to be lower than \$100, including \$1 for 3D printing, \$4 for OpAmp, \$20 for Arduino, \$40 for Raspberry Pi, \$10 for geophone, which can drop to \$20 or lower if massive produced.

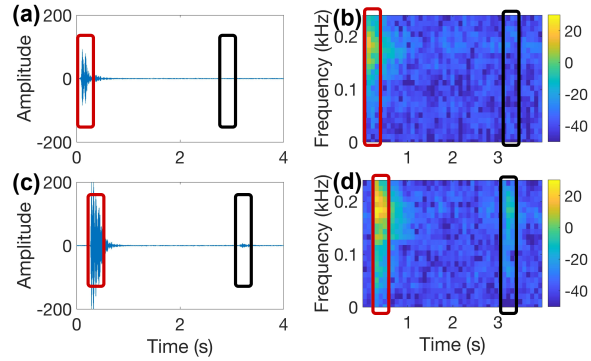


**Figure 7: An example of an energy bar (68g) put down and pickup events with labels marked in rectangular boxes. (a) and (b) show the time-domain and frequency-domain signals without physical augmentation, respectively. (c) and (d) show signals acquired with our physical augmentation.**

*4.1.2 Data Collection.* We select four products, including a plastic bottle of chocolate (220g) and three snack bars with different weights (68g, 36g, 20g), to evaluate the event detection performance. We consider them to fall into two categories based on mass: medium ($\geq 100g$) and light ($< 100g$). For each product, we put it down and pick it up 10 times on the centre of each shelf marked in Figure 6(a) and record the video as the ground truth of the event timestamp. The materials of the shelf and back panel are steel and cardboard.

*4.1.3 Baseline Methods and Evaluation Metric.* We compare *CPA* to three baselines: only apply physical augmentation to enhance SNR (Only PA); only apply cyber augmentation for event detection (Only CA); without physical and cyber augmentation (No Augm.). We take the F1 score as the metric to evaluate the performance of event detection. If the timestamp of an event (from the video data) falls into the groundtruth period, we consider this detected event is a true positive case (TP). Otherwise, this detected event counts as a false positive case (FP). If no event is detected at the groundtruth period, we count this as a false negative case (FN). The F1 score is calculated as: $F1 = 2TP/(2TP + FP + FN)$.

*4.1.4 Algorithm Implementation.* We empirically set the threshold and sliding window length for the energy-based event detection algorithm to be $\mu + 10\sigma$ and 50ms. We set the window length for the event binary classifier to be 20ms. We select the kernel-based SVM model as the classifier to train and detect events. We use the medium-weight product's interaction for autonomous pseudo label generation. We report the detection performance on all products as the overall evaluation of *CPA* and baselines.

### 4.2 Result and Analysis

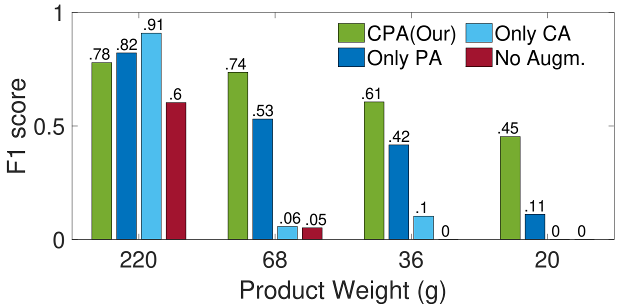


**Figure 8: The F1-score of *CPA* and baselines for pickup event detection. The colour of the bar represents the augmentation configuration: *CPA* (green), only physical augmentation (dark blue), only cyber augmentation (light blue), and without any augmentation (red).**

We first depict the vibration data in Figure 7 with and without physical augmentation to demonstrate the challenge of the low SNR problem and the efficiency of our physical augmentation solution. The signal is generated by putting down and picking up a 68g energy bar. Figure 7(a)(c) show time domain signal, and (b)(d) show frequency domain signals. The red rectangular boxes mark the time periods where put down events occur, and the black boxes mark pickup events. When the physical augmentation is applied, as shown in Figure 7(c)(d), we observe a higher signal energy for both pickup and put down events. Especially for the pickup event, when the physical augmentation is used, both time and frequency domain signals demonstrate distinguishable characteristics compared to the noise. Without the physical augmentation, we do not observe noticeable signal characteristic changes for pickup events, as shown in Figure 7(a)(b).

We report the pickup event's detection performance because 1) it is a more challenging problem than detecting put down events, and 2) it is a more frequent event that occurs in retail. Figure 8 shows the F1 score of the pickup event detection for each testing product. First, we can observe that the F1 score of *CPA* is higher than no augmentation, especially for the light products, where no augmentation yields a 0 F1 score for 36g and 20g products, and *CPA* still achieves 0.61 and 0.45. Overall, *CPA* achieves a 2.9× improvement compared to the no augmentation baseline across four products' pickup events.

The F1 score of *CPA* for three light-weight products are 0.74, 0.61, and 0.45, respectively, which are the highest in all evaluated methods. Compared with only physical augmentation, *CPA* achieves 0.21 (1.4 ×), 0.19 (1.5 ×), and 0.34 (4.1 ×) improvement, respectively. Compared with only cyber augmentation, *CPA* achieves 0.68 (12.9 ×), 0.50 (5.9 ×), and 0.45 (N/A). Physical augmentation demonstrates a significant improvement. The efficiency of cyber augmentation can be further improved by adapting a more representative classifier model or extending the pseudo label set to more products.

# 5 RELATED WORK

Many sensing modalities have been explored for event detection in the autonomous retail setting. The vision is one of the most widely used sensing modalities in autonomous retail stores. Load sensor, RFID, and piezo sensor are the modalities that have also been explored. All these modalities have advantages and disadvantages.

*Vision Based Event Detection.* The vision-based solutions rely on the video feed from multiple overhead cameras to detect customer activities and pickup and put down events [16]. They require the line of sight, and are susceptible to the occlusion caused by the customer's hand. This makes it hard to detect products that are relatively small in size, such as energy bar. The shelves or the other hardware in stores also cause occlusion in the camera view, especially when cameras are placed on the ceiling. Apart from this, high computation power is required to process the vision data from multiple overhead cameras.

*RFID Based Event Detection.* To overcome the occlusion in the vision, RFID has been explored with RFID tags placed on individual products to track customers' interaction with them [13, 15]. The RFID scanner placed either on the shelf, cart, or the checkout counter scans the item customer interacted with and bills them accordingly. Although this method is highly effective in product identification and inventory monitoring, it brings the continuous labour cost to place RFID tags on the products. Plus, the extra hardware, like shelves or carts with scanners, brings the additional setup cost.

*Load Sensor Based Event Detection.* The load sensors have also been explored for event detection and product recognition in the autonomous retail store setting. In this approach, shelves equipped with multiple load sensors (typically one for each product) are used to detect customers' interaction with different products [16]. This method requires special hardware, which brings the added deployment cost, which makes this solution less desirable for the existing retail stores, as they have to replace all the existing gondola shelves with new smart gondola shelves.

*Piezoresistive Smart Textile.* To convert existing shelves into smart shelves, the piezo sensor base smart shelf-liner are proposed for event detection in an automated retail store setting. In this approach, the smart shelf liner equipped with multiple piezoelectric material-based pressure sensors in the grid configuration are placed on the normal shelf to detect customers' interaction with products [5]. Although this method is highly effective in detecting the events and customers' interaction with the products, this method requires a dense deployment of piezo sensors, making this method expensive and less desirable for retailers.

# 6 DISCUSSION

*Cyber-Physical Augmentation Optimization.* The configuration (material, size, number, location) of the physical augmentation structure may impact the signal augmentation efficiency. For example, vibration wave propagation through different materials varies [7]. When placed at different locations with different densities/numbers, the effective vibration propagation path may vary, which would further impact system performance. In this paper, we explore a simple prototype of the arc structure, and the parameters (e.g., material, width, thickness) may not be optimized. Therefore, we plan to explore the optimization of the physical augmentation, including material, size and amount, in the future.

Cyber augmentation also can be further explored by utilizing a more representative classifier model to replace the kernel-based SVM model, such as neural network models. On the other hand, the selected features may also impact the end performance. We plan to explore more data-driven approaches to improve the feature representation for event and noise signal distinction, such as contrastive learning.

*Mitigate Impacts from Ambient Vibrations.* Multiple types of ambient events may induce vibration that can be detected by the vibration sensor, such as customer's footstep, door opening and closing, and the customer touch the gondola. These detected events are false positive cases for pickup/put down item event detection. To distinguish these false positive events, we plan to explore 1) classifier for different event recognition and 2) multi-modal sensing to provide robust labeling for customer-item interaction events.

Prior work on vibration-based human sensing has shown that different human activities would induce signals with distinguishable characteristics [6]. Therefore, it is feasible to recognize signals that are induced by pickup/put down items from those induced by other activities. However, our current cyber augmentation design is not sufficient to achieve pseudo labeling for different types of activities. Since there will be cameras installed in the autonomous retails, we could leverage camera data to achieve fine-grained cross-modal pseudo labeling for cyber augmentation.

*System Sensing Range.* We plan to further study the augmentation design and its impact on the system's sensing range. For example, what is the lower bound of the item that can be detected and how does the augmentation design configuration impact it? In this paper, *CPA* achieves 1 and 0.6 recall for 36g and 20g items, respectively. If we focus on the recall performance, we can consider the lower bound weight of the item that can be reliably detected in our experiment is 36 grams. We plan to define the acceptable detection rate based on the target application requirements and conduct fine-grained experiments to explore the lower bound of the system.

*Multiple Users Scenarios.* In the real-world scenario, there might be multiple users interacting with the same gondola at the same time. In the experiment, we observe that the duration of each vibration signal is around 0.1 seconds or less (might varies with gondola material and item), which means it is a low chance that their interaction induced signals will be overlapped. In this work, we check the feasibility of *CPA* when there is no signal overlapping. In the future, we plan to explore methods for signal separation, leveraging the information from other modalities. For example, the cameras in the store can provide complementary information, such as the number of customers interacting with the shame shelf, which could provide prior knowledge for the signal separation.

*Robustness to Inventory Changes.* The total load on the shelf may change over time. The vibration signals of the same event (e.g., pickup a candy bar) may change when the total load and load distribution change [10]. This means the event signals' data distribution may shift with an inferrable physical factor – total load on the shelf. We plan to model this physical phenomenon and design a physical and data-driven approach to mitigate the varying total shelf-load induced data distribution shift. For example, by applying a multi-task learning algorithm to 1) estimate shelf-load and 2) detect events or identify events.

In a real-world scenario, the customer may also return the item to another shelf, which changes the inventory on the shelf. We plan to solve this problem by fusing the vision data, which provides more information on the customer's identity, location, and items. We will also explore how to recognize the item using detected event signals.

## 7 CONCLUSION

In this paper, we present a vibration-based single-point sensing system for customer-product interaction detection for autonomous retail stores. To enable the detection of extreme low-SNR signals (light product pickup events), we introduce *CPA*, a cyber-physical augmentation scheme that enhances vibration data SNR with a physical arc structure and provides pseudo labels for event detection classifier leveraging properties of signals generated by different product pickup. We conduct real-world experiments to evaluate the performance of our proposed signal augmentation structure and event detection framework. *CPA* achieves up to 2.9× improvement for the detection of light-weight product pickup events.


## ACKNOWLEDGEMENT

This project was supported in part by AiFi Inc.


## REFERENCES


[1] 2023. Gondola Display. https://www.8thandwalton.com/blog/gondola/

[2] 2023. Our Retail Divisions. https://corporate.walmart.com/newsroom/2005/01/06/our-retail-divisions

[3] Philip R Dahl. 1976. Solid friction damping of mechanical vibrations. *AIAA journal* 14, 12 (1976), 1675–1682.

[4] James M Gere and Barry J Goodno. 2012. *Mechanics of materials.* Cengage learning.

[5] Gaffar Hossain, Ishtia Zahir Hossain, and Günter Grabher. 2020. Piezoresistive smart-textile sensor for inventory management record. *Sensors and Actuators A: Physical* 315 (2020), 112300.

[6] Zhizhang Hu, Yue Zhang, Tong Yu, and Shijia Pan. 2022. VMA: Domain Variance- and Modality-Aware Model Transfer for Fine-Grained Occupant Activity Recognition. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 259–270.

[7] Sylwester Kaliski. 2013. *Vibrations and Waves: Part B: Waves.* Elsevier.

[8] B Kang, CH Riedel, and CA Tan. 2003. Free vibration analysis of planar curved beams by wave propagation. *Journal of sound and vibration* 260, 1 (2003), 19–44.

[9] Dima Litvak, Yaniv Zigel, and Israel Gannot. 2008. Fall detection of elderly through floor vibrations and sound. In *2008 30th annual international conference of the IEEE engineering in medicine and biology society*. IEEE, 4632–4635.

[10] Mostafa Mirshekari, Jonathon Fagert, Shijia Pan, Pei Zhang, and Hae Young Noh. 2021. Obstruction-invariant occupant localization using footstep-induced structural vibrations. *Mechanical Systems and Signal Processing* 153 (2021), 107499.

[11] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J Mengshoel, Hae Young Noh, and Pei Zhang. 2017. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–31.

[12] S Patro and Kishore Kumar Sahu. 2015. Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462* (2015).

[13] Yacine Rekik, Evren Sahin, and Yves Dallery. 2008. Analysis of the Impact of the RFID Technology on Reducing Misplacement Errors at the Retailer. *International Journal of Production Economics* 112 (03 2008), 264–278. https://doi.org/10.1016/j.ijpe.2006.08.024

[14] F Rohde. 1953. Large deflections of a cantilever beam with uniformly distributed load. *Quart. Appl. Math.* 11, 3 (1953), 337–338.

[15] G. Roussos. 2006. Enabling RFID in retail. *Computer* 39, 3 (2006), 25–30. https://doi.org/10.1109/MC.2006.88

[16] Carlos Ruiz, Joao Falcao, Shijia Pan, Hae Young Noh, and Pei Zhang. 2019. Autonomous inventory monitoring through multi-modal sensing (aim3s) for cashier-less stores. In *Proceedings of the 6th ACM International Conference on Systems for*

*Energy-Efficient Buildings, Cities, and Transportation*. 395–396.
[17] Enrico Volterra and Eleftherios Charalampos Zachmanoglou. 1965. *Dynamics of vibrations*. Vol. 1. CE Merrill Books.
[18] Xiu-Shen Wei, Quan Cui, Lei Yang, Peng Wang, and Lingqiao Liu. 2019. RPC: A Large-Scale Retail Product Checkout Dataset. https://doi.org/10.48550/ARXIV.1901.07249
[19] Zhibo Yang, Xuefeng Chen, Xiang Li, Yongying Jiang, Huihui Miao, and Zhengjia He. 2014. Wave motion analysis in arch structures via wavelet finite element method. *Journal of Sound and Vibration* 333, 2 (2014), 446–469.
[20] Yue Zhang, Zhizhang Hu, Susu Xu, and Shijia Pan. 2021. AutoQual: task-oriented structural vibration sensing quality assessment leveraging co-located mobile sensing context. *CCF Transactions on Pervasive Computing and Interaction* 3, 4 (2021), 378–396.
[21] Yue Zhang, Shijia Pan, Jonathon Fagert, Mostafa Mirshekari, Hae Young Noh, Pei Zhang, and Lin Zhang. 2018. Occupant activity level estimation using floor vibration. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 1355–1363.
[22] Yue Zhang, Susu Xu, Laixi Shi, and Shijia Pan. 2020. Using Mobile Sensing to Enable the Signal Quality Assessment for Infrastructure Sensing Systems. In *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications*. 102–102.
[23] Yue Zhang, Lin Zhang, Hae Young Noh, Pei Zhang, and Shijia Pan. 2019. A signal quality assessment metrics for vibration-based human sensing data acquisition. In *Proceedings of the 2nd Workshop on Data Acquisition To Analysis*. 29–33.