1-2016

# IoT+Small Data: Transforming In-Store Shopping Analytics and Services

Meera RADHAKRISHNAN
*Singapore Management University*, meeralakshm.2014@phdis.smu.edu.sg

Sougata SEN
*Singapore Management University*, sougata.sen.2012@phdis.smu.edu.sg

Vigneshwaran SUBBARAJU
*Singapore Management University*, vigneshwaran@smu.edu.sg

Archan MISRA
*Singapore Management University*, archanm@smu.edu.sg

Rajesh BALAN
*Singapore Management University*, rajesh@smu.edu.sg



## Citation

RADHAKRISHNAN, Meera; SEN, Sougata; SUBBARAJU, Vigneshwaran; MISRA, Archan; and BALAN, Rajesh. IoT+Small Data: Transforming In-Store Shopping Analytics and Services. (2016). *2016 8th International Conference on Communication Systems and Networks: COMSNETS 2016, Bangalore, India, January 5-10 [COMSNETS Workshop: Wild and Crazy Ideas on the interplay between IoT and Big Data WACI]*. 7439946-1-6.

**Available at:** https://ink.library.smu.edu.sg/sis_research/3570

# IoT+Small Data: Transforming In-Store Shopping Analytics & Services

Meera Radhakrishnan, Sougata Sen, Vigneshwaran S., Archan Misra, Rajesh Balan
School of Information Systems, Singapore Management University

***Abstract*—We espouse a vision of small data-based immersive retail analytics, where a combination of sensor data, from personal wearable-devices and store-deployed sensors & IoT devices, is used to create real-time, individualized services for in-store shoppers. Key challenges include (a) appropriate joint mining of sensor & wearable data to capture a shopper's product-level interactions, and (b) judicious triggering of power-hungry wearable sensors (e.g., camera) to capture only relevant portions of a shopper's in-store activities. To explore the feasibility of our vision, we conducted experiments with 5 smartwatch-wearing users who interacted with objects placed on cupboard racks in our lab (to crudely mimic corresponding grocery store interactions). Initial results show significant promise: 94% accuracy in identifying an item-picking gesture, 85% accuracy in identifying the shelf-location from where the item was picked and 61% accuracy in identifying the exact item picked (via analysis of the smartwatch camera data).**

## I. INTRODUCTION

The retail segment already uses IoT & "Big Data" to optimize store-level operations, such as predictive inventory management and merchandise layout planning. These innovations are, however, *store-centric:* they do not focus on using a shopper's in-store behavior to optimize or personalize shopping-related services in real-time, *while the user is inside the store*. We believe that the joint real-time mining of sensor data, from store-deployed IoT devices and the personal mobile & wearable devices of an individual shopper, can transform the in-store shopping experience. In particular, this type of "Small Data Analytics" (detailed insights about an individual shopper) can (a) infer a shopper's in-store actions and product choices in real-time (even before the checkout counter), and (b) is cheap and easy to implement (requires no complex infrastructure support).

Examples of such novel services that exploit real-time knowledge of a shopper's in-store activities and product-level interactions include: (a) *Smart Reminder*: that reminds you, for example, pick up milk only if you walk past the milk section *without picking up milk*; (b) *recipeGuru*: that identifies and provides alerts if you pick up the wrong item for your stated recipe; and (c) *Recommender*: that uses knowledge of the products you've been picking so far to build a *dynamic, episode-specific* interest profile, and suggest complementary items (e.g., a bottle of wine to go with your selected cheese). Note that these are all significant improvements on the more vanilla versions of these use cases: (a) Reminders based purely on in-store location; (b) a shopping list that you must manually track on your smartphone, and (c) recommendations based purely on general, longer-term customer profiles.

Our vision engenders the following key question: "How can applications use infrastructure sensors, and wearable/mobile devices to unobtrusively obtain deeper insights on customer interactions?" We propose an architecture where low-cost BLE beacons+ embedded sensors are mounted on product shelves, and their data is fused with sensor readings from a smartwatch worn by a shopper. Using various micro-studies on how users interact with objects placed on shelves in our lab (to crudely mimic similar in-store interactions), we establish two promising principles: (a) to better identify item-level interactions, we must utilize correlation between infrastructure and wearable sensor data; and (b) the camera on a wrist-worn smartwatch can identify a specific product selected by a shopper, but must be intelligently triggered to conserve energy.

Our main empirical insights include:

- Using accelerometer and gyroscope features from the smartwatch, we can identify with 94% accuracy if an item was picked and with 85% accuracy, from where the item was picked. In addition, we also show that exploiting the correlation between the inertial sensors of the smartwatch and a door-mounted sensor allows us to reliably identify the specific display section with which the shopper interacts.
- Using a smatwatch's camera (pointing towards one's fingers), we can identify the exact item a person picked in 61% cases using a very basic image recognition algorithm; moreover, the best image is usually obtained closer to the mid-point of such a 'pick' gesture.

## II. RELATED WORK

Given the vast body of work on context and activity recognition via mobile & wearable sensing, we highlight recent work that is most closely related to our vision of using a combination of infrastructure & mobile sensing for understanding in-store shopping behavior.

*Monitoring shopper behavior using sensors:* Unobtrusive monitoring of shopper's behavior has always been of interest to researchers. ThirdEye [9] assumes that a shopper wears a smartglass, and uses such camera and Wi-Fi based location data to track the items that a shopper browses in a store. To reduce the overheads of continuous vision sensing, it uses inertial sensors to trigger the image capture only when the user spends some time gazing at a product. CROSDAC [10] explores a person-independent activity recognition technique, based on smartphone sensor and Wi-Fi based location data, to identify

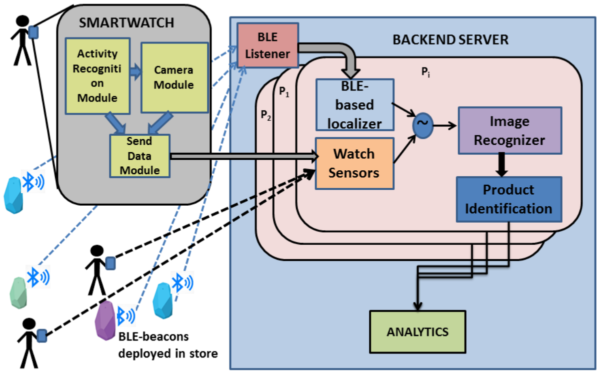



Fig. 1: Architecture of the system

the high-level shopping intent of users. ShopMiner [11] proposes an RFID-based system to infer the aggregated shopper interaction patterns with specific items in a physical clothing store. Zeng et. al [14] utilized the Channel State Information of Wi-Fi signals to infer a shopper's locomotive state (walking vs. standing) & location within a store. Lee et. al [7] proposed a framework for understanding a shopper's overall *in-mall* behavior, using Wi-Fi for store recognition and smartphone sensors for movement tracking. In contrast to these effort, we explore the possibility of combining IoT & wearable sensor data to identify an individual shopper's item-level interaction behavior within a store.

*Image based object identification:* Object detection and recognition, via smartphone-captured images, have been explored recently in prototypes such as: (a) Glimpse [4], which supports continuous recognition and tracking of traffic signs from a moving smartphone camera; (b) MobiMed [8], which identifies medication packages from images captured by a user's smartphone, and (c) FoodCam [5], which recognizes the food consumed by the user. Typically, deep learning using computationally-expensive convolutional neural networks [6], is employed for object recognition, but is performed offline on backend servers. Our exploration of image-based item recognition differs in that: (a) our images are captured opportunistically by a moving smartwatch camera, and are likely to exhibit motion blur and occlusion, and (b) we would ideally like to identify the items directly on the wearable device.

## III. Architecture and Study Details

In this section we discuss about our envisioned architecture of a future retail store. We then describe our setup to mimic this future retail store.

### A. High-Level Architecture of a future retail store

A future retail store should be able to address the following challenges: (i) Micro-activity recognition to identify shopper's interaction with individual items, (ii) BLE-beacon based fine-grained identification of the exact location from which the item was picked and (iii) adaptive image capturing and feature matching to identify the products being interacted with.

For realization of such a system and to address the challenges, we assume that BLE beacons are mounted on doors of racks in a shop. The BLE beacons have the capability not only to transmit signal at fixed intervals, but also are equipped with accelerometer and communication module. Using the communication module, the beacon transmits its accelerometer data to the back-end server. We also assume that shoppers entering the shop are wearing a smartwatch on their dominant hand. This smartwatch is equipped with a camera along with an accelerometer and gyroscope and installed with a custom application running to monitor shopping activity. For our application, the shopping activities included, picking item from a shelf, putting it in a trolley and interacting with shelves (opening and closing) in a store. However for different shops, different activities can be trained. Once the application on the watch determines that a shopping activity is taking place, it turns on the camera opportunistically to capture the image of item that the shopper is interacting with.

In the back-end, there is a module for listening to accelerometer data from BLE beacon. There is a separate module for listening to data from the smartwatch. Once data from a smartwatch reaches the server, a new sub-process is started for the person. This sub-process handles further incoming data from the smartwatch. As more data comes in, the RSSI of BLE beacon heard on the smartwatch are used to filter out beacons which are not in proximity of the user. Accelerometer data from BLE beacons which are in proximity is correlated with the accelerometer data from the smartwatch to determine which door has the user interacted with. Identification of the door that the user has interacted with helps in narrowing down the identification of possible items the user has interacted with (as a rack will have only a subset of objects present in the store). Next, based on image recognition, the server can identify the exact object the shopper interacted with. Figure 1 illustrates our architectural framework with the devices, server components and flow of the analytics pipeline.

In a real shop, there will be multiple users, so multiple sub-processes will be triggered at the same time to identify objects the shoppers are interacting with. With the output of the sub-process, various analytics can be performed. For example, identifying hot objects (which items are shoppers interacting more within a shop) or performing association rule mining with interacted object as opposed to current state of the art association rule mining on billed items.

### B. Experimental Design

We first describe the devices used in the data collection process. We classified the data collected for our experiments into two categories - (i) data coming from infrastructure devices and (ii) data coming from the wearable devices with an accelerometer and gyroscope built in. For the infrastructure devices, we used Estimote$^{TM}$BLE beacons [2], which was set (unless otherwise specified) to a default transmission power (-20 dBm) and a 101 msec advertising period. A lot of the latest BLE beacons are equipped with an accelerometer. Since we did not have access to the accelerometer of the beacon, we mounted a smartphone (Samsung Galaxy S3), installed

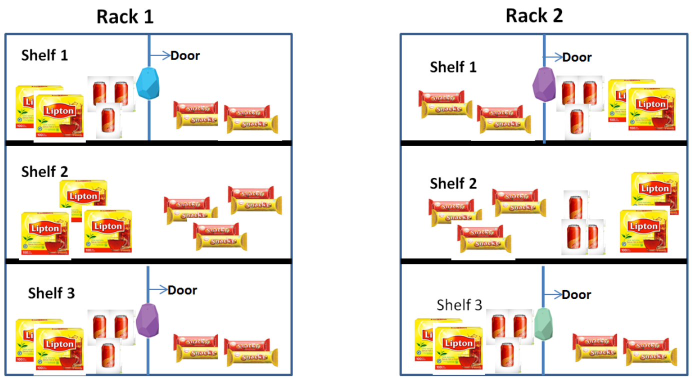



Fig. 2: Testbed Layout

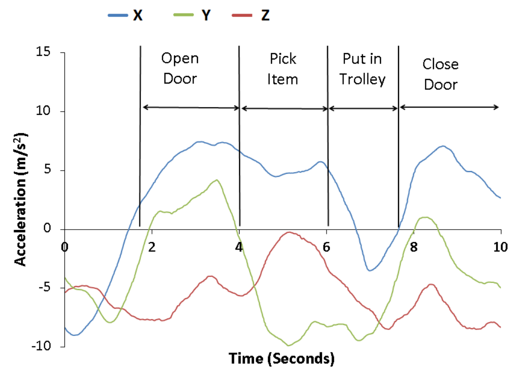



Fig. 3: Accelerometer reading for a sequence of hand activities

with an application that records accelerometer data, right next to the beacon horizontally. The accelerometer data from the smartphone was a proxy for the accelerometer data from the beacon. In the rest of the paper, we refer to this smartphone's accelerometer data as the beacon's accelerometer data. For the wearable device, we used the Omate TrueSmart smartwatch, which the participants wore around the wrist of their dominant hand and which was pre-installed with our custom data collection application which recorded video, collected accelerometer and gyroscope data and performed scanning for the BLE beacons. Both the smartwatch and the smartphone mounted next to the beacon recorded sensor data at a sampling frequency of 100Hz.

We next describe the location of our study. For our study, we emulated an aisle in a grocery store in the pantry area of our lab as shown in Figure 2. The pantry area in the lab has multiple racks and each of these racks has shelves at multiple levels. We used 3 shelves in 2 racks, which were 1.7, 0.85 and 0.5 meters from the ground level. The top and bottom shelves had hinged doors. In terms of distance between the two racks, the two racks are 1.1 meters away from each other. On each of the 3 shelves in both the racks, 3 category of items were placed - (i) boxes of Lipton tea (*Item-I1*)(ii) cold drink cans (*Item-I2*)and (iii) biscuit packets (*Item-I3*). We attached one BLE beacon each on the doors of the shelf. For our experiments, we kept the transmitting power of the beacons at -20dBm so that at this level the beacon couldn't be heard beyond 3 meters.

Finally, we describe the participants and the set up of the study. For our study, we recruited 5 participants from our lab (2 males, 3 females - all aged between 20 and 30) and who were almost in the similar height range (1.55 to 1.70 meters). Participants were asked to perform different activity sequences that are normally carried out while grocery shopping. The participants were asked to perform the activity sequence of (1) open the door of the shelf, (2) pick an item from the shelf, (3) put the item aside and (4) close the door of the shelf, all in a way similar to how they would have performed it in a grocery store. We also asked them to repeat the whole sequence to pick the items from the 3 different shelves at different heights of the rack. Each participant repeated this process 10 times. In total, we collected 30 sample sequences each from a participant. On an average, a round of item picking took approximately 10 seconds. The ground truth of the activities was collected by having a person shadowing the user and labeling the activities as the user performs it. All the devices used in our study were time-synchronized.

## IV. Shopping Activity Detection

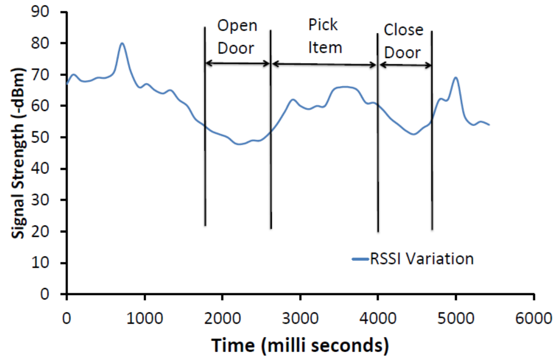
We next describe the experiments that we performed in our controlled environment.

### A. *Identifying "Item Picking" Gesture*

For a shopping study, we wanted to ensure that we could identify the picking action robustly. Picking action should be distinguishable from other similar actions or gestures such as putting items back/putting items aside, pushing/pulling a door etc. As mentioned previously, we collected 30 traces of accelerometer and gyroscope data from the smartwatch for a sequence of activities from 5 users. We first plotted the accelerometer data for one such sequence. Figure 3 shows how the accelerometer reading varies for the different activities in the sequence. We next extracted the features mentioned in [13] from the data. We used the decision tree implementation in Weka for our classification. A 10-fold cross validation was performed to see if we could distinguish *picking item* from other gestures. We found that we could achieve an accuracy of 88% when we ran a classifier for multi-class identification (picking, putting back, open door, close door, put aside). Since we wanted to identify picking gesture (so that we could capture image during the gesture), we labelled all non-picking classes as *others*. We performed a 10-fold cross validation for a binary classifier and found that the accuracy of identifying picking improved to 94%. To see if our model was user independent, we performed a leave-one-user-out cross validation and found a marginal 6% drop in accuracy. This indicated that picking could be easily distinguishable from other similar activities.

### B. *Identifying "Shelf-level Location of Item Picked"*

For the initial study, participants picked items only from the top shelf. We next wanted to see if there was any difference in the picking action, when a person picked an item from shelves




The BLE RSSI variation shows similar trend for 'Door Open' and 'Door Close' activities. So these gestures are not distinguishable just based on the beacon signals.

Fig. 4: Variation of BLE signal during shopping activity

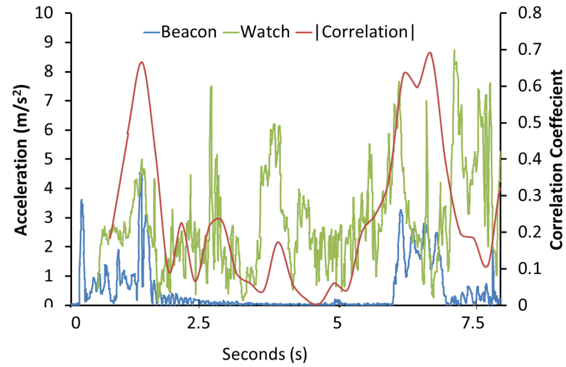



Fig. 5: Correlation between magnitude of accelerometer data of beacon and watch

which are at different heights. The 5 participants were asked to pick up items from the top, middle and bottom shelf 10 times. So for every individual, 10 *x* 3 picking activities were noted and were labeled as *top*, *middle* or *bottom*. We performed a 10-fold cross validation using the J48 (Decision tree) classifier in Weka. We found that using only the accelerometer and gyroscope data, we could identify which shelf a person picked the item from with **85**% accuracy. So based on this we can say that if only one item from one brand is kept in a shelf, then with 85% confidence we can claim that the person has picked the particular item. Again to ensure that our technique was generalizable, we performed a leave-one-user-out cross validation and the accuracy dropped by just 2% to 83%. This indicates that the approach is generalizable. However identifying the shelf from which an item is picked is a function of the height of the shopper. In our current study, the participants were in the height range of 1.5 to 1.65 meters. As a future work, we will perform the similar experiment with subjects of wider height range.

An alternate technique to identify which rack an item is being picked is by using the BLE beacons. Similar to what is shown in [12], when the person is near one shelf, the BLE beacon's RSSI heard on the phone should increase. For our study, we found that when a person was about 0.5m from a beacon, the RSSI was -50 dBm to -60 dBm.

### C. Improved Location Estimation via Infrastructure Sensors

As we showed in the previous sub section that the RSSI drops when the hand came closer to the beacon (See Figure 4). However just relying on RSSI to determine the rack of interest can lead to high false positives. For example, when a shopper walks past a rack, the received signal strength from beacon will peak. RSSI can be a good localizing factor, but we need to correlate the data from the inertial sensors on the watch with that on the beacon. In our experiment, when the user opened or closed the door of the top shelf, we observed a sudden spike in the accelerometer of the beacon attached to that door, whereas the reading from the accelerometer of the beacon attached to the door of the bottom shelf showed no variation. This suggested that using sensor-based features from both the BLE beacon and smartwatch would improve the accuracy of detecting which door the shopper has interacted with. On correlating the accelerometer signals from the smartwatch and the beacon, we found that there is a high correlation between the two signals during the 'Door Open' and 'Door Close' gestures, as shown in Figure 5. We plotted the magnitude of the watch accelerometer and beacon accelerometer for the different gestures in the activity sequence and found that during the door open and door close gestures, the correlation coefficient is always above a threshold, which we empirically determined as 0.5. Based on this,we can claim that independent of the user's height, based on correlating data from a user's device with an infrastructure sensor, door activity can be easily determined.

### D. Summary

To summarize our findings of determining shopping behavior using inertial and infrastructure sensors, we showed that we could identify if a person was picking an item with 94% accuracy in case of person dependent models and 88% when we had a person independent model. Thus with a minimal loss in accuracy, we can build a scalable item *picking* model. We also showed that we could identify using inertial sensors from a user's smartwatch, with 85% accuracy if a person was picking an item from the top shelf or bottom shelf, thus helping in determining the *rack of interest*. The *rack of interest* can be identified even more accurately when we correlate the sensor data from the inertial sensors on the user's devices with the inertial sensors mounted on the door.

## V. Capturing Image of Item

An alternate approach to identifying objects that a person interacted with is by capturing the image of the object using the cameras in the wearable devices. This section discusses our study of this approach using our experimental setup that was described earlier. While performing our experiments of picking items, we found that for a short period of time, the camera on the watch usually points towards the item that is being picked. We wanted to see if it is possible to capture a legible image of the item being picked from the camera. In the

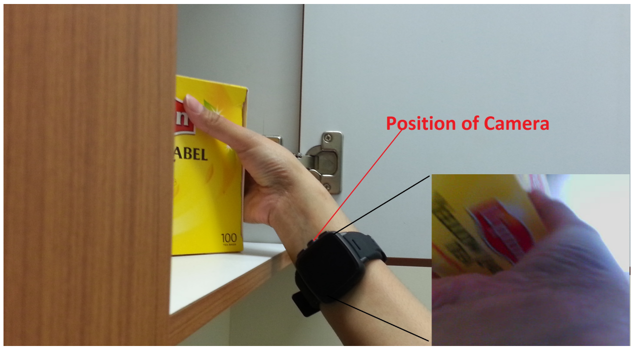



Fig. 6: Camera view and an image captured while an item is being picked

smartwatch that we used in our study, the camera is located on the side of the watch as shown in Figure 6. We found that if the watch was worn normally, we could not capture images of objects. However, if we rotated the watch to the position shown in Figure 6, we could easily capture the image of the item that was being picked. To study this further, we captured a video when the person was performing the picking activity and extracted all individual frames from the video. This was done for the 5 users, each one picking items from the three different racks. So in total, we obtained 150 such videos and amongst these, we found that the item which was picked was visible in all of the picking gestures indicating that if a watch has the camera in this position, it is inevitable that the image of the item will be captured.

However continuous video capture is not practicable because it will drain out the battery of the watch in a very short period of time. So instead of capturing video, we wanted to see if we could trigger the camera at a correct moment to capture the image of the item. To understand the variation, we plotted the probability of a captured image being 'useful' (i.e., provides a clear view of the item picked) as a function of the time when the image was captured, relative to the overall duration of the gesture. Figure 7 shows the plot of this probability as a function of the time, with the time being expressed as a percentage of the overall gesture duration. We can see that the probability of getting a useful image is highest around the first 20-80% of the time of the picking gesture. In terms of absolute time, this window is approximately 1.2 seconds, which is a fairly wide window and thus instead of capturing a video, even if a single image is captured, the image of the item will be procured and thus saving the battery of the watch.

### A. Feasibility of Item Recognition

As discussed previously, we could obtain the image of object at least once during a picking gesture. We next wanted to see if the captured image (extracted from the video frame) could be identified automatically by a image recognition software. We used the as-is implementation of SURF [3] algorithm in openCV and passed all the frames containing the *item-being-picked* during the gesture. Each image of the object obtained during the gesture was compared (by matching features) against the set of 3 objects (3 images of each object, taken from 3 different angles) which were present on the shelf. In total found that we could identify the correct item in 61% cases (vanilla baseline - 33%). When we analyzed the images that were identified wrongly, we realized that the wrong classification occurred due to (i) occlusion of the object — if the object is small, the fingers cover a major portion of the image. In our case, we found that in many of the frames containing the pack of biscuit, where part of the packet can be seen, the image recognition software mistook it for the tea box. (ii) blurriness — when the item is being picked, motion blur creeks in the image frame obtained from the video. This might result in mis-classification. (iii) Insufficient training of the recognition model — for our small study, we just used simple feature matching to recognise the objects. Even though these images were taken from different angles, they did not cover all possible angles that might be visible to the camera when the object is being picked. The accuracy should improve with a more sophisticated recognition model that can be trained carefully for this application.

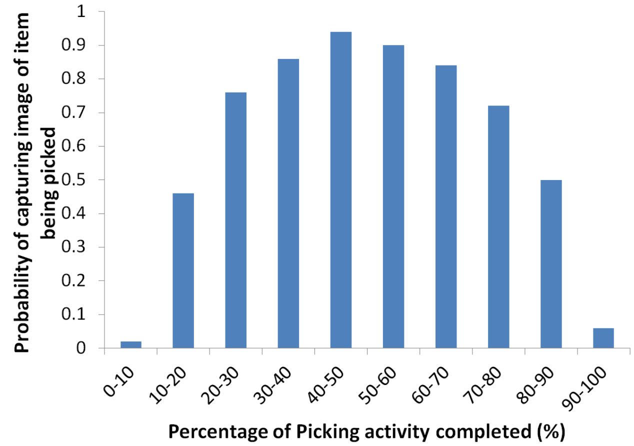



Fig. 7: Probability of capturing image of item being picked

To understand if images captured by the watch were identifiable by a commercial image recognition software, we ran some of the images on Clarifai [1], a commercial image recognition software which uses convolutional neural networks. We found that, even without supplying training data, the deep learning software was able to broadly tag the images obtained from our study. As a next step to our work, we plan to use a more robust image recognition algorithm for better image identification.

## VI. Discussion

Our initial results are promising, but admittedly conducted under a contrived setting: with the items being placed on the shelves in our office lab. There are many aspects of shopper behavior in real-world stores that we will have to consider to make this vision an eventual reality.

*Additional Shopper Interaction Gestures:* In our micro-studies, we have concentrated on "item picking" gestures, with the assumption that a picked item is of interest to the shopper and likely to be part of her shopping list. In reality, shoppers often browse a range of items, but only select a much smaller subset for actual purchase. To build an accurate profile (e.g., for the *recipeGuru* or *Reminder* applications), we need to detect the 'put back' gesture to filter out items that the shopper looked at but did not eventually select. Moreover, as such items could be put back at any point during the shopping session (e.g., a while after it was initially placed in the cart), the image-based item recognition process is needed to identify such returned items across a much wider range of shopper behavior. Additionally, in our initial exploration, we did not try to differentiate the 'pick up' gesture–e.g., to see whether such gestures exhibit microscopic differences between heavy vs. light, or compact vs. bulky objects. Such finer-grained gesture differentiation might provide additional priors on the type of item selected by a shopper.

*Possibilities for Enhanced Sensor Fusion:* The combination of wearable+ IoT sensor data offers many possibilities for further differentiated understanding of shopper behavior. For example, as show in Figure 4, the RSSI readings (on the smarwatch) are fairly similar for 'door opening' vs. 'door closing', making it hard to distinguish between these two gestural activities. However, the additional inertial sensing data from the smartwatch can be easily used to separate these two activities, as the hand movements effectively have opposite 'polarity'. Similarly, the temperature sensor readings on the smartwatch may be used to detect interactions with items in the freezer section (characterized by a sudden drop in ambient temperature).

*Additional Applications & Scenarios:* Real-time determination of the specific item being selected can be used for other types of consumer-specific alerts. Consumers today can use their mobile devices to obtain instant information (e.g., customer reviews, product ratings or price comparisons) from online sources. At present, such information retrieval typically requires manual input—the shopper must either upload a picture or a product specification to the online service. Real-time wearable+IoT analytics offers the possibility of making such retrieval unobtrusive. For example, if the item picked up by the shopper turns out to have ingredients to which the shopper is allergic, a *product alert* application can proactively alert the shopper to such inadvertent selections. Similar, more accurate tracking of the *numbers* selected, for a specific item, might alert the shopper to possible promotions and deals that she may be unaware of. For example, if a particular brand of apples has a "3 for $2" offer (with a unit price of $2), a *deal detective* application can automatically alert the shopper about the promotional offer, if it detects that she has selected only 2 apples.

## VII. Conclusion

Based on the studies conducted with users performing activities as in a normal grocery shopping in a simulated environment, this paper shows the practicality of using data sensed from the shopper's personal devices and multiple other IoT devices deployed in the store to perform fine-grained shopper in-store analytics. Results from a study (conducted in a lab environment that crudely replicate the shelves of a grocery store) show that, given a trace of the sensing data and the images captured, we are able to (i) identify the shopper's interaction with the products (picking gesture with an accuracy of 94%), (ii) the exact shelf from which the item was picked with 85% accuracy and also (iii) the exact item being picked with an accuracy of over 61%. We believe that novel analytics techniques that jointly harness such infrastructure-based IoT and shopper-specific wearable sensing data can lead to entirely new real-time, in-store, individual-specific shopping services and applications.

## Acknowledgement

This work was supported partially by Singapore Ministry of Education Academic Research Fund Tier 2 under research grant MOE2011-T2-1001 and partially by the National Research Foundation, Singapore under its Interactive Digital Media (IDM) Strategic Research Programme and the International Research Centre @ Singapore Funding Initiatives. All findings and recommendations are those of the authors and do not necessarily reflect the views of the granting agency, or Singapore Management University.

## References

[1] Clarifai. http://www.clarifai.com/. Accessed: 2015-10-30.

[2] Estimote beacon. http://estimote.com. Accessed: 2014-10-30.

[3] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359, 2008.

[4] Chen, T. Y.-H., Ravindranath, L. S., Deng, S., Bahl, P. V., and Balakrishnan, H. Glimpse: Continuous, Real-Time Object Recognition on Mobile Devices. *Proc. of Sensys'15*.

[5] Kawano, Y. and Yanai, K. Foodcam: A real-time food recognition system on a smartphone. *Multimedia Tools and Applications*, 2014.

[6] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Proc. of NIPS'12*.

[7] Lee, S., Min, C., Yoo, C., and Song, J. Understanding customer malling behavior in an urban shopping mall using smartphones. *Proc. of MSCSS'13 colocated with UbiComp'13*.

[8] Möller, A., Diewald, S., Roalter, L., and Kranz, M. Mobimed: comparing object identification techniques on smartphones. *Proc. of NordiCHI'12*.

[9] Rallapalli, S., Ganesan, A., Chintalapudi, K., Padmanabhan, V., and Qiu, L. Enabling physical analytics in retail stores using smart glasses. *Proc. of MobiCom'14*.

[10] Sen, S., Chakraborty, D., Subbaraju, V., Banerjee, D., Misra, A., Banerjee, N., and Mittal, S. Accommodating user diversity for in-store shopping behavior recognition. *Proc. of ISWC'14*.

[11] Shangguan, L., Zhou, Z., Zheng, X., Yang, L., Liu, Y., and Han, J. Shopminer: Mining customer shopping behavior in physical clothing stores with passive rfids. *Proc. of Sensys'15*.

[12] Subbaraju, V., Sen, S., Misra, A., Chakraborti, S., and Balan, R. Using infrastructure-provided context filters for efficient fine-grained activity sensing. *Proc. of PerCom'15*.

[13] Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., and Aberer, K. Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach. *Proc. of ISWC'12*.

[14] Zeng, Y., Pathak, P. H., Mohapatra, P., Xu, C., Pande, A., Das, A., Miyamoto, S., Seto, E., Henricson, E., Han, J., et al. Analyzing shoppers behavior through wifi signals. *Proc. of WPA'15*.