Evaluation of cupboard door sensors for improving activity recognition in the kitchen

Samuel Whitehouse[1], Kristina Yordanova[1,2], Stefan Lüdtke[2], Adeline Paiement[3], Majid Mirmehdi[1]

[1] University of Bristol, Bristol, UK
[2] University of Rostock, Rostock, Germany
[3] Swansea University, Swansea, UK

Abstract—Smart home systems are becoming increasingly relevant with every passing year, but while the technology is more available than ever, other issues such as cost and intrusiveness are becoming more apparent. To this end, we consider the types of sensors which are most useful for fine-grained activity recognition in the kitchen in terms of cost, intrusiveness, durability and ease of installation. We install sensors into a conventional residence for testing, and propose a system which meets the design challenges such an environment presents. We show that cupboard door sensors produce useful data about access to certain non mechanical processes and items, while being cheap and simple. We also show that they positively impact the activity recognition performance of our model through their addition, while providing information that we can make use of in future studies.

I. Introduction

Smart home monitoring is becoming an ever increasing focus in many different areas of life, such as health care, consumer technology and security. However, as the need for a system which is practical for real world, large scale deployment increases, the problems which are faced by such systems become ever more apparent and challenging. In general, people are easily dissuaded from in-home monitoring by high costs, visual/physical interference in their own homes and privacy concerns. In addition, ease and cost of installation become an issue when installing systems at scale, and varying domestic configurations require systems to be extremely flexible to meet the widest range of requirements.

Considerable work has already been done by the SPHERE project [7] with regards to evaluating the cost, longevity and ethical considerations of a home monitoring system. In our work, we use a system deployed in the SPHERE House [9], which we then augment with cupboard door sensors. We hypothesize that data on cupboard and drawer manipulation is of importance to recognising activities in a kitchen, as it infers access to key cooking related objects and ingredients. Additionally, it allows for more general awareness of the kitchen environment with a lower cost than that associated with sensing individual items. Normally, these interactions are missed by smart home sensing systems or inferred inaccurately from location and movement. In our work, we use a system of miniature snap-action switches in order to gain more accurate access to this information. We aim for this system to be low cost, easy to install, unobtrusive and able to run for long periods of time without maintenance.

We begin by presenting previous works in **Background**, including an overview of the SPHERE project. The explanation of our hardware choices, integration with the SPHERE system and our experimental setup is presented in **Method and Materials**. We evaluate each of the sensor groups in the kitchen in increasing activity recognition performance with our model, and present the experimental outcomes in **Results**. Finally, we discuss the project as a whole and potential future works in **Conclusion and Future Work**.

II. Background

In this section, we first present a background overview of different sensing systems and modalities which are commonly used in SMART home setups, before giving an overview of the SPHERE project and its current state.

A. Related work on sensing systems

Previous works have looked at evaluation of sensing systems; for example Lei et al. [5] performed fine-grained activity recognition for kitchen activities using RGB-D cameras. While they showed promising results, data used was collected in a laboratory setting and had the advantage of a controlled environment with scripted experiments. In reality, getting a clear view of the counter-top in the manner they did is a non-trivial issue for real residential environments.

In a recent work funded by Google [4], a general purpose sensor was created which could be introduced into any environment easily and trained to detect a range of different events. The sensor plugs in at the wall and is capable of sensing an entire room, with a wireless connection to a central processing device. It is compact, easy to install and requires no setup in terms of hardware. However, because it is not able to directly sense any particular device or object in a room, it is highly dependant on specific room setups, especially regarding the requirement for a plug socket in an appropriate place. In the presence of high background noise or multiple simultaneous events, sensor accuracy is likely to fall. Additionally, while some activities will produce the same or similar measurements regardless of environment, and can therefore be detected without specialised training for each location, other actions are more specific to environments or individuals and so would need to be trained by users, requiring a lot of data. While this is clearly a very versatile and user friendly device, it does not

provide the coverage afforded by other more varied sensing systems.

Zouba et al. [8] made use of a range of environmental sensors attached to objects in the environment along with visual sensing components. This system is installed in a laboratory setting, which is notably uncluttered, weakening the case for this system's real world suitability. No consideration is made for the price of the system, although the issues of installation and destructiveness are addressed. Ultimately this system is used to record a dataset of actions taken within the environment, but is not tested with regards to its recognition capability or the usefulness of the setup involved, although they posit that such a system could be used in recognition.

In some real world deployments, state changing sensors are attached to commonly used household objects [17], [3] and the data collected is used to perform activity recognition, either on the entire house or just a section of it. Systems such as these are limited by the number of objects they have tagged, and often lack coverage in certain areas. The solution to this is to tag as many objects as possible, but this would be completely impractical in a real residential environment where the set of items requiring tags is constantly shifting as items are replaced through normal usage. Additionally, maintaining such a system over a long period of time would be difficult, especially when considering power requirements for each of the devices.

In more practical systems, many different types of sensors are positioned throughout the environment with the aim of capturing data to classify a number of broad activities using Hidden Markov Models [2], [18]. Some evaluation is done in these works, showing that motion sensing outperforms other sensing modalities. However, this was primarily because the actions they were sensing were location-based and simply knowing that a participant was present at a location was enough to imply that a specific action was taking place. For distinguishing between kitchen-based activities, motion alone is less useful, since kitchen activities all take place in a similar area. There was generally good performance on classifying between some coarse-grained kitchen activities by the other environmental sensors, although they are not clear on the between-class separations.

Hnat et al. [12] evaluated some of the different types of sensors which could be realistically deployed into a residential environment. Their work did not consider the cost or useful-ness of certain sensors, but primarily focused on the practical side of deployment including some helpful considerations regarding power consumption, visual intrusiveness and the need to a system which does not require constant maintenance. Ultimately, some of the issues highlighted by their paper can be solved by reducing the number of sensors deployed in the system, for which an evaluation of the sensors in order to determine those which are most useful would be pertinent.

B. SPHERE

The SPHERE project [7] is an interdisciplinary project with the remit to provide an in-home monitoring solution in order to assist medical professionals in providing care to their patients.

This involves the combination of existing technologies in order to create a complete and functioning system, as well as developing new technologies to complement and improve this system.

The system created by SPHERE and broadly described in [9], [10] has been evaluated with the express intention of extensive residential deployment, including consultation to determine consumer acceptance. This system is able to perform activity recognition on activities of daily living in a real home-setting using a range of environmental and RGB-D sensors. However, while the system is capable of integrating additional sensors, the prototype version being rolled out for residential deployment does not consider interactions with certain environmental elements, such as cupboard doors, which can limit the system's capability for fine-grained kitchen-based activity recognition. The recognition of the detailed person's actions in the kitchen, however, could be essential in detecting nutrition-related medical conditions as well as problems caused by cognitive diseases, such as dementia.

(a) Exterior of the house.

(b) The left side of the kitchen.

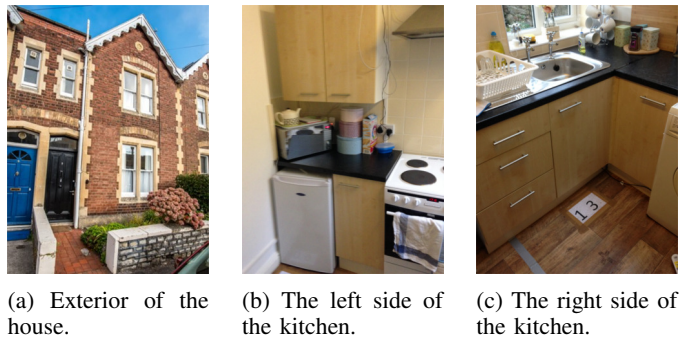(c) The right side of the kitchen.

Fig. 1: The SPHERE house in Bristol.

The SPHERE house in Bristol as shown in *Figure 1*, is a two-bedroom, terraced house owned by the University of Bristol close to the main campus. It is used for experimentation on new sensors and systems that can then be deployed in other houses. There are a number of different systems at work within the house which gather data about the current occupants. These systems are designed to be non-intrusive and automatic, with data being transmitted to the external SPHERE data hub for processing. Located in the kitchen is an RGB-D camera, electricity and water monitoring systems, and a range of environmental sensors monitoring light intensity, humidity, motion and temperature. During our project, we have augmented the system in the kitchen using our own system of switches and gateways, and integrate these changes into the house's network and data storage facilities.

III. METHOD AND MATERIALS

We begin by expounding on the hardware choices we considered, before explaining how our system was integrated into the SPHERE platform for our experiments. Finally, we outline how our experiments were conducted to collect the data required for evaluating the sensors.

A. Hardware Design

Miniature snap-action switches were chosen as the means of sensor cupboard door states. This simple binary data was collected and processed by simple, programmable development boards.

1) Sensors: When determining the hardware choice for the sensing units, several properties had to be considered. Firstly, any sensors used needed to be low cost, since a high cost would inhibit the possibility of large scale deployment. Initially, proximity sensors were considered since they could be installed at the back of the cupboards out of the way while still detecting the position of the cupboard door. However, these can be expensive for accurate models with sufficient ranges and additionally require a precise installation which may be difficult for untrained technicians. This is another important factor to consider in a large scale deployment. Data obtained from these sensors would also be noisy due to their continuous nature, susceptible to interference or obstruction and also can contain much more information than would actually be required.

Avoiding complexity was also important, since keeping the system simple reduces the opportunities for failures. For this reason, switches were chosen for their mechanical simplicity, and their production of binary data which is easier to process and transmit. These switches would need to be reliable and durable. Considering all of these requirements, miniature snap-action switches [13] are an optimal choice due to their low cost, mechanical stability and reliable activation at specific and repeatable positions. These were positioned flush with the frame of each cupboard (see Figure 2) to maximise the force acting on them from the cupboard doors. The same system was also successfully used for drawers.

Fig. 2: Miniature snap-action switches installed in a cupboard and a drawer.

2) Framework: Arduinos [1] were chosen to act as gateway components for the miniature snap-action switches, since they are low cost, low power, are easily programmable and have sufficient computing ability. Additionally, their digital pins come pre-equipped with a pull-up resistor which reduces the overall footprint size of the system.

During testing, only one Arduino was used connected to 5 switches. In the house kitchen however, two Arduinos were used connected to 9 switches in total and covering opposite halves of the kitchen. This scalability allowed us to avoid complications with wiring around the oven and other potential issues due to cables stretching across the kitchen, such as high latency and low signal strength. High gauge wire was used to mitigate some of the signal strength concerns. A plastic shell was used to protect the Arduinos from the kitchen environment, and they were placed at the back of a cupboard and behind the microwave to reduce exposure to kitchen occupants.

B. System Integration

The SPHERE house hardware infrastructure is based on the Next Unit of Computing (NUC) by Intel [14]. These fully functioning computers are extremely compact, making them highly suited to unobtrusive installation in a residence. They primarily act as gateways for other devices around the house, and are responsible for processing and relaying data.

The Arduinos from the cupboard door sensor system were connected to a NUC via two 10 metre USB cables. The NUC itself was already situated close to the kitchen making the installation straightforward. Originally, a wireless connection was considered to reduce physical location constraints, but due to the power requirements of the Arduinos and the lack of free electrical sockets, a USB carrying power and data was a better working solution. Due to the length of the USB cables, a powered USB hub was also added to aid in signal and electrical transmission.

C. Software Design

Software needed to be written from both the Arduinos and the NUC to process and relay data from the sensors into the wider system. This was performed using the Arduino IDE for programming the Arduino boards, Python for the NUC scripts and Message Queue Telemetry Transport (MQTT) for broadcasting data over the network.

1) Arduino Software: The software for the Arduinos was simple, initialising the digital pin input from the miniature snap action switches with the internal pull-up resistor enabled. This ensured stable states while the cupboards remained open or closed. Each Arduino continuously checks the state of each of its connected switches, compiling any changes into a short string and transmitting this over the USB connection to the NUC.

In addition to these event messages, a heartbeat is sent every 10 seconds to indicate liveness. Development of this software was made more convenient through the use of the Arduino IDE, using a variant of the C programming language.

2) NUC Software: All incoming data has a timestamp attached by the NUC to ensure synchronisation with the rest of the house data. In order to properly integrate with the existing smart house system, software on the NUC was written in Python allowing access to the Paho MQTT library. MQTT is a lightweight message passing system ideal for communication between networked machines, and is the primary protocol for high level data communication in the house [15]. A message

broker is responsible for logging and distributing messages, while client software is used by devices to send messages to the broker and subscribe to specific topics. Topics are given to messages to allow for filtering of the most relevant information by each machine.

In addition to being broadcast by MQTT, data collected was also stored in an SQL database and in a plain text log file. The software running on the NUC would read incoming messages from the Arduinos, unpack the data and distribute it via all communication methods to ensure parity across all data repositories. The program was daemonised to run as a service on the NUC and configured to run on start-up ensuring constant monitoring.

D. Experimental Set-up

In this section we describe the data collection setup and process, before explaining how the data was handled and processed.

1) Data Collection: To evaluate the performance of the cupboard sensors and their contribution to the recognition of everyday activities, we collected a sensor dataset, showing the preparation of meals in the kitchen of the SPHERE house.

The collected dataset contains sequences of individual human protagonists performing varied and complex activities in the SPHERE kitchen, without any predefined scripts. Additionally, no information was recorded regarding the contents of the cupboards or draws before or after the experiments. Thus, the dataset is a good example of natural human behaviour in a changing and unordered environment.

Each data collection event took place over the course of around two hours in the kitchen, involving 9 participants. The only instruction they received was to prepare a meal and/or a drink of their choice in the kitchen. This resulted in the collection of 15 unscripted meal preparation and consumption tasks. The meals/drinks included: pasta, ready meal, carrot sticks, rice and vegetables, toast, juice, tea, coffee, chicken and vegetables snack, rice and curry, macaroons, salad, and toasted cheese sandwiches. A total of 449 minutes were recorded with individual recording durations between 10 and 88 minutes.

The sensor network in the kitchen of the house collects data on *temperature*, *humidity*, *motion within the room*, and *water and electricity usage*. Apart from that, we included the *cupboard and room door sensors* to record changes in the state of the cupboards' and drawers' sensors. A head-mounted camera was used to record the actions of the participants to allow for annotation of the observations. The resulting dataset can be downloaded from [16].

2) Data Processing: The original sensor data was collected in JSON format. In order to make the data more usable, it was converted into a table with a separate column for each type of sensor and a column for the timestamp at which each reading was taken. Rows with the same timestamp were then combined as long as per sensor type there was only one unique value. As this new format produced undefined values for some sensors at a given time, any blank readings were replaced with the last known value for that sensor. The state of the most sensors

is being read at a heartbeat rate which varies from sensor to sensor, with some sensors also reading when a state changes. For that reason, we believe that this simple replacement of undefined values is sufficient. The resulting data contained identical observations for different action labels. To reduce the impact of this artefact on the model performance, a sliding window of 5 time steps with overlapping of 50% was used and the observations in this window were represented by the maximum value for each sensor in the window.

3) Annotation: To obtain the ground truth for the dataset, a head mounted camera was used to record the experiment. The video logs from the camera were later used to annotate the sensor data. In that manner, each data instance in the sensor data was assigned an action class (e.g. "put"), as well as a ground action (e.g. "put ingredients"). This annotation was later used for two purposes: for training the hidden Markov model and for evaluating the model performance.

4) Model: To evaluate whether the cupboard door sensors contribute to recognising the behaviour during the cooking tasks, we built a Hidden Markov Model (HMM), in which the number of hidden states was set to the number of action classes: *put tools*, *put ingredients*, *prepare*, *get tools*, *get ingredients*, *eat*, and *drink* plus an initial state.
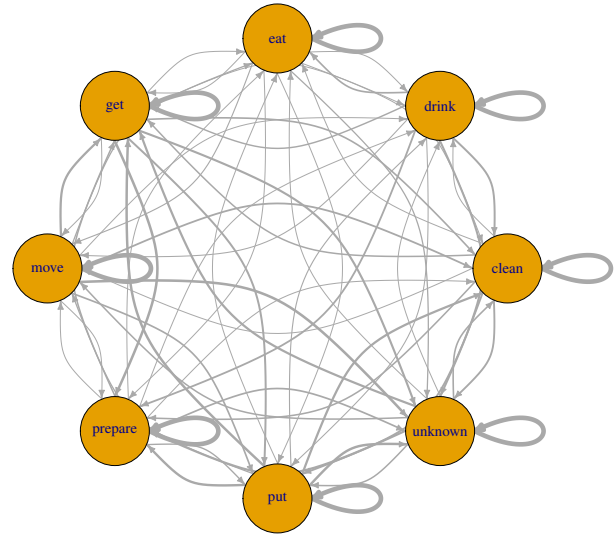

Fig. 3: The structure of the Hidden Markov Model showing the likelihood of transitioning between the different actions. Thicker lines indicate higher likelihood.

The transition model of the HMM consists of a transition matrix and priors for each state. Both have been estimated empirically from training data. For the transition matrix, the relative frequency of state transitions have been counted. The state priors are the relative frequencies of the states in the training data. To train the model, the first recording (which is the longest) was used for training and the remainder of the dataset for testing. Figure 3 shows the structure of the hidden Markov model.

To assess the performance of the cupboard sensors, we first computed the accuracy for all combinations of features ($2^{12} =$

4096 combinations) according to

$$Accuracy = \frac{\sum_C \lambda_C}{N}, \tag{1}$$

where C is the action class and N is the number of all classified instances. λ is the number of all correctly classified instances for a given class. In order to calculate λ, we used the classes as estimated by our model, and compared this to the action class labels provided in the annotation (our ground truth). Only the most likely class was considered for each estimation. Then to evaluate which features contribute the most to the model, the following procedure was performed for each feature $f \in F$, where F is the set of all available features: 1) the set of all possible feature combinations $\mathcal{P}(F)$ was generated; 2) the accuracy of recognising the executed activities $a(X)|X \subseteq F$ given the ground truth was computed, where $X \in \mathcal{P}(F)$; 3) the mean of all models that use f was compared against the mean of all models that do not use f; 4) if the mean accuracy of the models with f was lower than the mean accuracy of the models without f, we deduced that f may contain only noise.

We then investigated whether the cupboard sensors contribute to the model performance by applying a paired t-test to the results with and without the cupboard sensors.

IV. RESULTS

The mean accuracies of all 4096 feature combinations can be seen in Figure 4. The features showed accuracy between 0.269 and 0.433. This is to be expected as only one dataset was used for training the model. Nevertheless, the results show that there are feature combinations that perform considerably better than others. The best overall feature combination was ⟨"fridge", "kitchen cupboard top right", "PIR sensor", "warm water", "cold water"⟩ (accuracy 0.433). One of the cupboard sensors is also in the best feature combination, which already shows that the cupboard sensors contribute to the recognition of cooking activities.

sensor	with	without
kettle	0.375	0.385
fridge	0.371	0.389
kitchen cupboard top left	0.384	0.376
kitchen draw middle	0.382	0.378
kitchen draw bottom	0.384	0.376
kitchen cupboard top right	0.382	0.378
kitchen cupboard sink	0.384	0.376
temperature	0.376	0.384
humidity	0.352	0.408
PIR sensor	0.381	0.379
hot water	0.387	0.374
cold water	0.382	0.378

TABLE I: Mean accuracies with and without a given sensor.

Table I shows the mean accuracy when a sensor is used and when it is not used. It can be seen that the cupboard sensors all contribute to the accuracy (the accuracy with the cupboard sensors is higher than without). The small difference can be explained with the fact that some of the sensor combinations contained sensors such as temperature and humidity that

seriously reduced the performance of the feature combination. For that reason also the difference with and without a given sensor is very small.

To evaluate whether the difference in the accuracy with and without the cupboard sensors is significant, we used the paired *t-test*. The test t(4) had a t-value of 6.532 with a mean difference of 0.006. The results showed p-value of 0.003, which means that the difference is significant considering a 95% confidence interval. In other words, adding the cupboard sensors shows significant improvement in the performance of the HMM.

V. CONCLUSION AND FUTURE WORK

In this work, we proposed the use of cupboard door sensors as additional sensor modalities for activity recognition in home settings. To justify this claim, we instrumented the kitchen in the SPHERE house in Bristol with cupboard door sensors and showed that they do not reduce the activity recognition accuracy during cooking activities. In fact the results showed an improvement in accuracy when using the cupboard sensors.

In the cooking experiment, which was used to evaluated the cupboard sensors, we used simple activity classes such as "prepare" and "get ingredients". Since these are independent of location, this helps to explain the small difference between the accuracy with and without the cupboard sensors. However, due to the direct sensing capability of the sensors, we believe that they will provide invaluable additional information when reasoning about objects located at specific location. For example, opening the left top cupboard could indicate that a user has obtained a plate, while another cupboard could be more closely associated with canned foods. We plan to exploit this additional information in a more complex model that is able to reason about the objects in the environment and their manipulation through the user actions. In a previous work we proposed such a model and applied it to the annotation from the kitchen experiment [11]. We also used all available sensors to evaluate the performance of a Computational State Space Model (CSSM) [6] for the kitchen scenario. For future work, we plan to test the model also on the best feature set from the sensor data and thus better evaluate the effect of the cupboard sensors on the model performance.

Another interesting avenue of research to consider would be the use of different modelling paradigms. In this work, the model used to evaluate the accuracy with the different features was a HMM. In the future, we plan to compare the performance of this with that of the CSSM that makes use of the cupboard sensors and to see whether additional context information in the model combined with the additional information from the cupboard sensors improves the activity recognition performance. Additionally, the use of deep learning techniques for action recognition may allow for better use of our sensors, especially when temporal information is exploited as is the case with a Long-Short Term Memory (LSTM) network. Since our system is generic, this will allow us to easily adapt our data to other modelling paradigms.
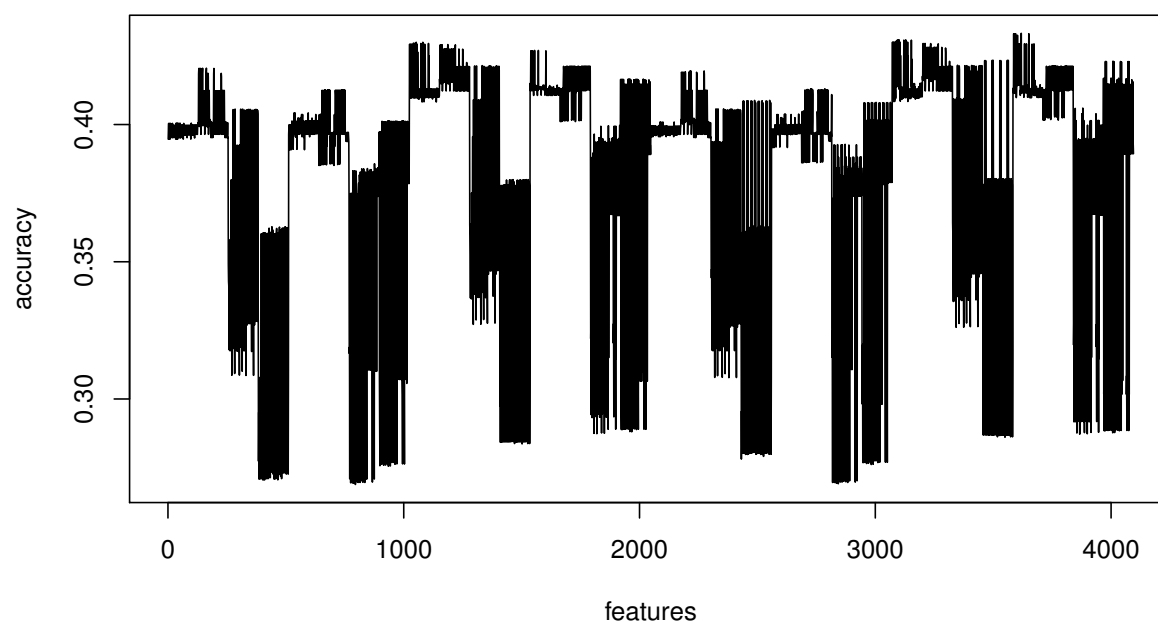
Fig. 4: Accuracies of all feature combinations for DT. Each feature index corresponds to a 12-digit binary number that represents which features are present. The order of these digits is the same as in Table I.

Finally, the cupboard sensors are not perfect and it is possible that they produce noisy observations. Some of the issues with the sensors could come from a mechanical perspective, since the precise switches used were not ideal for the irregular kitchen environment. Some of the switches were damaged by the force with which cupboards swung closed, or were simply deformed over time by repeat usage, and were then stuck in a closed position. We will seek to improve the mechanical issues with the system (such as difficulty of installation and sensor reliability) by using different types of switches and switching to a wireless system. This would bring the cupboard door sensors up to the same level as the rest of the SPHERE system, making them more suited to residential deployment and prolonged use.

REFERENCES

[1] Arduino. Arduino - Introduction. https://www.arduino.cc/en/Guide/Introduction, 2017. Online; accessed 2017-06-20.
[2] B. Logan et al. A Long-Term Evaluation of Sensing Modalities for Activity Recognition. In *UbiComp 2007*, pages 483–500.
[3] E. M. Tapia et al. Activity Recognition in the Home Using Simple and Ubiquitous Sensors. In *PERVASIVE 2004*, pages 158–175.
[4] G. Laput et al. Synthetic Sensors: Towards General-Purpose Sensing. In *CHI 2017*, pages 3986–3999, Denver, CO, USA.
[5] J. Lei et al. Fine-Grained Kitchen Activity Recognition using RGB-D. In *UbiComp 2012*, pages 208–211.
[6] K. Yordanova et al. What's cooking and why? behaviour recognition during unscripted cooking tasks for health monitoring. In *PerCom 2017*, pages 18–21.
[7] N. Zhu et al. Bridging e-health and the internet of things: The sphere project. *IEEE Intelligent Systems*, 30(4):39–46, 2015.
[8] N. Zouba et al. Assessing computer systems for monitoring elderly people living at home. In *IAGG 2009*, Paris, France.
[9] P. Woznowski et al. A Multi-modal Sensor Infrastructure for Healthcare in a Residential Environment. In *ICCW 2015*, pages 271–277.
[10] P. Woznowski et al. SPHERE: A sensor platform for healthcare in a residential environment. In *Designing, Developing, and Facilitating Smart Cities*, pages 315–333. Springer, 2017.
[11] S. Whitehouse et al. Recognition of unscripted kitchen activities and eating behaviour for health monitoring. In *TechAAl 2016*, pages 1–6.
[12] T. W. Hnat et al. The Hitchhiker's Guide to Successful Residential Sensing Deployments. In *SenSys 2011*, pages 232–245.
[13] Honeywell. Basic Switches. https://sensing.honeywell.com/switches/basic-switches, 2017. Online; accessed 2017-06-20.
[14] Intel. Intel NUC Products. http://www.intel.co.uk/content/www/uk/en/nuc/products-overview.html, 2017. Online; accessed 2017-06-20.
[15] D. Locke. MQ Telemetry Transport (MQTT) V3.1 Protocol Specification. https://www.ibm.com/developerworks/webservices/library/ws-mqtt/ws-mqtt-pdf.pdf, 2010. Online; accessed 2017-06-20.
[16] M. Mirmehdi, T. Kirste, S. Whitehouse, A. Paiement, and K. Yordanova. Sphere unscripted kitchen activities. University of Bristol, 2016. Online; accessed 2017-06-20.
[17] C. Pham and P. Olivier. Slice&Dice: Recognizing food preparation activities using embedded accelerometers. In *Aml 2009*, pages 34–43.
[18] T. van Kasteren et al. Accurate activity recognition in a home setting. In *UbiComp 2008*, pages 1–9.