Nonintrusive Occupant Identification by Sensing Body Shape and Movement

Nacer Khalil
University of Houston
nkhalil@uh.edu

Driss Benhaddou
University of Houston
dbenhadd@uh.edu

Omprakash Gnawali
University of Houston
gnawali@cs.uh.edu

Jaspal Subhlok
University of Houston
jaspal@cs.uh.edu

ABSTRACT

The ability to identify people has numerous applications including in smart buildings where the building can be customized to the needs of its occupants or for other applications such as in assisted living and customer behavior analysis in commercial settings. There are different methods used for occupant identification. Some are intrusive such as using cameras or microphone and others require the users to carry mobile gadgets to be identified. In this paper, we present a nonintrusive method to identify people by sensing their body shape and movement. Such information is derived from using ultrasonic sensors to measure the height and width as the occupant walks through the instrumental doorway. In fact, height and width are not unique to every occupant, but extracting a set of features from the variations in height and width makes identification possible. In this study, our system senses a stream of height and width data, recognizes the walking event when a person walks through the door, extracts features that capture a person's movement as well as physical shape. These features are fed to our clustering algorithm that associates each occupant with a distinct cluster. We deployed our system for 1 month. We found out that our approach achieves 95% accuracy with 20 occupants suggesting the suitability of our approach in commercial building settings. In addition, we found out that using girth to distinguish between occupants is more successful than using height.

CCS Concepts

•Computer systems organization → Special purpose systems; •Hardware → Sensor applications and deployments;

Keywords

Indoor Identification; Sensor Networks; Smart Buildings; Clustering; Machine Learning

1. INTRODUCTION

Identification and tracking are important for several Smart Building applications such as occupancy driven energy efficiency, tracking in assisted living, occupant behavior analysis and personalized comfort adjustment. Developing nonintrusive sensing technologies for the purpose of identification is challenging and has attracted interests of different research. Researchers have addressed this challenge with different approaches, including ultrasonic sensing using height [23], vibration of footsteps sensing [20, 18, 6, 9], pyroelectric infrared [3], as well as WiFi RF spectrum [26, 27]. They provide many useful insights on the non-intrusive sensing that could be useful in building smart spaces applications. Developing scalable nonintrusive cost-effective sensing technologies will enable the development of applications that improve the safety of people and save energy and maintenance cost, which will benefit both occupants and building owners. For instance, assisted living buildings for people with amnesia or Alzheimer's disease host between 15 to 20 patients per area. Cameras are used only in common areas making it difficult to track patient when they go outside the common areas which require an attendant to check on them every 30 minutes. Non-intrusive sensing technologies will give a real time view of where patients are as it can be installed in areas where cameras cannot be installed.

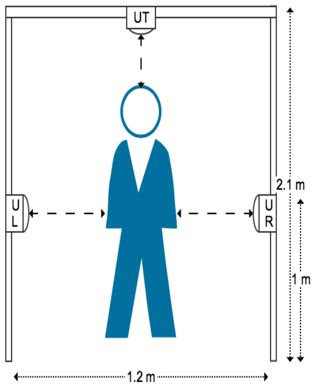
We introduce a nonintrusive indoor occupant's identification system that uses ultrasonic sensors in doorways. The sensing technique computes occupant's shape and movement which are used to identify people. The solution uses three ultrasonic sensors; one sensor is installed at the top of the door frame to measure a person's height, and two on the side of the door frame to measure the person's distance to the sides sensors. The system uses the side sensors to calculate the width of the person passing through the door. Using the height and width data, the system extracts a set of features to infer the occupant's body shape and movement. The identity is determined by a clustering algorithm using these features to associate features with occupants and thus the ability to identify is made possible. This solution is cost effective since it uses off the shelf ultrasonic sensors and is easily integrated into a door.

Clearly, such a system is subject to uncertainty as multiple people can have similar height and width and the system may fail to differentiate between them. We find that a combination of body shape and time under the door (thus movement) will increase the accuracy of the system. Our study shows that indeed these parameters are key to differentiating between people, thus, helping us achieve 95% accuracy for 20 people compared to the state-of-the-art.

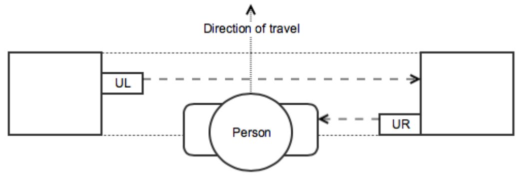
We present the system design, sensing techniques, filter design to eliminate noise from the measurements, and use a

BuildSys '16 November 15–17, 2016, Stanford, CA, USA

DOI: 10.1145/1235



(a) Figure showing the conceptual design of the sensor instrumented door frame.



(b) Top view of the doorframe showing direction detection as person walks through the door. UT has not been depicted in this figure to emphasize on the displacement between UL and UR

(c) Front view of the door frame with the ultrasonic sensors mounted.

Figure 1: Figure showing the sensor instrumented doorframe schematic figures and a photo of our current testbed.

clustering algorithm to identify occupants. We collected experimental data in a classroom of 20 people at the University of Houston over a period of one month. Our results show that it is possible to achieve an identification accuracy of 95%. Our contributions in this work are:

- We design a system that identifies occupants by combining computed girth with the time they spend walking under the door.
- We investigate the accuracy of parameters associated with height feature, namely the gait pattern bounce. Our results show that gait is more accurate than height.
- We compare the accuracy of different combinations of parameters in identifying occupants. The results show that the girth and time provide more accurate results that height and time. However, height could be used to scale up the system furthermore by categorizing people by their heights.

2. RELATED WORK

Occupant identification in buildings can be broken down to two approaches: systems in which the occupants carry a mobile gadget and systems that use sensors embedded in building to sense the occupants unobtrusively. Mobile gadgets include RFID-based wearables for tracking users when they cross a door [22], users' smartphone for identification and localization [24]. Motetrack uses RF signal strength and a set of beacons deployed in the building and infers location based on the RF signal strength [21]. BlueSentinel uses iBeacon protocol which is based on Bluetooth Low Energy (BLE) to infer the location [5]. These systems identify occupants with high accuracy but suffer from missing a user if she does not have the wearable. Wi-Fi-based systems such as Redpin and Ariel achieve high accuracy in identifying and detecting occupants' location [4, 13].
The second category of systems use sensors installed in the buildings to identify users. Different systems have been proposed that use facial, fingerprint, iris and hand geometry and achieved high accuracy [14, 12, 25, 19] but they raise privacy concerns and some require user's active interaction with the system. Among the systems that use nonintrusive sensors, the vibration of the footsteps has been proposed as a method to detect occupants [20, 18, 6, 9], it achieves very high accuracy rates however it cannot operate when multiple occupants are present in the area. WiFi RF signal has been used in previous work to identify people by harnessing the RF reflections on the occupants [1, 2]. In [26, 27], the authors use a Channel State Information (CSI) to infer gait for person identification. However, they do not scale to over six people with an accuracy of 80%. Moreover, measuring height using ultrasonic sensor is another biometric identifier in small populations [23]. Doorjamb uses height information, walking direction, and tracking information to identify users achieved a high accuracy rate within a population of 2-4 people [11]. However, our system leverages height and width to sensors to identify up to 20 people with an accuracy of 95%.

3. SYSTEM DESIGN

The system is designed to sense walkers' body shape and movement as they pass through the door. We design a door frame that implements three ultrasonic sensors (see figure 1(a) and 1(c)). One sensor is placed on top facing downwards referred to as UT (Ultrasonic Top), and other two sensors are placed on the sides horizontally at 1 meter high referred to as UR (Ultrasonic Right) and UL (Ultrasonic Left), which are at the right and left sides of the frame. The sensor-instrumented door is installed in room 219 at the Technology building at the University of Houston. The sensor-instrumented door sends the extracted features to the back end system to run the clustering algorithms and identify occupants.
The system includes four components: Sensing and calibration, event recognition, feature extraction, clustering and decision making. Figure 2 depicts the process. The ultrasonic ping sensors compute the distance between the sensor and the closest object (in this case a person) and generates three streams of readings per walker. When a walking event is detected the generated data is pre-processed and a set of features is computed and extracted. These feature instances are then used to build a clustering model for every user.

3.1 Sensing and Calibration

Sensors need calibration as signal acquisition and sam-

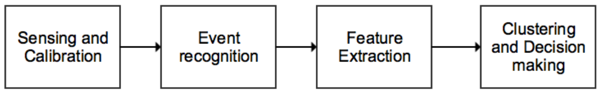


Figure 2: Sequence of operations for occupant identification in a building.

pling introduce errors in calculating the distance. For the sensor UT, the longest distance the ultrasonic pulse will travel is 4.2 m which is twice the height of the door frame as the ultrasound signal needs to go back and forth. Given the speed of sound of 341 m/s, we estimate that it will travel this distance (4.2 meters) in 12 ms, which represents the maximum possible delay. For the sensors UL and UR which are separated by a distance of 1.2 meters, the maximum expected delay is 7ms. The three sensors are sampled sequentially to avoid cross-talk between sensors. The way we sense a walking event is as follows: we send a beam from one sensor and wait for the reflection to reach the sensor, we compare the time it took with the maximum allocated time for the sensor and set the node to sleep for the rest of the allotted time. The purpose of this is to keep the sampling rate at a stable fixed rate regardless of the walker's size. This also will prevent crosstalk between three ultrasonic sensors as it allocates enough time for the beam to be sent and reflected back to the sensor before operating the next sensor. In fact, synchronization between sensors is key to avoiding crosstalk and reducing noise in distance measurements.

The measured delay from UT is converted to height using the following formula:

$d_{height}(t_{UT}) = d_{max\_height} - \frac{34.3}{2} t_{UT}$

where 34.3 is the distance, in cm, traveled by sound every 1 ms. The variable t is the delay in ms. We divide by 2 because the measured delay represents the time for the pulse to go back and forth. d_{max_height} refers to the maximum distance separating the sensor and the ground in the case of UT. The maximum distance measured by UT, UL, and UR in our testbed are respectively 212 cm, 124 cm,124 cm.

The measured distances are then converted to the width of the person passing through the door using the following formula:

$d_{width}(t) = d_{maxwidth} - \frac{34.3}{2} t_{UL} - \frac{34.3}{2} t_{UR}$

When there is no one under the door frame, then the width formula returns a negative figure, more exactly $-d_{max_width}$ because both distances computed from UL and UR will be equal to $d_{maxwidth}$. To avoid this problem, we first check if $\frac{34.3}{2} t_{UL} = \frac{34.3}{2} t_{ULR} = d_{maxwidth}$, if true we return 0, and if not we compute the width using the formula.

Our first design uses periodic polling with an interval of 29ms to simplify the implementation. Periodic polling is energy intensive and impacts the lifetime of the sensors, therefore is not the best way to operate in real world application. In a real world environment, we would add a motion sensor to optimize the operation to activate the sampling of the ultrasonic sensors when the motion sensor detects a person close to the door frame. Adding a motion sensor wouldn't significantly increase the overall cost of the system since an off-the-shelf costs as low as $5 and wouldn't add significant complexity to the system. In addition, activated motion sensing is less energy intensive than running three ultrasonic sensors continuously.

3.2 Walking Event Recognition

A Walking event refers to the stream of {UT,UL,UR} readings. Every time a person walks through the door frame, we receive a stream of data and the number of readings varies between 35 and 40 depending on the speed of the person. Figure 4(a) illustrates the data stream. The faster the person walks, the fewer the readings. This stream of data will contain noisy points and errors. These noisy points need to be corrected and recovered before processing the features. Since our testbed is in periodic polling, we get continuous data stream from the ultrasonic sensors. To detect a walking event, we look at the height as a detection mechanism. When there is no one under UT, we expect the maximum value. Algorithm 1 shows how we extract the walking event from a stream of data.

It has been shown that the average height of people in the United States is 169 cm with a standard deviation of 7.5 cm [17]. So a height interval of 3 standard deviations from the mean should statistically cover 99% of the walkers assuming height follows the Gaussian distribution. The lower end of the interval would be 146.5 cm and we chose 140 cm as a lower bound. The reason is that height measured by UT is not necessary the ground truth and many times it is lower because of how the person walks, especially if the walker is looking a bit downwards or holding a backpack or just using a smartphone. The walking event starts when the measured height is at least 140cm and stops when the height is less than 140 cm by allowing at most 4 consecutive points that are out of this interval. This last condition is chosen to prevent erroneous readings from making the system think the walker is no longer at the door and gives the impression we have multiple events. We chose 4 consecutive experimentally because it yields the most accurate walking event.

Algorithm 1 Extract Walking Event

```
procedure EXTRACT WALKING EVENT
missed ← 0
min_height ← 140
max_missed ← 4
queue ← FIFO Queue
walking_event ← empty FIFO Queue
    do
        reading ← Dequeue element from queue
        if reading.height > min_height then
            Enqueue reading to walking_event
            missed ← 0
        else if missed < max_missed then
            missed ← missed + 1
        else
            return walking event
        end if
    while queue not empty
end procedure
```

3.2.1 *Person Direction Recognition*

The width sensors UL and UR are displaced in parallel to the walking direction line. As the person is walking, one

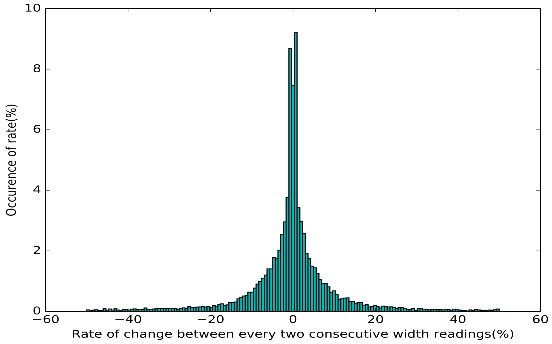


Figure 3: Histogram of the distribution of adjacent width measurements rate of change.

sensor is closer to her than the other and therefore the closest will be the one to detect her by returning a non-default (a default value means max width when no one is under the door) value indicating it came in contact with a person. Figure 1(b) shows how sensors UL and UR are displaced and how this displacement helps detect the person direction. This displacement helps not only with person direction recognition but also for more accurate width measurement as shown in Section 4.4.
The direction is set when setting up the door frame in the room. In our testbed, we positioned our door to have the sensor UR closer to the entrance, so if a person is entering the room, the first non-default reading will be from UR and when exiting the first non-default reading will be coming from UL. If a person intentionally rotates the door, then our system will start giving wrong directions, but in a real deployment, the sensors will be mounted on the door frame by drilling a hole into the door and such an intentional rotation would not be possible.

3.3 Noise Canceling and Correction

Once the walking event is recognized and its respective sensory data stream is detected, we pre-process the data to filter noisy points before the event is further processed. We filter out the readings that are outside the interval [0, Max Height] and more than 30% difference between adjacent points. The reason is that we have observed that the width varies approximatively by 15% and the latter is higher when a person is carrying a purse. Figure 3 is a Histogram that depicts the distribution of adjacent readings.
Once noisy points are identified, we need to either remove or recover them. Since the height measure uses only UT, then removing would not affect the overall height data. However, data from UL and UR are computed in pairs and removing one UR implies removing its equivalent UL or vice versa. For example, if we remove UL at t1 but not UR at t1, we will end up with more UR measures than UL measures and most importantly, we will end up with pairs that did not occur at the same time. Therefore, the best approach is to use linear interpolation to replace the noisy values. Therefore, after receiving the raw walking event data stream, we identify the noisy values and replace them using linear interpolation. The new stream of data is then used for feature extraction.

3.4 Main Features

To detect occupants, we first use the pre-processed stream of data to extract a set of features that will be used to detect and identify users. It is illustrated in Figure 4(a) and Table 4(b). We experimented with several features including max, min, average, bounce, girth, and time under the door. Girth and time under door gave the best results for occupant identification.

3.4.1 Girth

Girth is a circumference measurement around a person's waist. In order to compute the girth, we use the stream of width data to create 2 point clouds where y is the instance number multiplied by the distance traveled each sampling interval and x is equal to $\frac{width}{2}$. We generate two points, (x,y) and (-x, y). We then construct the convex hull for all the point clouds and calculate its perimeter using euclidean distance as the distance measure. Assuming an average speed of 5 km/h, a person will walk 3.6 cm every 29 ms. The pseudocode for constructing the girth is presented in Algorithm 2.

Algorithm 2 Girth Calculation Algorithm

procedure COMPUTE GIRTH
$distance\_walked\_per\_iteration \leftarrow 3.6$
$edge \leftarrow 0$
  $x_t \leftarrow \frac{width[t]}{2}$
  $y_t \leftarrow iteration_t * distance\_walked\_per\_iteration$
  $edge \leftarrow edge + \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}$
Return: $girth \leftarrow 2$ * edge
end procedure

[1]

3.4.2 Time

Since everyone has a different walking speed, we measure time indirectly by counting the number of interval time the person spent under the door. Given the sampling rate, s, (s=35Hz in the testbed), we calculate the time by dividing the number of height measurements $H = \{h_1, .., h_n\}$ by s and width measurements $W = \{w_1, .., w_n\}$ by s. We select the result that gives the max time spent under the door, t, is therefore given by:
$t = \frac{1}{s} max(|H|, |W|)$ We take the maximum because noise may alter the length of the H or W and therefore select the longest since it was the least affected by noise.

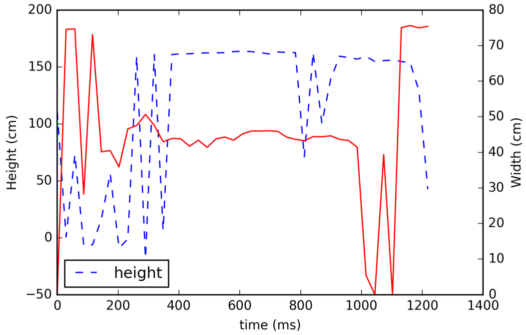
3.4.3 Bounce

Bounce is a gait measure of how a person bounces as she walks. Some people tend to bounce more than others when walking. We capture bounce from the height measurement by subtracting the minimum from maximum height. Given height measurements $H = \{h_1, .., h_n\}$, we set $Bounce = max(H) - Min(H)$.

3.5 Other Features

3.5.1 Maximum, Minimum and Average Height

From the stream of height measures, we compute the minimum, maximum and average height. To decide which of the three features is most appropriate for identification, we conducted a small experiment where the same person performs



(a) Plot showing an instance of Height and Width sensory readings as a function of time when a person walked through the door.

Feature	Value
Average Height	156.1
Maximum Height	163.8
Average Width	39.5
Maximum Width	42.1
Girth	83.1
Hand-Waist Distance	13.2
Bounce	19.7

(b) Table showing list of extracted features from the walking event height and width readings.

Figure 4: Figure showing an instance of a walking event height and width readings and the extracted Features.

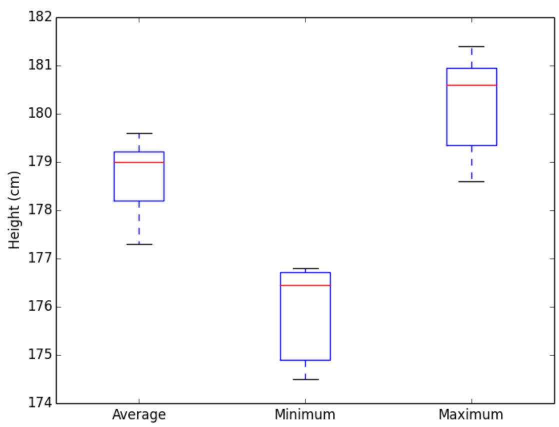


Figure 5: Box plot showing variance in average, minimum and maximum height for 7 trials.

7 walking events under the door. We compute for every event the minimum, maximum and average height. Figure 5 illustrates the data from the trials. Given a ground truth of 180cm, we can see that the maximum feature is the closest. However, we note that the average height is the feature with the least variance and this means it is more consistent for the same occupant. As consistency is important for identification, we choose average height.

3.5.2 Average Width

Once a person passes through the door we compute the person's width. The width measure is independent of the position of the occupant in the door i.e. if he is closer to one edge of the door as opposed to the other, the width measurement is still the same.

3.5.3 Body-Hand Distance (WH)

This feature captures how close a person's hands are to her body as she walks. As the walker swings her arms, the UR or UL sometimes measure the distance to the waist and sometimes to the arm. To compute this feature, we divide the measurements into 2 groups. The ones closer to by at most 10% of the minimum width and the others that are farther by at least 15%. We calculate the average of every group and return the difference. We do this for both sides using UL and UR and take the maximum. The reason for taking the maximum as opposed to the minimum or average is that you have people one one hand in the pocket.

3.6 Feature Selection

Given the feature set generated by the system, choosing a subset of features has been extensively discussed in the literature. In this study, we use two methods to find the find the most successful subset of features to use to identify occupants. First, we evaluate how would a feature set composed of one feature perform and then evaluate pairs of features' accuracy. This approach is very similar to the sequential search methods [15] where we start with one feature and add a second one to increase the goodness of our dataset.

In the second method, we perform feature selection using the Recursive Feature Elimination (RFE) algorithm [7] whose goal is to find a subset of features that maximize accuracy and increase the robustness of the identification. Since we have 3 degrees of freedom, namely height, width and time, we derive one feature from every degree of freedom to minimize interdependence between the features. Principal Component Analysis (PCA) has been widely used as a dimensionality reduction method that leverages the variance to measure the importance of features. In fact, we use PCA to generate a new set of features that are a combination of the input features to maximize the variance [10]. Therefore, we first run RFE on the experimental dataset presented in Section 4. We found that the 3 most important features are girth, time and bounce. Then, we use PCA to create a new set for building our model. To validate the model, we search for the important features by forming a model for all possible feature pairs and evaluate each model's accuracy.

3.7 Occupant Identification

If a feature computed for a given person is consistent across different instances of walking trials, this feature can be used as a unique signature to identify the person. However, the feature should also meet an additional requirement: a feature computed for different people should be

different. We process all sensor data stream and extract a set of features for the person. Since the system is unaware of the walker but rather tags her with a feature set, we treat this occupant identification problem as a clustering problem rather than classification, where every user will have her own cluster. In addition, the system does not need training to work and therefore is able to differentiate between people without prior information about the walkers.
Some of these features such as Girth and Bounce have been shown to be very consistent for the same user while other features failed. Other features such as Body-Hand distance did not vary much among different people and therefore are not be able to differentiate between different users. To address this issue, we decided to combine features into pairs and evaluate each pair to find the one that yields the highest identification accuracy.
To achieve high identification accuracy, we chose DBSCAN [8] as the clustering algorithm to use for multiple reasons:

- DBSCAN has been used in a widespread of application and proven to be a powerful clustering algorithm.
- DBSCAN has 2 parameters: *epsilon* that allows choosing how close the objects have to be. We chose a value of 2 because the precision of the ultrasonic sensor is 1 cm and since many of the features use 2 ultrasonic sensors as is the case of average width, then we expect a deviation of 2 cm on average. *Minpts* defines the minimum number of points to have in one cluster. We chose a value of four because we had occupants walk past the door for a minimum of four times. The value of four for minpts works best in our dataset.
- DBSCAN is very well suited for such a problem.

4. EVALUATION

We describe the setup used for evaluating the occupant identification system followed by evaluation results.

4.1 Testbed

The sensing testbed is composed of a door frame, three Parallax Ultrasonic ping sensors model 28015, an Arduino Uno board, and a Raspberry PI 2 Model B. Each door frame has 3 ultrasonic sensors attached to it. Figure 1 shows the testbed we built for this study. We sample ultrasonic delay values which are later converted to distance at a rate of 35Hz. We also have one Logitech C310 camera per door that is used to collect the ground truth.
The sensing is done sequentially by Arduino Uno and therefore the data is read at separate times. Since height measurement uses only UT, we do not consider the case of synchronizing UT with the other sensors as important.
The sensors are attached to a board that is attached to the frame as illustrated in Figure 1(c). However, this is not required by the system, it was designed for convenience but it would work similarly if the sensors were actually mounted inside the frame. The system is not susceptible to crosstalk because the sensor sampling is performed sequentially giving enough time for each sensor's signal to travel to the target and back. If installed in a wider door, this time parameters must be changed but could be computed given the dimensions of the door as shown in Section 3.1.
In order to compute the width of a person, we need both UL and UR readings to be at the same time in order to

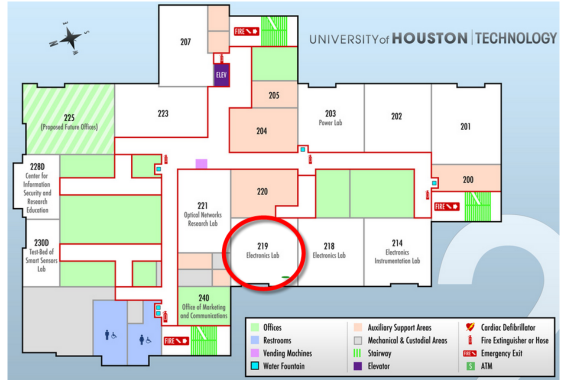


Figure 6: Figure showing the floorplan where lab room 219 (which was used for the experiment) is located.

have accurate width measure. To achieve this, we displace the sensors by 1.2cm on the walking direction in order to account for the temporal difference between the consecutive samplings of UL and UR. This specific displacement distance (1.2 cm) is chosen to account for the temporal difference in sampling. Assuming the walker walks at an average speed of 5 km/hr and having a sampling rate of 35Hz, the walker travels a distance of 1.2 cm every 8 ms. Also, the order to sampling the sensors is: $UT \rightarrow UL \rightarrow UR$ with a 8ms time difference between UL and UR. Thus, this displacement is crucial because though UL and UR are sampled at different times, the width measurements of the walker though taken at different times are the same points as if the person was standing.

4.2 Experimental Setup and Ground Truth

We conducted our experiment in lab room 219 (see Figure 6)in Building T2 at the University of Houston for a month. We recruited students from one of classes scheduled in the room as participants. Their age varies between 18 and 30 years old. We informed them about the purpose of the experiment and we asked them to walk naturally. The protocol was approved by the University of Houston Committee for the protection of Human Subjects. The door frame was at the entrance of the room and there was enough space for others to bypass it in case they don't want to participate. The camera was always recording. Whenever a person walks through the door, a walking event is recorded with the start and end time. Since we only keep records of when a person walks into the door frame, every minute, the video footage is processed and only the times when the walker(s) passed is extracted. We keep the video recording starting 3 seconds before the walking event start time. This extra time is added on purpose so that we can see the whole event when annotating the data.
Every time a person walks through the door, a stream of data {UT, UL, UR} as shown in Figure 4(a) and video is captured. Each stream is then converted to the set of features extracted from the dataset. After one month which marked the end of the experiment, we annotated the data manually by looking at the video footages for every event and marked the data with the person that walked.
The number of people that participated in this experiment is fifty three. However, many of these participants passed through the door only once or twice. We discard data for

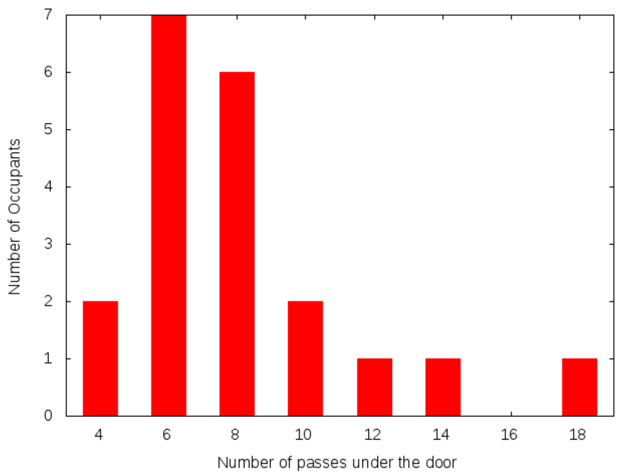


Figure 7: Histogram showing the number of participants by number of passes under the door frame during the study.

those participants from our dataset as the clustering algorithm expects at least four points per cluster. We decided to take top 20 people in terms of the number of walking events. This group averaged 7.5 passes per person, with a maximum of 17 passes and a minimum of 4; Figure 7 shows the distribution. Eleven participants were male. 9 were female. We did not measure the participants' true heights or width but it appeared to us that there was a fair distribution of body shapes.

4.3 Evaluation Metric

Since we model our system using an unsupervised method, training data is not required. Evaluation metrics such as Purity [16] have been proposed in literature. In fact, Purity is calculate as the ratio of the count of the most frequent label as a total number of labels in a particular cluster. However, this metric is unsuitable to evaluate our system because knowing how pure our clusters are does not indicate about how well our algorithm is able to identify occupants. We decide to evaluate it by dividing the dataset into a training and a testing dataset: 2/3 of the data for training and the remaining 1/3 for testing.
The issue with clustering is that it can create the correct number of clusters but may end up having different people in one cluster thus creating impure clusters. That said, we need to first find the feature pair that is able to generate the correct number of clusters and for those pairs, we then associate a cluster label with the class with the majority class. For instance if a cluster is composed of instances of different classes, then we label the cluster with the most frequently observed class in the cluster. Even if some clusters may be less pure, then this would affect the overall accuracy at the testing phase. Therefore the most successful feature pair would yield the purest clusters and highest accuracy.
Ideally, we should expect two instances of the same class to belong to the same cluster. In other words, we would like two different walking events of the same person to belong to the same cluster. All the instances of possible True/False Positives/Negatives are illustrated in table 1: We define accuracy as:
$accuracy = \frac{TP+TN}{TP+FP+TN+FN}$.

Table 1: Different outcomes from clustering

	Same cluster	Different cluster
Same Person	True Positive (TP)	False Negative (FN)
Different Person	False Positive (FP)	True Negative (TN)

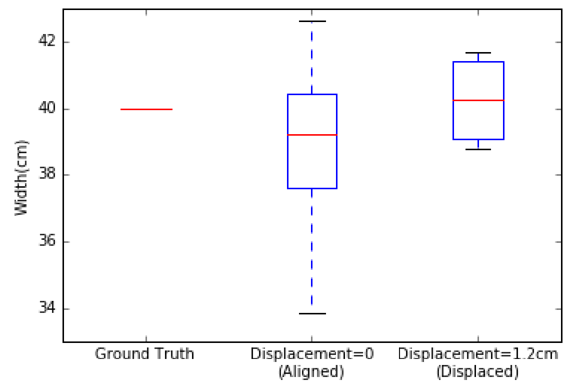


Figure 8: Box plot comparing width measurements when the sensors UL and UR are aligned Vs Displaced

4.4 Width Sensors Positioning Evaluation

The objective of this section is to evaluate if displacing the sensors UL and UR (as depicted in Figure 1(b)) or aligning them for measuring the walker's width is useful. We conducted an experiment where one person performed 6 passes through the doorway with the sensor aligned and repeated the same procedure having the sensors displaced.
With the width sensors were aligned, the occupant performed 6 passes under the door with the hands raised. The hands are raised to not bias the width measurements. The same operation is repeated but with the sensors UL and UR displaced by 1.2 cm as explained. The same occupant walked for 6 times. Figure 8 illustrates the result of the experiment.
We can see that displacing the sensors improves the accuracy of width measurement. However, if we install the sensors in an aligned fashion, we get more variation in width measurements. In addition, the average width for all 6 passes is 40.25cm and 38.50cm respectively for the displaced sensors and aligned whereas the ground truth is 40cm.
The reason why the aligned sensors generate more variation is because since we sample sequentially, after sampling from UR, the walker's position has changed by the time sample with UL. Therefore, both readings do not refer to the true width and depending on how fast the person walks or the direction (for example if he get's closer to one side as he walks), the width measurement will vary more.

4.5 Clustering with Single Features

We built a clustering model using one feature to evaluate how much accuracy can we get using one feature. Table 2 illustrates the result of using a single feature to identify occupants in a group of 20 people. None of the features was able to get 90% accuracy though clustering with girth appears to get close to 90%. Even though clustering with average height has been used in previous studies to identify occupants in smaller groups of 5 people [11], we find that it

Table 2: Accuracy achieved by clustering using different features.

Feature	Accuracy
Average Height	84.3%
Bounce	88.1%
Average Width	87.6%
Girth	89.5%
Time	82.6%
Body-Hand Distance	76.9%

Table 3: Accuracy (in %) achieved by clustering with feature pairs constructed from the features in row and column.

	Height	Width	Bounce	Time	Girth	WH
Height	84.3	89.5	89.5	90.5	93.2	86.4
Width		87.6	90.5	91.0	93.7	87.2
Bounce			88.1	87.6	94.7	89.4
Time				82.6	95.4	85.2
Girth					89.5	90.3
WH						76.9

is better to cluster with Bounce rather than average height to identify people as the identification accuracy is higher.

4.6 Clustering with Pairs of Features

We combined pairs of features and created a clustering model for every pair using the training set and test the accuracy of every model based on a pair of features. Table 3 illustrates the result. The pair (girth, time) achieved the highest accuracy. Clustering with Girth achieved the highest accuracy in single feature clustering, so pairing it with another feature seems to increase accuracy. However, time is not accurate, but combining it with girth is most accurate because because they do not commit the same mistake and people with close girth values appear to have different time values and vice versa. We should note that the pair (Bounce, Average Height) performs better than (Average Height, Average Width) which proves that it is better to use Bounce over Average Height. In Figure 9, the confusion matrix of clustering with the pair (Girth ,Time) is illustrated. The darkness of the color indicates the percentage of the trace of the i^{th} person ($P_i, i = 0..19$)which was recognized as the $i = j^{th}$ person ($P_j, j = 0..19$) Using the confusion matrix, we can observe how each person is identified and misidentified. The confusion matrix shows that most of the occupants are correctly identified all the time with the exception of person 2 and 9 which seem to confused with 2 other occupants. We also observe some misidentification for 14 and 16.

4.7 Clustering with Three features

Using the RFE feature selection algorithm along with PCA as shown Sections in 3.6 and **??**, the three most important features are Girth, Time and Bounce.
Afterward, we used Principal Component Analysis (PCA) and passed our data with only the features Girth, Time and bounce. PCA performs a linear transformation on the features and the resulting matrix has new latent variables based upon a combination of the old ones.
We, therefore, build our DBSCAN model using 2/3 of the data as we performed in the previous cases and test the accuracy using the resulting model. Our new model achieved an accuracy of 95.5% which is an increase of 0.1% over the

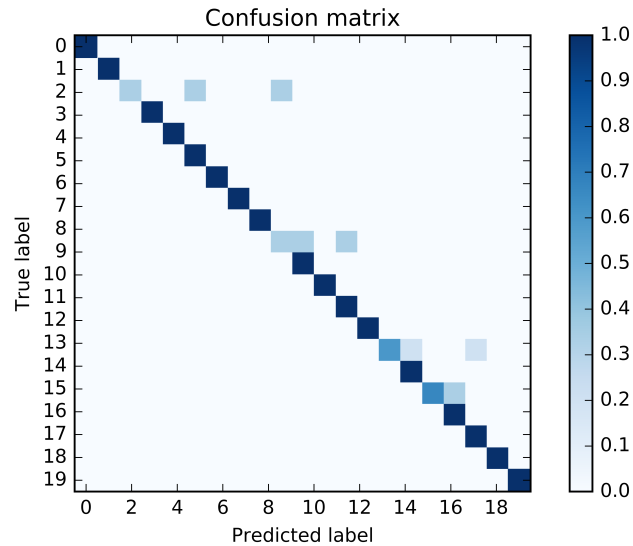


Figure 9: Confusion matrix showing the identification and misidentification of the occupants and how which occupants was identified as another.

previous model that uses only Girth and time.
We take 2 lessons from the following result:

- Bounce is the best feature that is derived from the UT, which means that bounce is better able to identify a person than maximum height as seen in previous studies [11]. In fact, clustering with Bounce only achieved 88.1% whereas average Height achieved 84.3%.

- Height does not significantly improve the overall accuracy and therefore can be omitted from the study because we would be able to reach the same accuracy with only 2 sensors as opposed to three.

Thus we recommend to use only the two main features Girth and time instead of including Bounce because the benefit from including it is small.

4.8 Accuracy with Larger Number of Occupants

To evaluate the efficacy of the technique as a function of the number of users, we calculated the accuracy of the technique in different population sizes. Figure 10 shows the accuracy as a function of the number of occupants. As expected, the accuracy decreases as the population size increases. We can observe that using the pair (Girth, time), we are able to differentiate between people with an accuracy of 97% in the case of 5 people. Our system achieves a slightly higher accuracy for the same number of people compared to systems such as Doorjamb [11] which achieves 93% accuracy and Pan et al. footstep induced identification system which achieves 96.5% accuracy[20]. Figure 10 shows the plot of accuracy as a function of the number of occupants for 2 clustering models. We show how the accuracy of clustering with (Girth, Time) changes as the number of occupants increases. We also compare it to clustering with Average Height. We observe that clustering with the (Girth, Time) not only achieves higher accuracy, but does maintain

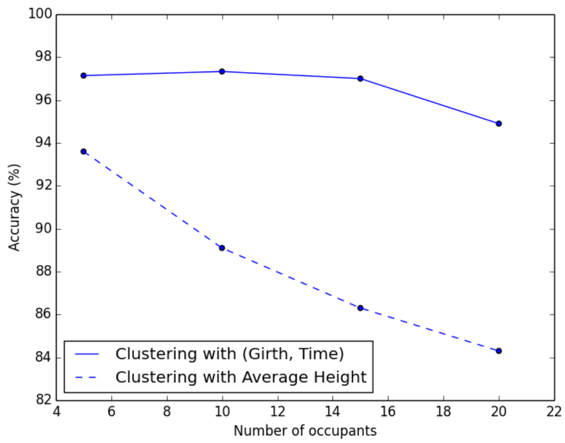


Figure 10: Plot showing identification accuracy of clustering with (Girth, Time) and clustering with Average Height as a function of the number of occupants.

higher accuracy compared to clustering with average height when we increase the number of occupants.

4.9 Robustness to Walking Angle

Girth is the most successful feature in differentiating between occupants as is shown in Table 2. One of the strengths of girth as a feature is its ability to not change with the direction of the walker. In fact, all width measure (minimum, maximum and average) suffer from the direction of the walker. The measures change drastically if a walker walks straight towards the door or at an angle relative to the door. However, girth does not appear to suffer from the direction because it represents a circumference of the person's waist and therefore is insensitive to the angle of walking when we make the measurements. We conducted an experiment where one walker walks through the door at the angles relative to the door of 0,45 degrees and 90 degrees. For each angle, the walker passed six times. For each pass we computed the walker's girth having the sensors UL and UR displaced. Figure 11 shows a box plot of the girth computed for every pass and every angle. We observe that the mean girth does not vary much regardless of the angle at which you walk. Also, most of the girth measures fall within less than 1 cm away from the mean. We conclude that the girth is not sensitive to the angle at which the person walks which make it a practical feature to differentiate between walkers.

4.10 Accuracy of Walking Direction

In an experiment with 30 walk-throughs (15 times each direction), we found that the system can determine the direction of walk-through at 90% accuracy for walking speed faster than 5 km/hr and 100% accuracy for slower walk-throughs.

5. PRACTICAL CHALLENGES

We discuss a set of challenges that we face in real world deployment of the door frame. We also discuss ways to tackle these issues.

Multiple entries In this study, we assume that only one occupant passes through the door at a time. The features

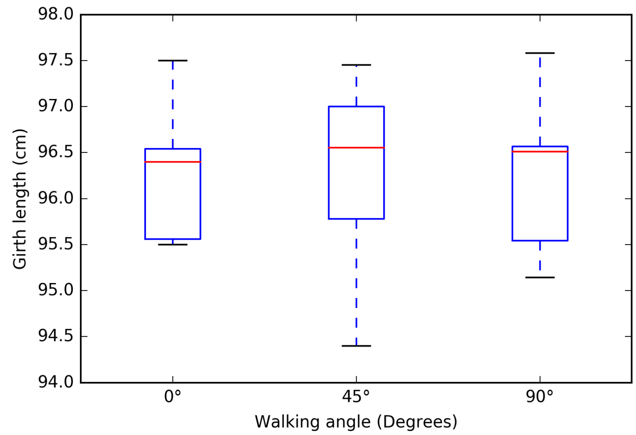


Figure 11: Box Plot showing Girth sensing distribution for three walking multiple angles relative to the door.

extracted assume that there is only one person. If more than one person walks through the door simultaneously, the whole stream will be seen as one person with an unusual width and time. To deal with multiple simultaneous entries per door, our system needs to disaggregate the data into multiple walkers. In the case of large doors that are designed for multiple entries, we can extend them by adding an extra UT to capture the second walker.

Higher number of users Previous methods have scaled their system up to 5-6 people whereas we were able to accurately identify up to 20 people. However, as the number of users increases past 20, the identification accuracy would decrease because the similarity between different features tends to become more probable. However, Height was not used for identification and therefore could be used to push further the number of people by grouping them by height and considering them as different groups.

Impact of belongings A person carrying a backpack or a woman holding a purse will be reflected in the data and drive the overall identification accuracy down. However, the bias arising from extra objects follows a pattern and can be removed. For instance, a purse can be detected by noting a higher body-hand distance on the hand carrying the purse compared to the other hand. In the case of a backpack, the height sensor will show a unique pattern showing a person carrying a backpack. These cases could be solved individually by identifying them first and pruning the data from such a bias.

Impact of Walking pattern If a person walks faster or slower, the data stream length will be impacted because she will spend more/less time under the door. However, in our studies we make two observations: (1) speed among the same person rarely varies outside of the average mean time $\pm$ sampling time, (2) it was observed that only three participants had different speeds. Sometimes the features from one subject created multiple (in our data up to three) clusters corresponding to different walking speeds and patterns for that subject. However, each cluster always was associated with a single subject. Thus there was never an ambiguity in mapping from a cluster to an individual. We believe this result holds in general but we have not performed experiments in other settings to confirm.

People with disability Our current deployment does not account for people with special disability. For instance people with wheelchairs would appear as having the same width and height and may fail to distinguish between them. Moreover, People with crutches usually walk slower. However, we may be able to detect such cases by observing a more square girth shape rather than a regular oval shaped one.

Low Power Sensing In the current setup, the sensors sample continuously and independently of whether there is an occupant. However, this would pose a problem in a real-world deployment because the current setup is energy inefficient. We suggest adding a passive motion detector and only when a person is detected, we activate the ultrasonic sampling. This would make the system much more energy efficient because only when a person approaches the door that we start sampling.

6. CONCLUSIONS

We designed and implemented a system that uses ultrasonic sensors in doorways to identify occupants in commercial buildings. The door measures body shape and movement as the features for identification purposes. We deployed the testbed for one month in a lab room at the University of Houston. Our system was able to identify people with an accuracy of 95% in a group of 20 people.

The sensors generate a stream of height and width measures whenever an occupant passes under the doorway. From this stream, we extract a set of features that capture body shape and movement of the occupants. We investigated different features and their combinations and found that the combination of the person's girth and her walking speed is the best way to distinguish among 20 people with an accuracy of 95%. We also found that clustering using bounce as a feature is more successful than average height in identifying occupants. However, bounce did not contribute significantly to increasing the accuracy when included as a third feature for clustering. On the hand, Height could be used to group occupants in order to scale up the system to larger populations.

7. REFERENCES

[1] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)*, 34(6):219, 2015.

[2] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller. 3d tracking via body radio reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, pages 317–329, 2014.

[3] I. Al-Naimi, C. B. Wong, P. Moore, and X. Chen. Advanced approach for indoor identification and tracking using smart floor and pyroelectric infrared sensors. In *Information and Communication Systems (ICICS), 2014 5th International Conference on*, pages 1–6, April 2014.

[4] P. Bolliger. Redpin-adaptive, zero-configuration indoor localization through user collaboration. In *MELT'08*, 2008.

[5] G. Conte, M. De Marchi, A. A. Nacci, V. Rana, and D. Sciuto. Bluesentinel: a first approach using ibeacon for an energy efficient occupancy detection system. In *BuildSys'14*, 2014.

[6] C. DeLoney. Person identification and gender recognition from footstep sound using modulation analysis. 2008.

[7] J. Doak. *An evaluation of feature selection methods and their application to computer security.* University of California, Computer Science, 1992.

[8] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[9] J. T. Geiger, M. Kneißl, B. W. Schuller, and G. Rigoll. Acoustic gait-based person identification using hidden markov models. In *Proceedings of the 2014 Workshop on Mapping Personality Traits Challenge and Workshop*, pages 25–30. ACM, 2014.

[10] A. Ghodsi. Dimensionality reduction a short tutorial. *Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada*, 2006.

[11] T. W. Hnat, E. Griffiths, R. Dawson, and K. Whitehouse. Doorjamb: unobtrusive room-level tracking of people in homes using doorway sensors. In *SenSys'12*, 2012.

[12] A. K. Hrechak and J. A. McHugh. Automated fingerprint recognition using structural matching. *Pattern Recognition*, 23(8):893–904, 1990.

[13] Y. Jiang, X. Pan, K. Li, Q. Lv, R. P. Dick, M. Hannigan, and L. Shang. Ariel: Automatic wi-fi based room fingerprinting for indoor localization. In *Ubicomp'12*, 2012.

[14] A. Lanitis, C. J. Taylor, and T. F. Cootes. Automatic face identification system using flexible appearance models. *Image and vision computing*, 13(5):393–401, 1995.

[15] H. Liu and L. Yu. Toward integrating feature selection algorithms for classification and clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 17(4):491–502, 2005.

[16] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval.* Cambridge University Press, New York, NY, USA, 2008.

[17] C. L. Ogden, N. C. for Health Statistics (US), et al. *Mean body weight, height, and body mass index: United States 1960-2002.* Department of Health and Human Services, 2004.

[18] R. J. Orr and G. D. Abowd. The smart floor: a mechanism for natural user identification and tracking. In *CHI'00 extended abstracts on Human factors in computing systems*, pages 275–276. ACM, 2000.

[19] S. Pan, A. Chen, and P. Zhang. Securitas: user identification through rgb-nir camera pair on mobile devices. In *SPSM'13.*

[20] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, and P. Zhang. Indoor person identification through footstep induced structural vibration. In *HotMobile'15*, 2015.

[21] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In *MobiCom'13*, 2000.

[22] J. Ranjan, Y. Yao, and K. Whitehouse. An rf doormat for tracking people's room locations. In *Ubicomp'13*, 2013.

[23] V. Srinivasan, J. Stankovic, and K. Whitehouse. Using height sensors for biometric identification in multi-resident homes. In *Pervasive computing.* Springer, 2010.

[24] K. P. Subbu and N. Thomas. A location aware personalized smart control system. In *IPSN'14*, 2014.

[25] C.-l. Tisse, L. Martin, L. Torres, M. Robert, et al. Person identification technique using human iris recognition. In *Proc. Vision Interface*, pages 294–299, 2002.

[26] Y. Zeng, P. H. Pathak, and P. Mohapatra. Wiwho: Wifi-based person identification in smart spaces. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 1–12. IEEE, 2016.

[27] J. Zhang, B. Wei, W. Hu, S. S. Kanhere, and A. Tan. Human identification using wifi signal. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 1–2. IEEE, 2016.