



A Look at Feet: Recognizing Tailgating via Capacitive Sensing

Dirk Siegmund^(✉), Sudeep Dev, Biying Fu, Doreen Scheller,
and Andreas Braun

Fraunhofer Institute for Computer Graphics Research (IGD),
Fraunhoferstrasse 5, 64283 Darmstadt, Germany
{dirk.siegmund,sudeep.dev,madapur.devraj,biying.fu,
doreen.scheller,andreas.braun}@igd.fraunhofer.de

Abstract. At many every day places, the ability to be reliably able to determine how many individuals are within an automated access control area, is of great importance. Especially in high-security areas such as banks and at country borders, access systems like mantraps or drop-arm turnstiles serve this purpose. These automated systems are designed to ensure that only one person can pass through a particular transit area at a time. State of the art systems use camera systems mounted in the ceiling to detect people sneaking in behind authorized individuals to pass through the transit space (tailgating attacks). Our novel method is inspired by recently achieved results in capacitive in-door-localization. Instead of estimating the position of humans, the pervasive capacitance of feet in the transit space is measured to detect tailgating attacks. We explore suitable sensing techniques and sensor-grid layout to be used for that application. In contrast to existing work, we use machine learning techniques for classification of the sensor's feature vector. The performance is evaluated on hardware-level, by defining its physical effectiveness. Tests with simulated attacks show its performance in comparison with competitive camera-image methods. Our method provides verification of tailgating attacks with an equal-error-rate of 3.5%, which outperforms other methods. We conclude with an evaluation of the amount of data needed for classification and highlight the usefulness of this method when combined with other imaging techniques.

Keywords: Capacitive sensing · Human machine interface
People counting · Mantrap portal · Tailgating

1 Introduction

Systems for safe separation of individuals have are of high importance at the entrances of security zones of all kind. They are used at access points of critical infrastructure, public transportation, event locations as well as in business and military areas with high security levels. Autonomous access control gates are being used more and more at accessing areas with a high security level. People

with permission pass through this designated transit space to access a secured area. The main advantage of these systems is foremost robustness against social engineering - meaning that an authorized person takes an unauthorized person into the secured area (tailgating).

So called mantrap portals (see Fig. 1a) provide a closed area with two doors, one as an entrance and another for leaving this area. Permitted subjects enter and close the portal, so that software can verify that only one subject is present in the transit space. To authorize them, biometric information, PINs, or passwords are used. After a successful verification, the system unlocks the final door to give access to the secured area.

Lower security-level areas and/or where enforcing of one-way traffic is needed however prefer to use turnstiles or access gates. Turnstiles allow only one person at a time to pass the secured area by using a mechanical ratchet mechanism. Disadvantages of this type of separation are its inconvenience for handicapped people and unauthorized tailgating. Optical turnstiles try to overcome these problems by using infrared beams to count individuals. But still, their mechanic functionality is easy to overcome, unaesthetic and inconvenient to some people. Therefore optical drop-arm turnstiles (see Fig. 1b) are increasingly used. They allow to pass individuals and handicapped people with less delay than all other methods [1]. Larger cabinets allow people to carry trolleys, bags and luggage. But even these systems are easy to overcome by tailgating or piggybacking as our review of the latest achievements shows (see Sect. 2).

This work provides a new technical viewpoint for the separation of individuals in access gates. We will show that capacitive in-door localization is a useful technique to be used in this application. Our technical approach investigates the appliance of capacitive sensing to be used in mantrap portals, turnstiles and drop-down turnstiles. We took into account that people purposefully attempt to overcome the system. Usability in this use-case is of high importance, as physical barriers are safety-relevant. Therefore, we will also evaluate how quick our technique allows users to pass through. We incorporate the use of capacitive sensors mounted on a grid in the floor. In the data classification, we evaluated two machine learning techniques to be used in this application. Furthermore, limitations regarding the distance and environmental effects are examined. Our related work, (see Sect. 2) gives an overview about previously developed systems for verification of a single person access verification and capacitive in-door localization. We will show their results obtained and outline their limitations. In Sect. 3 we will explain what hardware we use in our prototype and how people interact with it. We also introduce our sensing approach and explain the arrangement of the sensors. This includes our data-classification method. In our experiments, in Sect. 4.2 we give detailed information about the data acquisition process, the experiments and the examined use-case. Our results in the same Sect. 4.2 show the performance of our proposed method. We conclude with a comparison with other approaches.

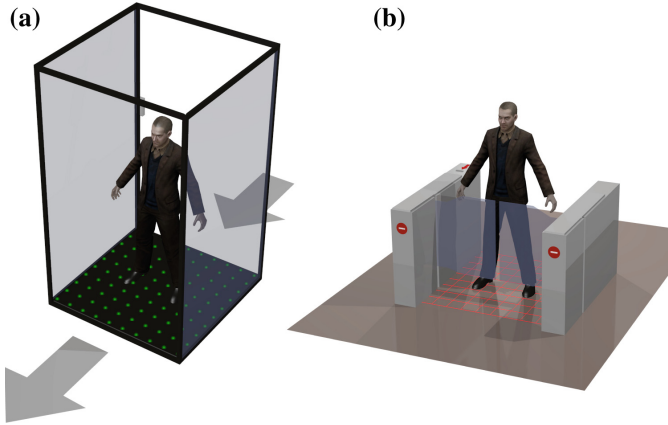


Fig. 1. (a) Mantrap portal (b) Drop-down turnstile

2 Related Work

In the field of secure entrance systems, a variety of different sensing technologies and computer-vision algorithms already exist. Most systems that allow only one person to pass at a time, are based on infrared beams [2], scales [3] or photo sensors [3]. Weight based systems must allow a certain range of weight or require an identity claim to pass the system, therefore they are not flexible to weight variations. Infrared beams are most commonly used at optical drop-arm turnstiles where they get mounted at waist level. They have the disadvantage that they are easily overcome by jumping or walking closely together. As a result of this, recent computer-vision systems have been developed that are based on different sensor data such as: thermal imaging, infrared RGB-D and color image sequences. These methods use a top-view perspective and pattern recognition to distinguish between one and more than one subject in an observed area.

A thermal imaging approach [4] showed disadvantages especially in test cases, where the attacker carried equipment (like a mirror or helmet) to hide themselves. In an evaluation using a test group that was trying to trick the system, Equal Error Rates (EER's) of 20.2% (all analyzed scenarios) and 7.9% (scenarios without equipment) were achieved. Another drawback of this system is the temperature dependence which leads to a malfunction at high ambient temperature. In a method using RGB-D images [5], a combination of change, blob detection and machine-learning is used to create a model of a single subject. It showed limitations in terms of the body height inside the chamber to the height of the installed cameras. A method using an image sequence of 21 frames was presented in a latter method [6]. It uses optical flow to make use of the effect of micro movements within the images over time. The classification results of this system show an overall error rate of 5.17%, evaluated in different attack scenarios. It has the disadvantage that people need to stand still for a certain time, which makes this solution not suitable for an application with high people

flow rates. Drawbacks of all these image based methods are: the time needed for verification and the camera position which allows subjects to hide on the floor or between the legs of a permitted person. The method shown here is intended to detect such manipulation attempts by using recent research results in the field of floor-based indoor-localization.

Precise localization within a closed building or apartment is of vital importance. It offers a wide range of new opportunities regarding smart home applications and efficient energy control. Floor-based indoor localization systems have the advantage of unobtrusive detection or sensing of the surrounding. They are mostly applying pressure-based measurement principles or capacitive sensing technologies. Feng et al. [7] offered a proof-of-concept study applying fiber optics for localization. Their grid displacements of optical fibers are distributed on the floor. Due to the pressure applied to the fibers through step motions, the signal throughput changes and thus allows the system to localize a person within the measuring area. However, the disadvantage of such a system is its maintenance. Exposed to external forces, the sensors can be easily damaged. To overcome the limitations of a pressure based system, various floor-based capacitive sensing have been announced with promising results in recent years. SenseFloor [8] introduced by Steinhage et al. worked with modular capacitive measurement units to detect the presence and location of a person walking within the sensing environment. This system is unobtrusively sensing the presence of individuals. Another floor-based indoor positioning system using grid layout, instead of modular setups called Capfloor [9] was introduced by Braun et al. in the year 2011. The advantage of this system is its easy maintenance, since a malfunctioning sensor can easily be replaced instead of replacing a whole floor module. The resulting system is more efficient in terms of power consumption while maintaining the precise localization ability [10]. Overall, active capacitive measuring systems are more efficient for remote sensing and therefore more robust against pressure.

Floor-based systems encouraged us to develop a system based on active capacitive sensing embedded in the floor to solve the problem of recognizing tailgating issues. However, most research work on capacitive indoor localization systems offer only low resolution due to the large size of measuring electrode used, which is inadequate for our targeted application. TileTrack as introduced by Valtonen et al. in [11] can locate a standing human with at least 15 cm accuracy by using 9 separately controllable 60 cm \times 60 cm tiles placed in a 3 \times 3 m square area. However, TileTrack as well as Capfloor are not able to locate people standing close to each other, which is needed in our use-case. A far better resolutions as ordinary capacitive indoor localization systems is therefore needed. To overcome this limitation we built a system using better resolution cells. The classification method used in their methods is furthermore hardly transferable to our method. As single values getting analyzed in Capfloor and TileTrack, this method seems costly with higher resolution.

3 Capacitive Sensing Grid

As discussed in the previous chapters, imaging methods have shown good results on detecting tailgating using top-view mounted cameras, but lack efficiency in several other attack scenarios (e.g. hiding on the floor). Furthermore, they are not applicable to locations with high flow-rates, where fast verification is required. We incorporate a novel approach using a grid of capacitive sensors for detecting and classifying capacitive resistance on the floor to recognize tailgating and provide easier access for handicapped people. Capacitive sensing has different properties that influence their applicability. Especially the sensibility to humidity and distance are challenging. In the following sections, we show how we face this limitation. It contains 1. a short review about sensing theory (see Sect. 3.1) 2. our sensor-grid hardware (see Sect. 3.2) and 3. our method for classification of the collected data (see Sect. 3.3).

3.1 Capacitive Sensing Theory

Capacitive sensors are proximity sensors that detect nearby conductive objects by creating an electric field [12]. The technology is based on the capacitive coupling that takes the capacitance produced by the human body to an electrode as an input. This way, it is possible to detect and measure anything that is conductive or has a dielectric difference from air. The measured capacitance is a function of the distance (d) of the object to the electrode, the area of capacitive plates (A), and dielectric the constant (ϵ_r) of the material between object and electrode; Therefore:

$$C = \frac{A}{d} \cdot \epsilon_0 \cdot \epsilon_r \quad (1)$$

Capacitive sensing is divided into three categories based on their modes of operation (shunt mode, loading mode, transmit mode). We use loading mode sensors [13], because they have a large range and are easy and cheap to implement. Loading mode sensors deliver a constant current to the attached measurement electrode. The time needed to charge the electrode up to 80% and discharge the electrode down to 20% is measured. If an object approaches the electrode, the capacitance becomes larger and the time needed to charge/discharge the electrode rises.

3.2 Our Sensor-Grid Hardware

We propose the use of loading mode sensors because they provide a continuous signal for analysis compared to others. The sensor consists of a microchip providing UART (Universal Asynchronous Receiver Transmitter) for communication. A MSP430 micro-controller is used to process the sensor values to binary format. A sensor requires between 1 to 2mA at 5V, therefore a 5V USB connector is sufficient as a power supply. All sensors are connected in a chain to the UART data-bus which is read at baud rate of 115,200 per second. As the form of the



Fig. 2. Our capacitive sensor with UART interface.

electrode has a high impact on the distance and sensitivity, we have conducted several tests in order to find the right shape (see Sect. 4.1) (Fig. 2).

We assumed having a square area of 800×800 mm as the area to be analyzed. We propose using sensors mounted in the floor of the transit area, located in a grid used for the alignment. The sensors are mounted in the middle of each cell at a distance of 100 mm between each sensor. In Fig. 3 our hardware is shown as a wooden prototype. The grid is placed inverted on the wooden board, so that the sensors are facing the ground. External plywood pieces with specific dimensions are attached to support the whole structure. It stops the breakage of sensors and problems with the wire cabling. As the sensors are not visible to the subjects, they will have no direct impact on their general functionality. The front plate consists of a medium-density-fiberboard with a thickness of 12 mm.

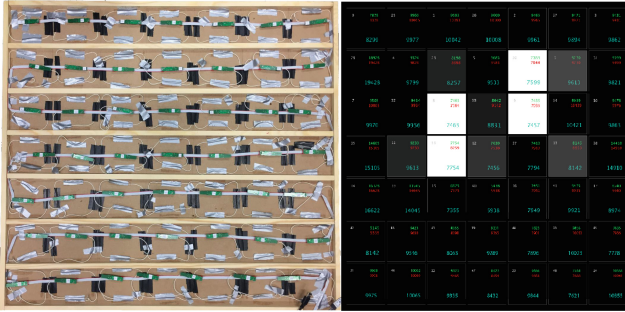


Fig. 3. Sensing grid prototype with cable-shape electrodes.

The sensing distance of the capacitive sensors depends on following factors: (1) sensor diameter, (2) sensor design (with/without GND electrode), (3) material of the medium to be detected and (4) the size of the developed body. On one hand, a bigger electrode increases the range of the sensor and reduces the effect of noise in the signal. On the other hand, a higher sensor-range causes indifferences between sensors. Reading sensors at the same time, which are arranged close by, results in wrong capacitance values. The range of sensors used is therefore limited by the density of the sensors on the grid. In our hardware we increased the range of the sensors by reading them in order like a chessboard. Only diagonal neighbored sensors are read at the same time. Consequently the maximal

distance between two sensors is reduced to around 140 mm. For choosing electrodes, we evaluated the use of a cable ring and a solid copper plate. Our results are shown in Sect. 4.1.

3.3 Data Analysis

Each sensor receives an edge like signal from the timer that indicates its individual capacitance by counting the number of edges in a defined time window. The timer turns the electrode consecutively to charging and discharging mode (see Fig. 4).

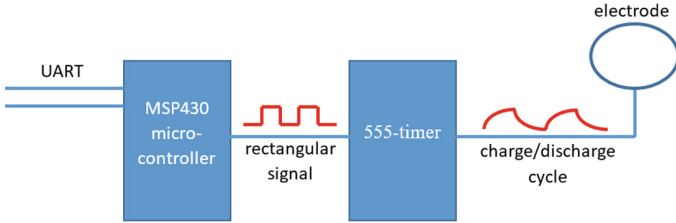


Fig. 4. Left: Sensor measuring process and sensor layout Right: Detail of visualization of measured capacitance.

The counted edges in a period of 0.5 s get transmitted to the UART bus. Every sensor has its specific ID. All even IDs start measuring the electrode while all uneven ones wait and vice versa. This ensures that sensors that are conducting a measurement do not influence adjacent sensors. The classification software receives a data package, containing single numerical values of each sensor, in intervals of 0.5 s. Even, when there are no objects close to the sensing area, the measured values show some differences in delta and amplitude. These are caused by environmental noise and differences in the electronic parts used (e.g. slightly different size of electrode). In order to eliminate the environmental effects, we calculate a baseline value for each sensor. We use the following equation to update the baseline constantly over time:

$$b_n = a \cdot b_{n-1} + (1 - a) \cdot x_n, \quad (2)$$

where x_n is the current sensor value and b_n is the currently updated baseline value for this sensor. How much the current sensor value influences the baseline will be determined by the factor a .

After subtracting the measured value from the baseline to get the actual signal strength, we normalize the value to a range between 0 and 1. We use a min-max algorithm where min and max are the minimal and maximal measured values over a period of time, ranging from time index $\{i = 1..N\}$ using the following equation:

$$\overline{x_n} = \frac{x_n - b_n}{\max(\{x_i - b_i, i = 1..N\}) - \min(\{x_i - b_i, i = 1..N\})}$$

We propose the use of machine learning for classification of the measured values. We chose SVM with a linear kernel and AdaBoost using REAL boosting as classification methodology. The feature vector, used for training, contains the values of the normalized sensor output. In Sect. 4.2 we describe our experiments about the influences of the size of the feature vector on classification accuracy. In these experiments we accumulated the sensor output over time (t) of all sensor to the feature vector receiving $fv = t_0, t_1, \dots$. We separated our data into two classes, one subject and more than one subject. In order to detect malfunction of the sensors or transmission, we visualized the sensor output as shown in Fig. 3.

Based on the different locations where mantrap portals and drop-arm turnstiles are being used, we defined two test scenarios. In case of mantrap portals, subjects can be asked to stand on an exact position which is marked on the ground. When using drop-arm turnstiles, people usually expect that there is no such position guideline. We therefore acquired data for both scenarios:

1. The access allowed subject is positioned at a marked position
2. The access allowed subject is encouraged to stand freely at a random position

We used a test group of 15 subjects with varying feet-size (between 38 to 48) and body-weight. All recordings are made over a period of 6 s (1×49 values per 0.5 s), resulting in 12×49 values per recording. One subject acts as an authorized person while the 2nd person acts as the attacker.

We evaluated two attack scenarios:

1. The attacker positions himself randomly on different challenging positions (on the edge, close to the each other ...)
2. The attacker positions himself randomly as in scenario 1 and lifts one foot from the ground

The evaluation is performed by training separate classifiers for the two attack scenarios, using Adaboost [14] and SVM classifier. We used the collected feature-vectors of an attack-scenario (two subjects at the same time) and of the different single subjects (see Table 1 #1 or #2) for training. We used data collected in 3 s (6×49 unsigned integer values) as one-dimensional feature-vector for training and tests.

4 Experiments and Results

We performed tests in order to ensure that our considerations about the chosen hardware are correct. Therefore, we performed laboratory tests about the measuring distance of the capacitive sensors using different electrodes. To ensure that different soles of shoes do not influence the measurement we performed empirical tests which are described in Sect. 4.1.

Table 1. Quantities of acquired test-data.

#	1st subject at pos.	2nd subject at pos.	Data-sets
0	Marked	None	1080
1	Random	None	1800
2	Marked	Random	1080
3	Random	Random	1800
4	Marked	Rand. + foot lifted	1764
5	Random	Rand. + foot lifted	1764

4.1 Sensor Range and Robustness

We evaluated our hardware accordingly to Valtonen et al. [11] with two different electrodes and with ten persons of different sizes.

We considered the electrodes types: loop and copper plate to be used as electrode. Connecting wires are made as small closed loops to act as an electrode, with one end of the wire connected to the sensor. On the other hand, copper plates in the size of 50×70 mm are getting used in comparison. We performed test in range and sensibility using a wet bottle of water as conductive object. We measured the sensor value for different distances and subtracted it by the baseline value. The noise showed for both electrode types, values between 0 and 20 measured over a time period of 2 min. Our results (see Table 2) show that the sensing range for the copper plate is higher than for the loop. The measurable distance is of around 100 mm which leads us to the assumption to chose a horizontal and vertical distance of 100 mm between the sensors as best layout compromise.

Table 2. Measured capacitance of electrodes at different distances and SNR of copper plate.

Object	Distance	Cable loop (SNR)	Copper plate (SNR)
Bare foot	20 mm	7744 (54.83)	1300 (45.26)
Bare foot	40 mm	10420 (58.02)	4900 (48.01)
Bare foot	60 mm	11940 (62.16)	7500 (49.98)
Bare foot	80 mm	12425 (68.22)	8560 (51.65)
Bare foot	100 mm	12680 (75.35)	8810 (54.23)
Bare foot	150 mm	12770 (89.12)	9030 (59.01)
None	-	13100	9225

In addition, we measured the Signal-To-Noise Ratio (SNR) of the system when the test subject was standing at different distances from the receiver with

and without shoes. The SNR depends upon the distance of the feet from the electrode, its type and size. We calculated the SNR using following formula:

$$SNR = 20 \cdot \log \frac{Signal\ Range(SR)}{Noise\ Range(NR)} \quad (3)$$

$$SR = \max(\{x_i, i = 1..N\}) - \min(\{x_i, i = 1..N\}) \quad (4)$$

$$NR = \max(\{y_i, i = 1..N\}) - \min(\{y_i, i = 1..N\}) \quad (5)$$

where y_i is the sensor value for a sensor without any feet on it and x_i with feet, over an amount of 200 sensor values (N).

In order to verify that there is no impact of different shoes on the measured values, different kinds of them, like sports shoes, sneakers and woodland shoes are considered. The size of the shoes varied from euro size 38 to 48. The test procedure was done by recording the conductive sensor value of different kinds of shoes individually. We compared those values with the values of the same subjects in bare foot (see Table 3).

Table 3. Capacitance of different shoes with copper plate electrode.

Type	Material	Sole height	Sensor value
Sport shoe	Plastic	21 mm	1993
Sneaker	Plastic	23 mm	2043
Woodland	Leather	34 mm	1845
Boots	Plastic	38 mm	1859
Bare foot	-	0 mm	2096

We noticed that the SNR decreased somewhat linearly with the distance between the feet and the electrode increases. Comparing the wire electrode with the copper plate, results the copper plate in better SNR then the loop electrode, which stands in contrast to the results of Valtonen et al. [11].

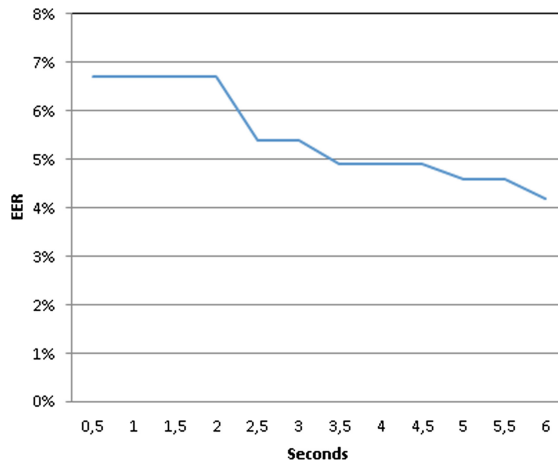
4.2 Results

The performance of our method depends on individual user properties, which are: feet position, behavior, feet size and their individual capacitance. We have chosen a quantitative evaluation with a test-group in order to prove our test-setup with realistic data. The evaluation process is carried out by classifying the data into four individual folds and taking 75% (3/4) of the data as training and 25% (1/4) of the data into the test. We calculated the false acceptance rate (FAR), false reject rate (FRR) and EER for all scenarios. We regard the verification case as a closed set, because it only differentiates between ‘one subject’ and ‘more then one subject’.

Table 4. EER of examined scenarios.

ID	1st pos.	2nd position	AdaBoost	SVM
1	Marked	All feet on ground	0%	3.4%
2	Random	All feet on ground	0%	5.2%
3	Marked	One foot lifted	5.4%	9.1%
4	Random	One foot lifted	7.1%	10.3%

Our results show that, the first attack scenario gets always recognized correctly, independently from the position of the access allowed subject (see Table 4). Only in the second scenarios, where the attacker lifts one foot, the performance decreased considerably. Classifying the data with linear multi-class SVM, resulted in significantly higher error rates in all attack scenarios.

**Fig. 5.** EER in respect to time (one foot lifted).

The time needed for data collection is an important factor of usability. Therefore, we performed tests with a varying measuring period. The length of the feature vector decreased respectively. We started using a feature vector of size 1×49 (1/2 s) and ended using the complete data-set of 1×294 as feature vector. As demonstrated in Fig. 5 a change of the error rate of 2.5% is shown within 0.5 (6.7% EER) and 6 s (4.2% EER).

We compared our method with results of other camera-based methods, namely: (1) Optical Flow [6] (2) Thermal Imaging [4] and (3) RGB-D [5]. As in their work, evaluation was carried out on attack scenarios, where the attacker used additional objects in order to hide himself, only scenarios without any objects were used for comparison. The ROC curve (see Fig. 6) shows the here presented method, without positioning guideline for the access allowed subject,

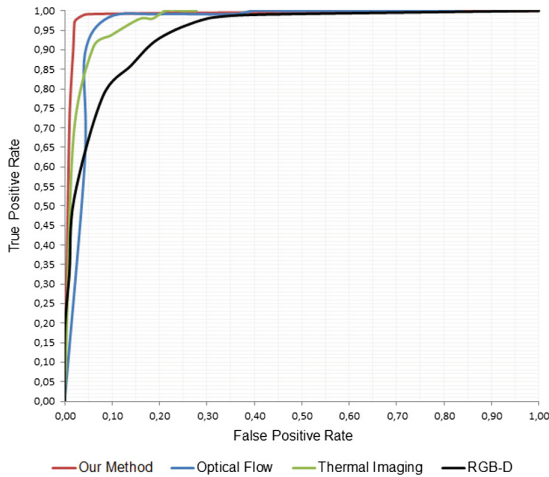


Fig. 6. Results in comparison with camera-image methods.

at random feet position, in comparison with the other approaches. We can conclude that the here presented methods shows a better performance in comparison, nevertheless, might a combination of one of the camera-based methods with our approach be useful in some cases.

5 Conclusion

We presented a novel approach for identifying attacks in an autonomous access control system by using a grid of capacitive sensors. We explored suitable sensing techniques and its corresponding sensor hardware by combining well-performing aspects of known methods in the field of in-door localization. We identified attack scenarios, in which attackers tried to pass through our system and explored machine-learning classification strategies to identify them. Our evaluation proved the layout and performance of the proposed sensor-grid, even under different environmental conditions. The performance of our method was defined in empirical testing, where we achieved good results in test-cases with feet of all subjects on the ground, even when only data of 0.5s was classified. We assume that a combination of our method with an image based approach focusing on movements like [6] will provide even higher security. Our method is vulnerable in cases where people are standing on one foot only but it might be complicated for an attacker to lift one foot while standing still.

References

1. McGovern, M.: Upgrading your entrance lanes (2013)
2. Edam, B.: Mantrap portal solution, July 2016
3. Amrehn & Partner EDV-Service GmbH: Bioporta, July 2016
4. Siegmund, D., Handtke, D., Kaehm, O.: Verifying isolation in a mantrap portal via thermal imaging. In: 2016 International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 1–4, May 2016
5. Siegmund, D., Wainakh, A., Braun, A.: Verification of single-person access in a mantrap portal using RGB-D images. In: XII Workshop de Visao Computacional (WVC), November 2016
6. Siegmund, D., Fu, B., Samartzidis, T., Wainakh, A., Kuijper, A., Braun, A.: Attack detection in an autonomous entrance system using optical flow. In: 7th International Conference on Crime Detection and Prevention (ICDP 2016), pp. 1–6. IET (2016)
7. Feng, G., Yang, Y., Guo, X., Wang, G.: A smart fiber floor for indoor target localization. *IEEE Pervasive Comput.* **14**(2), 52–59 (2015)
8. Steinhage, A., Lauterbach, C.: Monitoring movement behavior by means of a large area proximity sensor array in the floor. In: BMI, pp. 15–27 (2008)
9. Braun, A., Heggen, H., Wichert, R.: CapFloor – a flexible capacitive indoor localization system. In: Chessa, S., Knauth, S. (eds.) *EvAAL 2011*. CCIS, vol. 309, pp. 26–35. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33533-4_3
10. Braun, A., Wichert, R., Kuijper, A., Fellner, D.W.: Capacitive proximity sensing in smart environments. *J. Ambient Intell. Smart Environ.* **7**(4), 483–510 (2015)
11. Valtonen, M., Maentausta, J., Vanhala, J.: Tiletrack: capacitive human tracking using floor tiles. In: *IEEE International Conference on Pervasive Computing and Communications, PerCom 2009*, pp. 1–10. IEEE (2009)
12. Baxter, L.K.: *Capacitive Sensors: Design and Applications*. Wiley, New York (1996)
13. Smith, J.R.: Field mice: extracting hand geometry from electric field measurements (1996)
14. Viola, P., Jones, M.: Fast and robust classification using asymmetric Adaboost and a detector cascade. In: *NIPS, Vancouver, British Columbia, Canada*, vol. 2001, pp. 1311–1318 (2001)