Gaussian Process Person Identifier Based on Simple Floor Sensors

Jaakko Suutala[1], Kaori Fujinami[2], and Juha Röning[1]

[1] Intelligent Systems Group, Infotech Oulu
P.O. Box 4500, 90014 University of Oulu, Finland
E-mail: {jaska, jjr}@ee.oulu.fi
[2] Department of Computer, Information and Communication Sciences
Tokyo University of Agriculture and Technology, Japan
E-mail: fujinami@cc.tuat.ac.jp

Abstract. This paper describes methods and sensor technology used to identify persons from their walking characteristics. We use an array of simple binary switch floor sensors to detect footsteps. Feature analysis and recognition are performed with a fully discriminative Bayesian approach using a Gaussian Process (GP) classifier. We show the usefulness of our probabilistic approach on a large data set consisting of walking sequences of nine different subjects. In addition, we extract novel features and analyse practical issues such as the use of different shoes and walking speeds, which are usually missed in this kind of experiment. Using simple binary sensors and the large nine-person data set, we were able to achieve promising identification results: a 64% total recognition rate for single footstep profiles and an 84% total success rate using longer walking sequences (including 5 - 7-footstep profiles). Finally, we present a context-aware prototype application. It uses person identification and footstep location information to provide reminders to a user.

Key words: Person identification, machine learning, floor sensors, context-awareness

1 Introduction

Providing context-aware services to the user by means of smooth human-computer interaction requires natural and transparent ways to identify and locate users [1], [2], [3].

We present an approach to person identification based on human motion. More specifically, we concentrate on a person's style of walking, which presents behavioral characteristics of biometrics and is very natural because no additional action is required of the user. In this work, binary switch sensors are used to detect a person s walking sequence. A binary switch sensory system consisting of an array of 300 sensors was installed on the surface of the floor. Each 10 cm x 10 cm binary switch senses weight affecting its surface. A 3 m^2 floor area was covered to collect data and to recognize walkers.

User identification is based on statistical machine learning. We use a Gaussian process (GP) classifier [4] with specific features extracted from the footstep profiles produced by the sensor array as well as features calculated between consecutive footsteps.

The usefulness of the GP as a fully Bayesian kernel method relies on the ability to model uncertainty of data, which leads to automatic determination of hyperparameters (e.g., the importance of different features). It also produces conditional posterior probabilities of class labels (i.e., the degree of belonging to a certain class), which allows extensions and different post-processing capabilities of the classifier. Feature extraction itself is based on the standard methodology of image processing, as the sensor array can be presented as a binary image. Different features are extracted from the binary image based on single footstep profiles (e.g., length and width calculated from connected components in the binary image) and from the sequence of walking (e.g., step length and duration between consecutive footstep profiles). Along with these typical walker identification features we examined more specific ones that were calculated from the time-integrated signal. In practice, the binary signals were summed over time to form a grey-level image, and then features such as mean, standard deviation, and the center of mass were extracted from the connected components.

Data sets were collected from nine different subjects, including 20 walking sequences for each person. Each person wore their own shoes. In addition, four persons walked at three different walking speeds (slow, normal, fast) and also with two different pairs of shoes and without shoes. In the GP classification we examined the identification based on single footstep profiles as well as the identification using information from walking sequences (i.e., including multiple footstep profiles). The importance of different features were also analyzed.

This paper presents a simple yet modular floor sensor system that is able to identify persons based on their walking. It also describes an accurate classification method that is simultaneously able to analyze and choose the most important features and produce posterior probability of ID labels for post-processing. Furthermore, the effect of changes in walking speed and footwear is analyzed for the first time.

2 Related Work

Various floor sensor settings have been used to model human behavior for identification, tracking, and other purposes in smart and interactive environments as well as in health care. In the early works by [5] and [6], footstep identification was based on a small area of ground reaction force (GRF) sensors using nearest-neighbor and hidden Markov model (HMM) methods, respectively. In the work by [7], human GRF-based authentication system was developed for use as part of a surveillance system. Recently, a sensor installation, collection of a large data set and experiments with a person verification scenario were presented in [8]. They used a GRF sensor with geometric and holistic features along with a support vector machines classifier.

In [9], electromechanical films (Emfi) that measures dynamic pressure changes on the floor surface were used for person identification. A comparison of different methods (e.g., support vector machines and neural networks) was done. In addition, classifier fusion techniques were applied to combine different feature sets and walking sequence information to achieve a more reliable recognition system.

UbiFloor [10], uses simple ON/OFF switch sensors, and identification is based on features of both single footsteps and walking calculated from five consecutive footsteps on the floor. The sensor arrangement differs from our work, but the use of simple binary sensors is most similar to ours from the application viewpoint. A multi-layer perceptron (MLP) neural network was used as a classifier. [11] developed a high-resolution low-cost pressure sensor mat made of resistive switches. They also performed person identification based on sequential features such as stride length, gait period, and heel-to-toe ratio along with an Euclidean distance measure as a classifier.

A sensor approach similar to this work was established by [12]. However, they concentrated on human tracking applications based on Markov chain Monte Carlo methods. [13] presents a system that also uses binary ON/OFF sensors in which over 65,000 pressure switches in an area of 4 m^2 give a very high resolution to the modeling of the details of single footstep profiles as an image of footprints. The floor was tested by detecting humans and robots and discriminating between them. [14] reported the use of a beneath-the-floor accelerometer and tactile sensors to model footsteps and footprints in order to recognize gender. [15] covered the floor of an interactive space with hexagonal pressure-sensitive floor tiles to detect the presence of users.

Besides identification and tracking, force plates have been used to detect and classify simple human body movements, such as crouches and jumps as well as standing up and sitting down [16]. A lot of work has also been done in medical research domains, including [17], where pattern recognition methods were used to classify different gait-related injuries based on GRF sensor measurements.

In summary, this work present a unique sensor approach to person identification ([12] uses similar sensors but they are used in a tracking application). We also extract novel features from the floor and analyze the importance of individual features as well as the effect of walking speed variations and different footwear, which are typically not included in the other studies related to floor-based identification. Our approach has a direct possibility of combining sequential information from multiple footsteps based on the classifier s posterior probability outputs. This is quite similar to [9], except that further post-processing is not needed to get the confidence of labels.

3 Binary Switch Floor Sensor System

VS-SF55 InfoFloor sensor system made by Vstone Corporation (in Japan) [18] was installed in our research laboratory. The system contains 12 blocks of 50 cm x 50 cm sensor tiles. Each tile includes 25 10 cm x 10 cm binary switch sensors. A 3 m^2 area was covered by altogether 300 sensors (see Fig. 1). The sensors use diode technology and are able to detect over 200-250 g/cm^2 weight affecting the surface. Data were collected from each sensor using a 16 Hz sampling rate and sent to a PC via an RS-232 serial interface. In the PC, a multi-threaded TCP-IP server was implemented to share raw sensor data with client applications.

Compared with other floor sensor technologies (e.g., Emfi [9]), the advantages of using this kind of floor sensor system are low cost, easy installation, and little need for pre-processing to get the data (e.g., for positioning and identification). Moreover, the sensor floor utilized in this paper is designed to be modular, which allows the sensor area to be able extended incrementally. On the other hand, compared with cameras, audio, or RFID technology, floor sensors are more stable, i.e, they do not suffer from environmental changes. A drawback is that only very limited information is obtained from the binary floor compared with cameras or other floor sensor technologies (e.g., Emfi and GRF sensors). This is very challenging, especially in complex recognition tasks such as person identification, where discrimination between different persons can depend on very detailed differences in persons walking styles. One aim of this work was to be able to extract such useful and discriminative information from this limited, yet practical, sensor system.

Fig. 1. Arrangement of binary sensor tiles.

4 Discriminative Bayesian Classification: Gaussian Processes

Discriminative learning is a very effective way to train mappings from multidimensional input feature vectors to class labels. Kernel methods, in particular, have become state-of-the-art, due to their superior performance in many real-world learning tasks. Along with the popular support vector machines (SVM) [19], Gaussian processes (GP) [4] have recently been given much attention in the machine learning community.

Although the SVM method has many favorable properties, such as good generalization by finding the largest margin between classes, the ability to handle non-separable classes via soft-margin criteria, non-linearity modeling via explicit kernel mapping, sparseness by presenting data using only a small number of support vectors, and global convex optimization with given hyperparameters, it lacks some properties. One drawback of SVM is that it is directly applicable only in two-class problems. Thus, there have been various attempts to generalize it for multi-class classification. The simplest and most popular methods are based on multiple binary classifiers using one-vs.-one or one-vs.-rest approaches as well as error correcting output codes and directed acyclic graphs, to name a few [20].

Another problem is the choice of a good model, which is very important in kernel-based discriminative learning. This is due to the fact that a good solution is usually dependent on a number of hyperparameters (which control the properties of kernel mapping). In SVM, the hyperparameters (and possibly the good subset of features) need to be found using ad-hoc methods such as cross-validation or other search-based methods. When the number of hyperparameters or the number of features increases, the search space can become very large. Finally, SVM cannot directly give a confidence measurement as an output, it only gives a decision as an unscaled distance from the margin in the feature space. Posterior distribution over predicted class labels is a very important property in many pattern recognition systems in order to be able to implement some post-processing tasks (e.g., rejecting unreliable examples, combining multi-modal sensor data, combining sequential data, etc.). There have been some attempts to extended SVM to give probabilistic outputs (see [21], for example). However, this method needs to train another mapping to the SVMs output after the training based on parametric sigmoid mapping. This makes the method more complicated and possibly another validation data set needs to be optimized for post-processing mapping.

To tackle these problems, we apply a Bayesian approach to kernel-based learning via Gaussian process priors. We use the multi-class approach presented in [22], which approximates a complex posterior probability by maximizing the variational lower bound. By using a multinomial probit likelihood model, it is possible to derive a full multi-class classifier as a combination of multiple regression models. These regression models are coupled via the posterior mean estimates of another set of auxiliary variables, which gives a statistically dependent multi-class model. Add-hoc post-processing is not needed. In addition, predictive distribution over unknown examples provides direct confidence measurement as a conditional posterior probability of class labels.

During the training phase of the classifier, Gaussian processes provide a possibility to optimize the hyperparameters by maximizing the marginal likelihood via gradient-based optimization routines [23], [4] or by setting a prior distribution for the hyperparameters and employing sampling methods such as importance sampling to get posterior expectations [22]. We follow the approach used in [22], which uses exponential distribution and a gamma distribution placed on its mean to form a conjugate pair. Furthermore, by applying a radial basis function (RBF) kernel with individual length scale parameters to each feature dimension, we can determine the importance of each feature when optimizing the hyperparameters (i.e., automatic relevance detection (ARD)). This is used to increase the accuracy of person identification as well as to analyze features in different practical settings. One drawback of GPs is that all the training data are needed in the classification phase. When a large data set is used, some sparse approximation methods need to be applied [22]. In this paper, a full model is used due to its capability of real-time performance in the prototype application.

5 Person Identification Based on Floor Sensors

5.1 Feature Extraction

As in typical pattern recognition systems, we need to extract some higher level features from the raw data to be able to perform accurate identification. The binary switch sensor floor forms a matrix where each sensor tile can be presented as a pixel in the image. This allows us to apply standard image processing techniques to detect footsteps and to extract features. We use two kinds of presentations: binary and grey-level images.

A binary image is detected by summing up the sensor values over time, and then thresholding each positive value to one. The summing is performed over each walking sequence. A binary image gives us a direct way to detect the position of each footstep in a sequence. This is done by labeling the 8-connected components of the image. Furthermore, when collecting each individual image in a sequence, we are able to detect the starting and ending time of each connected component for feature extraction.

In addition, the integrated image (i.e., sequence of summed sensor matrices) is saved without thresholding. This matrix presents a grey-level image in which each pixel forms a duration value over the sequence and provides a possibility to extract a rich set of features from the connected components. A "duration map" is presented as a grey-level image in Figure 2, where a brighter value means more time is spent in that position. A binary image can be calculated by thresholding grey-level sensor values larger than 0 to 1.

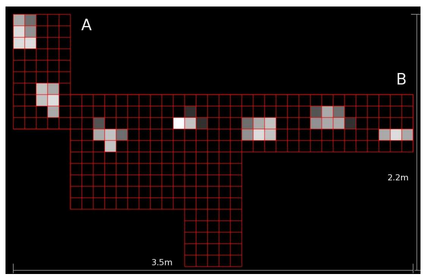


Fig. 2. Grey-level image calculated from sensor measurements of a walking sequence. In addition, the size of the sensor area is illustrated.

Feature extraction is based on the connected components found in the binary image. The features can be divided into two categories: micro and macro-level features. Micro-level features are extracted from each footstep using both the binary and grey-level presentations. This feature set includes features such as the sum of binary pixels in a single footstep profile. Minimum, maximum, mean, and standard deviation values are also extracted from the grey-level component. All these features describe the shape of the "duration map" inside a single footstep profile. To describe the spatial properties of shape, convolution filters, familiar from image processing, are used. We apply four different 3x3 line detection filters and four different 3x3 sobel gradient filters (see, for example, [24] for details). After filtering, the values inside the connected components are summed. Also, the length and width of the footstep, the compensated center of masses, and the duration of the footstep are calculated. Macro-level features present useful information between consecutive footsteps. We use Euclidean distances between the center of mass points of adjacent footsteps as well as individual distances in the longitudinal and transversal walking directions. They are closely related to step length measurement used in gait analysis. Finally, the duration between the starting times of consecutive footsteps is calculated. Macro features are always calculated against the previous footstep in a sequence. A total of 28 features were extracted and are presented in Table 1. It is also straightforward to modify the footstep detection and feature extraction techniques for a real-time application, which is discussed in section 7.

Table 1. Spatial, statistical, and time-related features derived from each footstep profile (1-20) as well as between consecutive footstep profiles (21-28).

Number	Name	Description
1.	sum_{bin}	Number of activated pixels (i.e. sensor tiles) in this footstep profile
2.	sum_{grey}	Sum of grey-level pixel values
3.	min_{grey}	Minimum grey-level value
4.	max_{grey}	maximum grey-level value
5.	$mean_{grey}$	Mean of grey-level pixels
6.	std_{grey}	Standard deviation of grey-level pixels
7.	sum_{vline}	Sum of grey-level component filtered with 3x3 line mask (vertical)
8.	sum_{hline}	Sum of grey-level component filtered with 3x3 line mask (horizontal)
9.	sum_{lline}	Sum of grey-level component filtered with 3x3 line mask (left diagonal)
10.	sum_{rline}	Sum of grey-level component filtered with 3x3 line mask (right diagonal)
11.	sum_{bgrad}	Sum of grey-level component filtered with 3x3 gradient mask (ball of the footstep)
12.	sum_{rgrad}	Sum of grey-level component filtered with 3x3 gradient mask (right side of the footstep)
13.	sum_{hgrad}	Sum of grey-level component filtered with 3x3 gradient mask (heel of the footstep)
14.	sum_{lgrad}	Sum of grey-level component filtered with 3x3 gradient mask (left side of the footstep)
15.	$length_{bin}$	Maximum length of connected binary pixels (longitudinal direction of walking)
16.	$width_{bin}$	Maximum width of connected binary pixels (transversal direction of walking)
17.	com_{bin_x}	Center of mass of connected binary pixels (longitudinal direction of walking)
18.	com_{bin_y}	Center of mass of connected binary pixels (transversal direction of walking)
19.	com_{grey_x}	Center of mass of connected grey-level pixels (longitudinal direction of walking))
20.	com_{grey_y}	Center of mass of connected grey-level pixels (transversal direction of walking)
21.	$duration_{inside}$	Duration of footstep (i.e., activated tiles over time)
22.	$distance_{bin}$	Euclidean distance from previous footstep (using binary center of mass)
23.	$distance_{grey}$	Euclidean distance from previous footstep (using grey-level center of mass)
24.	$duration_{between}$	Duration from the previous footstep (to beginning time of this footstep in milliseconds)
25.	$distance_{bin_x}$	Longitudinal distance from previous footstep (using binary center of mass)
26.	$distance_{bin_y}$	Transversal distance from previous footstep (using binary center of mass)
27.	$distance_{grey_x}$	Longitudinal distance from previous footstep (using grey-level center of mass)
28.	$distance_{grey_y}$	Transversal distance from previous footstep (using grey-level center of mass)

5.2 Person Identification: Single Footsteps and Walking Sequences

We derive two kinds of person identification methodologies based on the multi-class Gaussian process classification and features presented in the previous section. The first one is a conventional classification scenario where we use posterior distribution of class labels predicted from a single footstep profile to make the decision. In this case we use micro-features as well as macro-features related to the previous footstep. This scenario is useful in situations where the decision has to be made as quickly as possible.

On the other hand, if we want more accurate recognition, we can use classification information from multiple adjacent footstep profiles by combining the posterior distribution of class labels. This scenario gives a recognition based on a sequence, which in this case is one walking sequence (5-7 footsteps) on the floor. As GP classification provides posterior over class labels, we can use summation and product rules to combine the outputs. This kind of rule has been shown to be simple, yet powerful, in many information fusion problems [25]. The advantage is that we can use a conventional training phase and an arbitrary number of examples in a sequence to make the final decision. If $P(\omega_k|x_i)$ represents the posterior probability of class labels $(1 \dots n)$ conditioned on unknown example x_i and S is the total length of a sequence, the final decision can be calculated using the sum (Eq. 1) and product rule (Eq. 2), as follows:

$$\omega_c = \operatorname*{argmax}_{k=1}^{n} \Big[\sum_{i=1}^{S} P(\omega_k|x_i) \Big] \tag{1}$$

$$\omega_c = \operatorname*{argmax}_{k=1}^{n} \Big[\prod_{i=1}^{S} P(\omega_k|x_i) \Big] \tag{2}$$

The disadvantages of using this kind of scenario are related to optimization. Due to the fact that the model is trained on single footsteps, it does not use information of sequences to find a global optimum. In addition, the choice of combination rules in our scenario is more ad-hoc and experimental compared with approaches where sequential information is directly learned from the data. However, this simple approach is able to use the information of walking sequences at some level to be able to produce more accurate decisions, as is shown in the results section. A comparison with more advanced models, such as sequential kernels and other sequential classifiers, is left for future work.

6 Results

6.1 Data sets

To test the identification methods presented here, we collected a large data set. The data set included walking sequences of nine different subjects. The test group consisted of two female and seven male subjects, and each wore their own shoes (which were indoor sandals in this case). They were told to walk their natural walking speed over the sensor floor (from A to B in Figure 2) 20 times. To get as natural a data set as possible, the starting foot or the absolute position of each footstep in the sequence was not constrained in any way. Each sequence included 5-7 footstep profiles, depending on the stride length of the subject. Altogether 1143 footstep profiles were collected from the nine walkers.

In addition, to examine the effect of different walking styles (i.e., walking speed) and footwear on identification, we collected more data from four subjects. To study variations in walking speed, we recorded additional sequences in which the subjects were told to walk slower and faster than usual. Both settings were performed 10 times. To test the effect of different footwear, 20 sequences of subjects wearing their own outdoor trackers and no shoes at all were collected. Combining this data set with the footsteps of the four persons collected earlier gave us 1981 footstep profiles for studying the effect of variation in walking speed and footwear.

A total of 2597 footstep profiles were collected in these sessions. To test and analyze the usefulness of the features and the classification method as well as the modeling capability of the features and adaptation of the classifier to novel data, we split the data set into different subgroups. The standard nine-person data set included 20 sequences of normal walking speed and sandals for studying the extracted features and the capability to perform multi-class classification using Gaussian processes. To analyze the effects of variations on the extracted features more precisely, the footstep profiles of four persons were divided into three subgroups: standard (including walking at normal speed and with sandals), footwear (including three different footwear at normal speed), speed (including three different speeds with sandals on). The aim of these data sets was to be able test how well the extracted features can handle variations in the data set and which features have the best discriminative power in these settings.

Furthermore, we split the four-person data set into 12 subgroups: sandals (including all the data from sandals), without sandals (all the data except from sandals), trackers (including data from outdoor shoes), without trackers (including all the data except from trackers), without shoes (including the session without shoes), shoes (including the session with shoes), normal (including normal speed), not normal (including slow and fast walking), slow (including slow walking), not slow (including normal and fast walking), fast (including fast walking), not fast (including slow and normal walking). These data sets were used to examine the generalization capability of the classifier and the need for adaptation when the test data set includes differently distributed (in this case walking speed and footwear) data. These are very important when building practical applications. A summary of the data set categories is presented in Table 2.

Table 2. Summary of different data set categories used in the person identification experiments

Number	Name	Description	Number of examples	number of sequences
1.	9 persons standard	Normal walking speed with sandals	1143	180
2.	4 persons standard	Normal walking speed with sandals	527	80
3.	Footwear	Normal walking speed with footwear variations	1516	240
4.	Speed	Slow, normal, and fast walking speed with sandals	992	160
5.	Sandals	All the data with sandals	992	160
6.	Without sandals	All the data without sandals	989	160
7.	Trackers	All the data with trackers	441	80
8.	Without Trackers	All the data without trackers	1540	240
9.	Shoes	all the data with shoes	1433	240
10.	Without Shoes	All the data without shoes	548	80
11.	Normal	All the data with normal speed	1516	240
12.	Without normal	All the data without normal speed	465	80
13.	Slow	All the data with slow speed	248	40
14.	Without slow	All the data without slow speed	744	180
15.	Fast	All the data with fast speed	215	40
16.	Without fast	All the data without fast speed	755	180

6.2 Person Identification

In this section we present the recognition result of using the nine-subject data set described in Section 6.1. We split the data set so that 2/3 were used for training and 1/3 for testing, and all the features were scaled between 0 and 1. Variational GP approximation was achieved using 10 iterations, simultaneously learning the hyperparameters of the RBF kernel [20], [22]. This was repeated 10 times on randomly chosen training and test sets. All the tests were implemented with Python programming language and the GP models were trained with an R language variational Bayesian GP package [26].

Furthermore, sequential recognition was tested by combining the GP outputs using similarly trained models and fixed sum and product rules. Table 3 presents the average total identification (and standard deviations) of single footstep profiles as well as combined recognition rates. The classifier is able to classify correctly 64% of the individual footsteps, which shows the complexity of the data set obtained from the simple binary switch sensors. Using the fixed combination rules increases accuracy and the product rule outperforms the sum rule in this data set, showing an 84% success rate. The results show that to achieve a high success rate, sequential information is needed.

Table 3. Total identification accuracies of recognizing nine different walkers

	GP (single examples)	GP (sum rule)	GP (product rule)
Accuracy (%)	64.23 (3.27)	82.33 (6.59)	84.26 (6.69)

6.3 Feature Analysis of Footwear and Walking Speed Variations

This section presents the results of analyzing the effect of different footwear and walking speed variations. Moreover, we rank the individual features based on their relevance in the identification method to determine which are the best and worst ones. To our knowledge, this is the first time both footwear and walking speed changes are analyzed in the context of floor sensors. These are very important issues when building a practical identification system.

We used the different four-person data sets presented in Table 2, where we summarize the total success rates (accuracy) as well as the most relevant features (mrf) and least relevant features (lrf) (cf. Table 1 for the order number of the features). Table 4 presents the results using standard data sets and footwear/speed variations. Looking at the accuracies, the total number of persons in a classification has a large impact (nine persons vs. four persons.). Secondly, footwear variation slightly decreases accuracy compared with the standard data set (4.36 percent units). Walking speed decreases accuracy much more (10.50 percent units). In all the data sets, the most important features are related to walking sequence (i.e., $distance_{bin}$, $distance_{grey}$, $duration_{between}$) and the duration of footsteps. The least relevant features change, but are always related to micro-features. These results indicate that when using limited binary sensors, the use of features carrying sequential information is very important. The average length scales of each feature in the nine-person data set are presented in Figure 3. A smaller value means the feature is more important in the classification decision. The walking sequence features are the most important, but footstep shape features (e.g., calculated by the convolution filters) have a large impact, too (e.g., features 8, 10 and 14)

Similar experiments are shown in Table 5. Now the test set contains variations (i.e., footwear and walking speed) that are not included in the training data set. This is the most complex approach presented in the paper. Clearly, a large decrease in total accuracies can be seen when comparing the results with those in Table 4. This indicates that it is important to collect and use all available information for training if these variations are assumed to happen. Similarly, it can be concluded that speed variations have a larger negative impact on accuracy compared with footwear variations. Interestingly, the same features as in the above data sets have the most relevant information for identification, on average.

Table 4. Total identification accuracies and feature ranking using different datasets. The data sets are described in Table 2 and the features are presented in Table 1. The three most relevant features (mrf) and least relevant features (lrf) are shown

Dataset	Accuracy (%)	mrf	lrf
9 persons standard (1.)	64.23 (3.27)	21.,24.,23.	2.,28.,20.
4 persons standard (2.)	81.45 (1.62)	21.,23.,24.	16.,20.,3.
Footwear (3.)	77.09 (1.22)	24.,21.,22.	12.,11.,4.
Speed (4.)	70.95 (2.20)	21.,23.,24.	3.,19.,20.

7 Prototype Application: Context-aware Reminder

A prototype application was built based on single footstep identification. A multi-class Gaussian process classifier was learned from the training data set of four laboratory

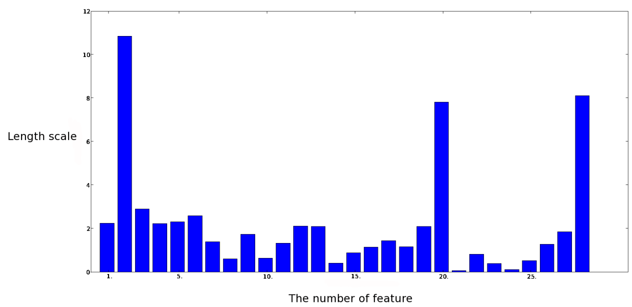


Fig. 3. RBF kernel length scales of each feature using a nine-persons data set. The horizontal axis presents the feature number from Table 1 and and the vertical axis describes the importance of the feature, where a smaller length scale value means the feature is more important.

Table 5. Total identification accuracies and feature ranking using different data sets. The data sets are described in Table 2 and features are presented in Table 1. The three most relevant features (mrf) and least relevant features (lrf) are shown

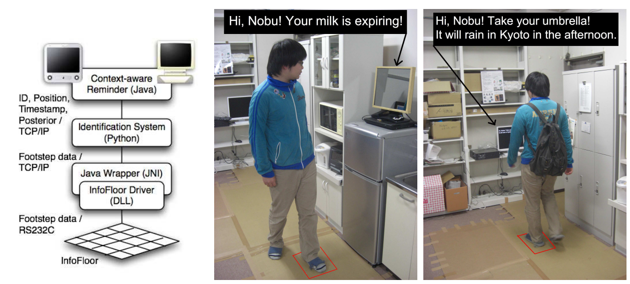
Train	Test	Accuracy (%)	mrf	lrf
Without sandals (6.)	Sandals (5.)	59.68	24.,23.,21.	15.,13.,5.
Without trackers (8.)	trackers (7.)	59.49	23.,24.,21.	13.,16.,9.
Shoes (9.)	Without shoes (10.)	59.85	21.,24.,23.	26.,1.,14.
Normal speed (11.)	Without normal (12.)	48.60	21.,27.,24.	5.,7.,3.
Without slow (14.)	Slow speed (13.)	57.66	21.,23.,24.	5.,3. 12.
Without fast (16.)	Fast (15.)	41.01	21.,23.,24.	1.,20.,11.

members. In addition, the position of each footstep was calculated using the center of mass in the binary image. This very simple method is able to locate one person at a time. In the future, more advanced tracking methods will be applied to detect the positions of multiple simultaneous walkers.

The prototype was implemented as a distributed system consisting of three different levels, were each level provides information via TCP/IP socket communication. The TCP/IP-based approach was chosen to leverage existing libraries for rapid prototyping, which requires language independence. The first level provides raw sensor data, which is read by the identification system on the second level. The lower-level implementation consists of a Windows DLL (VC++) for InfoFloor driver and Java TCP/IP server software. The identification system extracts features from the raw data and sets the identity based on GP as well as position and the time stamp information of each example. In this application feature extraction needs to be implemented in real time. This was done by monitoring the starting and ending times of connected binary events on the floor, and when these were detected micro-level features were calculated from the sensor area of the footstep using both binary and grey-level presentation. After that macro-level features were calculated based on the detection information from the previous footstep. If a certain time period (e.g., 5 seconds) expired without any events, it was assumed that the person has left the sensor area and the detection phase is starting over again. The second level was implemented with Python language, as presented in the results section. The recognition software worked in real time and it took no more than 20 ms to process the raw data into identification prediction (using the model trained on four-person data). The time between two adjacent footsteps was approximately 500ms. The third level is an application that reads identified events from the identification system. Along with side information about the context of the environment, it provides reminders to a user. The client program was implemented with Java. The components of the software architecture are presented in Figure 4(a). In this application scenario the user interface is implemented with two displays. The first one is located above the refrigerator and the second one is located near the entrance to a "smart room" (see Figures 4(b) and 4(c)). The scenario, which assumes side information, is as follows:

1. *Nobu bought a bottle of milk a week ago and put it into the refrigerator. One week later, when he is passing in front of the refrigerator, it notifies him of the expiring status of the milk. Here, a mirror display is installed on the fridge, and the fridge is capable of determining the status of the contents.*
2. *Nobu, a Tokyo resident, is going on a trip to Kyoto. Although the weather is fine in Tokyo, the weather forecast says it will be rainy in Kyoto. The "smart room" knows his schedule, i.e. date and location, as well as the identify of the person and the walking direction. When he is leaving the room, a display installed at the entrance recommends him to take an umbrella with him because of the forecast.*

This prototype application shows a simple approach to using naturally obtained person identification information, recognized from walking (along with the side information), in a context-aware system.



(a) Architecture of the prototype.

(b) Context-aware reminder above the refrigerator.

(c) Context-aware reminder at the entrance of the smart room.

Fig. 4. Software architecture and scenarios in the prototype application.

8 Conclusions

In this paper we presented a floor sensor system based on binary switches as well as methods for recognizing a persons identity based on sensor measurements collected from the floor. In addition we showed a prototype application that uses the information of a walkers identity and the position of footsteps to provide context-aware reminders for daily life. For the recognition purposes, a set of useful features were extracted from the raw measurements. The measurements are presented as binary and grey-level images, which allow us to use basic image processing methods to derive higher-level features. A variational Bayesian approximation of a multi-class Gaussian process (GP) classifier is used to identify the walkers. As a Bayesian method the GP gives the posterior distribution of predicted class labels. This information was used to combine the classifier outputs of multiple footsteps using conventional classifier combination rules. This provides a simple approach to recognizing a sequence of walking in an application where a more accurate decision is needed. The total recognition rates of nine different subjects using individual footsteps as well as walking sequences were 64% and 84%, respectively. This is a very promising result using simple binary switch sensors.

Furthermore, GPs provide a flexible solution to model selection (e.g., the choice of hyperparameters). We used a kernel that is able to weigh each feature's dimensions differently through hyperparameters. This provides automatic relevance detection (ARD), where the most important features get more weigh in a similarity measurement. ARD was used to train an accurate model and to analyze the importance of individual features. We analyzed the effect of different footwear and variations in walking speed on identification accuracy. This kind of analysis is missing from most of the previous studies using floor sensors. In our experiments we found that both of these variations have an impact; walking speed variations have a larger negative impact. Moreover, the most relevant features in all the tested data sets were related to distance and duration between footsteps as well as the duration of a single footstep profile.

Acknowledgments

This work was supported by The Ministry of Education, Culture, Sports, Science and Technology in Japan under a Grant-in-Aid for Division of Young Researchers, and InfoTech Oulu Graduate School.

References

1. I.A. Essa. Ubiquitous sensing for smart and aware environments: Technologies towards the building of an aware home. *IEEE Personal Communications, October 2000, Special issue on networking the physical world,*, pages 47–49, 2000.
2. B. Brummit, B. Meyers, J. Krumm, A. Kern, and S. Shafer. Easyliving: Technologies for intelligent environments. In *2nd International Symposium of Handheld and Ubiquitous Computing (HUC)*, pages 12–29, New York, USA, 2000. Springer-Verlag.
3. A. Pentland. Smart rooms. *Scientific American*, 274:68–76, April 1996.
4. C.E. Rasmussen and C.K.I. Williams. *Machine Learning for Gaussian Processes*. The MIT Press, USA, 2006.
5. R.J. Orr and G.D. Abowd. The smart floor: A mechanism for natural user identification and tracking. In *Proceedings of 2000 Conf. Human Factors in Computing Systems (CHI)*, pages 275–276, The Hague, Netherlands, 2000. ACM Press.

6. M.D. Addlesee, A. Jones, F. Livesey, and F. Samaria. ORL active floor. *IEEE Personal Communications*, 4(5):35–41, 1997.
7. P. Cattin. *Biometric Authentication System Using Human Gait*. PhD thesis, ETH-Zürich, Institute of Robotics, Switzerland, 2002.
8. R.V. Rodriguez, R.P. Lewis, J.S.D. Mason, and N.W.D. Evans. Footstep recognition for a samrt home environment. *International Journal of Smart Home*, 2(2):95–110, 2008.
9. J. Suutala and J. Röning. Methods for person identification on a pressure-sensitive floor: Experiments with multiple classifiers and reject option. *Information Fusion Journal, Special Issue on Applications of Ensemble Methods.*, 9:21–40, 2008.
10. J-S. Yun, S-H Lee, W-T Woo, and J-H Ryu. The user identification system using walking pattern over the ubifloor. In *Proceedings of International Conference on Control, Automation, and Systems (ICCAS)*, Gyeongju, Korea, October 2003.
11. L. Middleton, A.A. Buss, A. Bazin, and M.S. Nixon. A floor sensor system for gait recognition. In *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)*, pages 171–176, 2005.
12. T. Murakita, T. Ikeda, and H. Ishiguro. Human tracking using floor sensors based on the markov chain monte carlo method. In *Proceedings of Seventeenth International Conference on Pattern Recognition (ICPR)*, pages 917–920, Cambridge, UK, August 2004.
13. H. Morishita, R. Fukui, and T. Sato. High resolution pressure sensor distributed floor for future human-robot symbiosis environment. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1246–1251, EPFL, Lausanne, Switzerland, October 2002.
14. K. Sudo, J. Yamato, and A. Tomono. Determining gender of walking people using multiple sensors. In *Proceedings of the IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 641–646, December 8-11 1996.
15. K. Eng, R.J. Douglas, and P.F.M.J. Verschure. An interactive space that learns to influence of human behaviour. *IEEE Transaction on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 35(1):66–77, January 2005.
16. R. Headon and R. Curwen. Recognizing movements from the ground reaction force. In *Proceedings of the Workshop on Perceptive User Interfaces*, pages 1–8, Orlando, Florida, USA, November 15-16 2001.
17. M. Köhle and D. Merk. Clinical gait analysis by neural networks: Issues and experiences. In *Proceedings of the IEEE Symposium on Computer-Based Medical Systems*, pages 138–143, Maribor, Slovenia, 1997.
18. Vstone corporation. URL: http://www.vstone.co.jp/e/etop.html.
19. N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, UK, 2000.
20. B. Schölkopf and A. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, USA, 2001.
21. J. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In A. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans, editors, *Advances in Kernel Methods - Support Vector Learning*, pages 61–74. MIT Press, 1999.
22. M. Girolami and S. Rogers. Variational bayesian multinomial probit regression with gaussian process priors. *Neural Computation*, 18:1790–1817, 2006.
23. D.J.C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, UK, 2003.
24. R.C Gonzalez and R.E. Woods. *Digital Image Processing*. Prentice Hall, 2002.
25. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.
26. N. Lama and M. Girolami. vbmp: Variational bayesian multinomial probit regression for multi-class classification in R. *Bioinformatics*, 2008.