Detection of Human Movements with Pressure Floor Sensors

Martino Lombardi, Roberto Vezzani, and Rita Cucchiara

Softech-ICT, University of Modena and Reggio Emilia, Modena, Italy
{martino.lombardi, roberto.vezzani, rita.cucchiara}@unimore.it
WWW home page: http://imagelab.ing.unimore.it

Abstract. Following the recent Internet of Everything (IoE) trend, several general-purpose devices have been proposed to acquire as much information as possible from the environment and from people interacting with it. Among the others, sensing floors are recently attracting the interest of the research community. In this paper, we propose a new model to store and process floor data. The model does not assume a regular grid distribution of the sensing elements and is based on the ground reaction force (GRF) concept, widely used in biomechanics. It allows the correct detection and tracking of people, outperforming the common background subtraction schema adopted in the past. Several tests on a real sensing floor prototype are reported and discussed.

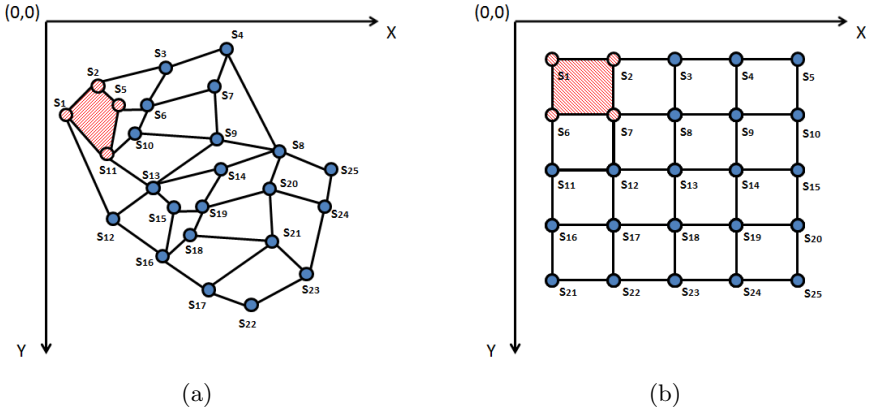
Keywords: Human-Computer Interaction, Sensing Floor, Pressure Analysis, Center of Pressure, Ground Reaction Force

1 Introduction

In the last years, the research on non-invasive human-computer interaction systems has attracted a wide interest. Therefore, a lot of systems based on video cameras [2], depth sensors [7], wearable devices [3], and sensing environments [5] have been proposed for interactive media applications. In particular, the adoption of *sensing floors* plays a key role in the development of sensing environments thanks to two significant properties: low invasiveness and high invisibility (i.e., the sensing layer is invisible to the users and the floor appears similar to traditional floors to avoid the "observer effect").

Their applications are manifold in several fields, including both public and private environments. For example, smart buildings can include sensing floors to detect the presence of people and to automatically switch on/off the lighting or the heating systems. In the e-health field, these devices can be used to detect dangerous situations such as an elder falling or getting out of his/her bed. Furthermore, sensing floors can be used for people counting or to monitor crowd movements during public events, exhibitions, and so on. In comparison with other traditional technologies such as video cameras, sensing floors provide less information but have two undoubted advantages. First, they are completely privacy compliant as above mentioned. In fact, it is not feasible to recognize

and identify users from floor data only. Installations on very private places such as toilets or bedrooms are allowed. Second, sensing floors data are not affected from occlusions, a typical issue of visual camera systems.



(a) (b)

Fig. 1. Schema of two different placement of 25 sensors with (a) a random distribution or (b) a grid. Examples of Floor Cells with 4 sides and 4 vertices are colored in red.

Since 1997, when J. Paradiso [12] at MIT presented the first example of a sensing floor, several prototypes have been proposed and designed. The adopted sensors exploit different physical characteristics, such as the pressure as measurable quantity, and the proximity effect related with the electrical properties of a human body. A complete analysis of these differences can be found in [15] and in [20].

Regardless of the technology adopted to build the sensing floor, this paper aims at focusing on the data model and the corresponding processing algorithms. The prevalent approach has been borrowed from the image processing field. Each sensing element is related to a pixel of an image. The pressure applied on top of a sensor is translated to a corresponding pixel intensity. In the following, we refer to this approach as PIM (*Pressure Image Model*). At each sampling instant, the sensing floor generates a sort of *pressure image*, where each *pixel* corresponds to a spatial portion of the floor and the *pixel value* is related to the pressure applied on the top of it. Consecutive temporal pressure images can be collected as frame sequences to generate a *pressure video*, whose analysis allows to detect and determine spatio-temporal events on sensing floors [9].

Two different types of image can be generated, depending on the physical sensor capabilities, i.e., *binary* and *grey-level* pressure images, respectively. Binary pressure images are generated by sensing floors made using matrix of switch sensors, such as those proposed in [10] or in [19]. When a switch sensor is activated, the corresponding pixel value is triggered to on. Binary images only provide information related to people or object positions on the sensing area.

Additional information related to the object/person weight or to the dynamical interaction with the floor are lost.

On the contrary, grey-level pressure images are generated by floors made using regular distribution of continuous sensors, such as those described in [18] and in [1]. In this case, each sensor response and the corresponding pixel values are changing as a function of the applied pressure. Grey-level images are characterized by a higher information content.

Despite its simplicity, PIM allows to use common video processing and analysis techniques to provide people detection and the further classification of their behaviors [9]. In particular, background subtraction techniques based on Gaussian or median distributions of the pixel values are used to extract the foreground regions, position-based trackers are employed to follow sensed people in the course of time, while machine learning classifiers are implemented for the high level action or interaction analysis as in [16],[8] and [11]. As a consequence, sensor floors can be adopted as input devices in a plethora of applications, spanning from multimedia content access to surveillance, from entertainment to medical rehabilitation.

Nevertheless, two major drawbacks characterize traditional approaches. First, their implementation is not straightforward in absence of a regular spatial distribution of the sensors. Second, symmetric and pixel-wise statistical models (such as Gaussian or median distributions) are not suitable for floor sensors.

In gait-postural analysis, the exploitation of force platforms made of piezoelectric sensors, capacitance gauges, strain gauges, or FSR can be considered a common practice. Kinesiologists can estimate the *Ground Reaction Force* (GRF) and the *Center of Pressure* (COP) of a person standing or moving on these measuring instruments. The first is the vector sum of the normal components of the forces exerted on the top of the measuring platform. The second is the point location of the vertical GRF vector and represents a weighted average of all the pressures over the surface of the area in contact with the platform [21]. The temporal analysis of GRF variations and COP displacements allows to detect people (see [20], [13], [4] and [6]).

For these reasons, starting from the concepts of COP and GRF, the aim of this work is to describe a model which overcomes the above mentioned drawbacks.

The paper is structured as follows: in Section 2 we introduce a COP based data model and we propose a possible implementation of a detection and tracking algorithm developed on the basis of it. Section 3 describes the experimental setup, the obtained results and a comparison with the traditional PIM model. Finally, conclusions and future works are drawn in Section 4.

2 The COP MODEL

2.1 Data Model

As a reference technology, we have adopted the sensing floor solution described in Lombardi et al. [9]. The device is composed of a sensing layer covered by a grid

of ceramic tiles. The sensing layer is obtained by disposing on the ground plane a set of sensing elements. Differently from the PIM approach, a free distribution of the sensing elements is allowed and is not imposed to follow a grid. Each sensor s_i is thus identified by its real position on the floor (X_s, Y_s), instead of its indexes on the grid. The tiles coverage has been included to preserve the integrity of the sensors and, at the same time, to diffuse the pressures exerted on a single point of the floor to a neighbor area.

At each capturing interval t, all the sensors $s \in S$ provide a corresponding discrete response $V_s(t)$. The whole area covered by the set of sensors can be partitioned into convex polygonal cells, having the sensors as vertices. The type of polygonal decomposition strictly depends on the sensor layout and the spatial resolution. Two very different cases are reported in Figure 1. If the sensors are placed with a regular grid distribution as usual, a possible decomposition is composed by rectangular cells as reported in Figure 1(b). The obtained areas are called **Floor Cells** hereinafter, and are represented by the set of sensors placed at the vertices: $FC_i \subset S, (i = 1, 2, \cdots, M)$.

The state of a Floor Cell FC_i is represented by a $3D$ point P_i $(i = 1, 2, \cdots, M)$ as follows:

$$P_i(t) = \begin{bmatrix} P_i^x(t) \\ P_i^y(t) \\ P_i^z(t) \end{bmatrix} = \left[\frac{\sum\limits_{\forall s \in FC_i} X_s \cdot V_s(t)}{\sum\limits_{\forall s \in FC_i} V_s(t)}, \frac{\sum\limits_{\forall s \in FC_i} Y_s \cdot V_s(t)}{\sum\limits_{\forall s \in FC_i} V_s(t)}, \sum_{\forall s \in FC_i} V_s(t) \right]^T . \quad (1)$$

The first two coordinates P_i^x, P_i^y are respectively equivalent to the plane coordinates of the COP, while the third one P_i^z is the intensity of the GRF associated to the floor cell.

When no pressures is exerted on a floor cell FC_i, the location P_i^{eq} of the 3D point is only influenced by the sensor calibration and the dead weight of the tiles. Instead, when a person walks on the floor, the corresponding pressure moves the point P_i from the equilibrium state P_i^{eq} toward a new position $P_i(t)$. From Eq. 1, the projection of P_i on the ground plane falls within the FC_i convex-hull.

At time t, a floor cell FC_i can be considered in an excited state (i.e., a person or an object is located on the corresponding region) if the following condition occurs:

$$||N_i \cdot (P_i(t) - P_i^{eq})||_2 \geq TH \quad (2)$$

where TH is an application defined threshold, $|| \cdot ||_2$ is the Euclidean norm, N_i indicates the following normalization matrix:

$$N_i = diag\left(d_{x,i}^{-1}, d_{y,i}^{-1}, d_{z,i}^{-1}\right). \quad (3)$$

The normalization matrix is defined for each floor cell FC_i and takes into account the geometrical extent of the cell itself. $d_{x,i}$ and $d_{y,i}$ are the dimensions

of the FC_i convex-hull, while $d_{z,i}$ is the maximum variation of the GRF intensity and is estimated during a calibration phase of the capturing board. As a consequence of the random spatial distribution of the sensors, even a single person may trigger more than one floor cell. Thus, a cluster of neighbor floor cells will switch to a non-equilibrium condition for each person located on the sensing floor. The temporal analysis and tracking of these clusters allows to detect and track people on the floor, as detailed in the following section. A visual example

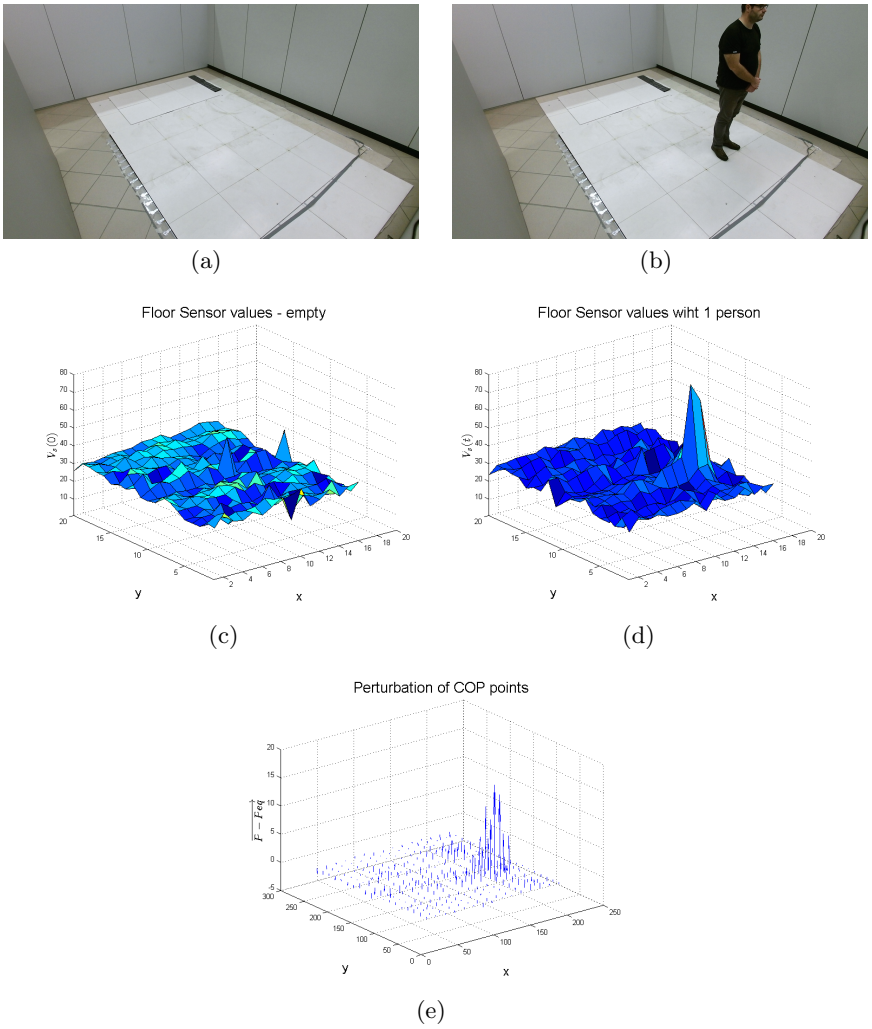


Fig. 2. Visual example of the COP vectors. (a) and (b) pictures of the empty floor and with a walking person. The sensor values captured at the equilibrium (c) and with the walking person (d). On the right (e), a plot of the vectors $P_i(t) - P_i^{eq}$ evaluated in equation 2.

is reported in Figure 2. The sensor values captured at the equilibrium and with a walking person are reported on the left and center graphs. On the right, a plot of the vectors $P_i(t) - P_i^{eq}$ evaluated in equation 2 are shown.

2.2 People Detection and Tracking

Let $C_j(t) = \{FC_k\}$ be a cluster of neighbor floor cells which are simultaneously excited at time t. Eq. 2 filters out contributions due to the noise and assures that the cluster has been generated by a person on the floor. His position $(B_j^x(t), B_j^y(t))$ on the floor at time t can be estimated from the 3D points P_i associated to the floor cells included in the cluster as in equation (4):

$$\begin{aligned} M_j(t) &= \sum_i P_i^z(t) \\ B_j^x(t) &= \frac{1}{M_j(t)} \textstyle\sum_i P_i^x(t) \cdot P_i^z(t) \\ B_j^y(t) &= \frac{1}{M_j(t)} \textstyle\sum_i P_i^y(t) \cdot P_i^z(t) \end{aligned}, \tag{4}$$

The set of clusters is obtained with a connected component labeling of all the excited cells. Two FCs are defined as connected if their intersection contains at least one sensor. For a uniform grid distribution, this assumption is similar to the 8-connection of pixels (See Fig. 1(b)).

Successive detections of the same person are temporally tracked with a nearest neighbor matching based on positions only. The main purpose of the tracking step is the recovering of people's positions in some short temporal slots, during which the floor is not able to detect them. For example, when a person changes the front feet during a walking or, more clearly, during a jump, the pressure exerted on the floor is null.

Given the detections D_i at frame t and the current set of tracks T_j, we first compute the Euclidean distance matrix $\Gamma(i,j) = D_2(D_i, T_j)$. The detection-to-track association is provided using the schema proposed in [14]. For each frame, some detections may be assigned to tracks, while other detections and tracks may remain unassigned. The assigned tracks are updated using the corresponding detections. The unassigned tracks are marked invisible. Finally, unassigned detections begin new tracks. Each track keeps count of the number of consecutive frames where it remained unassigned. If the count exceeds a specified threshold, the tracking algorithm assumes that the object left the floor and it deletes the track.

Since the applications described in this paper are not required to work in very crowd situations, the implemented tracking algorithm does not handle groups (i.e., people closer to each other than a foot step) nor abrupt position changes (e.g., people leaping around). The reader can refer to [17] for more complex tracking schemes, if required by the application.

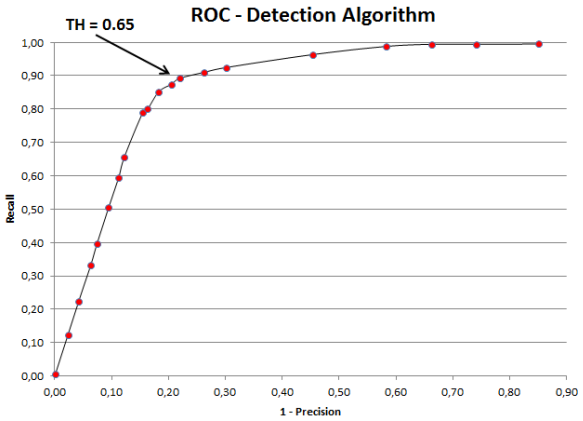


Fig. 3. ROC curve at different thresholds of the detection algorithm proposed in Section 2.2 on a calibration sequence with a person of 45 Kg on the sensing floor.

3 Experimental Evaluation

3.1 Experimental Setup

To evaluate the proposed method, we have exploited a *Florimage Device* distributed by Florim Ceramiche SpA [1] and described in [9]. The device is covered by tiles of $600\,mm \times 600\,mm$, thin enough ($4.5\,mm$) to allow the sensing elements below them to capture the presence of walking people. The sensors are distributed on a regular grid of 16 rows by 20 columns. The 320 sensing units covers a rectangular area of 4 square meters.

The experimental environment is also equipped with a Microsoft Kinect sensor. The acquisition of the floor data and of the Kinect sensor are synchronized with an external trigger (set to work at 10Hz in these experiments). The people detection and tracking capabilities of the Kinect subsystem have been exploited to automatically generate the ground truth, composed by the set of people positions on the floor. The Extrinsic calibration parameters of the Kinect device with respect to the floor have been estimated to transform the 3D coordinates of people feet joints into floor coordinates.

Using the experimental setup, 6 data sequences[2] of walking people have been acquired, involving 1 to 7 individuals. People weights are also ranging from 45 Kg to 100 Kg.

The algorithm has been evaluated counting the number of True Positive (TP), False Positive (FP) and False Negative (FN) detections. A detection provided using the algorithm reported in Section 2.2 is counted as a TP if the Kinect sensor provides a corresponding person position closer than 25 cm. The precision

[1] `http://www.slim4plus.it/en/floor-sensor-system/`
[2] Dataset available at `http://imagelab.ing.unimore.it/go/sensingFloor`

Table 1. Detection algorithm results

	Seq1		Seq2		Seq3		Seq4		Seq5		Seq6	
	PIM	COP	PIM	COP	PIM	COP	PIM	COP	PIM	COP	PIM	COP
Nframes	2353		1936		1504		1060		1311		1901	
Precision	0.79	**0.91**	0.53	**0.87**	0.77	**0.95**	**0.79**	0.76	0.79	**0.89**	0.71	**0.87**
Accuracy	0.86	**0.92**	0.53	**0.86**	0.78	**0.96**	**0.83**	0.81	0.81	**0.91**	0.65	**0.84**
TP	1245	1204	1333	1253	1287	1286	699	699	1028	1027	1793	1921
FP	326	121	1161	192	393	65	183	218	280	126	737	285
TN	861	1024	18	479	87	173	210	220	157	271	88	314
MD (cm)	20.68	**20.44**	24.56	**20.37**	17.37	**14.63**	19.70	**19.02**	**15.40**	15.69	21.03	**19.11**

Table 2. Tracking algorithm results

	Seq1		Seq2		Seq3		Seq4		Seq5		Seq6	
	PIM	COP	PIM	COP	PIM	COP	PIM	COP	PIM	COP	PIM	COP
Nframes	2353		1936		1504		1060		1311		1901	
Precision	0.84	**0.94**	0.57	**0.90**	0.80	**0.96**	**0.82**	0.79	0.83	**0.91**	0.75	**0.90**
Accuracy	0.88	**0.94**	0.56	**0.88**	0.81	**0.97**	**0.85**	0.83	0.85	**0.92**	0.68	**0.86**
TP	1232	1203	1305	1249	1284	1286	698	699	1026	1027	1768	1920
FP	239	82	999	143	326	48	157	190	213	105	598	223
TN	905	1027	35	482	95	174	217	221	173	272	106	318

$Pr = TP/(TP + FP)$ and accuracy $Ac = TP/(TP + FN)$ metrics are also computed as usual.

Thanks to the regular distribution of the sensing elements on a grid, $M = 285$ floor cells FC_i of 4 sensors each have been generated using the schema reported in Figure 1. For each floor cell FC_i, the coordinates of the equilibrium state point P_i^{eq} were estimated by averaging a short sequence of sensor values captured with the empty floor. The threshold TH of Equation 2 has been set to 0.65 by maximizing the precision and recall of the detection algorithm proposed in Section (2.2) on a calibration sequence. The corresponding ROC curve obtained at different threshold values is reported in Figure 3.

3.2 Reference PIM method

As a baseline, we have implemented a PIM based detection and tracking algorithm, following the recommendations provided in [9].

Let $I(x, y, t)$ be the pressure image obtained at time t in the PIM based approach. Each pixel intensity is proportional to the value $V_s(t)$ captured by the corresponding sensor. In order to filter the contributions due to the dead weight of the tiles coverage, each pressure images is pre-processed using a background subtraction approach as follows:

$$I^{\dagger}(x, y, t) = I(x, y, t) - I_{eq} \tag{5}$$

where I_{eq} is the pressure image acquired when no people are walking on the sensing floor. Peaks on the image $I^{\dagger}(x, y, t)$ that are higher than a threshold are

considered as activated (i.e., generated by a person or an object moving on the sensing floor). The detections $D(t)$ are obtained through a mean-shift clustering of the activated sensors. Similarly to the COP based system, a nearest-neighbor tracking algorithm is included in the processing chain (see Section 2.2).

3.3 Quantitative Results

We tested and compared the proposed method and the baseline PIM algorithm on the six sequences described in section 3.1. Table 1 and Table 2 report the values of all the estimated performance parameters for each sequence, with or without the tracking stage, respectively. In each table, the best results in term of precision and accuracy are highlighted. The precision and accuracy obtained with the COP model are higher than those obtained with the PIM one, except for the 4-th sequence.

The mean distance between the detection and the ground truth positions is also reported in the last row of Table 1. The closest mean distances are those obtained with the COP based detection algorithm.

4 Conclusion and future works

In this paper, we proposed a new data model for storing and processing information acquired by sensing floor. Due to the customary regular distribution of the sensor units, their values are usually stored as pressure images and processed with common computer vision algorithms [10, 9]. These methods are difficult to apply to a general sensor distribution.

Instead, the proposed COP model does not assume a regular grid distribution of the sensing elements and is based on the ground reaction force (GRF) concept, widely used in biomechanics. It allows the correct detection and tracking of people, outperforming the common background subtraction schema exploited in the past. Several tests on a real sensing floor prototype confirm the validity of the model and the outperforming capabilities on people detection and tracking.

As future work, we plan to deeply address the Floor Cell creation step, taking into account different shapes and sizes of the floor cells. In addition, we want to extent the method to the case of overlapping cells, which requires a new definition of connection between cells for the clustering task. Finally, a more sophisticated detection and tracking algorithm based on both the module and the direction of the COP vectors will be handled.

Acknowledgments

This work was supported by Florim Ceramiche S.p.A. (Italy)

References

1. Anlauff, J., Großhauser, T., Hermann, T.: tactiles: A low-cost modular tactile sensing system for floor interactions. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. pp. 591–594. ACM, New York, NY, USA (2010)
2. Betke, M., Gips, J., Fleming, P.: The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. IEEE Transactions on Neural Systems and Rehabilitation Engineering 10(1), 1–10 (March 2002)
3. Blake, J., Gurocak, H.: Haptic glove with mr brakes for virtual reality. IEEE/ASME Transactions on Mechatronics 14(5), 606–615 (Oct 2009)
4. Headon, R., Curwen, R.: Recognizing movements from the ground reaction force. In: Proceedings of the 2001 Workshop on Perceptive User Interfaces. pp. 1–8. PUI '01, ACM, New York, NY, USA (2001)
5. Hsu, J.M., Wu, W.J., Chang, I.R.: Ubiquitous multimedia information delivering service for smart home. In: International Conference on Multimedia and Ubiquitous Engineering. pp. 341–346 (April 2007)
6. Jung, J.W., Sato, T., Bien, Z.: Dynamic footprint-based person recognition method using a hidden markov model and a neural network. International Journal of Intelligent Systems 19(11), 1127–1141 (2004)
7. Lai, K., Konrad, J., Ishwar, P.: A gesture-driven computer interface using kinect. In: IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI). pp. 185–188 (April 2012)
8. Leusmann, P., Mollering, C., Klack, L., Kasugai, K., Ziefle, M., Rumpe, B.: Your floor knows where you are: Sensing and acquisition of movement data. In: 12th IEEE International Conference on Mobile Data Management (MDM). vol. 2, pp. 61–66 (June 2011)
9. Lombardi, M., Pieracci, A., Santinelli, P., Vezzani, R., Cucchiara, R.: Human behavior understanding with wide area sensing floors. In: Proceedings of the 4th International Workshop on Human Behavior Understading (HBU2013). Barcelona, Spain (Oct 2013)
10. Middleton, L., Buss, A., Bazin, A., Nixon, M.: A floor sensor system for gait recognition. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies. pp. 171–176 (Oct 2005)
11. Murakita, T., Ikeda, T., Ishiguro, H.: Human tracking using floor sensors based on the markov chain monte carlo method. In: Proceedings of the 17th International Conference on Pattern Recognition. vol. 4, pp. 917–920 Vol.4 (Aug 2004)
12. Paradiso, J., Abler, C., Hsiao, K., Reynolds, M.: The magic carpet: physical sensing for immersive environments. In: Extended Abstracts on Human Factors in Computing Systems. pp. 277–278 (1997)
13. Qian, G., Zhang, J., Kidane, A.: People identification using floor pressure sensing and analysis. IEEE Sensors Journal 10(9), 1447–1460 (Sept 2010)
14. Rangarajan, K., Shah, M.: Establishing motion correspondence. CVGIP: Image Understanding 54(1), 56 – 73 (1991)
15. Rangarajan, S., Kidane, A., Qian, G., Rajko, S., Birchfield, D.: The design of a pressure sensing floor for movement-based human computer interaction. In: Proc. of the 2nd European conference on Smart sensing and context. pp. 46–61. EuroSSC'07, Springer-Verlag, Berlin, Heidelberg (2007)
16. Rimminen, H., Lindstrom, J., Linnavuo, M., Sepponen, R.: Detection of falls among the elderly by a floor sensor using the electric near field. IEEE Transactions on Information Technology in Biomedicine 14(6), 1475–1476 (Nov 2010)

17. Smeulder, A., Chu, D., Cucchiara, R., Calderara, S., Deghan, A., Shah, M.: Visual tracking: an experimental survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 36(7), 1442–1468 (Jul 2014)
18. Srinivasan, P., Birchfield, D., Qian, G., Kidané, A.: A pressure sensing floor for interactive media applications. In: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology. pp. 278–281. ACE '05, ACM, New York, NY, USA (2005)
19. Valtonen, M., Maentausta, J., Vanhala, J.: Tiletrack: Capacitive human tracking using floor tiles. In: IEEE International Conference on Pervasive Computing and Communications. pp. 1–10 (March 2009)
20. Vera-Rodriguez, R., Mason, J., Fierrez, J., Ortega-Garcia, J.: Comparative analysis and fusion of spatiotemporal information for footstep recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(4), 823–834 (April 2013)
21. Winter, D.: Human balance and posture control during standing and walking. Gait & Posture 3(4), 193 – 214 (1995)