Human Behavior Understanding with Wide Area Sensing Floors

Martino Lombardi, Augusto Pieracci, Paolo Santinelli, Roberto Vezzani, and Rita Cucchiara

Softech-ICT, University of Modena and Reggio Emilia
Via Vignolese, 905 - 41125 Modena - Italy
{martino.lombardi,augusto.pieracci,paolo.santinelli
roberto.vezzani,rita.cucchiara}@unimore.it
http://imagelab.ing.unimore.it

Abstract. The research on innovative and natural interfaces aims at developing devices able to capture and understand the human behavior without the need of a direct interaction. In this paper we propose and describe a framework based on a sensing floor device. The pressure field generated by people or objects standing on the floor is captured and analyzed. Local and global features are computed by a low level processing unit and sent to high level interfaces. The framework can be used in different applications, such as entertainment, education or surveillance. A detailed description of the sensing element and the processing architectures is provided, together with some sample applications developed to test the device capabilities.

Keywords: Sensing floor, human behavior understanding, OSC, natural interfaces.

1 Introduction

The creation of sensing environments constitutes one of the major trends in interfacing humans with virtual worlds, computers and cyberspaces. Natural and new human computer interfaces overcome now standard interfaces: for instance, human actions and events can be automatically captured and decoded by distributed sensors, without the need of direct interaction with input devices such as pads, mouses, keyboards, and so on. Entertainment and edutainment applications, for example, got a great benefit from hands-free visual devices, such as the well known Microsoft Kinect, webcams or smart cameras.

Vision based devices are suitable for behavior recognition of single users, capturing their details and body parts with a suitable resolution. However, they have some drawbacks for monitoring multiple people or covering wide areas. This last case, in fact, requires a network of sensors. In addition, the placement constraints and the handling of occlusions due to furniture, objects and other people usually impose the adoption of multiple views, especially in indoor environments. Furthermore, privacy issues strongly limited the usability and also the user acceptability in some cases.

Following the same objective of hands-free devices, we propose an innovative sensing floor called "Florimage device", that is able to capture and measure the pressure field exerted by people or objects on the tiles. Instead of optical images, the floor is able to generate a "pressure image", where each pixel corresponds to a spatial portion of the floor and the pixel value is related to the pressure applied by people or object onto the floor itself. The technology we developed is modular and scalable (See Figure 1(a)). This result is obtained adding a sensing layer below the ceramic tiles. The solution proposed is cheap enough to allow the coverage of wide areas and the sensing elements do not change the design or the appearance of the floor.

The proposed sensing floor allows a plethora of applications, spanning from entertainment to surveillance, from multimedia content access to medical rehabilitation. In this paper, and in particular in Section 4, we will describe some sample applications we developed to test the system capabilities.

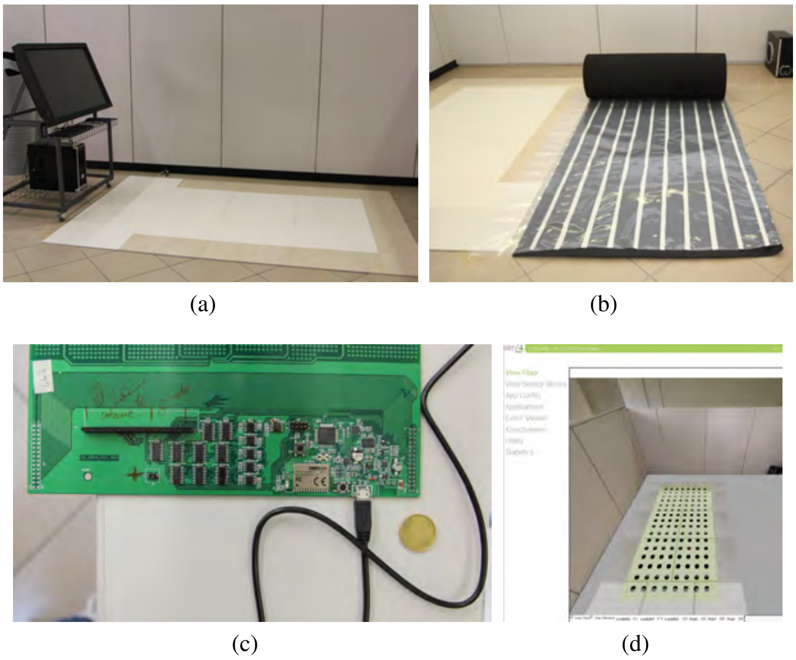


Fig. 1. Pictures of the installed prototype. a) the prototype floor, b) the contact stripes, c) the capturing board and d) a virtual reconstruction of the environment.

Some preliminary experiments have been described in [1], where an initial prototype of the "Florimage device" was exploited to monitor indoor environments for surveillance applications. In this paper we provided a detailed characterization of the sensor response (the *floor image*), the description of the data processing system and the methodology to understand the human behavior by pressure images. In particular, two processing layers have been defined, working at a low and high level respectively. The two layers communicate through the Open Sound Control (OSC) protocol [2, 3], that spreads data and events detected from the low level processing to each client application.

In Section 4 we will present the following five sample applications:

- *Human Media Player*: the person location on the floor controls a media player setting the current position in the media item;
- *Memory game*: the player can turn the cards of the well known memory game by jumping on the corresponding tile;
- *Music edutainment*: the floor is divided in small portions that emulate a piano key, allowing several edutainment games in the music field;
- *Foot painting*: the user paints original artworks or allow to reveal a hidden painting by changing his position on the floor;
- *Intrusion detection*: safe regions could be defined and the sensing floor will detect people inside them.

Some prototypes of sensing floors have been developed and proposed in the past. A representative subset of them is described in Section 2. However, none of them fulfills all the following requirements, which are mandatory to generate a reliable solution:

- *low cost*: the cost of the sensing elements should be comparable to traditional floors;
- *high scalability*: the sensors should be integrated into a hierarchical network, in order to allow the coverage of narrow rooms as well as wide areas;
- *transparent design*: for design issues, the sensing layer must be invisible to the users and the floor should appear like traditional floors;
- *high reliability and durability*: breakable and fragile elements should be avoided or limited to protected packages;
- *temporal and spatial resolutions*: they should be high enough to allow people detection and tracking, even in presence of multiple targets.

The "Florimage device" is still under developing within the long-term project "Florimage: the floor is an Image" funded by Florim S.p.A and Regione Emilia Romagna. The project aims at creating a flexible and general purpose platform for human behavior understanding and interaction.

2 Related Works

Several prototypes of sensing floors for human-based action detection and identification have been presented in the past. The adopted sensors exploited different physical characteristics; among the others, the pressure as measurable quantity, and the proximity effect related with the electrical properties of a human body are the mainly used.

Lee Middleton et al. [4] developed a prototype of a floor sensor mat as a gait recognition system. The sensor consists of individual switches arranged in a separated pair of wires by foam a deformable material. The design shows simplicity and scalability even if switches do not provide a response commensurate with the strength of the applied force. Chen-Rong Yu et al. [5] proposed a localization system to accurately estimate human position. It performs single person and multiple people tracking in a home environment. The Condensation algorithm is exploited to locate residents' position via multi-camera and sensory floor approaches. Domnic Savio et al [6] developed a smart carpet that can be laid on the floor. The sensor set forms a self-organizing sensor network. To identify the footstep, clustering algorithms based on Maximum Likelihood Estimate and Rank Regression have been applied. The proposed approach is scalable and commercially viable even if the binary nature of the embedded nodes does not provide a response commensurate with the applied weight. Miika Valtonen et al [7] described an unobtrusive two-dimensional human positioning and tracking system based on a low-frequency electric field. The capacitance between multiple floor tiles and a receiving electrode is measured. The method is based on the fact that humans well conducts a low-frequency signal. The proposed method does not provide any information related with the human weight; it is not possible to detect any kind of non conductive objects, regardless of the weight; also the conditioning circuits are very complex. In the work by Prashant Srinivasan et al. [8] a portable high-resolution pressure sensing floor prototype able to detect pressure information about the human interaction with the system is described. The pressure sensing floor consists of several sensor mats; each of them is composed of a 42x48 grid of pressure sensors with size of 48.8 cm x 42.7 cm. The sensor elements of the mat are made using a pressure sensitive polymer between conductive teaks on sheets of Mylar. The sensor elements change the resistance with the applied pressure. Jan Anlauff et al. [9] presented a prototype of a floor surface based on sensing elements made out of conductive black art paper. The sensing elements are grouped into modules forming a grid of resistors able to measure quasi-static forces. The proposed approach is a low cost alternative for spatially resolved tactile sensing. The employed signal conditioning system is based on the matrix arrangements of the sensing elements but the proposed solution suffers of mutual interference between different sensors in the matrix. Finally, in the work by Rishi Rajalingham et al. [10] a probabilistic approach to the tracking and estimation of the lower body posture is presented. Their sensing floor has limited sizes, it is not easily scalable and commercially viable. It employes off the shelf resistive force sensors. It consists of a 66 array of rigid tiles, 30 cm on each side. Tiles are equipped with four resistive force sensors, which are located at the corners. An array of six small-form-factor computers is used for data processing.

3 Hardware Architecture

The modularity of the system is one of the key aspect of the proposed sensing floor. As depicted in Figure 2, the system is composed of five main items: the sensors, the capturing boards, the communication network, the processing unit, and the client applications.

The sensors are obtained by introducing a sensing layer below a commercial floating floor. The sensing stratum replaces the polymeric layer which is usually adopted to cobble surfaces using floating floors. Each element is located in the crosses of a regular grid. A board equipped with micro controllers is connected to each matrix of sensors (Base Node). The sensed values are sampled and digitized at a constant frame rate and

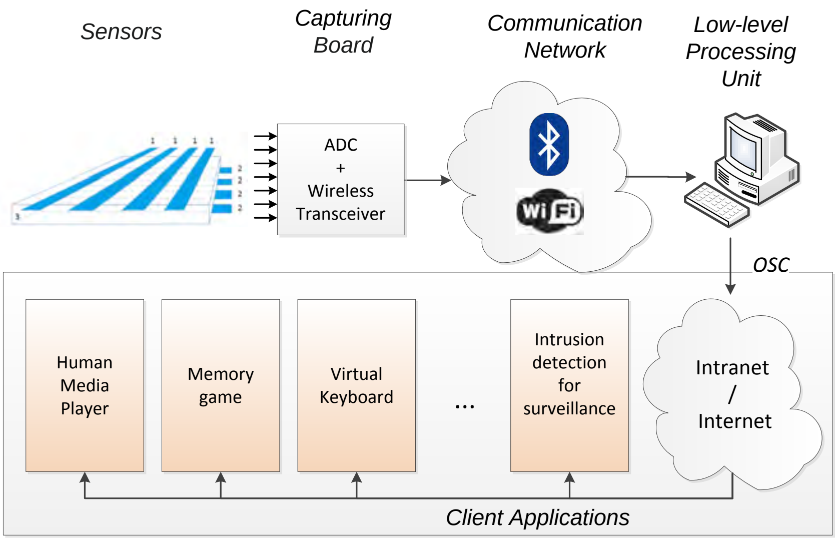


Fig. 2. The five main components of the Florimage architecture.

sent via Blue-tooth to a processing unit for their analysis. A hierarchical communication network is also proposed for scalability reasons and it can be used when the number of BNs exceeds the capabilities of a single processing unit. A low-level data processing analyzes the data, filters noise and errors, detects actions and events. The obtained information are sent to client applications using the Open Sound Control (OSC) protocol [2], as detailed in Section 4.

Differently from other solutions which have been proposed in the past [11, 8, 12], each sensing element looks like a normal floating tile. Thus, it inherits some key advantages: low cost, easy tile replacement, high resilience, high stability, feasible relocation of the floor. The proposed approach allows to make wide sensing areas at affordable costs. Some pictures of the prototype floor and the embedded board used to capture and transmit the data are reported in Figure 1.

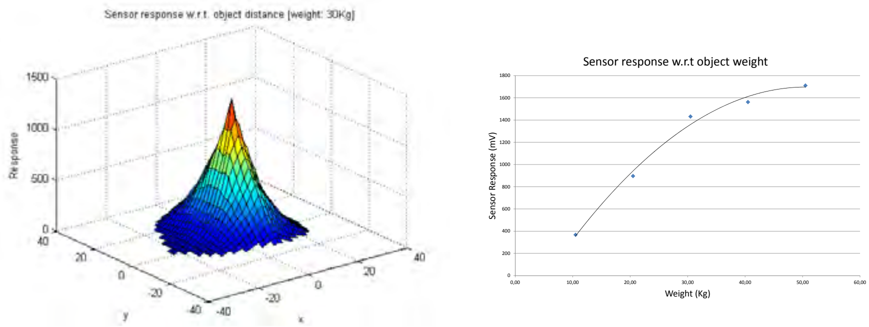
3.1 Sensing elements

The key element of the proposed sensors is represented by a commercial paving technology, called floating floors. A floating floor does not need to be nailed or glued to the sublayer and thus it might be constructed over a sub-floor or even over an existing floor. It consist of a polymeric, felt or cork layer holding up a laminate floor. Drowning commercial pressure sensors such as FSR or piezoelectric elements into the padding layer allows to reach very high performances in terms of measurement precision and reliability. However, the solution becomes too expensive and thus unfeasible in wide areas.

We propose to replace the traditional polymeric layer with a sensing element made using a sandwich structure. A conductive polymer (3) is put between aluminum stripe electrodes, lengthwise (1) and crosswise (2) (see Figure 2), generating a grid of sensors, one on each cross. When a pressure is applied on the top of the tiles, the rough surface of the conductive rubber is compressed onto the electrodes surface [13]. As a result, the contacting area between rubbers and electrodes is increased whilst the resistance between them is proportionally reduced. The resistance at each cross is acquired using the board shown in Figure 4(b). We exploited very slim ceramic tiles for the higher layer. The tiles are 600mmx600mm large and thin enough (4.5 mm) to allow the sensing elements below them to capture the presence of walking people.

The sensor response has been characterized placing objects with an increasing weight on a tile and measuring the resistance of the sensor located below the object. Figure 3 plots the sensor response as a function of the applied weight. Even if the response is not linear, it can be reasonably approximated with a polynomial expression of the second order. The polynomial coefficients can be calibrated given the specific floor. However, in the described applications we are interested in founding reasonable thresholds only, which is guaranteed by the monotonic behavior of the response shown in the graph.

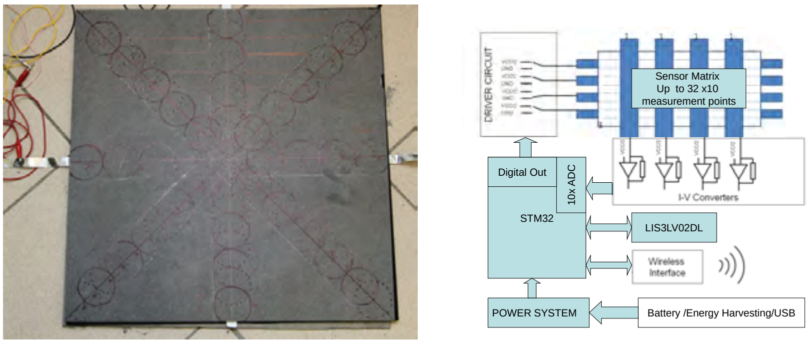
In addition to esthetic and design reasons, the ceramic tiles play an important role in the sensor device, since they act as a blurring filter, spreading the pressure of the objects on a wide area. As a consequence, each sensor captures the weights applied on a neighborhood of its position. To test and measure this behavior of the tiles, we have placed the weights at different positions with respect to the sensing node. A picture of the tile used for the experiment is reported in Figure 4(a). The object positions have been marked on the tile and the sensing node has been placed in the middle of the tile. Figure 3(b) plots the sensor response as a function of the sensor-to-object distance.



(a) (b)

Fig. 3. Plot of the sensor response w.r.t. a) the position of the object and b) the object weight.

Each Base Node is equipped with a capturing board which measures the electric resistance at each grid cross, digitizes and transmits the values to the processing unit. In Figure 4(b) the block diagram of the electronic part is shown. The frames are sent to the processing unit which runs the system application through a wireless communication



(a) (b)

Fig. 4. a) the portion of floor and b) the schema of the electronic device used to characterize the sensor response.

interface. More details on the capturing board and the communication network can be found in [1].

4 From pressure to Human Behavior Understanding

The software part of the framework has been split into two different layers. A low-level pre-processing step collects the data coming from different sensor nodes, filters the noise, computes feature vectors and detect some simple events as described in [1]. Filtered data as well as the detected events are sent to the high-level applications. The intermediate protocol used between the low and the high level layers is OSC, as shown in Figure 2.

The processing unit collects, processes and stores the data from all the sensors. Thanks to the regular distribution of the sensors below the floor, the processing unit firstly generates a “floor image”, in which each sensor corresponds to a pixel and the spatial neighborhood relations are preserved. Subsequent sensor samples are handled as frames in a video, allowing the adoption of common image and video processing techniques to extract high level information.

Since the recognition of the people behavior using video cameras has been deeply addressed in the past using spatio-temporal descriptors[14, 15], we applied the same approach to the sensor data.

In video analysis, spatio-temporal 3D patches can be extracted selecting rectangular regions and aggregating successive frames cropped on them. To simulate the same approach, we selected a single tile of the floor and we collected the NxM values of the sensors below it during a short time interval (2 seconds, corresponding to 20 frames). The obtained matrix $I(x, y, t)$ is processed in order to generate a feature vector for the classification step. We adopted a feature vector composed by aggregated measures which do not depend on the spatial and temporal position of the selected event within the analyzed block. In other words, the feature vector does not depend on the absolute position of a person within the analyzed floor tile. Let $M(t)$ be the mean value of I at time t, $x(t) = \sum_{x=1}^{M} \sum_{y=1}^{N} (x \cdot I(x, y, t))$ and $y(t) = \sum_{x=1}^{M} \sum_{y=1}^{N} (y \cdot I(x, y, t))$ the coordinates of the barycenter $b(t) = (x(t), y(t))$. $C(t)$ is the covariance matrix of the vectors $\{x, y, I(x, y, t)\}$. Let be $\{e_1(t), e_2(t), e_3(t)\}$ the three eigenvalue of the matrix $C(t)$. The features $\Phi = \{\phi_1, \ldots, \phi_{15}\}$ are extracted as average, min or max of the previous defined values as follows:

$$
\begin{array}{l}
\phi_1 = \frac{1}{n} \cdot \sum_{t=1}^{n} M(t) \\
\phi_2 = \frac{1}{n} \cdot \sum_{t=1}^{n} \left(M(t) - \phi_1\right)^2 \\
\phi_3 = \frac{1}{n} \cdot \sum_{t=2}^{n} |M(t) - M(t-1)| \\
\phi_4 = \frac{1}{n} \cdot \sum_{t=2}^{n} \left(|M(t) - M(t-1)| - \phi_3\right)^2 \\
\phi_5 = \frac{1}{n} \cdot \sum_{t=2}^{n} ||b(t) - b(t-1)|| \\
\phi_6 = \frac{1}{n} \cdot \sum_{t=1}^{n} (||b(t) - b(t-1)|| - \phi_5)^2 \\
\{\phi_7 \ldots \phi_9\} = \frac{1}{n} \cdot \sum_{t=1}^{n} \{e_1(t) \ldots e_3(t)\} \\
\{\phi_{10} \ldots \phi_{12}\} = \min_t \{e_1(t) \ldots e_3(t)\} \\
\{\phi_{13} \ldots \phi_{15}\} = \max_t \{e_1(t) \ldots e_3(t)\}
\end{array}
\tag{1}
$$

where ϕ_1 and ϕ_2 are the mean and variance of $M(t)$ over the temporal interval; ϕ_3 and ϕ_4 are the mean and variance of the $M(t)$ variations. ϕ_5 and ϕ_6 take into account the movement of the barycenter. Finally, ϕ_7 to ϕ_{15} evaluate the average, the minimum and the maximum three eigenvalues of the covariance matrix.

The obtained feature vector has been provided as input to a set of supervised classifiers trained to detect the presence of a person or a fixed object on top of the tiles. If a person is detected, the same feature vector is evaluated to further classify its posture or actions. We developed a collection of Randomized Trees classifiers for both the tasks. A qualitative and quantitative analysis is reported in [1], that shows the very high accuracy of the approach. For example, the *jumping event* is detected with an accuracy close to 96%.

4.1 OSC communication protocol

The low level processing unit communicates with the client applications using the Open Sound Control protocol (OSC) [2] over UDP. Open Sound Control (OSC) is a digital media content format for streams of real-time audio control messages. It can be considered as a sort of evolution of the MIDI protocol. Although its origin, OSC had a lot of applications outside of audio technology, founding use in domains such as control and robotics [3].

Each application client starts with a request to the floor device, indicating its IP address and port. As a result, the floor device will send an acknowledgment as well as some useful details about the sensor (spatial extension, number of nodes and sensors, sampling frequency, and so on). After that, the client can ask the floor to receive the processed data or the original sensor values. Each type of required information can be enabled or disabled by the client through specific OSC commands.

A list of the implement OSC messages is reported in Table 1. Among the others, the following server-to-client messages have been used in the described applications. The *baricenter* message is sent after each frame capture and it indicates the baricenter of the response field. If a single person is walking on the floor, the baricenter will provide an estimation of his position.

Client to Server	
/florimage/request/register [ip] [port]	register the client [ip],[port] to the server
/florimage/request/datastream/baricenter	
/florimage/request/alive	The client should periodically send this command to keep the communication alive
/florimage/request/getconfig	The server will answer to this command by sending the configuration parameters
Server to Client	
/florimage/datastream/sensors [data]	Stream of the sensor data
/florimage/datastream/baricenter [x] [y]	Baricenter of the pressure field
/florimage/datastream/featurevector [data]	Stream of feature vectors ϕ_i
/florimage/event/person [x] [y]	Person detected at the position (x, y)
/florimage/event/jump [x] [y]	Jumping person detected at the position (x, y)

Table 1. List of the main OSC commands for the communication between the sensing floor and the client applications.

5 Interactive Applications

We have developed some interactive applications in order to test the system capabilities and to show the wide range of uses allowed by the platform. The five demo applications have been introduced in Section 1. All of them are based on the OSC messages generated by the low level processing unit. In this section we provide some additional details on the application logic and potential target.

Human Media Player The core of this application is a multimedia content. In the case of a video, the position of a user on the floor controls the media position. To this aim, we exploited only one coordinate of the *baricenter* OSC command filtered with a smoothing function to obtain a fluid playback.

$$mediapos(t) = mediapos(t-1) \cdot \lambda + medialength \cdot \frac{baricenter.x}{floor.length} \cdot (1-\lambda) \quad (2)$$

where the coefficient λ is the smoothing ratio. A screenshot of the media player application is reported in Figure 5(a).

Memory game This application is under development within a children education project. The well known *memory game* is controlled by the floor. The children should jump on a tile to flip the corresponding card. As a edutainment game, the drawing reported on the cards are collected from famous historical, television or political couples. A screenshot of the memory game application is reported in Figure 5(b).

Music edutainment In this application, the sensing floor has been divided into 81 squared cells and a key of a virtual piano has been assigned to each cell. The complete *datastream* has been used to compute the mean sensor response within each cell.

The values over a sensitivity threshold trigger a Midi note event, whose velocity is proportional to the sensor response and the duration is fixed. Walking and jumping on the floor will allow the user to play songs or to simply generate a music feedback. The corresponding screenshot is shown in Figure 5(c).

Foot Painting This application will be probably located in a showroom of out project partner. A pixelated famous painting is initially hidden. Each part of the painting is associated to a sensing element. The detection of a stimulus over the sensor will reveal the corresponding portion of the painting. The user will be encouraged to walk all around the floor to recover the entire image. After a period of inactivity of the sensor, the corresponding part of the painting is reverted to the hidden state. Two of the adopted images are shown in Figure 5(e) and Figure 5(f).

Intrusion detection and Surveillance The implementation of a people detection for surveillance applications has been described in [1], where some basic actions were classified in addition to the detection of the intruders. Figure 5(d) reports a screenshot of the surveillance dashboard, where the intrusions are highlighted with alarm events.

6 Conclusions and Future Works

In this paper we described an innovative sensing floor that can be used as input device in a plethora of applications, such as entertainment, surveillance, multimedia content access. The low cost of production, the high reliability and high scalability make the proposed solution very promising also from the commercial point of view, allowing wide installments in both private and public spaces. Preliminary tests on a laboratory prototype highlighted their sensing capabilities, which are high enough to allow the detection of people and their behavior.

Acknowledgements

This work was supported by Florim Ceramiche S.p.A. (Italy) and within the Regional Operational Programme POR FESR 2007-2013 of Softech-ICT. The foot painting and the edutainment applications are under developing in collaboration with the Italian Company Vision-e s.r.l.

References

1. Vezzani, R., Lombardi, M., Pieracci, A., Santinelli, P., Cucchiara, R.: Sensing floors for privacy-compliant surveillance of wide areas. In: Proceedings of the 10th IEEE International Conference on Advanced Video and Signal-Based Surveillance, Krakw, Poland (August 2013)
2. Wright, M., Freed, A.: Open sound control: A new protocol for communicating with sound synthesizers. In: International Computer Music Conference, Thessaloniki, Hellas, International Computer Music Association (1997) 101–104

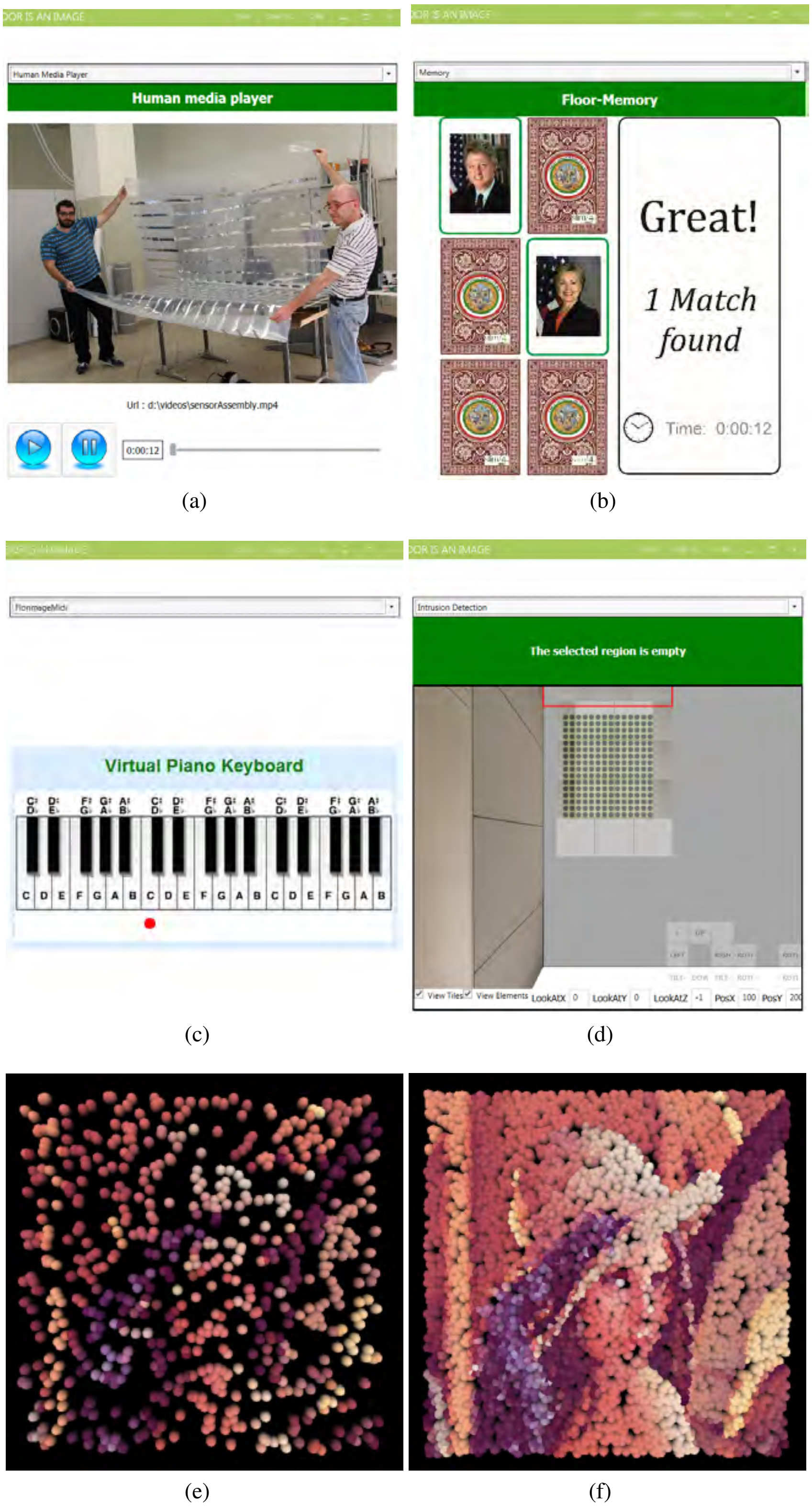


Fig. 5. Screen shots of the four sample applications described in the paper: a) human media player, b) memory game, c) music edutainment, d) intrusion detection, e-f) foot painting.

3. Schmeder, A., Freed, A., Wessel, D.: Best practices for open sound control. In: Linux Audio Conference, Utrecht, NL (01/05/2010 2010)
4. Middleton, L., Buss, A., Bazin, A., Nixon, M.: A floor sensor system for gait recognition. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies. (2005) 171–176
5. Yu, C.R., Wu, C.L., Lu, C.H., Fu, L.C.: Human localization via multi-cameras and floor sensors in smart home. In: IEEE International Conference on Systems, Man and Cybernetics, 2006. Volume 5. (2006) 3822–3827
6. Savio, D., Ludwig, T.: Smart carpet: A footstep tracking interface. In: Int'l Conf. on Advanced Information Networking and Applications Workshops. Volume 2. (May 2007) 754–760
7. Valtonen, M., Maentausta, J., Vanhala, J.: Tiletrack: Capacitive human tracking using floor tiles. In: IEEE International Conference on Pervasive Computing and Communications. (March 2009) 1–10
8. Srinivasan, P., Birchfield, D., Qian, G., Kidané, A.: A pressure sensing floor for interactive media applications. In: Proc. of the Int'l Conf. on Advances in computer entertainment technology. (2005) 278–281
9. Anlauff, J., Groß hauser, T., Hermann, T.: tactiles: a low-cost modular tactile sensing system for floor interactions. In: Proc. of the 6th Nordic Conf. on Human-Computer Interaction: Extending Boundaries, ACM (2010) 591–594
10. Rajalingham, R., Visell, Y., Cooperstock, J.: Probabilistic tracking of pedestrian movements via in-floor force sensing. In: Canadian Conference on Computer and Robot Vision. (31 2010-June 2) 143–150
11. Paradiso, J., Abler, C., Hsiao, K.y., Reynolds, M.: The magic carpet: physical sensing for immersive environments. In: Extended Abstracts on Human Factors in Computing Systems. (1997) 277–278
12. 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. EuroSSC'07, Berlin, Heidelberg, Springer-Verlag (2007) 46–61
13. Weiss, K., Worn, H.: The working principle of resistive tactile sensor cells. In: Proc. of IEEE International Conference on Mechatronics and Automation. Volume 1. (2005) 471–476 Vol. 1
14. Laptev, I.: On space-time interest points. Int. J. Comput. Vision **64**(2-3) (September 2005) 107–123
15. Gorelick, L., Blank, M., Shechtman, E., Irani, M., Basri, R.: Actions as space-time shapes. IEEE Trans. Pattern Anal. Mach. Intell. **29**(12) (December 2007) 2247–2253