

Received June 13, 2017, accepted July 22, 2017, date of publication August 3, 2017, date of current version March 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2735809

A Multi-Sensor Fusion Scheme to Increase Life Autonomy of Elderly People With Cognitive Problems

GUSTAVO HERNÁNDEZ-PEÑALOZA, (Member, IEEE), ALBERTO BELMONTE-HERNÁNDEZ, MARCOS QUINTANA, AND FEDERICO ÁLVAREZ, (Member, IEEE)

Visual Telecommunication Applications Group, Signals, Systems and Radio Communications Department, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Corresponding author: Gustavo Hernández-Peñaloza (ghp@gatv.ssr.upm.es)

This work was supported by the European Project ICT4LIFE through the European Unions Horizon 2020 Research and Innovation Programme under the Grant 690090.

ABSTRACT Elderly people care is a major challenge for the smart-cities of future. This represents a valuable opportunity to develop scalable applications to cover the special needs in terms of health monitoring and accessibility for people with cognitive impairments. In this paper, a complete system to support daily activities of elderly people based on a multi-sensor scheme is presented. This system is intended to be deployed not only at home, but also at crowded places, such as daily care centers. A multi-layer architecture is drawn to ensure system modularity and interoperability of heterogeneous data with concurrent services. The proposed system includes a set of algorithms for data gathering and processing to detect abnormal events in the considered scenarios. The experiments performed in real scenarios have led to a good performance of the algorithms proposed as well as high accuracy in event detection for both environments.

INDEX TERMS Abnormal behavior detection, healthcare, WSN, multi-modal fusion, depth sensors.

I. INTRODUCTION

The number of people living with dementia worldwide today is estimated at 44 million, set to almost double by 2030 and more than triple by 2050 [1]. Cognitive impairment, however, is further deteriorated in individuals with other diseases, such as Parkinson [2].

There is a growing interest in the employment of Information and Communication Technologies (ICT) as a response to the healthcare system requirements for aging people [3]–[5]. Most of the healthcare applications are mainly focused on the active homes monitoring which represents a considerable advance as the patients autonomy can be extended [3], [6]. Additionally, the multi-disciplinary health teams can obtain a continuous vital signs and behavior tracking which yields to appropriate diagnosis or treatment [8].

However, to guarantee the sustainability of these systems subject to special needs of elderly in terms of accessibility and mobility, several issues must be faced in organizational, medical, stakeholders and infrastructural aspects [8], [9].

Furthermore, elder people activities must not be only constrained to home care, but it must be considered as part of a global monitoring in their society [8], [10]. In this

sense, smart cities play a key role, as some activities can be included within the range of services aiming for improving the citizens' quality of life [10]. The impact that healthcare services for smart cities can have into population has been quantified in up to 12 billion in 2020 [9]. In the context of integral healthcare applications, multiple opportunities from a research perspective arise.

In the recent years, advances in ICTs have allowed the development of low-cost devices for monitoring complex activities. As an example, camera deployments for people tracking [11], Wireless Sensor Network (WSN) based technologies permit non-intrusive monitoring [6]. Additionally, inclusion of e-health sensors in wearable devices have increased the potential of full-time monitoring applications [16]. However, there exist several constrains associated to each technology, such as occlusions and lighting changes in camera-based systems; accuracy in wireless devices among others [17]. To tackle these factors, an interesting alternative is given by fusing the information from multiple sensors.

In this paper, a multi-modal sensor fusion scheme for health monitoring is presented. This system considers the medical, ethical and functional issues for data retrieving and

synchronization in Parkinson Disease (PD) patients. Additionally, the system is intended to detect abnormal events by the proper fusion of the measurements gathered from the sensors. The system is able to detect several persons simultaneously and associate the data from the sensors considered to an individual. This is a significant advantage of the proposed system as it allows to extend the scope of detection and tracking to multiple patients.

Several works have been presented for integrated healthcare using ICT-based solutions. A complete review of WSN systems for healthcare applications is drawn in [5]. Moreover, traditional health monitoring systems rely on the use of vision sensors [11]. Recently, implementation of depth sensors have permitted to analyze body movements in a detailed manner by the employment of interesting features such as the joint skeleton detection [7]. Probably the most similar work to the one presented here is [4]. In this paper, a complete architecture for homecare monitoring is described. In this architecture, information from infrared sensors, microphones and depth cameras is centralized in a gateway for homecare. However, the system presented in this work aims to be modular as it is able to detect events even with partial information. Furthermore, healthcare services are intended to be deployed and work in several scenarios. The system proposed is capable of monitoring several pedestrians and individualize the events detected for analysis which is called in this paper as person re-identification.

The main contributions of this paper are:

- A multi-sensor modular scheme for indoor people monitoring.
- A set of modules for the implementation of tracking and sensing functionalities for detection of abnormal events in PD patients.
- A framework to associate measurements gathered from diverse sensors to the corresponding person.
- A collaborative system that allows the appropriate fusion of multiple sensors.

The remainder of this paper is organized as follows: in Section II, the general architecture of the proposed system is outlined. Furthermore, in Section III and IV-A, the main modules of the proposed system are presented. In Section V, the methodological aspects of the experiments, as well as setup and results of the described modules are drawn. Finally, conclusions and future work are detailed in Section VI.

II. SYSTEM ARCHITECTURE

From a global perspective, the inclusion of healthcare systems into smart cities services must address several issues, specially in terms of data privacy [5]. A robust codification process must be performed in order to guarantee that patients identity is preserved. Consequently, the overall architecture of the proposed system is illustrated in Figure 1. This architecture is mainly composed of a low level and a high level subsystem. The former is mainly devoted to the data acquisition and low level sensing whereas the latter is in charge of the data processing and inference extraction. This

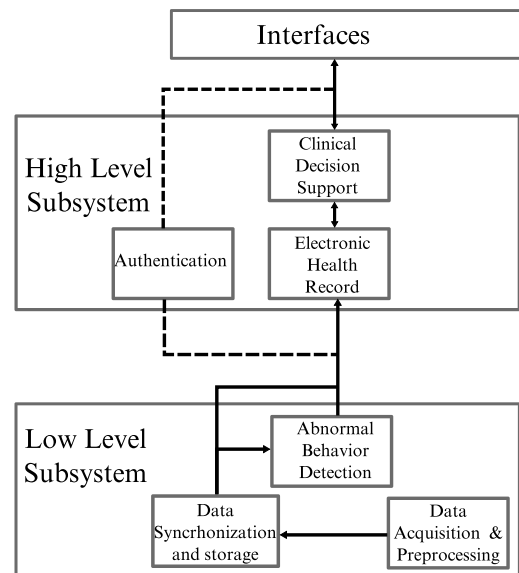


FIGURE 1. General architecture of healthcare systems and its integration into smart cities services (High Level Subsystem).

modular architecture involves the physical implementation of sensor systems in the places to be monitored. The aim of this system is to be integrated into a smart city in a Software as a Service (SaaS) distribution model [13]. Additionally, the high level subsystem can access information from the smart cities for decision-making procedures using a Platform as a Service (PaaS) model [15].

The low level layer is formed by (a) the Data Acquisition and Processing (DAP) module, (b) the Data Synchronization and Storage stage and (c) the Abnormal Behavior Detection (ABD) module. DAP is responsible for the proper acquisition and filtering of the information gathered from diverse sensors. Further, Data Synchronization and Storage stage aims to securely save data as well as synchronization tasks for later processing. Finally, methods for the detection of events associated to people with cognitive problems are implemented in ABD. These methods rely on fused information from the data gathered in previous modules.

The high level Subsystem consists of: (a) an Electronic Health Record (EHR), (b) a Clinical Decision Support System (CDS) and (c) an authentication component. EHR is the module where the medical related information is stored. CDS contains the recommendation engine to support the decision making of health professionals. Finally, the authentication module endows the system with security layers by adding headers to protect the identity of users. In next section, details on the subsystems of the architecture are provided. This work is part of a research project aiming to develop a system to extend life autonomy of elderly people with Parkinson Disease. The experiments as methodological aspects were defined in collaboration with the Asociación Parkinson Madrid (APM).¹ The novel concept of this paper is given

¹Asociación Parkinson Madrid (APM) <http://www.parkinsonmadrid.org>

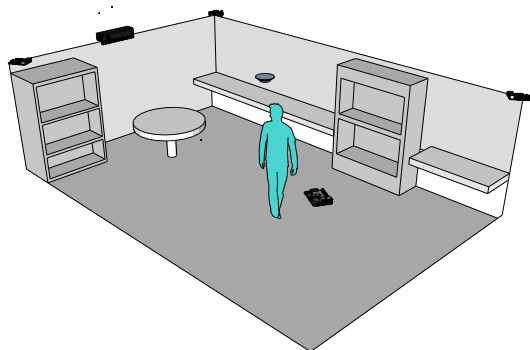
TABLE 1. Data extracted from the selected devices. Zenith camera coverage depends on the lens employed. Kinect default resolution is 640 × 480 at 30 Frames Per Second (FPS).

Device	Sensor	Feature
Bracelets and Belts	Accelerometer	ax,ay,az
	Gyroscope	gx,gy,gz
	Bio-sensors	Heart Rate, Temperature
WSN	IEEE802.15.1	RSSI (dBm)
	IEEE802.15.4	RSSI (dBm)
Depth Sensor	Skeleton joints	6 * 25 * x, y, z
	RGB Image	up to 30 fps at 1280 × 1024
	Infrared	black and white object movement
Zenith Camera	RGB Image (up to 360°)	up to 1280 × 960 at 15 fps

by the low level Subsystem modules and functionalities for PDs patients care as well as its implementation. Therefore these aspects will be detailed whereas the high level modules will be briefly addressed and reviewed in Future Work. The remainder module of this architecture comprises user interfaces that are out of the scope of this paper.

III. LOW LEVEL SUBSYSTEM

The low level subsystem comprises the modules for data acquisition and processing algorithms to extract the information used for ABD. It is divided into three modules. The first includes the data collection from diverse sensors and devices in addition to some brief pre-processing tasks (DAP). The second stage involves the data synchronization, labeling and storage. In the third stage, several processing algorithms for synchronization, association and data fusion are performed in two senses: (a) to reduce the amount of data sent to higher layers by sending only processed information and (b) to allow the system to make real-time decisions by containing the whole set of algorithms that permit the extraction of useful information from sensors. In a global overview, the multi-sensor approach of the low level subsystem for elderly people monitoring is presented in Figure 2.

**FIGURE 2.** Experiment setup for the multi-modal approach used in low level subsystem. This proposal includes Kinect, Zenith camera, WSN and eHealth bands.

A. DATA ACQUISITION AND PRE-PROCESSING (DAP)

The devices used in the system retrieve data from several sensors to extract information about elderly people activity and behavior. This data must be collected and stored before

applying processing algorithms to make inferences or generate alarms yielding to useful reports for the medical teams. A high synchronization level is required to label and classify the entire dataset. Sensors deployed over diverse scenarios provide multiple data types, sampling rate, resolution and accuracy. An overview of devices and features extracted is depicted in Table 1. Specifically, the following sensors are considered:

1) DEPTH SENSOR

Kinect camera V2 is a multi-sensor device developed by Microsoft [19]. It is able to provide up to 30 frames per seconds (fps) for several types of data formats as described in Table 1. Both images and skeleton joint points are labeled with the corresponding timestamps in the DAP module in the same manner as other sensors. The 640 × 480 resolution at 30 *frames per second* (fps) was chosen in the scope of this project. Data retrieved from this sensor will be employed for several behavioral (daily motion, etc...) and physical related problems (festination, freezing).

2) ZENITH CAMERA

Traditional vision sensors have played an important role in general person detection and tracking applications. Additionally, the use of special lens can increase the range of detection which represents an important advantage for large range applications. In the scope of this work, a fish-eye camera by Vivotek [12] with a 1280 × 960 and a sampling of 15 *fps* will be employed for both person identification and tracking.

3) WIRELESS SENSOR NETWORK (WSN)

A WSN consists of a group of self-organized sensing nodes deployed across the area to be monitored. There is a wide range of applications where these devices have been used [5], [6]. According to Table 2, the main goal of these devices is the daily tracking of PD patients. For this purpose, a set $\mathbf{v} = \{v_{i,\lambda} \mid i = 1, \dots, I; \lambda = 1, \dots, \Lambda\}$ of Access Points (AP) located at cartesian coordinates $p_i = (x_i, y_i) \in \mathbb{R}^2$ is defined. Furthermore, let Λ be the number of sensors that each node is equipped with. In this work, two sensors from mentioned wireless technologies are considered: IEEE802.15.1/Bluetooth v4.0 [21], IEEE802.15.4/XBee [22]. Network nodes are able to obtain the Received Signal Strength Indicator, denoted

TABLE 2. A range of abnormal events to be detected with the corresponding functions that must be implemented to reach it. In left column the sensors involved in the multi-modal detection are described.

ABD	Functionality	Sensors
Daily Motion	Trajectories	WSN, Kinect, Zenith
Falls detection	Acceleration Bracelets and Belts	Bracelets, Belts and Kinect angular speed
Parkinson Freezing	Movement assessment	Bracelets, belts and cameras
Parkinson Festination	Movement assessment	Bracelets, belts, cameras and WSN.
Dream motion	Movement assessment Patient leaving the home	Bracelets and WSN. WSN and camera.

by $RSSI$, from the monitored target. Consequently, $RSSI_{i,\lambda}$ describes the $RSSI$ measurements gathered by node i using technology λ .

4) HEALTH BRACELETS AND BELTS

Several smart devices were chosen for the testing purposes. Firstly, a commercial health band, developed by Microsoft, with a sampling rate of up to 45 samples per second has been adopted [18]. This bracelet is equipped with accelerometer, gyroscope, Galvanic Skin Response (GSR), Heart Rate (HR), skin temperature sensor and bluetooth transmitter. Furthermore, another commercial eHealth band endowed with the aforementioned sensors and magnetometer has been also adapted to the system [20]. Additionally, this band is IEEE802.15.4 XBee/Zigbee compatible [22]. The reason to evaluate both bands is to perform multiple tasks taking advantage of the wireless monitoring capabilities that WSNs are able to estimate.

In concrete, IEEE802.15.4 sampling rate is up to 30 messages per second. This information is gathered by the WSN for position estimation. Moreover, sampling rate for Bluetooth $RSSI$ extraction is 2 seconds. The DAP module is in charge of receiving measurements from bracelets (to obtain HR, temperature and GSR measurements) and the WSN ($RSSI$ from the nodes deployed). The measurements are labeled with both timestamps from bracelet devices and local synchronization of the gateway node where samples are collected.

B. DATA SYNCHRONIZATION AND STORAGE

Because of the nature of sensors employed, a synchronization stage is required in order to allow proper data treatment. Accuracy of timestamps labeling guarantees the similarity of measures to be analyzed. For this reason, this module will simultaneously add the common timestamps to each measurement. Additionally, the multiple sampling rates of devices considered can yield to comparison errors. Therefore, time-windows are created to process the data from multiple sensors.

The first step is to quantify the delay introduced by every sensor. Empirical experiments allowed to obtain the Zenith delay by performing an action (e.g. applause) and calculating the lag between the action realized and the processed timestamps (stored). 100 repetitions of this process yield as result that the camera delay is 2.6 seconds.

Additionally, images from cameras (RGB) are processed to extract the useful features and discarded to guarantee privacy protection. Data extracted will remain stored until the ABD access this information. Once data and events are forwarded to the high level subsystem, the local information is deleted.

Moreover, in order to ensure data privacy, a policy for storage in a non-human readable format is adopted. This policy involves a serialization process into binary files. The subsequent module (ABD), will initially perform a de-serialization task in order to extract the data.

Kinect delay depends on the move speed, therefore for slow-motion events it can be neglected whereas for fast moves, the delay is up to 200ms [40].

Finally, delays introduced in bluetooth transmission have been extensively studied [42]. In concrete, for the devices employed in this work, it has been proved to be 200 ms

The second step is to define a reference for synchronization. In this work, the reference time-system is defined by using the Kinect sensor frame rate. Accordingly, the samples of all sensor devices must satisfy the following condition:

$$ts_{B_i} - \frac{1}{v_{fps}} \leq ts_{K_j} < ts_{B_i} + \frac{1}{v_{fps}} \begin{cases} i = 1, \dots, N; \\ j = 0, \dots, 5; \end{cases}$$

where ts_{B_i} and ts_{K_j} denote the band and Kinect timestamp respectively. Furthermore, the v_{fps} parameter refers to the Kinect frame rate, which is obtained by counting the number of frames per second to be processed. This condition is applied for inertial data of the N devices available and the skeletons detected (in the range 0-5) by the depth sensor.

Moreover, a smoothing filter is applied to the information gathered from all sensors to minimize the impact of outliers and some abnormal moves. In concrete, the well known Butterworth filter is applied [41]. This is a low-pass filter that aims to reduce high frequency variations.

C. ABNORMAL BEHAVIOR DETECTION (ABD)

In this section, the main activities to be detected as well as the multi-sensor scheme and algorithms developed are described. Firstly, it is important to highlight that this work is part of a research project which aiming is to extend the autonomy of elderly people with cognitive problems (see Acknowledgments). Therefore, general activities such as daily motion and fall detection are combined with specific targeted activities for people suffering from PD as shown in Table 2.

Appropriate fusion of sensors involved will increase the accuracy of the ABD.

In the next subsections, details on the functions implementation as well as the ABD decision making will be provided.

1) DAILY MOTION

Based on data from previous modules, processing algorithms are mainly focused on the fusion of sensors data to track PD patients. Non-intrusive tracking is an emerging research topic whose goal is to achieve the best accuracy in the final estimation. In this project, three sensors for tracking algorithms are considered using both wireless signals and cameras. The first system uses WSN ability to obtain the RSSI parameter from Bluetooth and XBee transmitters. By the application of advance processing techniques it is possible to estimate the position. The second system uses RGB image information to track people based on the detection of movement in the image. The third functionality is implemented by the extraction of the joint skeleton positions and estimate the orientation and position of the object sensed.

2) WIRELESS SENSOR NETWORK TRACKING

WSNs represent a valuable option for multiple-person positioning and tracking due to its ability to perform non-intrusive detection and overcome the problem of visual occlusions. Furthermore, WSN does not require additional hardware to wireless transmitters for RSSI parameter detection. However, the large variance of these measurements raises an important challenge for the accurate estimation process. For the proposed system, several nodes have been deployed along scenarios considered to gather the RSSI parameter from bracelets IEEE802.15.1Bluetooth and IEEE802.15.4XBee devices.

Once the RSSI is collected by all nodes, the positioning algorithm can be applied and afterward, a filtering stage to smooth the estimation is carried out to calculate the final trajectory. In a general statement, the algorithm is composed by the following steps:

- *Fingerprinting calibration.* This technique performs an initial RSSI dataset by collecting data from bracelet wireless sensors at certain predefined locations. These locations are obtained by splitting the room into several cells [24]–[26].
- *Fingerprinting Kernel density Estimation.* This algorithm selects the cell that fits better between the fingerprint dataset and the current measure [27]–[29].
- *Orientation estimation.* In this stage, the heading angle to estimate the device orientation in the 360 degrees is obtained [30].
- *Velocity estimation.* Fusing information from accelerometers and positions estimated, the person velocity can be achieved by using this algorithm [31].
- *Particle Filter Tracking.* Filtering of the estimated positions to determine the final trajectory applying a real movement model [32].

Fingerprinting is a technique commonly used to characterize the signal propagation for indoor environments [44].

Due to the hard conditions, reflection and obstacles, it is difficult to fit the propagation attenuation into a valid model. Fingerprinting tackles this problem by creating a RSSI datasets which are obtained by splitting the scenario into a set $\mathbf{c} = \{\mathbf{p}_l | l = 1, \dots, L\}$ of cells with geometrical center at $\mathbf{p}_l = (x_l, y_l) \in \mathbb{R}^2$ and collecting RSSI values for a predefined period in each cell c_l .

Moreover, taking into account the fingerprint dataset, *Kernel Density Estimation* is applied to estimate a position in the scenario. The algorithm employed in this work is called Nadaraya–Watson Kernel Regression [29]. This algorithm estimates the position by comparing the similarity degree of the fingerprinting dataset with the WSN RSSI measurements. The position $\hat{\mathbf{p}}$ is obtained by using the following equation:

$$\hat{\mathbf{p}} = \sum_{c=1}^N w_c \mathbf{p}_c \quad (1)$$

where w_c denotes the weights, with its corresponding center position \mathbf{p}_c . These weights are obtained by using Gaussian Kernel functions [28] as follows:

$$w_c = \frac{K\left(\frac{\mathbf{RSSI} - \mathbb{E}[\mathbf{RSSI}_c]}{h}\right)}{\sum_{j=1}^L K\left(\frac{\mathbf{RSSI} - \mathbb{E}[\mathbf{RSSI}_j]}{h}\right)} \quad (2)$$

where $K(\cdot)$ is the Kernel function employed, L is the number of total cells, \mathbf{RSSI} is the vector containing the current RSSI measures from WSN, $\mathbb{E}[\mathbf{RSSI}_c]$ denotes the vector with mean RSSI values from fingerprinting dataset and h is the bandwidth parameter for the Kernel function used. Further details as well as examples and public datasets of this process can be found in [39]

Particle Filter is employed for tracking purposes by applying the following movement model:

$$\begin{aligned} x_k &= x_{k-1} + v_{x_k} \cos(\theta) dt \\ y_k &= y_{k-1} + v_{y_k} \sin(\theta) dt \end{aligned} \quad (3)$$

where $\mathbf{p}_k = (x_k, y_k)$, v_{x_k} and v_{y_k} denoted the x , y axis velocity respectively, whereas θ represents the orientation. To use this model, the mentioned parameters (x , y , θ and v) must be previously estimated. The latter can be calculated from accelerometer, gyroscope and magnetometer sensors. The orientation angle for the movement model is known in the same manner as the heading angle or yaw in navigation and it can be obtained by applying the following equation:

$$\theta = \arctan\left(\frac{2q_1q_4 + q_2q_3}{2q_1^2 + 2q_2^2 - 1}\right) \quad (4)$$

where q_x are the values of the quaternion vector obtained from accelerometer, gyroscope and magnetometer sensors using Magdwick Filter [30]. This filter is able to fuse sensor measures to calculate the quaternion vector that is a orientation representation using complex numbers.

The other parameter needed is the velocity of the monitored people. This information can be estimated using accelerometer measurements. The three axis values can be fused into a single value that provides information of the acceleration behavior as a function over time:

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2} \quad (5)$$

Using an adaptive threshold peak detector algorithm [33], the maximum and minimum peaks of this curve can be obtained. In addition, Weinberg formula [31] can be employed to calculate the step length of people walking. With this information, the velocity by time interval is also estimated as follows:

$$Step = v * dt = K_1 \sqrt[4]{a_{max} - a_{min}} \quad (6)$$

where $0.3 < K_1 < 0.7$ is a constant value that depends on the patient height, age and a_{max} , a_{min} are the maximum and minimum values obtained from Equation (5). Once these parameters are computed, the particle filter can be utilized to iteratively estimate the position with a Kernel-based algorithm. The particle filter adopts Monte Carlo methodology to find the solution of filtering problems. The steps of the mentioned algorithm are:

- Generate P random particles with a Gaussian distribution across the considered scenario.
- Predict the new state based on the movement model selected (4).
- Calculate the weights according to the distance of the particles to the prediction.
- Re-sampling particles.
- Estimate the new state for the system.

This algorithm will return the trajectory of individuals wearing bracelets around the room. The final estimation presents high accuracy in terms of Minimum Squared Error (MSE) in comparison to the real path. Further details can be found in Section V.

3) ZENITH CAMERA TRACKING

Zenith cameras have gained relevance as suitable option for large area tracking systems. The main advantage of these devices is given by the comprehensive detection range as they offer the widest-field-of-view possible. Computer vision algorithms are employed to detect the movement on the frame and decide whether the objects detected correspond to people or not [34]. Moreover, a tracking algorithm is used over these detected persons for trajectories estimation and finally, the conversion of pixel images to real world coordinates is carried out. The entire procedure for person detection and tracking using Zenith cameras is described as follows:

- *Camera calibration.* This process is performed by using a chessboard pattern to obtain the intrinsic, distortion and extrinsic parameters of the camera chosen.
- *Area calibration.* Selection of several points on the floor to get the parameters for the pixel to world coordinates conversion.

- *Histogram of Oriented Gradients training.* In this stage, the Algorithm training to detect people in the image is carried out [34].
- *Undistorted images.* Procedure executed to eliminate the omni-directional distortion effects introduced by the 360 degrees lens [35].
- *Image movement detection.* This algorithm is utilized for background subtraction of the image; giving as a result the possible motion objects in the image.
- *Blob classification.* Based on the probable motion objects detected, applying this algorithm allows to verify if this object is a person or not [34].
- *Kalman Filter Tracking.* Implement the well known Kalman Filter tracking to persons detected along the images collected [36].
- *Detection Assignment.* To allocate the most probable distribution of trackers along the images.
- *Pixels to real world coordinates conversion.* Transform the tracked pixels route to real world coordinates in the reference system.

The entire process is drawn in Figure 3, where an image of every step of the Zenith sensor is shown.

Camera Calibration: This process is required in case the camera intrinsic and extrinsic parameters are unknown. It consists in obtaining several images from a chessboard pattern to detect the corners in different positions and orientations. An example of the mentioned procedure is illustrated in Figure 4. Harris corner detector is employed to extract the points from the image and therefore to estimate the intrinsic, distortion and extrinsic camera parameters using Zhengyong Zhang calibration [43].

Once the camera calibration is performed, *the area under monitoring must be calibrated.* This process allows to convert from pixels to real spatial coordinates. To carry out this task, several known points on the floor must be marked and a camera image must be taken. The calibration algorithm consists of selecting the pixel points where the real world coordinate markers are.

The next step is to train the algorithm for people detection. In this work, the Histogram of Oriented Gradients (HOG) is chosen [34]. HOG algorithm is a feature descriptor that counts occurrences of gradient orientation in localized portions of an image. This algorithm relies on the training of a Supported Vector Machine (SVM) with images of persons across the area under monitoring. As a result, a hard decision about the objects detected in the crop image as persons is made.

In order to mitigate the effects of eye-fish lens, *an algorithm to obtain undistorted images* is executed [35]. This algorithm contains functions to compensate radial and tangential image deformity by using a predefined distortion model.

Subsequently, the *Kalman Filter* is applied to the persons detected to obtain their optimal pixel position. In this paper, the Kalman Filter works in combination with the Munkres or Hungarian algorithm [37] for people track

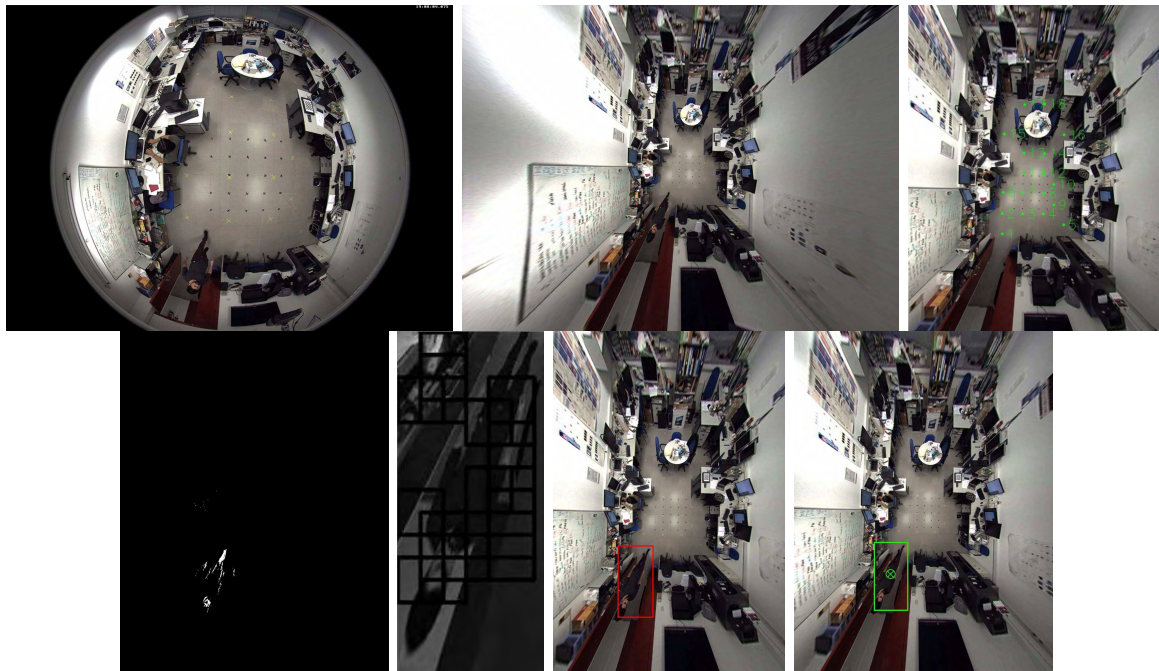


FIGURE 3. Description of the Zenith camera algorithm process. *from left to right in top row: (1) Original eye-fish image. (2) Camera calibration, (3) Area calibration. Bottom row, from left to right: (4) Background subtraction. (5) Blob classification. (6) Person detection. (7) Person tracking.*



FIGURE 4. Example of calibration images with chessboard pattern in different positions and orientations.

assignment. The latter algorithm is responsible for the association of trackers detected in a certain frame with the ones of the previous and subsequent ones based on pixel distance between frames.

Finally, the conversion from pixels to real coordinates in a common system for all the sensors considered is performed by using the *Perspective-n-Point* formula. As a result, the real 3D coordinates are computed as follows:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (7)$$

where s is a scale parameter and u, v are the 2D projections. On the right side of Equation (7), the first matrix contains the intrinsic camera parameters: focal lengths f_x, f_y , and optical center c_x, c_y matching the 2D pixels with the 3D real world coordinates $[X, Y, Z, 1]^T$. Furthermore, the second matrix, comprises the rotation r_{xx} and translation t_x values that change the coordinate reference to the desired system.

4) KINECT CAMERA TRACKING

Kinect camera incorporates three depth sensors and an algorithm for skeleton detection. This algorithm can be used to extract the person position in real 3D coordinates with respect to the camera focus [38].

Furthermore, the calibration process of this device can be performed by transforming the real coordinates $[X, Y, Z]^T$ from Equation (7) to the desired common system reference. The process is summarized as follows:

- To get the calibration points an image is taken from Kinect camera to the floor markers.
- Using the open SDK from Kinect [23] the real world coordinates from camera coordinate system can be obtained by each selected point in the calibration.
- To move the coordinate system to the common coordinates, it is required to obtain the rotation matrix and translation vector:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t \quad (8)$$

Both, the rotation matrix and the translation vector are obtained by applying a singular value decomposition (SVD) over the selected points.

$$U, S, V = \text{svd}((A - \bar{A})' * (B - \bar{B})) \quad (9)$$

$$R = V * U' \quad (10)$$

$$T = -R * \bar{A}' + \bar{B}' \quad (11)$$

where A is a $(n \times 3)$ matrix with n denoting the number of selected points and the column vector X, Y, Z contains

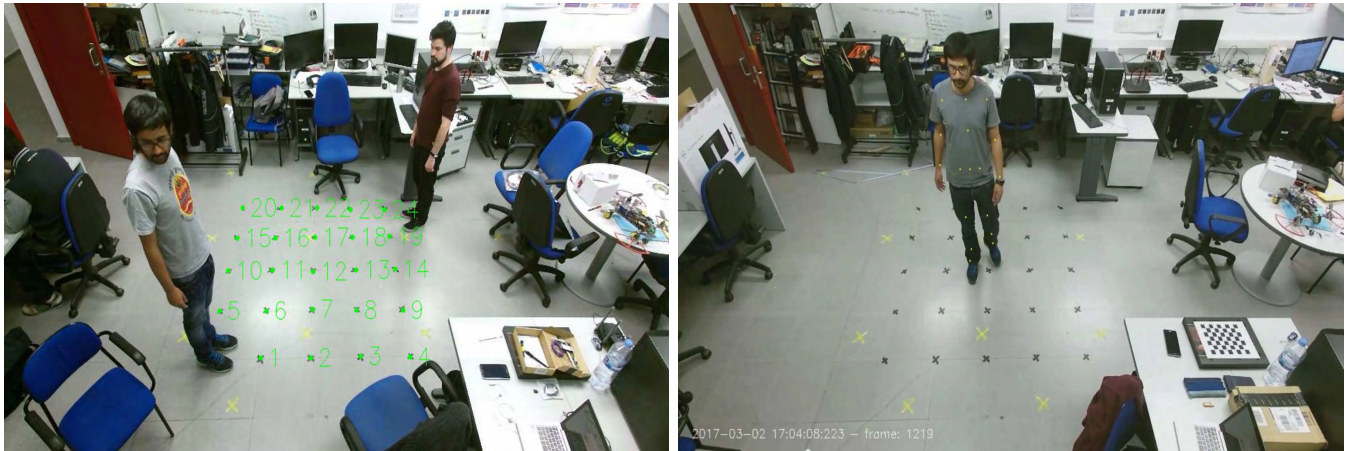


FIGURE 5. Kinect procedures description. On the left, the calibration markers are drawn. On the right side, Joint skeleton detection is shown (yellow dots) for a person walking.



FIGURE 6. Example of trajectories extraction from WSN (left), Kinect (center) and Zenith camera (right). The red line shows the real path whereas the green line represents the estimated path. The WSN estimation accuracy is lesser than the one obtained from cameras. However, WSN coverage area is the largest of the sensors chosen. Finally, the Kinect range of detection constraints the potential applications. In the experiment, WSN is composed of 4 nodes located at the top corners of the room. Kinect sensor is located at 500, 300 in horizontal and vertical axis respectively and Zenith Camera is located in the room center at ceiling.

the coordinates of the depth sensor. Moreover, B is a $(n \times 3)$ matrix with the x, y, z coordinates of the selected points in the reference coordinate system. \bar{A} and \bar{B} are the centroid values of x, y, z for all selected points.

D. PERSON IDENTIFICATION AND MULTI-MODAL FUSION

ABD is a fundamental part of the proposed system since this module is in charge of the data analysis to extract trajectories from data sensors and make inferences about elderly people actions. To reach such objective, a multi-modal scheme is proposed to fuse the information shown in Table 2.

However, in order to combine the information from the sensors, a previous *identification* stage must be performed. The reason is that multiple factors such as target occlusions, lighting changes, signal attenuation, reflections, among other effects, affect the tracking algorithms yielding to estimation errors. In fact, for scenarios such as the ones considered in this

work, affluence of several persons is assumed and therefore a classification tool to associate the routes estimated with the corresponding person is required. Furthermore, the identification task must be performed for each sensor which implies that several matching levels from sensors data are needed.

In this work three main inference levels are proposed:

- *Track-to track comparison*: using the trajectories from Bracelets, Kinect camera and Zenith camera.
- *Orientation comparison*: based on heading estimation from bracelets and estimated orientation from Kinect and Zenith camera using the trajectories.
- *Accelerometer measures comparison* with wrist body point from Kinect camera.

Because of the system modularity, the identification process must be performed even if there is no data gathered

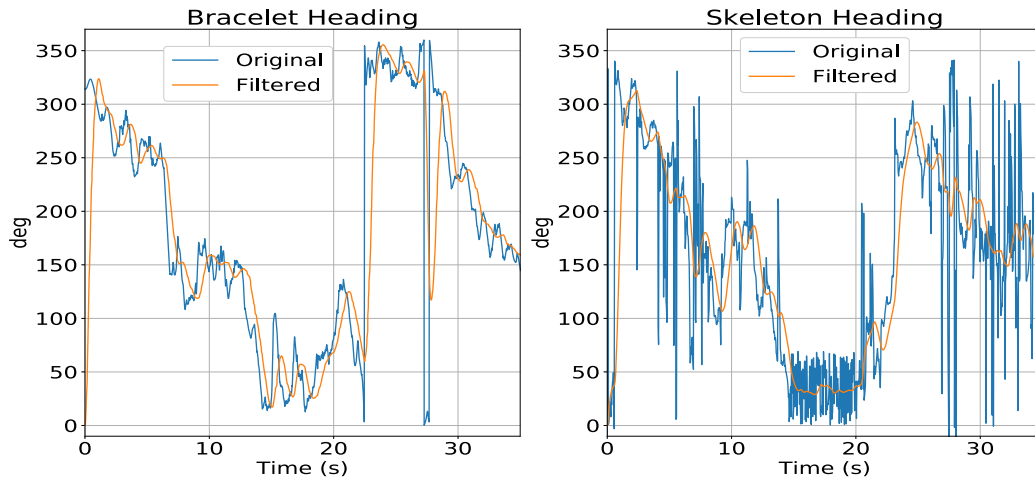


FIGURE 7. Comparison of the estimated heading for bracelet and skeleton. The blue line is the raw data whereas the orange line represents the signal filtered. The skeleton noise is higher than the bracelet due to two main factors: (a) changes between 360 degrees and 0 are reflected in spite that variations are small; and (b) the short-term angle variation in the skeleton joints is large.

from all sensors due to occlusions or coverage. Conversely, if measurements from Bracelets, Kinect and Zenith cameras are available, the three comparison methods can be applied to associate the estimated trajectories to a person in the scene.

1) TRACK-TO-TRACK COMPARISON

Assuming that data gathered have been synchronized in the previous stages of the low level subsystem, the goal of the Track-to-Track procedure is to detect the routes with the higher similarity. The metric employed is the Minimum Square Error MSE along the route duration and can be expressed as follows:

$$\begin{aligned} & \text{minimize } t \\ & \text{subject to } f_i(\mathbf{x}) \leq t \quad i = 1, \dots, m. \end{aligned} \quad (12)$$

$$f_i(\mathbf{x}) = \frac{1}{M} \sum_{j=1}^M \sqrt{(x_j^i - x_j^j)^2 + (y_j^i - y_j^j)^2} \quad (13)$$

where x, y are the axis coordinates of a predefined route, m represents the number of routes in a scene (frame or bracelet time interval). M is the number of trajectory points taken into account in the estimation (corresponds to the time interval where synchronization is performed). The cost Matrix can be created with the MSE combinations between the mentioned sensors. As an example, if Kinect estimates two trajectories and Zenith camera three, the cost matrix is determined as:

$$C = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \end{bmatrix} \quad (14)$$

where every t_{ij} is obtained by solving the optimization problem (12). In this paper, Munkres algorithm [37] is applied to the matrix (14). If all conditions are satisfied, the final assignment can be carried out and therefore link the trajectories for a determined person.

2) ORIENTATION COMPARISON

Orientation comparison is proposed as second level of inference for person identification. Bracelets are equipped with sensors that directly provide the heading orientation. Moreover, there are several techniques to obtain the orientation based on the movements of the object tracked [38], [45].

The comparison will be performed using the metric described in (12). However, the level of comparison in this case is higher. The reason is that the orientation of the bracelets is not absolute when these devices lack of magnetometer. Therefore, a comparison of the orientation with the route rotated in the 360 degrees must be introduced in this process. The filtering process of the obtained curves is then executed to obtain the MSE. Analogously to the previous inference level, Hungarian algorithm [37] is utilized for the assignment process.

3) ACCELERATION COMPARISONS

The third inference level relies on the Bracelets accelerometer and wrist acceleration from Kinect body joints.

The process consists of extracting the data from every frame in Kinect raw. In concrete, wrist points (6,18 for left and right wrist respectively) from skeleton body must be compared to the acceleration values (ax, ay, az) extracted from the bracelets. Depending on the technology employed, the inertial measurements can include the gravity force (expressed in g). Therefore, initial normalization procedure must be performed before Kinect body and bracelets analysis. The association process can be then carried out as described in previous inference levels.

4) MULTI-MODAL FUSION

Once the identification procedure is performed, the measures can be fused in a multi-modal scheme to improve the accuracy of estimations. A fusion algorithm can be used to

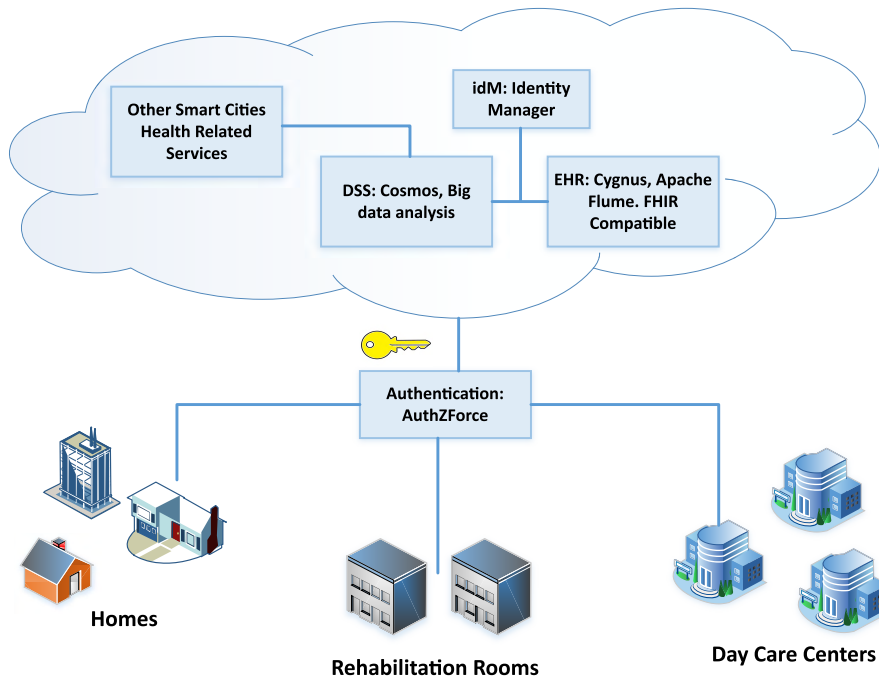


FIGURE 8. Details of high level subsystem modules and technologies integrated into a smart city infrastructure.

achieve a high level of accuracy estimation using information from cameras, WSN and bracelets. In this paper, Kalman Sensor Group Fusion algorithm (KSFG) is employed [29]. This algorithm exploits the Kalman filter properties for multi-sensor data fusion estimating a final x, y coordinates based on several trajectories. In particular, for the case of the sensor considered in this work, the filter has the following form:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (15)$$

H is the state matrix for Kalman filter. Furthermore, R matrix includes sensor error covariance measurements. These values are appended to the main diagonal:

$$R = \begin{bmatrix} \sigma_{x_1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{y_1}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{x_2}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{y_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{x_3}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{y_3}^2 \end{bmatrix} \quad (16)$$

Finally, the observations vector z collects the x and y coordinates estimated from each sensor:

$$z = [\hat{x}_1, \hat{y}_1, \hat{x}_2, \hat{y}_2, \hat{x}_3, \hat{y}_3]' \quad (17)$$

As an output of the KSFG stage, a high accuracy daily motion estimation is obtained.

IV. HIGH LEVEL SUBSYSTEM

A. SMART CITIES CONTEXT

As previously described in Section II, the high level subsystem is intended to access information from the monitored places and using data from the city in a (PaaS). The reason to select this option is that it includes not only the remote use of software (as in Software-as-a-Service) but a complete application development and distribution platform and therefore, a service providing remote utilization of cloud computing.

This platform technique streamlines development by eliminating the need to customize the code to run on different platforms. In this work, a proposal for the mentioned architecture is presented in Figure 8. This platform relies on the use of an open technology modules for specific back-end services. The subsystem comprises the EHR, CDS and Authentication services. The complete stack of services for eHealth systems is described in [50]. Two security modules are considered in this architecture. The AuthZForce GE provides an API to obtain authorization decisions based on authorization policies, and authorization requests from Policy Enforcement Point. The API follows the REST architecture style, and complies with XACML v3.0. XACML (eXtensible Access Control Markup Language).

Moreover, Identity Management GE covers a number of aspects involving users' access to networks, services and applications, including secure and private authentication from users to devices, networks and services, authorization and trust management, user profile management, privacy-preserving disposition of personal data,

Single Sign-On (SSO) to service domains and Identity Federation towards applications.

B. ELECTRONIC HEALTH RECORD (EHR)

The EHR is the collection of medical and behavioral data of patients. Information such as Health annotations, medications, frequency of ABDs generation, information from sensors (low level Subsystem) among others is stored in the EHR. This data is used by the recommendation engines in CDS to provide inferences to the related stakeholders. As a result, General practitioners, Neurologists and Physicians will have a more complete health status of the patient for appropriate decision making about patients treatment.

There exist some extended standards for EHR (also known as Virtual Medical Record) [47], [48]. However, for the presented work, the Fast Healthcare Interoperability Resources (FHIR) standard is adopted. The reason to choose this standard is the interoperability between legacy health care systems (as the low level Subsystem), making the process of delivering the information to multiple stakeholders simple. The FHIR API [48] implements the operations for interaction with the medical information stored based on a HTTP RESTful. This implementation is fully compatible with the first-catalogue Generic Enable (GE) [49]. The main advantage of this GE is that it allows the interoperability of the modules developed with the Next Generation Services Interface (NGSI).

C. CLINICAL DECISION SUPPORT (CDS)

There is a large list of proposed commercial and research CDS systems which are continuously evolving by applying complex algorithms for the recommendation processes. These systems have been classified into multiple groups:

- *Machine Learning*: These methods rely on known techniques such as Neural Networks, Supported Vector Machines. The goal is to train the algorithms to learn from the data.
- *Knowledge Based*: These systems contain knowledge modules that provide the rules for decision-making based on the expertise of professionals.
- *Graphical Representation*: these methods allow the visualization of data to verify the data behavior and trends. The extraction of inferences is controlled by the data evaluator
- *Data Mining*: This approach consists in analyzing the data through linguistic techniques such as Natural Language Processing (NLP) [46].
- *Hybrid Approaches*: These techniques employ several of the aforementioned methods combined for decision-making.

The main advantage of the proposed system is that it allows the integration of multiple data sources to provide recommendations. In Table 3, data inputs are described, as well as some initial outputs are outlined. External sources such as smart city services (weather, traffic and environmental data) can be used to increase the system Knowledge (e.g. avoid places

with large pollution, etc). The proposed platform to support these services is based on Cosmos ecosystem. This GE is intended to deploy means for analyzing both batch or/and stream data to provide insights on such a data revealing new information that was hidden. Batch data is stored in advance, and latency is not extremely important when processing it.

V. EXPERIMENTAL RESULTS

A. METHODOLOGY

The methodological aspects of experiments described in this work have been designed according to the project research book.² To summarize, the experiments were performed in several sessions with a total of 18 patients. The control group has been composed of elderly people without cognitive problems (APM volunteers). Additionally, the experimental group was formed by patients that were classified into three groups according to they the illness stage as: initial, moderate or severe. Diverse aspects concerning gender balance as well as related diseases. This classification is of particular interest for the high level subsystem. However, for the physical detection of abnormalities it is neglected. Furthermore, PD patients were informed on the procedure as well as the type of data that was going to be collected. However, participants were not notified neither hypothesis nor expected results to ensure unbiasedness of their behaviors and consequent of the results.

B. EXPERIMENTS AND RESULTS

The first experiment was intended to show the identification and individualization process. It is important to highlight that bracelet-result orientation is not absolute with regard to cardinal system, there exist a phase-change with regard to Kinect orientation which is absolute with camera focus as origin. It is drawn in Figure 7 where the vertical axis denotes orientation moves (angles) over time for both bracelets and Kinect sensor. After calibration process, the orientation changes must be similar in both cases.

The second experiment consists in abnormal events detection by the use of the bracelets as described in Section III. The analysis of accelerometer measures from bracelets provide information about strong changes in the movement. Gyroscope produces data about the velocity of turning (angular speed). Therefore, falls events can be detected by the combination of these two sensors as shown in Figure 9. In this figure, at top image the raw measurements can be observed. The left side figure is noisy, therefore a smoothing stage is applied using a Butterworth filter on the right side. Moreover, the bottom figure depicts how the skeleton is detected.

Furthermore, another abnormal events can be detected by the proper combination of inertial and visual sensors. In particular, in freezing event detection, the goal is to find Freezing of Gait (FOG) [45] patterns in the time evolution of accelerometer measurements. In Figure 10, a time-lag corresponding to a FOG is shown. Notice that for this purpose, the signal from accelerometers must not be filtered.

²<http://www.ict4life.eu>

TABLE 3. Data sources and outputs of the Clinical Decision Support System.

Data Source	Inputs	Output
Bracelets and Belts	Accelerometer Gyroscope Bio-sensors	Gait Evolution Heart Rate, Temperature
ABD: Patient leaving home	WSN estimation	Warning generation.
ABD: Festination	evolution in bracelets and camera movements	Warning generations, number of festination episodes over time.
ABD: Falls Detection	peaks in bracelets and kinect movements	Warning generations, number of falls over time.
ABD: Daily Motion	Movement evolution	Time spent performing daily activities, apathy.
Interfaces	Medical Annotations	Comorbidities, General Practitioner notes, medical history.
Interfaces	Elderly people interactions	i.e. Time surfing on interfaces, difficulties to use it, among others.
External sources	Information that support medium-term decisions PaaS	Weather conditions can explain why patients stay at home, etc. Traffic information can be useful to schedule activities, etc.

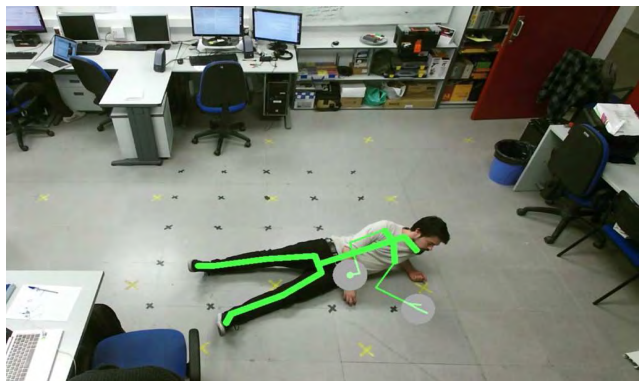
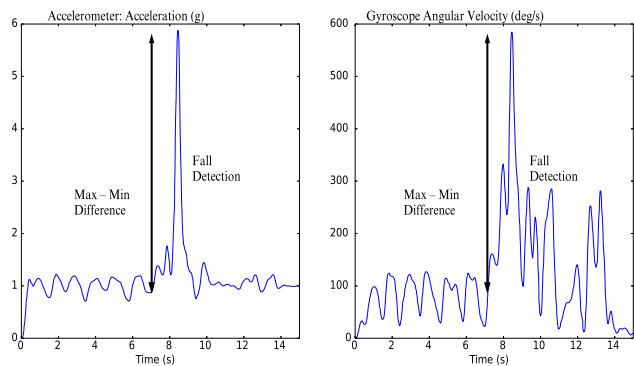


FIGURE 9. Example of fall detection module using fusion of sensors. In top graph, accelerometer variations for fall inference are outlined. In bottom, Kinect skeleton fall detection is drawn on the body.

In Figure 11, the location of the patient in the home is shown for a day. The living room area was the area covered by the whole set of sensors. The detection was mainly obtained through the WSN (via bracelet). The accuracy of the detection is high as the resolution is not so large (a room). Kinect and Zenith camera were deployed in the living room.

Finally, numerical results of the activities detected by the sensors separately are drawn in Table 4. Patients were recorded while walking for around 1 minute. Additionally, the improvements reached by the multi-modal fusion are detailed. From top to bottom:

- A comparison of diverse position estimation methods for WSN in a route across the room employed for the experiments using fingerprinting technique. These

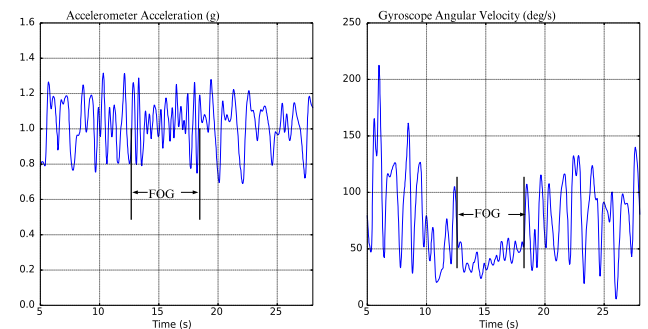


FIGURE 10. Accelerometer and Gyroscope data for freezing detection. On the (right side), the Freezing of Gait is appreciated whereas on the (left side), accelerometer data is noisy and therefore it can not be observed.

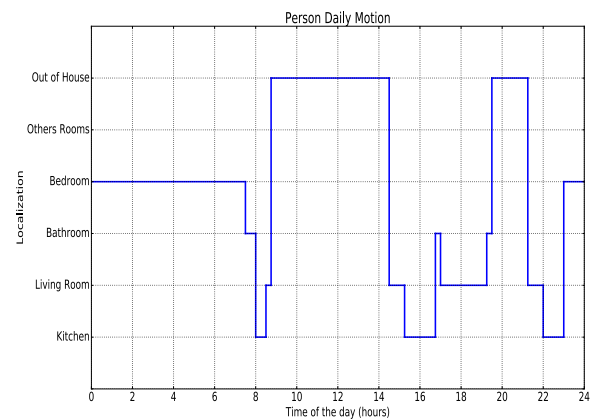


FIGURE 11. Evolution of patient location within home over a 24-hour period. The time the patient is outside is not monitored by the system.

values show the error in centimeters of the estimated positions with respect to the real patient location. The employment of Kernel functions reaches better estimation results. The reason is that these functions attain a better wireless signal propagation modeling than linear functions.

- A comparison of tracking methods for position estimation is shown. It can be observed the significant error reduction by applying these techniques. The Kalman filter mitigates the impact of outliers, specially for WSN

TABLE 4. Algorithms Accuracy. Comparison of different positioning and tracking methods and the associated average error. Person identification success percentage based on the number of person presented in the monitored room.

WSN: Average error (centimeters) in fingerprinting technique using different methods					
	Euclidean distance	Mahalanobis distance	Weighted K-NN	Gaussian Kernel	Exponential Kernel
802.15.1 Bluetooth	150.57	142.58	140.32	128.15	137.67
802.15.4 XBee	110.71	104.56	98.36	81.73	87.5
WSN Average error (centimeters) using Tracking methods					
	Kalman Filter		Extended Kalman Filter	Particle Filter	
802.15.1 Bluetooth	120.02		108.33	85.11	
802.15.4 XBee	78.47		63.82	54.75	
Kinect Average error (centimeters) in skeleton Tracking					
	Direct Kinect Skeleton			Smoothed Skeleton Tobit Kalman Filter	
2D Calibration	36.45			29.67	
3D Calibration	43.91			34.34	
Zenith Average error (centimeters) in Tracking people					
	Kalman Filter	Extended Kalman Filter (Only in combination with WSN)		Particle Filter (Only in combination with WSN)	
	45.89	43.98		39.07	
Tracking Fusion Average error (centimeters)			Identification Accuracy		
WSN + Zenith	36.98		1 person		98%
WSN + Kinect	32.71		2 persons		94%
Kinect + Zenith	31.54		3 persons		90%
WSN + Kinect + Zenith	30.11		+3 persons		≤89%

estimation, due to use of previous estimations reduces the error of noisy measurements.

- Estimation accuracy for two methods in Kinect devices is compared. There are two scenarios considered: (a) *2D calibration*: this method consists in use a fix plane (e.g. vertical axis $z = 0$) for the calibration process. (b) *3D calibration*: This method employs points in all coordinates. As it can be observed, the former method results are better as the main interest is in obtaining the patient position location in the room without the particular interest of the patient height.
- Tracking accuracy of several methods using the zenith camera (deployed at roof) is compared. The results obtained by using Particle Filter improves the results compared to the rest of mentioned methods. It is important to highlight that the error is similar to the one of Kinect sensor. However, for the sake of simplicity, results are shown only for areas covered by all sensors.
- Results of multi-modal process are drawn. On the left column, the accuracy improvement for diverse sensor configurations (ensuring modularity) fused is shown. Finally, the identification process results for cases (up to 4 individuals in the scene simultaneously) are described.

From a technical perspective, the main limitation of the presented system is given by the search of the best tradeoff between coverage and accuracy of the sensors considered. Kinect sensor has a high precision, however, the maximum distance is around 4m and 120 degrees view. Furthermore, WSN detection has a large coverage (up to 10m bluetooth and 40m XBee). Nonetheless, accuracy of estimations is lower than Kinect and Zenith camera as shown in Table 4. Finally, Zenith cameras have a wide coverage range (up to 20m), however, occlusions and angle distortion decrease the

tracking performance compared to the accuracy obtained with the Kinect sensor.

Finally, conversely to tailored solutions for healthcare monitoring which are generally expensive, most of the devices composing the system presented are well known and of general purpose. The Kinect sensor is available to be acquired for around 150USD. Moreover, Zenith camera employed in this work is around 250USD, and the bands cost 100USD [20] and 250USD [18] respectively. There is a wide range of WSN technologies and the prices are around 100USD. Finally, the Kinect sensor requires at least 1 USB 3.0 port and 8GB RAM.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, a complete multi-modal scheme for elder people monitoring was presented. The system architecture was outlined as well as the main modules such as DAP, ABD of the low level subsystem. The key concept presented in this work is the multi-modal approach, which includes a set of sensors deployed to detect abnormal events in scenarios associated to PD patients. It has been shown that accuracy of the estimations for the events considered is high as well as the usefulness of this approach for multiple environments (e.g. home, rehabilitation therapy). Results confirm the modularity of the proposed system as the system performance was good even when not all sensors were available. As a conclusion, it has been demonstrated that the multi-modal approach allows to improve the accuracy of locations estimated which is of special interest for the PDs patients' monitoring. Additionally, WSN in combination with the trajectories can be used to detect anomalies such as the patients leaving the house.

Moreover, the fusion of bracelet sensors and skeleton body points allows to perform an activity analysis. The individu-

alization process described in this paper endows the system with a multi-user functionality which is interested for places with concurrent patients such as daily centers. ABD module permits to identify unexpected behaviors that might require immediate attention. The time-line evolution of these events will allow professionals to better understand the progress of the illness and make the proper decisions.

As future work, several activities can be integrated to this approach. As an example, an analysis of medical information can be employed for the implementation of algorithms for multi-modal Fusion and analysis. Additionally, action points to fuse physical activities with subjective patients information can increase the data available for the creation of inferences in the CDS.

REFERENCES

- [1] M. Prince, W. Prina, M. Guerchet, and E. Albanese, "Dementia and risk reduction an analysis of protective and modifiable factors," Global Observatory Ageing at King's College London, Alzheimer's Disease Int., Tech. Rep., 2014.
- [2] J. Kulisevsky and J. Pagonabarraga, "Cognitive impairment in Parkinson's disease: Tools for diagnosis and assessment," *Movement Disorders*, vol. 24, no. 8, pp. 1103–1110, 2009.
- [3] M. Sajjad et al., "Leukocytes classification and segmentation in microscopic blood smear: A resource-aware healthcare service in smart cities," *IEEE Access*, vol. 5, pp. 3475–3489, 2016.
- [4] H. Medjahed, D. Istrate, J. Boudy, J. L. Baldinger, and B. Dorizzi, "A pervasive multi-sensor data fusion for smart home healthcare monitoring," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Taipei, Taiwan, Jun, 2011, pp. 1466–1473.
- [5] A. Alaiad and L. Zhou, "Patients' adoption of WSN-based smart home healthcare systems: An integrated model of facilitators and barriers," *IEEE Trans. Prof. Commun.*, vol. 60, no. 1, pp. 4–23, Mar. 2017.
- [6] A. Mathur, T. Neue, and M. Rao, "Healthcare WSN: Cluster elections and selective forwarding defense," in *Proc. 9th Int. Conf. Next Generat. Mobile Appl. Services Technol.*, 2015, pp. 341–346.
- [7] R. Lun, C. Gordon, and W. Zhao, "Tracking the activities of daily lives: An integrated approach," in *Proc. Future Technol. Conf. (FTC)*, San Francisco, CA, USA, 2016, pp. 466–475.
- [8] V. Vimarlund, N. G. Olive, I. Scandurra, and S. Koch, "Organizational effects of information and communication technology (ICT) in elderly homecare: A case study," *Health Inf. J.*, vol. 14, no. 3, pp. 195–210, 2008.
- [9] K. Rybova and J. Slavik, "Smart cities and ageing population-implications for waste management in the Czech Republic," in *Proc. Smart Cities Symp. Prague (SCSP)*, Prague, Czech Republic, 2016, pp. 1–6.
- [10] M. N. K. Boulos and N. M. Al-Shorbaji, "On the Internet of Things, smart cities and the WHO healthy cities," *Int. J. Health Geogr.*, vol. 13, no. 1, p. 10, 2014.
- [11] F.-Á. Moreno, J. A. Merchán-Baeza, M. González-Sánchez, J. González-Jiménez, and A. I. Cuesta-Vargas, "Experimental validation of depth cameras for the parameterization of functional balance of patients in clinical tests," *Sensors (Basel)*, vol. 17, no. 2, p. 424, 2017.
- [12] *Technical Specifications*. Accessed: 2016. [Online]. Available: <http://www.vivotek.com/fe8180/#specifications>
- [13] E. F. Z. Santana, A. P. Chaves, M. A. Gerosa, F. Kon, and D. Milojicic. (2016). "Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture." [Online]. Available: <https://arxiv.org/abs/1609.08089>
- [14] G. Merlino, D. Bruneo, F. Longo, A. Puliafito, and S. Distefano, "Software defined cities: A novel paradigm for smart cities through IoT clouds," in *Proc. IEEE 12th Int. Conf. Ubiquitous Intell. Comput., IEEE 12th Int. Conf. Auto. Trusted Comput., IEEE 15th Int. Conf. Scalable Comput. Commun. Assoc. Workshops (UIC-ATC-ScalCom)*, Beijing, China, Aug. 2016, pp. 909–916.
- [15] C. Doukas and F. Antonelli, "A full end-to-end platform as a service for smart city applications," in *Proc. IEEE 10th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Larnaca, Cyprus, Oct. 2014, pp. 181–186.
- [16] A. Costanzo, A. Faro, D. Giordano, and C. Pino, "Mobile cyber physical systems for health care: Functions, ambient ontology and e-diagnostics," in *Proc. 13th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Las Vegas, NV, USA, Jan. 2016, pp. 972–975.
- [17] Y. Zhang, F. Wang, and S. Yu, "Use Fukunaga-Koontz transform to solve occlusion problems in multitarget tracking," in *Proc. Chin. Conf. Pattern Recognit.*, Nanjing, China, 2009, pp. 1–5.
- [18] *Microsoft Band Specifications*. Accessed: 2015. [Online]. Available: <https://support.microsoft.com/en-us/help/4000325/band-hardware-specifications>
- [19] *Kinect Sensor Specifications*. Accessed: 2014. [Online]. Available: <https://developer.microsoft.com/en-us/windows/kinect/hardware>
- [20] *Hexiwear Specifications*. Accessed: 2017. [Online]. Available: <http://www.hexiwear.com/hardware/>
- [21] *Bluetooth Technology*. Accessed: 2016. [Online]. Available: <https://www.bluetooth.com/develop-with-bluetooth>
- [22] *IEEE802.15.4 Wireless Personal Area Networks WPAN Task Group 4. 960*. Accessed: 2010. [Online]. Available: <http://www.ieee802.org/15/pub/TG4.html>
- [23] *Kinect SDK. Windows Development Centre for Kinect SDK*. Accessed: 2016. [Online]. Available: <https://developer.microsoft.com/es-es/windows/kinect>
- [24] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, "Locating user equipments and access points using RSSI fingerprints: A Gaussian process approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.
- [25] J. Y. Zhu, A. X. Zheng, J. Xu, and V. O. K. Li, "Spatio-temporal (S-T) similarity model for constructing WIFI-based RSSI fingerprinting map for indoor localization," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Busan, South Korea, Oct. 2014, pp. 678–684.
- [26] S. Doğan, F. Yılmaz, and S. Ertürk, "On the error performance bounds of positioning using RSSI-based fingerprints," in *Proc. 23rd Signal Process. Commun. Appl. Conf. (SIU)*, Malatya, Turkey, May 2015, pp. 1325–1328.
- [27] J. Kim and C. Scott, "Robust kernel density estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Las Vegas, NV, USA, Sep. 2008, pp. 2529–2565.
- [28] V. Akbarzadeh, C. Gagné, and M. Parizeau, "Kernel density estimation for target trajectory prediction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Hamburg, Germany, Sep. 2015, pp. 3449–3456.
- [29] A. Kushki, K. N. Plataniotis, A. N. Venetsanopoulou, and C. S. Regazzoni, "Radio map fusion for indoor positioning in wireless local area networks," in *Proc. 7th Int. Conf. Inf. Fusion (FUSION)*, 2005, p. 8.
- [30] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Zürich, Switzerland, 2011, pp. 1–7. doi: 10.1109/ICORR.2011.5975346.
- [31] H. Ngoc-Huynh, P. H. Truong, and G.-M. Jeong, "Step-detection and adaptive step-length estimation for pedestrian dead-reckoning at various walking speeds using a smartphone," in *Sensors*, A. M. Sabatini Ed. Basel, Switzerland: NCBI Published, Sep. 2016. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/2759817#>, doi: 10.3390/s16091423.
- [32] N. Wagle and E. Frew, "A particle filter approach to WiFi target localization," in *Proc. AIAA Guid., Navigat., Control Conf.*, Toronto, ON, Canada, 2010, pp. 2287–2298.
- [33] F. Scholkmann, J. Boss, and M. Wolf, "An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals," *Algorithms*, vol. 21, no. 4, pp. 588–603, 2012.
- [34] B. Li, C. Yang, Q. Zhang, and G. Xu, "Condensation-based multi-person detection and tracking with HOG and LBP," in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Hailar, China, Jul. 2014, pp. 267–272.
- [35] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Proc. IEEE Int. Conf. Intell. Robots (IROS)*, Beijing China, Oct. 2006, pp. 5695–5701.
- [36] J. Gui, T. Zhang, and L. Liu, "The application of Kalman filtering in the system of image difference moving objects tracking," in *Proc. Int. Conf. Comput. Inf. Sci.*, Chengdu, China, 2011, pp. 83–86.
- [37] T. D. Gottschalk, "Concurrent implementation of Munkres algorithm," in *Proc. 5th Distrib. Memory Comput. Conf.*, 1990, pp. 52–57.
- [38] A. Lorusso, D. W. Eggert, and R. B. Fisher, "A comparison of four algorithms for estimating 3-D rigid transformations," in *Proc. Brit. Mach. Vis. Conf.*, 1995, pp. 237–246.

- [39] A. Belmonte-Hernández, G. Hernández-Peñalosa, F. Álvarez, and G. Conti, "Adaptive fingerprinting in multi-sensor fusion for accurate indoor tracking," *IEEE Sensors J.*, vol. 17, no. 15, pp. 4983–4998, Aug. 2017.
- [40] S. Yokota, D. Chugo, H. Hashimoto, and K. Kawabata, "Visual impression to robot motion imitating human-study on delay motion," in *Proc. 9th Int. Conf. Hum. Syst. Interact. (HSI)*, Portsmouth, U.K., 2016, pp. 435–439.
- [41] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [42] R. Rondón, K. Landernäs, and M. Gidlund, "An analytical model of the effective delay performance for Bluetooth low energy," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–6.
- [43] Z. Zhang, "A flexible new technique for camera calibration," in *Proc. IEEE Trans. Pattern Anal. Mach. Intell.*, Nov. 2000, pp. 1330–1334.
- [44] S. Kokalj-Filipovic, L. Greenstein, B. Cheng, and M. Gruteser, "Methods for extracting V2V propagation models from imperfect RSSI field data," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC Fall)*, Boston, MA, USA, Sep. 2015, pp. 1–5.
- [45] A. Kita, P. Lorenzi, R. Rao, and F. Irrera, "Reliable and robust detection of freezing of gait episodes with wearable electronic devices," *IEEE Sensors J.*, vol. 17, no. 6, pp. 1899–1908, Mar. 2017.
- [46] G. Parthasarathy, A. Olmsted, and P. Anderson, "Natural language processing pipeline for temporal information extraction and classification from free text eligibility criteria," in *Proc. Int. Conf. Inf. Soc. (i-Soc.)*, Dublin, Ireland, 2016, pp. 120–121.
- [47] *Standard Practice for Content and Structure of the Electronic Health Record (EHR)*. Accessed: 2013. [Online]. Available: <https://www.astm.org/Standards/E1384.htm>
- [48] *Fast Healthcare Interoperability Resources (FHIR) Standard*. Accessed: 2017. [Online]. Available: <https://www.hl7.org/fhir/toc.html>
- [49] *FIWARE Catalogue of Generic Enablers Related to EHR*. Accessed: 2017. [Online]. Available: <http://fistarcatalogue.fiware.eng.it/enablers/ehr/documentation>
- [50] M. S. Neofytou et al., "Electronic health record application support service enablers," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Milan, Italy, Aug. 2015, pp. 1401–1404.



GUSTAVO HERNÁNDEZ-PEÑALOZA (M'07) received the degree in telecommunication engineering from Universidad Santo Tomás in 2007 and the M.Sc. degree in telecommunication technologies, system and networks from the Universidad Politécnica de Valencia, in 2009. He is currently pursuing the Ph.D. degree with Universidad Politécnica de Madrid (UPM). From 2010 to 2013, he was an Associated Research Fellow with the Universidad de Valencia. He is currently with the Research Group in the Visual Telecommunications Applications Group, UPM. He has been participating with different technical developments in several national and EU projects.



ALBERTO BELMONTE-HERNÁNDEZ received the degree in telecommunication engineering and the master's degree, involving in communication systems, from the Universidad Politécnica de Madrid (UPM) in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree. He was with EVERIS SPAIN SL using Liferay programming tool and JAVA programming language to deploy web services and web pages. He is currently with the Research Group in the Visual Telecommunications Applications Group, UPM. His main interests include new communications technologies, Internet of Things, sensors, and wireless communications.



MARCOS QUINTANA received the degree in computer science engineering from UC3M in 2008, the Master Erasmus Mundus degree in color in informatics and multimedia technology in 2011, and the Researcher Aptitude in the Ph.D. Program called Communications Technologies and Systems in 2012. He developed his first year with University Jean Monnet, France, and the specialization semester in multimedia technologies with the Gjøvik University College, Norway. He is involved in different national projects concerning the improvement of transportation systems. His current research interests include object detection, classification, and tracking in different scenarios.



FEDERICO ÁLVAREZ (M'07) received the degree (Hons.) in telecom engineering and the Ph.D. (*cum laude*) degree from the Universidad Politécnica de Madrid (UPM), in 2003 and 2009, respectively. He is currently an Assistant Professor with UPM. Since 2003, he has been with the Research Group in the Visual Telecommunications Applications Group, UPM. He has been participating with different managerial and technical responsibilities in several national and EU projects, being coordinator of five EU projects in last six years.

He has authored or co-authored 60+ papers and several books, book chapters, and patents in the field of ICT networks and audiovisual technologies. He had participated in national and international standardization, including DVB and CENELEC TC206. He is a member of the program committee of several scientific conferences.

...