3-D Object Localization in Smart Homes: A Distributed Sensor and Video Mining Approach

Farzad Amirjavid, Petros Spachos, *Member, IEEE*, and Konstantinos N. Plataniotis, *Fellow, IEEE*

***Abstract*—Tracing objects in 3-D space provides information while can be used to analyze, model, predict, and recognize daily activities in the smart home environment. In this paper, we introduce a novel approach for object localization in the smart home. Our proposed method does not require the use of sensors attached to objects, so in the data collection step, the objects may move freely in the home environment. Fuzzy logic techniques are utilized to model the localization information. Specifically, the proposed framework integrates the distributed information streams obtained from multiple sensors including the visual sensors. We discuss the use of robotic assistants as part of an integrated smart home environment. Simulation results indicate that the proposed solution provides improved localization performance over the state of the art methods and introduces an intuitively pleasing robot guiding solution.**

***Index Terms*—Adoptive learning, data mining, distributed information system, fuzzy logic, image processing, object localization, robot, sensor, smart homes.**

I. Introduction

SMART home is a home like ambient environment where the embedded sensors observe the human activities of daily living (ADLs). With advances in technology, activity recognition in the smart home is attracting more interests for better technological life assistance service. Essentially, activity recognition system is a pattern recognition system. It models the activities by comparing numerical features of smart home sensors/images at learning time [1] and recognizes the learned activities in runtime. Smart home software applies artificial intelligence techniques, analyzes the home observations, and perceives the activities' patterns in a data-driven manner. Finally, it provides automatic technological life assistance services to its resident [2].

One way to observe the activities is to trace the ADL objects such as dishes, spoons, and books that are used to accomplish the ADLs. By modeling the objects movements, a smart home can recognize the activities and predict the future events [3]. In this regard, *object localization* is a key issue for activity recognition while the increase of accuracy with localization information enhances the precision with object tracking [4] and decreases error rate in activity recognition [5]. An accurate object localization system may support to identify some specific nonnormal events in the smart home environment such a knife being on the ground, whereas it is not supposed to be there. An accurate object localization system is required to detect above-mentioned such events.

Typically, the localization approaches observe the environment by using a hardware technology, and then they apply a software method to process the collected data. Improvement in any of the mentioned steps may enhance the accuracy of the object localization process [6].

Since the current problems with existing approaches concern both the data collection hardware and localization method, therefore, we focused on the improvement of these parameters. Mostly, they require installing wireless senders or Radio-frequency identification (RFID) tags on the objects, which limits the functionality of those objects. For instance, we cannot heat a pot, which is tracked by use of some attached sensors/tags. In plus, it is needed to charge frequently the batteries with these sensors, so it is not very economical to apply the localization methods, which employ many sensors. Furthermore, installation of sensors/tags on the tiny daily living objects such as the knife, spoon, and similar objects is not very practical. Regardless of the mentioned hardware constraints, the inaccuracy with their discovered localization estimations should be reduced (a few centimeters) to be proper for activity recognition purposes.

In this work, we intend to improve the accuracy factor comparing to the available localization approaches, so that we can localize tiny objects under a few centimeter error rates and with more certainty. Therefore, we propose enhancements in both system observation and data interpretation stages. We present an experiment in which the robots may autonomously infer their tasks in the home environment as a part of the localization system. The objects will be found by their visual appearance and will be localized through an adoptive learning algorithm [7], which processes the sensor data and the visual data. We analyze the positioning error and accuracy rate with our approach, and we compare the results with others' works in field of indoor object localization.

Because, we removed the requirement to attach the sensors/tags on the objects, so the objects are freely applied in the accomplishment of ADLs. Instead, they are tracked by their visual specifications, but the camera's view on the scenes may be covered by other objects. Therefore, the objects are observed from different angles of view for localization job.

Manuscript received March 12, 2016; revised July 12, 2016, October 31, 2016, and December 13, 2016; accepted January 21, 2017. Date of publication March 6, 2017; date of current version May 2, 2018.

F. Amirjavid and K. N. Plataniotis are with Multimedia Laboratory, The Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: farzad.amirjavid@utoronto.ca; kostas@ece.utoronto.ca).

P. Spachos was with the Multimedia Laboratory, The Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. He is now with the School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada.

Digital Object Identifier 10.1109/JSYST.2017.2669478

The other parameter, which we thought is the "cost of implementation." With our proposed method, we require some limited number of wireless transceivers to be independently present in the home environment. The transceivers are graphically recognizable by the visual sensors, and they are present in the smart home scenes. Their functionality is to provide distance measurement criteria as additional information to ease the processing of the visual information. Therefore, this approach does not require maintenance of as many sensors as the number of daily living objects there are in the home, but some distance measure sensors (experimentally six sensors).

Because our proposed observation methodology takes multiple information sources, which point to a single scene, then we designed a particular model to conclude the distributed information. Within this process, some hypotheses about the objects' locations are produced, and the uncertainty with the current information and the imprecision in distance measurement are determined, and finally, the position estimations, which are more accurate will be produced as the localization knowledge. Therefore, in this paper, the terms "information uncertainty," "measurement imprecision," and "position accuracy" are the abstract concepts and will have specific formalizations.

We, first, contribute to object localization in smart home by replacing the role of sensors being attached on objects with the visual sensors, which results in more efficiency and precision in object localization. Second, we contribute to distributed information systems by proposal of a methodology, which wraps the information streams from multiple observation points up and produce more certain and precise estimations about the objects' positions.

This paper is organized as follows. We review the related works about object localization in Section II. In Section III, we state the particular problem that we will focus on. In the next Section IV, the concerning mathematical definitions and algorithms are presented. In Section V, we discuss an experiment, in which we subject simulation of robotics application for object localization in the smart home. Finally, in Section VI we conclude our work and discuss the future works.

II. Related Works

Object localization and object tracking in the smart home provides rich information for activity recognition. Regarding the data collection method in indoor environments, we recognized two major groups of the existing works. First are the works, which find the objects using their graphical appearance and visual features and the second group are the ones, which find the objects by installing sensors on the target objects.

A. Visual Localization Approaches

For taking advantage of visional features of the activities, image analysis process is frequently performed to detect the simple motions in the observed scenes [8], [9] and to track human [11]–[16]. Particularly, in [1] the role of visional information for tracking the human in the ambient environment is discussed. In [16], ten visual sensors (cameras) are applied to localize five relatively large objects. Although these works are defined in the ambient environment, they avoid merging the visional localization information to nonvisional one. They focus mainly on tracking of humans and relatively large objects, but do not take the role of ADL objects into account.

B. Sensor-Attached Localization Approaches

The major sensor-attached localization approaches apply RFID tags [5] and wireless signal senders [6], [10], [17]. In some particular works, the objects' localization information is employed to recognize the human ADLs [18], [19], in which they try to estimate the location of human and objects using the embedded binary (on–off) sensors such as infrared motion sensors in home rooms. In [18] and [19], this is presumed that any actuation of the binary sensors is a spatial data, though they do not distinguish the entity of the moving objects. In [20], they tried to identify the objects by installing bluetooth sensors on the objects and tracked them by analyzing the received signals' features. In [2], [5], and [21], they attached RFID tags to the objects and localized them by processing the received signal strength indicator (RSSI) features of the signals, which are sent from RFID tags and received on RFID antennas. In [4] and [22], wireless-based approaches for indoor object localization are introduced. Another example refers to [23], in which they apply Zigbee network sensors to track the objects. The major limitation of the latter mentioned approaches concerns to sensors that should be attached and carried on the ADL objects such as dishes, spoons, knives, etc. Although sensor-attached approaches achieve the success rate of over 80%, their positioning error scale is usually more than 1 m. To improve these rates, some researchers increased the number of sensors [21]. On the one hand, these objects are often tinier than the sensors while on the contrary carrying less than four sensors leads to noticeable imprecision with localization data. Sometimes, the home resident should wash the objects, but not the sensor (because it is an electrical device). Warming up the food or any other processes with the objects may damage the sensors. Furthermore, while holding the objects, the resident should care to avoid covering the sensors by hands.

In this paper, we compare the results of our work with others [4], [5], [16], [21], [22], who let access to their data and details of their experiment. The works are compared regarding the inaccuracy rate, the experiment environment dimension and root-mean-square error (RMSE). Another parameter is dependency of the localization approach to the number of wireless nodes (topology), which are employed to localize the objects [6].

III. Problem Statement and Data Preparation Approach

A. Problem Statement

In the home environment, the scale of object localization positioning error reduces naturally to some centimeters. For example, in order to detect a knife on the ground, a system with over 1-m inaccuracy rate will not give reliable alerts to the resident. In order to reduce the inaccuracy scale to a few centimeters, we require to improve both data observation and data interpretation

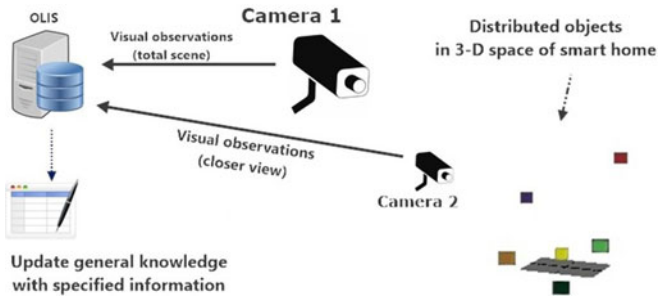

Fig. 1. Observing distributed objects with two far and close cameras.

methodologies. The main objective is to improve the current accuracy rates of object localization in the smart home.

The proposed approach detects objects using their graphical appearance. It applies robots as parts of adoptive machine learning system, so the system enhances the localization information in home environment by getting more information and details. The model employs fuzzy logic to transform the data streams into knowledge. It integrates the information coming from different sources in order to reduce the uncertainty and vagueness in knowledgebase. The context, in which the object localization information system (OLIS) automatically infers the robot duties, is simulated, and we discussed the procedure by which OLIS uses the robots information to update the existing localization knowledge. To present practicality of the proposed approach, we conducted an experiment to localize the objects in a smart home scenario. Then, we compare the achieved localization accuracy to accuracy rates with other similar works [4], [5], [16], [21], [22]. OLIS does not need to use the sensors with ADL objects. In a data-driven manner, it measures the uncertainty and imprecision with existing localization information to adopt the knowledgebase. We refer to the latter feature by the term "adoptive learning" and we propose it for the first time in the field of object localization in 3-D space.

B. Data preparation

Color and graphical features provide significant information to detect objects [24]–[26]. In [4] and [6], geographical positions of objects are defined as the closeness likelihood to some particular fixed points in the environment. To accomplish this task, OLIS detects objects using their graphical features and finds their position by the use of some graphical indicators that we already embedded in the scene. These graphical indicators are the wireless nodes, which frequently (one per second) estimate distances between each other. OLIS inputs video and sensors data streams. Then, it applies fuzzy logic to extract the knowledge from videos' and sensors' streams as shown in Fig. 1.

In Fig. 1, the cameras are the visual sensors that observe the distributed graphically recognizable objects (GROs) at different distances and angles. They pass the video streams to OLIS. Inside some of the GROs, which are in appearance of colorful cubes, we embedded wireless sensors; they find each other and calculate their metric distances through RSSI processing [27]. We call these sensors the observed network structure (ONS) elements. We presume to distribute the ONS elements in three dimensions of ambient environment space, and their primary role is to declare the actual distances between each other to OLIS. Therefore, OLIS finds the ONS elements and the GROs by their graphical specifications. In the next step, OLIS combines the two data streams and localizes the objects using triangular rules. Eventually, the knowledge will be indicated regarding the ONS elements' positions. This is called relative position [6]. For this task the sine rule [29] is applied.

The visual sensors produce 2-D pictures, and these images do not provide information regarding the distances of objects to the observation point. However, we can measure the approximate angles between the objects. In this regard, ONS elements are the distance comparison means, which let us estimate the spatial dispersion of GROs in graphically watched scenes. In this paper, multiple cameras point to a single scene out while the closer cameras provide more accurate information rather than the farther ones. In this case, two information sets differ in the scale of details they provide. In Fig. 1, we presume the distributed objects in the environment, are observed with two cameras. First one is far from the scene and captures general visual information from the entire scene; however, the closer camera focuses on a more specific region. Meanwhile, it misses coverage of some objects. In the next section, we will discuss that the prerequisite for merging two information streams is that data stream of closer camera matches the general characteristics of the remote camera's observations. In this context, we justify application of robots in the smart home by assigning them the objective of *uncertainty reduction*. In the current work, we define "decrease of uncertainty and imprecision" as the fundamental "objective and task" to apply the robots in the smart home. In this way, a robot may be assigned automatically the task to provide auxiliary/complementary observation from the target scene, so that as the result the uncertainty with available localization knowledge decreases. Therefore, in Fig. 1, we presume a robot carries the camera #2 and it is supposed providing visual information for closer views.

IV. System Model

The inputs to localization system are the video streams and wireless sensor data streams while the OLIS production is a matrix, which indicates the distance of the GRO from any of the wireless nodes (ONS elements). Every object is assigned two sorts of properties: "distance" to each of the ONS element, and "angle degrees" between GROs. In our proposed model the "distance" is the absolute information, which does not vary if the watching standpoints change. On the other hand, the angle degrees are the relative information, which changes if the viewing standpoint displaces.

In the proposed methodology for object localization the applied hardware provides initial data for object localization process to the OLIS. Then, it models the data and updates the knowledgebase with the new paradigm. This procedure terminates whenever it achieves the minimum required certainty with the current localization model. We introduce the applied definitions and propositions in prior to present the localization procedure in OLIS.

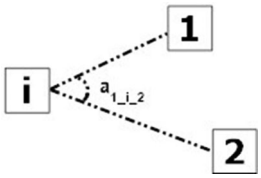

Fig. 2. Angle $a_{1\_i\_2}$ is the angle that is formed around node i at vertex and nodes 1, 2 at sides.

A. Initial Data Matrices

OLIS inputs two data streams. First are the distances between wireless nodes in the scene and second are the angles between nodes, which are observed by the camera(s).

The wireless nodes, which are called ONS elements, declare the distances between each other one per second to OLIS. Considering there are totally n ONS elements in the smart home, then the matrix holding ONS elements distance information is in form of the following:

$$\mathrm{ONS}_{\mathrm{distance}}= \begin{bmatrix} & \mathrm{N}_1 & N_2 & \dots & N_n \\ N_1 & 0 & d_{12} & \dots & d_{1n} \\ N_2 & d_{21} & 0 & \dots & d_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ N_n & d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}. \quad (1)$$

The data in (1) comes directly from embedded ONS elements. Initially, each wireless node contains physically a microcontroller and by processing the RSSI measure, it frequently generates the concerning data [27]. On the other hand, OLIS applies the "angles matrix," which indicates the angles' degrees between the GRO, including ONS elements and the target GRO in form of the following:

$$\mathrm{Angle} = \begin{bmatrix} & \mathrm{N}_1 & \mathrm{N}_2 & \dots & \mathrm{N}_\mathrm{i} & \dots & \mathrm{N}_\mathrm{n} \\ \mathrm{N}_1 & 0 & \mathrm{a}_{1\_\mathrm{i}\_2} & \dots & \mathrm{N/A} & \dots & \mathrm{a}_{1\_\mathrm{i}\_\mathrm{n}} \\ \mathrm{N}_2 & \mathrm{a}_{2\_\mathrm{i}\_1} & 0 & \dots & \mathrm{N/A} & \dots & \mathrm{a}_{2\_\mathrm{i}\_\mathrm{n}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathrm{N}_\mathrm{i} & \mathrm{N/A} & \mathrm{N/A} & \mathrm{N/A} & \mathrm{N/A} & \mathrm{N/A} & \mathrm{N/A} \\ \dots & \dots & \dots & \dots & \dots & \dots & \\ \mathrm{N}_\mathrm{n} & \mathrm{a}_{\mathrm{n}\_\mathrm{i}\_1} & \mathrm{a}_{\mathrm{n}\_\mathrm{i}\_2} & \dots & \mathrm{N/A} & \dots & 0 \end{bmatrix}. \quad (2)$$

In (2), we indicate the angles formed by the Node i at the vertex; for instance, $a_{1\_i\_2}$ shows the angle forming between sides: $\mathrm{Node}_1\_\mathrm{Node}_i$ and $\mathrm{Node}_i\_\mathrm{Node}_2$ as illustrated in Fig. 2.

OLIS processes the visual information and calculates the data of (2) frequently [28]. Considering n nodes are at each scene then $(n \times n)$ will be the size of the angles matrix.

B. Information Matrices

OLIS processes the initial data and extracts localization information from it. The first extracted information is the "GRO distances" matrix, which indicates the estimated distances between the ONS elements and the target GRO. This matrix is as similar as the one introduced in (1); however, it will be inferred by the sine law that is an equation relating the lengths of the sides of any shaped triangle to the sines of its angles [29]. OLIS determines the distance between the target GRO and the ONS element i regarding the element j by the following equation:

$$d_{\mathrm{GRO}\_i\_j} = d_{ij} \times \frac{\sin\ i\hat{j}o}{\sin\ i\hat{o}j} \quad i,j = 1:n, \quad i \neq j. \quad (3)$$

To estimate the distance of GRO to element i, OLIS applies other ONS elements (counted by j), and finds the distance ($\mathrm{GRO\_Node}_i$) using the Sine rule. Therefore, considering n ONS nodes are in the scene, then $n-1$ values will be generated as the hypothetic distances between Node i and the GRO, whereas we can apply $n-1$ ONS nodes as distance estimation criteria. In the consequence, OLIS fuzzifies the $n-1$ hypothetic distances between Node i and GRO. In other words, OLIS classifies the $n-1$ distance units in fuzzy clusters and represents each group by a set of centroids. Therefore, a possibilistic space to represent the GRO position will be formed. The products of the fuzzy clustering process are two matrices: one indicates the cluster centroids (c) and the other one specifies the "influence range values." We refer to the latter matrix by s. In order to fuzzify the data points, we propose to employ a modified version of subtractive clustering algorithm [30]. The main reason for selecting the subtractive clustering algorithm is that it does not require a prior expert knowledge about the object locations [3], [30], so it estimates the centroids automatically.

By this algorithm, the potential measure (pot) per each data point for being the centroid is computed. The point, which is the closest to all data points is $\min(\mathrm{pot})$. The data points around each centroid are assigned to the concerning cluster and subtracted from further processing. In the next step, the potential measure of the recently selected centroid is deduced from all remaining data points, and the process continues to find the other centroids. Considering there are k data points in a dataset $d = \{x_1, x_2, .., x_k\}$, then the algorithm finds the potential function of being the new centroid is as the following:

$$\mathrm{Pot}_i = \sum_{i,j=1}^{k} e^{-\alpha \|x_i - x_j\|^2} \quad \alpha = \frac{4}{r_a^2}. \quad (4)$$

In (4), the Euclidean distance between data points is computed, and they are put in the clusters, which have the cluster radius of r_a and the centroid is in (c). Another product of clustering procedure is the s matrix, which indicates the influence range of each centroid. OLIS finds the s value on d by the employment of the following equation:

$$s = (\max(d) - \min(d)) \times 4.53 \times 10^{-2} \times e^{2.355 \times r_a}. \quad (5)$$

In (5), we can see the s value is dependent on the range of the dataset and the desired influence range $0 \leq r_a \leq 1$. In Fig. 3, we present the centroid estimation procedure.

The algorithm in Fig. 3 is modification of "subtractive clustering" algorithm [30]. The modifications are removal of the squash factor effect (to increase efficiency) and presentation of a clear definition for matrix s.

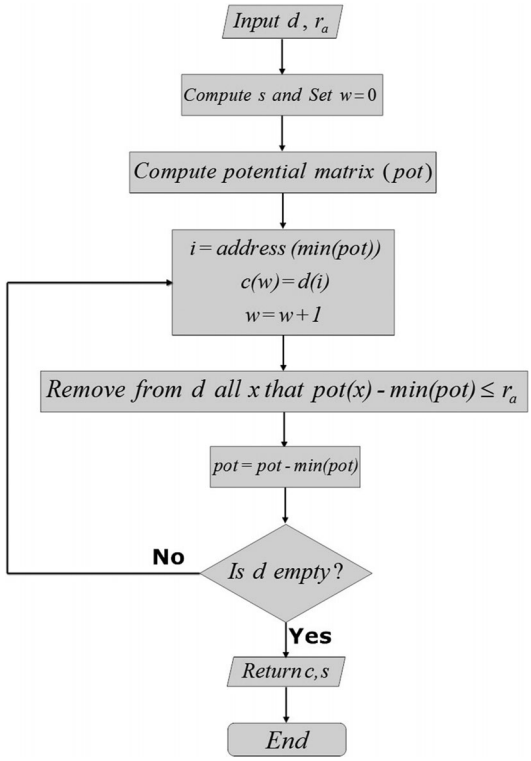

Fig. 3. Fuzzy centroid estimation procedure.

The information in matrix c provides criteria for evaluating the closeness of hypotheses to the GRO position, but it does not provide a mean to reject the erroneous hypotheses. In [3], an operator to calculate the least possible values for object positions is introduced. This is called absolute fuzzy symmetry of the centroid matrix: $\text{sym}(c) = 1 - c$. The fuzzy symmetry values let us view the world in a range of the impossible values to very possible values regarding the target GRO position. This range explains the GRO location by a possibilistic space. The set of centroids may be traversed by a fitting curve, which simultaneously avoids the impossible values [3], [31]. The coefficients for a polynomial $p(x)$ of degree n that is the best fit (in a least-squares sense) for the data in y. Here, x is the matrix, which has the most possible and the least possible positions for GRO and y is the matrix that gives highest possibility rank (+1) to the c ; and the lowest possibility rank (-1) to the $\text{sym}(c)$

$$x = [c_{\text{GRO}}; \text{sym}(c)_{\text{GRO}}] \quad y = [\text{ones}(n); -1 \times \text{ones}(n)]. \quad (6)$$

The coefficients in p are in descending powers and the length of p is $n + 1$. The p is as the following:

$$p(x) = p_1 x^n + p_2 x^{n-1} + \ldots + p_n{}^x + p_{n+1} + e_i \ (i = 1 : n). \quad (7)$$

The curve, which traverses x and maps it to y, is p. The symbol P_{GRO} represents a possibilistic space regarding the GRO position. To calculate the p, the model in (7) can be written as a system of linear equations

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ 1 & x_3 & x_3^2 & \ldots & x_3^n \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{bmatrix} \times \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \ldots \\ p_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \ldots \\ e_n \end{bmatrix}. \quad (8)$$

In (8), m is the size of x. y_i is an element of the $\vec{y}$. Similarly, the $\vec{e}$ and $\vec{p}$ are the concerning vectors. The (8) when using pure matrix notation is written as: $\vec{y} = x\vec{p} + \vec{e}$. The vector of estimated polynomial regression coefficients (using ordinary least squares estimation) is: $\vec{p} = (x^T x)^{-1} x^T \vec{y}$.

P_{GRO} represents a bipolar possibilistic space regarding the GRO position. It inputs the hypothetic positions and outputs how much "closed" the hypothetic positions regarding the GRO location are. Therefore, to evaluate trueness of any hypothetic position, we load P_{GRO} with hypothetic values h. The result $P_{\text{GRO}}(h)$ is the possibility measure indicating closeness to P_{GRO}. In this regard, z is the matrix, which indicates the likelihood degrees for all hypotheses as the actual location of the GRO: $z = P_{\text{GRO}}(h)$. The matrix z is, in fact, a set of fuzzy ranks, about the correctness measures of any hypothesis as the likely right position of the GRO.

The scenes may be observed from different standpoints and per each point specific information matrices may be gained. In order to achieve accuracy with knowledge in the knowledgebase, OLIS evaluates the information uncertainty and imprecision in a data-driven manner and decides whether to apply the new information. Considering n ONS elements create $n - 1$ hypotheses as possible positions for the GRO, then $z = P_{\text{GRO}}(h)$, will be a $1 \times n - 1$ matrix. By the following, we indicate the evaluated average of the position precision

$$\overline{PP}_{\text{GRO}} = \frac{\sum_{i=1}^{i=n-1} z_i}{n}. \quad (9)$$

$\overline{PP}_{\text{GRO}}$ is an index indicating the Euclidean distances between a group of hypotheses and the existing model in knowledgebase. We will refer to this parameter by the term "information precision." OLIS uses this criterion to evaluate how much the existing model p_{kb} (the model in the knowledgebase) matches the information from the new standpoint $\overline{PP}_c$ (current hypotheses). Watching the similar scene from a closer view causes the $\overline{PP}_c \geq \overline{PP}_{\text{kb}}$. In this case, the new observation in new standpoint will be modeled and checked if it leads to better understanding of the world. Precision and *Imprecision* $(1 - \overline{PP})$ are the products of evaluation process. An accurate model causes relatively higher imprecision value, whereas it produces high z value to the single unique true hypothesis, but smaller z values to the remaining $n - 2$ hypotheses. OLIS applies another criterion to evaluate the information, which we will refer to it by the term "information uncertainty."

By applying the (5), OLIS measures the dissimilarity between the hypothetic positions. The more centroids are similar (or closed), then the s matrix will include smaller values; but if the hypotheses are dissimilar, then the s matrix will contain

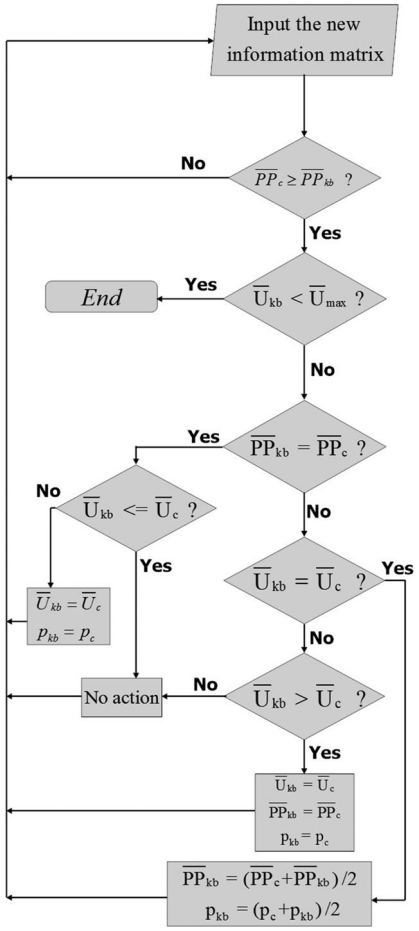

Fig. 4. Algorithm to merge the information and update OLIS knowledgebase.

higher values. We define the average value of s elements as the uncertainty measure and we define it in form of the following:

$$\overline{U}_{\mathrm{GRO}\_i\_j} = \frac{\sum_{k=1}^{n} S_{\mathrm{GRO}\_i\_j}(k)}{n}. \tag{10}$$

The uncertainty matrix is the sum of the s matrix elements. If the $\overline{U}$ value is high then OLIS infers that the information includes high uncertainty. Uncertainty can be the result of viewing a scene from far distances. By high uncertainty we mean although the data includes valid hypotheses, it also includes far acceptable assumptions, so further investigations may be required. In this case, a robot may automatically infer the task to go around the vagueness area.

In (10), the uncertainty measure included with information is formalized. In there, the uncertainty is formalized as the *radius of the clusters*. $\overline{U}$ is the uncertainty measure, which is the product of the modeling process, particularly in centroid estimation procedure. When the clustering method selects small clusters for data points, then $\overline{U}$ value decreases and machine states can be clearly classified by analysis of the training data. On the other hands, if the training data includes noise or the applied clustering procedure could not clearly identify the machine states, then large clusters indicating vagueness on machine states will be formed and the $\overline{U}$ value increases.

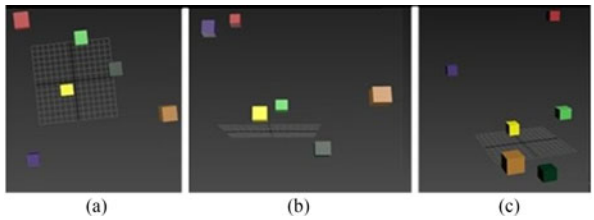

Fig. 5. Three views on distribution of objects in 3-D space.

C. Updating the Knowledgebase

To keep OLIS knowledgebase updated with more precise and certain information, we define three general conditions we face in information processing and per each status, OLIS performs a reaction. The first status is that the current information and existing knowledge indicate similar information (or precision measure to object location); in other words, they point to the same location for GRO out. The second category is that the two information matrices indicate different information with similar uncertainties, and the third status is that the information matrices show dissimilar information at dissimilar certainty values. Fig. 4 shows the proposed algorithm to update the OLIS knowledgebase.

In Fig. 4, $\overline{PP}_{\mathrm{kb}}$ is the existing GRO position precision in knowledgebase, $\overline{U}_{\mathrm{kb}}$ is the existing GRO position uncertainty in knowledgebase, $\overline{PP}_c$ is the current GRO position precision, and $\overline{U}_c$ is the current GRO position uncertainty. $\overline{U}_{\max}$ is the maximum desired GRO position uncertainty in knowledgebase. p_{kb} refers to existing position (knowledge) in knowledgebase and p_c is the current position (knowledge). Fig. 4 shows that for the case, OLIS faces similar information; it keeps the current GRO position matrix, but updates the uncertainty matrix ($\overline{U}$). For the second case, OLIS updates the GRO position matrix by taking an average of GRO position information; and in case, both information and certainty measures are different, the OLIS updates both GRO and $\overline{U}$ matrix by the one, which holds less uncertainty. We need a limit to identify two values as approximately equal or unequal. In our experiment, we propose to consider the limit of "1%" as the maximum rate of difference, which is required to identify two values as approximately equal.

V. Experimenting Localization of a GRO

In this case study, we localize the GRO in a small environment. The objective is to demonstrate the accuracy of the proposed adaptive learning model in small rooms.

A. Simulation Environment

Six objects are the subject of the current localization problem. Five objects are presumed to be wireless ONS elements [27] and one object (yellow) is targeted to be localized. The GRO does not send any wireless signal and the only way to perceive it is the graphical appearance. Fig. 5 shows 3-D schema of the ONS elements and GRO distribution from three views.

Fig. 5 shows the ONS elements distributed in three dimensions of space. The experiment space is $2 \times 2 \times 2$ m. The visual sensor is an AXIS M1004-W camera. The processor with server

TABLE I
ACTUAL DISTANCES BETWEEN THE OBJECTS (IN CENTIMETER SCALE)

	Orange	Dark Green	Light Green	Yellow	Red	Violet
Orange	0	45	39	47	80	50
Dark Green	45	0	40	36	84	51
Light Green	39	40	0	37	56	52
Yellow	47	36	37	0	55	20
Red	80	84	56	55	0	55
Violet	50	51	52	20	55	0

TABLE II
HYPOTHESES ABOUT DISTANCE OF YELLOW OBJECT TO ONS ELEMENTS

Orange	Dark green	Light green	Red	Violet
37	50	26	30	18
42	31	30	60	19
82	50	60	48	13
52	28	28	51	19

computer is Intel(R) Core(TM) i7-4500U CPU @ 1.88 GHz and includes 8 GB RAM.

B. Methodology

The ONS elements are in colors of orange, red, violet, light green, and dark green. Table I, shows the actual position of target GRO (yellow object).

In this experiment, the target is to localize the *yellow object* regarding the positions of ONS elements. The volume of this object is one cubic centimeter. The methodology is that according to the algorithm in Fig. 4, we surveyed the problem in three steps to achieve the target $U_{\max} = 0.0075$. At first, we localize the yellow object by viewing the scene from the standpoint (a) in Fig. 5. Then we observe the same scene from an entirely different perspective, but similar approximate distance (view b in Fig. 5). At the third step, the scene is observed from a different angle but closer distance (View (c) in Fig. 5).The localization process terminates whenever OLIS achieves the $U_{\max}$.

The images in each scene are analyzed, and angles between nodes are calculated. OLIS estimated the GRO position at each step. For instance, the estimated GRO position at the first standpoint is shown in Table II.

Information in Table II points to the position of the GROs in space. The object position will be represented by a set of centroids accompanied with a vector, which indicates the sigma values that specify the range of influence (s). Fuzzy subtractive clustering algorithm will produce the fuzzy position of the yellow node [c matrix in (10)]. The product of the clustering process at influence range of 0.9 on the yellow spatial data is indicated in Table III.

Therefore, Table III shows the fuzzy distance of the yellow object regarding other nodes in the scene. A fuzzy (7) indicates the fuzzy space for the yellow object. The fuzzy equation matrix concerning to the location of the yellow GRO is as the following[1]:

TABLE III
FUZZY DISTANCES BETWEEN YELLOW NODE AND OTHERS

Orange	Dark green	Light Green	Red	Violet
52	28	28	51	19
82	50	60	48	13
37	50	26	30	18
		Sigma values		
14.31	7	10.81	9.54	1.9

$$\mathrm{p}_1 = [\,16.8511 \quad -25.2766 \quad 7.2784 \quad 0.4736\,]. \tag{11}$$

The polynomial curve fits all three mentioned locations and avoids the least possible areas (estimations). By this model, the closer positions to the yellow object get a higher rank and the farther objects take the lowest grades for fuzzy space of the yellow object. For example, an arbitrary point $(80, 50, 60, 60, 10)$ will be evaluated by the current rank

$$\mathrm{Eval} = [\,0.4991 \quad 1.0267 \quad .8972 \quad 0.8972 \quad 0.7864\,]. \tag{12}$$

The matrix in (12) indicates the closeness degree of yellow object to the arbitrary position $(80, 50, 60, 60, 10)$. This possibility degree is 0.684433. This is although the closest rank to the actual position, that is: 0.8697833. According to data of Table III, the uncertainty with available knowledge is at the rate of 0.04356, which is greater than $U_{\max}$. In this regard, a robot task may be assigned to go around the region and to watch the area so that OLIS may receive complementary information to reduce the uncertainty rate. The estimation of GRO position by processing the graphical image resulted from viewing at second standpoint is as follows:

$$P_2 = [\,16.5779 \quad -24.8669 \quad 7.3048 \quad 0.4921\,]. \tag{13}$$

The knowledge in here is adoptable by getting an average on two p matrices. The new P matrix is represented in the following:

$$P_{1,2} = [\,16.7145 \quad -25.07175 \quad 7.3916 \quad 0.48285\,]. \tag{14}$$

This model gives the uncertainty rate of 0.0528, which is still greater than the $U_{\max}$. Therefore, in the following, the GRO position at third standpoint is calculated as follows:

$$p_3 = [\,14.6574 \quad -21.9862 \quad 6.2492 \quad 0.5397\,]. \tag{15}$$

p_3 gives uncertainty rate of 0.07, which is still greater than the $U_{\max}$. Therefore, according to the update algorithm, the current model (p_3) merges with the previous model ($p_{1,2}$) in the knowledgebase. The following represents the ultimate model, which is the outcome of observing a single scene at three different standpoints:

$$p_{1,2,3} = [\,14.5472 \quad -21.8208 \quad 6.1826 \quad 0.5455\,]. \tag{16}$$

The model in (16) produces the uncertainty rate of 0.0073, which satisfies the target $U_{\max}$. In Table IV, we show how the

[1]In order to normalize to normalize the distance values; we considered the value '200' as the maximum distance.

TABLE IV
EVALUATION OF ALL EXPERIMENT HYPOTHESES WITH $p_{1,2,3}$

Standpoint No	Orange	Green	Light green	Red	Violet	Possibility degree
1	52	28	28	51	19	0.8122
1	82	50	60	48	13	0.6719
1	37	50	26	30	18	0.8281
2	47	27	26	55	19	0.8099
2	34	65	54	59	15	0.739
2	59	41	20	19	20	0.789
2	16	32	21	48	26	0.8168
3	48	35	32	44	20	0.8352
3	55	43	31	44	21	0.8196
3	46	34	34	44	22	0.8417
Real position	47	36	37	55	20	0.8185
Arbitrary point	80	50	60	60	10	0.6455

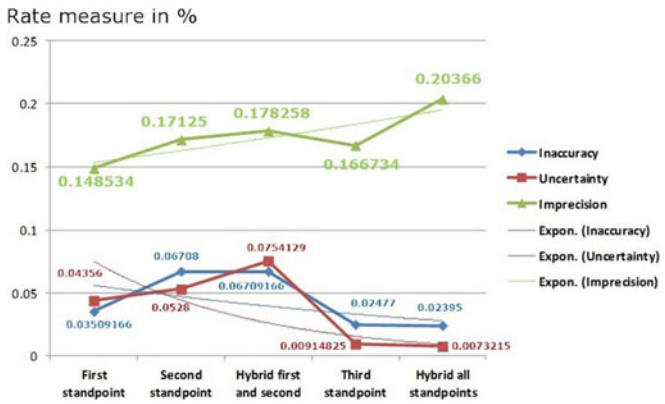

Fig. 6. Uncertainty, imprecision, and inaccuracy rates in experiment I.

final model distinguishes between the far acceptable hypotheses and the more actual ones.

We expect that OLIS assigns a smaller possibility degree to the hypothetic areas, where are farther than the actual position of the GRO, and it gives a better (higher) rank to the areas that are closer to the real location of the GRO. In Fig. 6, we show the achieved uncertainty rates, imprecision rates $(1 - pp)$, and inaccuracy rates at each step of the object localization.

In Fig. 6. the exponential trend curves indicate the tendency with inaccuracy, imprecision, and uncertainty measures while localization knowledge in five experimented steps updates. Inaccuracy rate is the relative Euclidean distance between the actual position and the estimated position:

$$\text{Inaccuracy}_k = \frac{\sum_{j=1}^{N_h} \sum_{i=1}^{n} \frac{\sqrt{(d_{\text{GRO},i} - h_{j,i})^2}}{\max(d_{\text{GRO},i}, h_{j,i})}}{i \times j}, \quad k = 1:5. \tag{17}$$

Inaccuracy is computed as the relative distance between the true position $d_{\text{GRO},i}$ (information in Table I) and the hypothetic position $h_{j,i}$ (information in Table IV). In (17), N_h refers to the number of hypotheses and k counts the produced models. In Fig. 6, the trend curves indicate that uncertainty and inaccuracy rates are correlated; however, the imprecision is negatively correlated to the inaccuracy rate. Therefore, in this experiment, whenever OLIS calculates a more accurate model,

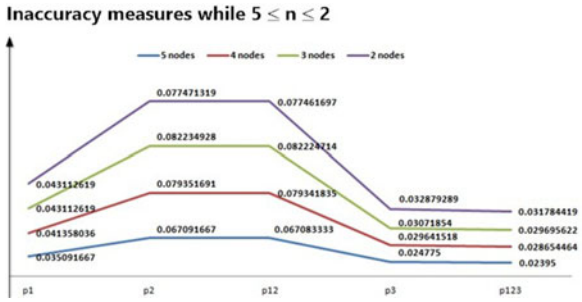

Fig. 7. Inaccuracy measures in experiment I.

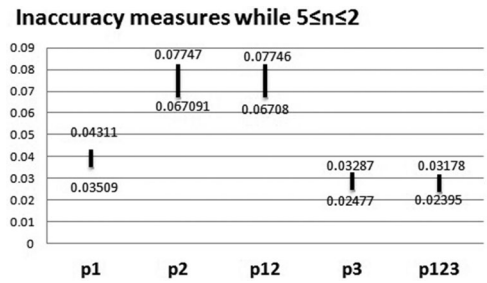

Fig. 8. Inaccuracy ranges in experiment I.

then the difference between position imprecision and the uncertainty rates increases. In this regard, targeting achievement of uncertainty/imprecision < 0.04 will result in similar results to searching the Umax, which can be used as an alternative objective function for OLIS robots. In Fig. 7, we review the inaccuracy rates, when the quantity of nodes decreases (the effect of topology change).

In Fig. 7, we show the system inaccuracy measures, the wireless network topology changes. In Fig. 8, we show the range of inaccuracy measures.

In Fig. 8, we represent briefly the information of the Fig 7. The accuracy rate stays around 97% if the number of ONS elements decreases. In the work of [6], which is in scale of large wireless sensor networks, the accuracy rate stays around 80%, whereas the quantity of nodes change from 10% (100 nodes out of 1000 nodes) to 90% (900 nodes out of 1000 nodes). This indicates the information will remain reliable if major part of the observation network change.

VI. COMPARISON WITH SIMILAR APPROACHES

The introduced approach makes it feasible to localize the objects while being applied in the accomplishment of daily living activities because it does not require attaching sensors directly on the sensors. The proposed prototype observes "visual" and "distance" specifications in the smart home. The visual information is provided through visional sensors, and it is used to calculate the angle degrees between the objects. On the other hand, the distance sensors give specific measures to estimate the real distances between observed objects. Because there are several objects in each scene, then the position each object might be defined regarding the positions of multiple other objects, as a result per each target object several localization hypotheses

TABLE V
COMPARING LOCALIZATION EXPERIMENT

Approach	Inaccuracy rate	Experiment dimension	$\overline{ARMSE}$	No. nodes
Wireless network [4]	80%	42 × 46 m	2.5 m	93
RFID tag [5]	82%	5 × 6 m	3 m	9
RFID tag [22]	83%	5 × 5 m	1.5 m	4
RFID tag [21]	93%	60 × 60 cm	15 cm	4
Visional approach [16]	93%	15 × 8 m	75 cm	10
Current approach	97%	2 × 2 × 2 m	2 cm	5

could generate. The proposed model in Section IV refines the hypotheses and based on the available knowledge it achieves the relatively most certain and precise estimation regarding the position of the objects in the home environment. In this section, we compare the outcome of our proposed approach with other similar works.

One parameter that we apply to compare the current work with others is the accuracy of the localization estimations, which means how much the real-object position corresponds to what the OLIS estimates. Using the RMSE factor, we compare the two approaches outcomes. This factor indicates the positioning error and the $\overline{\text{ARMSE}}$ is the average RMSE among the hypotheses [4], [6]

$$\overline{\text{ARMSE}}_k = \frac{\sum_{j=1}^{N_h} \sum_{i=1}^{n} \sqrt{d_{\text{GRO},i} - h_{j,i}}}{i \times j}, \quad k = 1:5. \tag{18}$$

RMSE indicates the precision scale or positioning error scale of an approach. Note that the RMSE measure provides an indication of both bias and variance of the estimator. In Table V, we compared the proposed approach with others [4], [5], [16], [21], [22], particularly in static positioning, in regard of inaccuracy rate, experiment dimension, $\overline{\text{ARMSE}}$ factor, and number of nodes.

In Table V, we compared some approaches that let us access to data and some details of experiments. With [4], [5], [21], and [22] the objects carry one or more signal senders.

VII. CONCLUSION AND FUTURE WORKS

In order to localize the objects used in daily living activities, we proposed a new approach, which lets human apply objects freely, so the particular constraints caused by sensors physical entities such as, sensors' size, weight, battery live, and none-resistance against the heat are eliminated. We localized the objects using their visual appearance, and all applied sensors do not intervene directly in the accomplishment of daily activities. Besides the proposed hardware methodology for object localization, we also introduced a software technique that combines information streams coming from visual sensors and wireless sensors. This approach automates the update process of localization information. To make the smart home localization knowledgebase we introduced and formalized "information precision and uncertainty" as the two main decision-making factors for this artificial-intelligence-based software process. We justify application of robots in the smart home as a mechanism to increase certainty and accuracy with localization information. We achieved the accuracy and positioning error rates, which let localize ADL objects in home. One noticeable feature of the model is that it reduces the uncertainty to a few centimeters. We localized an object, which is as tiny as one cubic centimeter and smaller than an RFID tag. Therefore, little and small objects may be localized accurately with our proposed model. Besides, we simulated how to apply robots in the smart home as a part of the information system. OLIS infers their tasks automatically, and their observations assist the smart home information system. The result of this work is compared with other works in field of indoor object localization. In the future works, we will track the moving objects, and we will track the objects for recognition of daily living activities by use of robots.

REFERENCES

[1] C. Wu *et al.*, "Multiview activity recognition in smart homes with spatio-temporal features," in *Proc. 4th ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Aug. 2010, pp. 142–149.

[2] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Trans. Syst. Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 790–808, Nov. 2012.

[3] F. Amirjavid *et al.*, "Data driven modeling of the simultaneous activities in ambient environments," *J. Ambient Intell. Humanized Comput.*, vol. 5, no. 5, pp. 717–740, 2013.

[4] A. Kushki, K. N. Plataniotis, and A. N. Venetsanopoulos, "Intelligent dynamic radio tracking in indoor wireless local area networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 3, pp. 405–419, Mar. 2010.

[5] H. Noguchi *et al.*, "Object localization by cooperation between human position sensing and active RFID system in home environment," *SICE J. Control, Meas. Syst. Integr.*, vol. 4, no. 3, pp. 191–198, 2011.

[6] L. Shancang, X. Wang, S. Zhao, J. Wang, and Ling Li, "Local semidefinite programming-based node localization system for wireless sensor network applications," *IEEE Syst. J.*, vol. 8, no. 3, pp. 879–888, Sep. 2014.

[7] Z. Chen *et al.*, "Unobtrusive sensing incremental social contexts using fuzzy class incremental learning," in *Proc. IEEE Int. Conf. Data Mining*, Nov. 2015, pp. 71–80.

[8] Z. Pan, J. Lei, Y. Zhang, X. Sun, and S. Kwong, "Fast motion estimation based on content property for low-complexity H.265/HEVC encoder," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 675–684, Sep. 2016.

[9] Z. Pan *et al.*, "Fast reference frame selection based on content similarity for low complexity HEVC encoder," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 516–524, 2016.

[10] Q. Hao, F. Hu, and Y. Xiao, "Multiple human tracking and identification with wireless distributed pyroelectric sensor systems," *IEEE Syst. J.*, vol. 3, no. 4, pp. 428–439, Dec. 2009.

[11] N. Bo Bo *et al.*, "Real-time multi-people tracking by greedy likelihood maximization," in *Proc. 9th Int. Conf. Distrib. Smart Cameras*, Seville, Spain, Sep. 2015, pp. 32–37.

[12] S. Zhang *et al.*, "Tracking multiple interacting targets in a camera network," *Comput. Vis. Image Understanding.*, vol. 134, pp. 64–73, 2015.

[13] Z. Wang and H. Aghajan, "Tracking by detection algorithms using multiple cameras," in *Distributed Embedded Smart Cameras (Part III)*, 1st ed. New York, NY, USA: Springer, 2014, pp. 175–188.

[14] M. Bredereck *et al.*, "Data association for multi-object tracking-by-detection in multi-camera networks," in *Proc. 6th Int. Conf. Distrib. Smart Cameras*, Oct. 2012, pp. 1–6.

[15] K. Ozcan, A. K. Mahabalagiri, M. Casares, and S. Velipasalar, "Automatic fall detection and activity classification by a wearable embedded smart camera," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 2, pp. 125–136, Jun. 2013.

[16] K. Leung *et al.*, "The UTIAS multi-robot cooperative localization and mapping dataset," *Int. J. Robot. Res.*, vol. 30, no. 8, pp. 969–974, 2011.

[17] X. Jiang *et al.*, "Feature adaptive online sequential extreme learning machine for lifelong indoor localization," *J. Neural Comput. Appl.*, vol. 27, pp. 215–225, 2016.

[18] Y. Sahaf, "Comparing sensor modalities for activity recognition," Master's thesis, Washington State Univ., Pullman, WA, USA 2011.

[19] B. Minor *et al.*, "Data-driven activity prediction: Algorithms, evaluation methodology, and applications," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Sydney, NSW, Australia, Aug. 2015, pp. 805–814.
[20] D. Schwarz *et al.*, "Cosero, find my keys! Object localization and retrieval using Bluetooth low energy tags," *RoboCup 2014: Robot World Cup XVIII*, (Lecture Notes Comput. Sci. Series), vol. 8992, 1st ed. New York, NY, USA: Springer, 2015, pp. 195–206.
[21] A. A. Nazari Shirehjini, A. Yassine, and S. Shirmohammadi, "An RFID-based position and orientation measurement system for mobile objects in intelligent environments," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1664–1675, Jun. 2012.
[22] B. Kaluza *et al.*, "An agent-based approach to care in independent living," *Lecture Notes in Computer Science*, 1st ed. Berlin, Germany: Springer, Nov. 2010, pp. 177–186.
[23] P. da Fonseca and P. Rosa, "Tracking objects in a smart home," in *Proc. Comput. Intell. 11th Brazilian Congr. Comput. Intell.*, Sep. 2013, pp. 574–579.
[24] X. Li and K. N. Plataniotis, "Circular mixture modeling of color distribution for blind stain separation in pathology images," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 150–161, Jan. 2017.
[25] Y. Zheng *et al.*, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *J. Intell. Fuzzy Syst.*, vol. 28, no. 2, pp. 961–973, 2015.
[26] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 166–176, Jun. 2015.
[27] P. Spachos and D. Hantzinakos, "Scalable dynamic routing protocol for cognitive radio sensor networks," *IEEE Sensors J.*, vol. 14, no. 7, pp. 2257–2266, Jul. 2014.
[28] Mathworks.com, "Measuring angle of intersection," [Online]. Available: http://www.mathworks.com/help/images/examples/measuring-angle-of-intersection.html. Accessed on: Dec. 12, 2016.
[29] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematicals Tables*, North Chelmsford, MA, USA: Courier Corporation, 1970.
[30] S. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intell. Fuzzy Syst.*, vol. 2, no. 3, pp. 267–278, 1994.
[31] N. Chernov, *Circular and Linear Regression*. Boca Raton, FL, USA: CRC Press, 2011.

Farzad Amirjavid is a Postdoctoral Fellow with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. He is interested in the extension of fuzzy logic applications in the design of intelligent systems. His research interests also include the design of smart 3-D positioning systems, IoT systems, distributed video processing systems, and predictor systems. He has been involved in projects regarding the machine learning of real-world problems such as activity recognition systems and wireless ambient environments.

Petros Spachos (S'09–M'14) received the Diploma degree in electronic and computer engineering from the Technical University of Crete, Crete, Greece, in 2008, and the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2010 and 2014, respectively.

He was a Post-Doctoral Researcher with the University of Toronto from September 2014 to July 2015. Since August 2015, he has been an Assistant Professor with the School of Engineering, University of Guelph, Guelph, ON, Canada. His research interests include wireless networking and network protocols with a focus on wireless sensor and cognitive networks. He is involved in protocol design, real world experimentation, and performance analysis.

Dr. Spachos is a member of the ACM.

Konstantinos N. (Kostas) Plataniotis (S'90–M'92–SM'03–F'12) is a Professor and the Bell Canada Chair in Multimedia with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. He is the co-founder and inaugural Director-Research for the Identity, Privacy and Security Institute (IPSI), University of Toronto, and he served as the Director of the Knowledge Media Design Institute (KMDI), University of Toronto, from January 2010 to July 2012. His research interests are knowledge and digital media design, multimedia systems, biometrics, image and signal processing, communications systems, and pattern recognition. He authored *WLAN Positioning Systems* (2012) and *Multi-Linear Subspace Learning: Reduction of Multidimensional Data* (2013).

Dr. Plataniotis is a Registered Professional Engineer in the Province of Ontario. He is a Fellow of the Engineering Institute of Canada. He has served as the Editor-in-Chief of IEEE SIGNAL PROCESSING LETTERS and as Technical Co-Chair of the IEEE 2013 International Conference in Acoustics, Speech and Signal Processing. He was the IEEE Signal Processing Society Vice President for Membership (2014–2016). He is the General Co-Chair for the 2017 IEEE GlobalSIP Conference, the General Co-Chair for the 2018 International Conference on Image Processing (ICIP-18), and the General Co-Chair for the 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2021).