# A Big Bang-Big Crunch Type-2 Fuzzy Logic System for Machine Vision-Based Event Detection and Summarization in Real-world Ambient Assisted Living

Bo Yao, Hani Hagras, *Fellow*, *IEEE*, Daniyal Alghazzawi, *Senior Member*, *IEEE* and Mohammed J. Alhaddad, *Member*, *IEEE*

***Abstract*—The area of Ambient Assisted Living (AAL) focuses on developing new technologies which can improve the quality of life and care provided to elderly and disabled people. In this paper, we propose a novel system based on 3D RGB-D vision sensors and Interval Type-2 Fuzzy Logic based Systems (IT2FLSs) employing the Big Bang Big Crunch (BB-BC) algorithm for the real time automatic detection and summarization of important events and human behaviours from the large-scale data. We will present several real world experiments which were conducted for AAL related behaviours with various users. It will be shown that the proposed BB-BC IT2FLSs outperforms the Type-1 FLSs (T1FLSs) counterpart as well as other conventional non-fuzzy methods, and the performance improvement rises when the amount of subjects increases.**

***Index Terms*—Interval Type-2 fuzzy logic systems, 3D machine vision, event summarization.**

## I. Introduction

The World Health Organization (WHO) estimated that in 2050, there will be 1.91 billion people aged 65 years and over worldwide [1]. Hence, recently, there have been increased interests in Ambient Assisted Living (AAL) technologies due to the increase of ageing population, shortage of caregivers and the increasing costs of healthcare [2], [3], [4]. Employing advanced machine vision based systems for behaviour and event detection as well as event summarization in AAL applications can help to increase the level of care and decrease the associated costs. However, the great expansion of deploying and utilizing video sensors can lead to massive amounts of redundant video data which require high associated costs related to data storage in addition to the human resources spent on watching or manually extracting key video information. This problem is becoming increasingly obvious as the number of video cameras in use is estimated to be 100 million worldwide [5] and the estimated number of in-use cameras is 5.9 million in United Kingdom owning the largest amount of camera in the world.

This project was supported by the NSTIP strategic technologies program in the Kingdom of Saudi Arabia – Project No. (12-INF2259-03). The authors also acknowledge with thanks Science and Technology Unit, King Abdulaziz University for technical support.

Bo Yao and Hani Hagras are with the Computational Intelligence Centre, School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester, CO43SQ, UK. (e-mail: byao@essex.ac.uk)

Hani Hagras is with the Computational Intelligence Centre, School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester, CO43SQ, UK. (e-mail: hani@essex.ac.uk)

Daniyal Alghazzawi is with the Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia.

Mohammed J. Alhaddad is with the Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia.

Many intelligent applications of AAL and healthcare have been reported based on behaviour and activity recognition. Recognizing and categorizing human behaviour into one of several behaviour classes falls under the generic class of pattern recognition problems which aims to determine the mapping between behavioural feature space and action categories. To describe and represent the behaviour of the human subject, behavioural features are used and can be captured by different types of sensors such as cameras, 3D sensors, RFID sensors, and wearable sensors, etc. In [6], a method was introduced to analyse the behaviour of watching TV for diagnosing health conditions. In [7], researchers have proposed an algorithm to analyse risk of falling down for elderly users according to the walking patterns. Wan et al. [8] developed a behaviour recognition system to prevent the wandering behaviour of dementia patients. Lin et al. [9] utilized RFID sensors to detect if a dementia patient has approached an unsafe region in order to avoid potentially injurious situations. Barnes et al. [10] presented a low-cost solution to realizing an intelligent remote telecare system to assess the lifestyle activity data of the elderly. However, the system is simple and is functional limited. Hoey et al. [11] introduced a cognitive rehabilitation system using AAL technologies to help the elderly with dementia. Another cognitive orthotics system [12] analyses a model of the everyday activity plan according to multi-level events for the purpose of cognitive orthotics. However, unlike our system, extendable recognition for complex behaviour and activity together with the summarization of the frequency, duration, timestamp and the user information is not implemented in these conventional systems.

Machine vision based behaviour recognition and summarization in real-world AAL is a very difficult task due to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



the high levels of encountered uncertainties caused by the large number of subjects, behaviour ambiguity between different people, occlusion problems from other subjects (or non-human objects such as furniture) and the environmental factors such as illumination strength, capture angle, shadow and reflection, etc. Dynamic models of behaviour characteristics can be constructed by utilizing statistics-based algorithms, for example Conditional Random Fields (CRF) [13] and Hidden Markov Model (HMM) [14]. However, the accuracy was not satisfactory. Dynamic Time Warping (DTW) is another classic algorithm [15] for behaviour recognition. However, DTW only returns exact values and thus is inadequate for modelling the behaviour uncertainty and activity ambiguity. To handle the high-levels of uncertainty associated with the real-world environments, Fuzzy Logic Systems (FLSs) have been employed. Various linguistic summarization methods based on Type-1 FLSs (T1FLSs) have been proposed where [16], [17] employed T1FLSs for fall down detection. These type-1 fuzzy-based approaches perform well in predefined situations where the level of uncertainty is low. But these methods require multi-camera calibration which is inconvenient and time-consuming. In [18], [19] T1FLSs were used to analyse the input data from wearable devices to recognize the action. However, such wearable devices are intrusive and could be uncomfortable and inconvenient as the deployment of wearable devices is invasive for the skin and muscles of the users. T1FLS was used in [20], [21] to analyse the spatial and temporal features for efficient human behaviour recognition. However, there are intra- and inter- subject variations in behavioural characteristics which cause high levels of uncertainty. In [20], [22], [23], IT2FLS performed much better than T1FLS in human event detection and summarization.

The contribution of this paper is that we employed a 3D Kinect video camera and we proposed and developed a novel linguistic video summarisation system which is capable of robustly detecting and summarising the important events of several human subjects within an ambient assisted living environment. The proposed robust framework for behaviour recognition is based on interval type-2 fuzzy logic system whose membership functions and rule base were automatically constructed from the raw input data and were automatically optimised by the Big Bang Big Crunch (BB-BC) optimisation algorithm [24], [25]. Our system outperforms the traditional and type-1 fuzzy counterparts and was successfully deployed and used in real world environments.

The rest of the paper is organized as follows. Section II presents the overview of the hardware platform and the system Graphical User Interface (GUI). Section III introduces the proposed BB-BC based IT2FLS for the behaviour recognition and event linguistic summarization. Section IV presents the experiments and results. Finally conclusions are presented in section V.

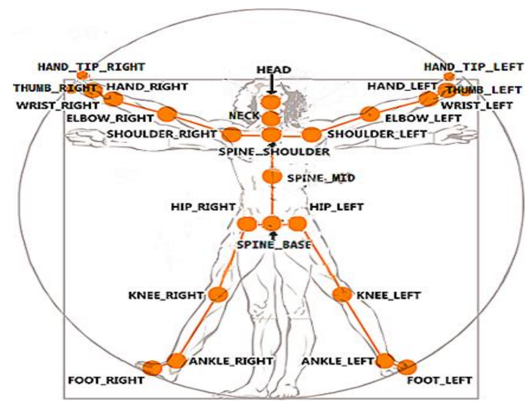
## II. Overview of the Employed Hardware Platform of RGB-D Sensor and the GUI of the Proposed System

The Kinect is one of the most popular RGB-D sensors in recent years. It has been applied in the fields of intelligent environments and robotics as an affordable but robust replacement for various types of wearable sensors, expensive distance sensors and conventional 2D cameras. It has been successfully used in various applications including object tracking and recognition [26] as well as 3D indoor mapping and human activity analysis [27]. We use Kinect v2 shown in Fig. 1a as the 3D sensor and its skeleton tracker to obtain the 3D data which describes the skeleton joints of the user in the scene shown in Fig. 1b.

Our system detects six behaviours which are crucial for AAL activities which are falling down, drinking/eating, walking, running, sitting and standing in ambient assisted living environment.

The GUI of the proposed system has two parts where the first

(a)




(b)

Fig. 1. Main RGB-D sensor and its 3D skeleton tracker (a) Kinect v2, (b) 3D Skeleton and joints of Kinect v2

part is used to continuously capture the 2D/3D sensing data and analyse behaviours of the human subjects in real-time. Since this event detection is connected to the back-end event database, once an activity is detected, the system will summarize the relevant details of an event and store it into the back-end server.

Meanwhile, if the detected event is an urgent emergency, a

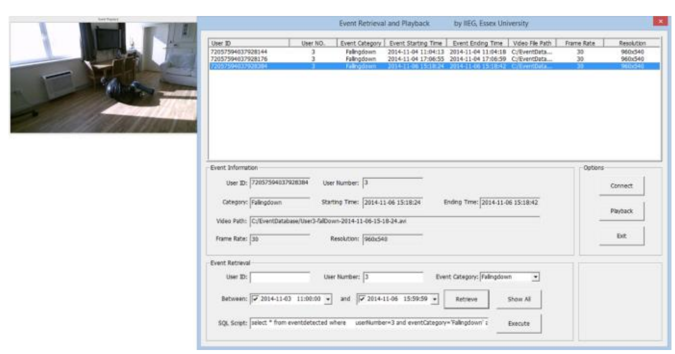


Fig. 2. The front-end GUI for the event search, linguistic summarisation and video

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



warning message will be sent to relevant caregivers so that instant action can be taken. The second part of the GUI is shown in Fig. 2 and it deals with the event retrieval, linguistic summarisation and playback. An example has been given in Fig. 2, where the user has selected searching the event category "Fallingdown" from the target behaviour list and inputted further refinement of the retrieval criteria, the particular subject number as well as a fixed time period to retrieve the relevant events from the back-end event database server. Finally, the retrieved events with details to the front-end GUI. The results of event retrieval are depicted in the list showing the relevant activities which have previously been detected and stored. The details of the selected event in the retrieval list is shown in the event information section, and the retrieved events can be used to play back the video matching the sequences the user wants to browse.

## III. The Proposed BB-BC based Interval Type-2 Fuzzy Logic System for the Event Detection and Linguistic Summarization of Video Monitoring

### *A. Overview of Type-2 Fuzzy Logic Systems*

The IT2FLS (shown in Fig. 3a) uses the interval type-2 fuzzy sets [28] (shown in Fig. 3b) to represent the inputs and/or outputs of the FLS. In the interval type-2 fuzzy sets all the third dimension values are equal to one [28], [29]. The use of interval type-2 FLS helps to simplify the computation of the type-2 FLS [29]. More information regarding the interval type-2 FLS and its benefits can be found in [28], [29], [30].

### *B. Overview of the Proposed System*

Fig. 4 provides an overview of our proposed system. There are two phases in the system which are the learning phase and the recognition phase. In the learning phase, the training data for each behaviour category are collected from the real-time Kinect data captured from the subjects in different circumstances and situations. Then behaviour feature vectors based on the distance and angle feature information are computed and extracted from collected Kinect data so as to model the motion characteristics. From the results of the features extraction, the type-1 fuzzy Membership Functions (T1MFs) of the fuzzy systems are then recognized/ known/discovered via Fuzzy C-Means Clustering (FCM) [31]. After that, the type-2 MFs are produced by using the obtained type-1 fuzzy sets as the principal membership functions which are then blurred by a certain percentage to create an initial Footprint of Uncertainty (FOU). Then, with the learned membership functions, the rule base of the type-2 fuzzy system is constructed automatically from the input feature vectors. Finally, a method based on the BB-BC algorithm is used to optimize the parameters of the IT2FLS which will be employed to recognize the behaviour and activity in the recognition phase.

It should be pointed that we generated initial fuzzy sets and rules for the FLSs which we then optimized via the BB-BC approach as such initial fuzzy sets and rules provided a good starting point for the BB-BC to converge fast to the optimal position. If we started from random fuzzy sets and rules, the BB-BC will take very long time to converge to optimal values.

During the recognition phase, the real-time Kinect data and HD video data are captured continuously by the RGB-D sensor monitoring the scene. From the real-time Kinect data, behaviour feature vectors are firstly extracted and used as input values for the IT2FLSs-based recognition system.

In our fuzzy system, each behaviour model is described by the corresponding rules, and each output degree represents the likelihood between the behaviour in the current frame and the trained behaviour model in the knowledge base. The candidate behaviour in the current frame is then classified and recognized by selecting the candidate model with the highest output degree. Once important events are detected by the optimized IT2FLS, linguistic summarization is performed using the key information such as the output action category, the starting time and ending time of the event, the user's number and

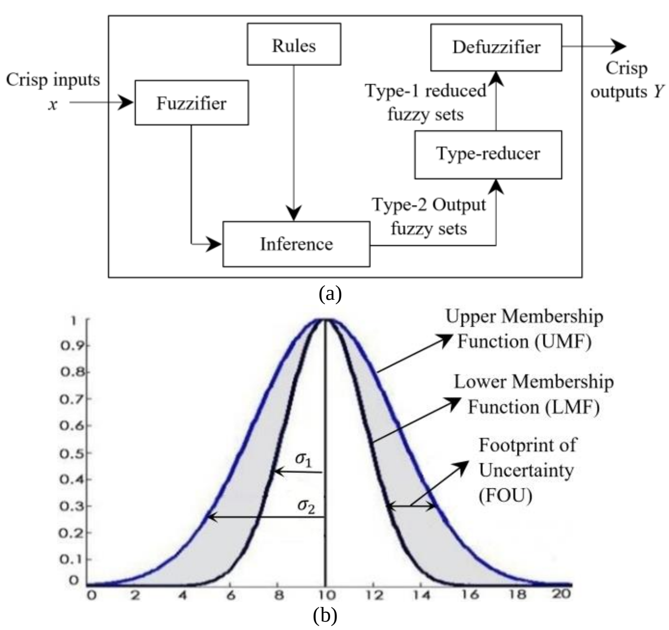



Fig. 3. (a) Structure of the type-2 FLS. (b) An interval type2 fuzzy set.

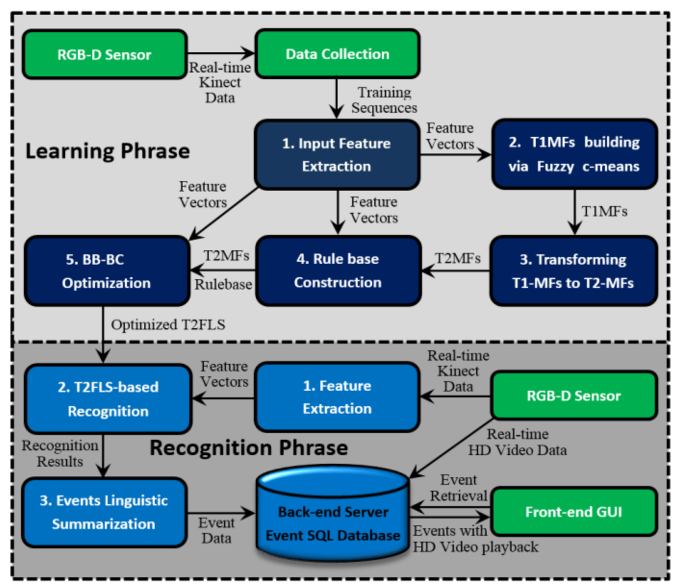



Fig. 4. Overview of our proposed system.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



identification, and the relevant HD video data and video descriptions. After that, the summarized event data is efficiently stored in our back-end server of event SQL database from where the users can access locally or remotely by using our front-end Graphical User Interface (GUI) system and perform event searching, retrieval and playback. The details of the employed phases are discussed in the following subsections.

### *C. Learning Phase*

#### *1) Fuzzy c-means*

The Fuzzy c-mean (FCM) algorithm developed by Dunn [32] and later improved by Bezdek [31] is an unsupervised clustering method to classify the unlabelled data by minimizing an objective function. In this paper, the FCM is used to compute the clusters of each feature to generate the type-1 fuzzy membership functions for the fuzzy-based recognition system.

#### *2) Feature Extraction*

- *Joint-angle Feature Representation*

For each frame, the skeleton is a sequence of graphs with *15* joints, where each node has its geometric position represented as a 3D point in a global Cartesian coordinate system. For any three different 3D points $P_1$, $P_2$, and $P_3$, an angle feature $\theta$ is defined by these three 3D joints $P_1$, $P_2$ and $P_3$ at a time instant. The angle $\theta$ is obtained by calculating the angle between the vectors $\overrightarrow{P_1P_2}$, and $\overrightarrow{P_2P_3}$ based on the following equation:

$$\theta = cos^{-1}\left(\frac{\overrightarrow{P_1P_2} \times \overrightarrow{P_2P_3}}{\left|\overrightarrow{P_1P_2}\right| \left|\overrightarrow{P_2P_3}\right|}\right) \quad (1)$$

- *Joint-position Feature Representation*

In order to model the local “depth appearance” for the joints, the joint positions are computed to represent the motion of the skeleton. For distance, between joint *i* and joint *j*, the arc-length distance is calculated:

$$D_{ij} = \left\|P_i - P_j\right\| \quad (2)$$

where $\|\cdot\|$ is the Euclidean norm.

- *Posture Representation*

To perform efficient behaviour recognition, an appropriate posture representation is essential to model the gesture characteristics. In this work, we use Kinect v2 to extract the 3D skeleton data which comprises 3D joints which are shown in Fig. 5. After that, based on the 3D joints obtained, we compute the posture feature using the joint vectors as shown in Fig. 5. In the applications of AAL environments, the main focus is to understand the users' daily activities and regular behaviours to create ambient context awareness such that ambient assisted services can be provided to the users in the living environments. Therefore, in our current application scenarios of ambient assisted living environments, we recognize and summarize the following behaviours: drinking/eating, sitting, standing, walking, running, and lying/falling down to provide different ambient assisted services. The purpose of choosing these behaviours as target categories is that they are common and important behaviours and activities in AAL environments. We detect fall-down event so that our system can send a warning message to the nearby caregivers or other people who can help [33] [34]. Our system summarises the frequency of the drinking activity to ensure that the user drinks enough water throughout the day to avoid dehydration [35] [36]. By a daily summarization of the sitting and lying duration and frequency, healthcare advice would be provided if the user remains inactive/active most of the time [6], [10]. The detection results of running demonstrate a potential emergency happening [37], [38]. From the detection results of standing and walking, our system obtains the location and trajectory of the subject so that services, such as wandering prevention, can be provided to dementia patients [8]. Also, the risk of falling down can be reduced by analysing the pattern of standing and walking [7]. Furthermore, cognitive rehabilitation services can be provided to help the elderly with dementia by summarizing this series of daily activities [11], [12]. Moreover, in intelligent environments, the electric appliances can be intelligently tuned and controlled according to the user's behaviour and activity to maximize their comfort and safety while minimizing the consumed energy. To achieve the robust recognition and summarization of behaviours in AAL environments, we use the angles and distance of the joint vectors as the input features which are highly relevant when modelling the target behaviours in AAL environments. The identified behaviours are extendable to enlarge the recognition range of the target behaviour by adding the needed joints.

As most behaviours in daily activity such as drinking, eating, waving hands, taking pills, etc., are related to the upper body, in this work in order to recognize behaviour and activity, we

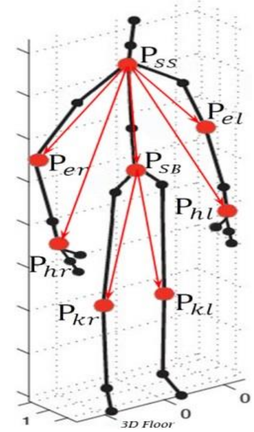



Fig. 5. 3D feature vectors based on the Kinect v2 skeleton model

focus on the following joints: spine base ($P_{sb}$), spine shoulder ($P_{ss}$), elbow left ($P_{el}$), hand left ($P_{hl}$), elbow right ($P_{er}$), hand right ($P_{hr}$). Since our algorithm is highly extendable, more joints can easily be added and utilized for more application scenarios. Based on the discussion above, the pose feature is obtained by calculating the joint-angle feature and joint-position feature of the selected joints, as given in the following procedure:

(1) Compute the vectors $\overrightarrow{P_{ss}P_{el}}$, $\overrightarrow{P_{ss}P_{hl}}$ modelling the left arm, and $\overrightarrow{P_{sc}P_{er}}$, $\overrightarrow{P_{sc}P_{er}}$ modelling the right arm.

(2) Angle features of the left arm $\theta_{al}$ can be obtained by calculating the angle between vectors $\overrightarrow{P_{ss}P_{el}}$, $\overrightarrow{P_{ss}P_{hl}}$ based on Equation (1). Similarly, angle features of the right arm $\theta_{ar}$ can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



be computed by applying the same process on $\overrightarrow{P_{ss}P_{er}}$, $\overrightarrow{P_{ss}P_{hr}}$.

(3) Based on Equation (2), position feature $D_{hl}$, $D_{hr}$ of the vectors $\overrightarrow{P_{ss}P_{hl}}$, $\overrightarrow{P_{ss}P_{hr}}$ can be obtained. In order to recognize activities, the status (3D position and angle) of the spine of the human subject is modelled in a way which is invariant to orientation and position, as shown below:

(4) Compute the vector $\overrightarrow{P_{ss}P_{sb}}$, modelling the entire spine of the subject, and $\overrightarrow{P_{ss}P_{kl}}$, $\overrightarrow{P_{ss}P_{kr}}$ modelling the left knee and right knee. Compute the angle $\theta_{kl}$ between $\overrightarrow{P_{ss}P_{sb}}$ and $\overrightarrow{P_{ss}P_{kl}}$ by using Equation (1). Similarly, the angle $\theta_{kr}$ can be obtained by applying Equation (1) on the vectors $\overrightarrow{P_{ss}P_{sb}}$ and $\overrightarrow{P_{ss}P_{kr}}$. Then, the bending angle $\theta_b$ of the body can be modeled, which is used mainly for analysing the sitting activity.

$$\theta_b = \max(\theta_{kl}, \theta_{kr}) \tag{3}$$

(5) In order to recognize the lying/falling down activity, we compute the distance $D_f$ between the 3D coordinates Spine Base $P_{sb}$ to the 3D Plane of the floor in the vertical direction.

(6) We compute the movement speed of the human by analysing $P_{sb}^{i-1}$ and $P_{sb}^{i}$ which are the positions of the joint $P_{sb}$ in two successive frame *i-1* and frame *i*. The speed $D_{sb}$ can be obtained by applying Equation (2) on $P_{sb}^{i-1}$ and $P_{sb}^{i}$. The movement speed $D_{sb}$ is mainly utilized for analyzing the common activities: falling down, sitting, standing, walking, and running.

For each tracked subject at a certain frame, the motion feature vector is obtained:

$$\mathrm{M} = (\theta_{al}, \theta_{ar}, D_{hl}, D_{hr}, \theta_b, D_f, D_{sb}) \tag{4}$$

For simplicity, we also denote each feature in M using the following format:

$$\mathrm{M} = (m_1, m_2, m_3, m_4, m_5, m_6, m_7) \tag{5}$$

As we can see, our system is a general framework for behaviour recognition which can be easily extended to recognize more behaviour types by adding more relevant joints into the feature calculation.

Each of these antecedents is represented by fuzzy sets which are Low, Medium, and High. The output of the fuzzy system is the behaviour possibility which is represented by two fuzzy sets which are Low and High. The type-1 fuzzy sets shown in Fig. 6 have been obtained via Fuzzy C Means (FCM)-based algorithm. Specifically, in our training dataset, we use the FCM-based algorithm to process all the data of each antecedent separately. In this FCM-based algorithm, the centres clustered by FCM will be used as the Means $m_k$ of our Gaussian membership functions $(m_k, \sigma_k, x) = \exp\left(-\frac{1}{2}\left(\frac{x-m_k}{\sigma_k}\right)\right)$, where $k = 1,\ldots, p$; $p$ is the number of antecedents. And $\sigma_k$ were obtained by $\sigma_k = (m_k - m_{k-1})/3$, where $k = 2,\ldots, p$; $p$ is the number of antecedents. And for $k = 1$, $\sigma_k = (m_{k+1} - m_k)/3$. The type-1 output membership functions were designed by expert knowledge.

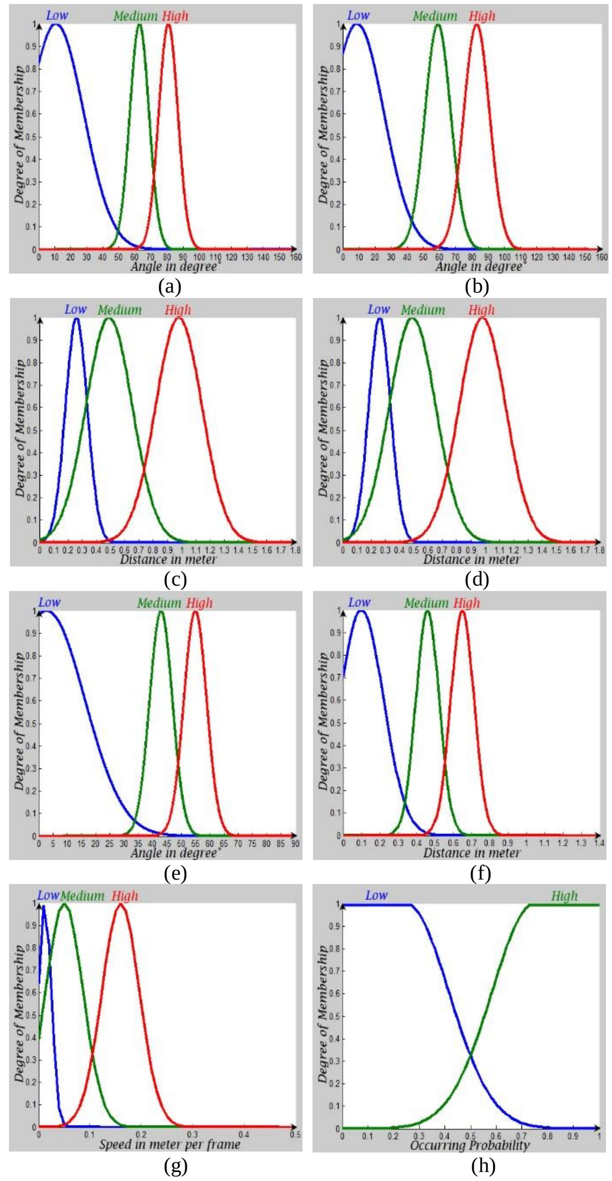



Fig. 6. Type-1 membership functions constructed by using FCM, (a) Type-1 MF for $m_1$ (b) Type-1 MF for $m_2$ (c) Type-1 MF for $m_3$ (d) Type-1 MF for $m_4$ (e) Type-1 MF for $m_5$ (f) Type-1 MF for $m_6$ (g) Type-1 MF for $m_7$(h) Type-1 MF for the Outputs

- *Occlusion problems and Tracking State Reliability*

For most available 3D motion capture devices in the market, the hardware system provides the level of the tracking reliability of the 3D joints. Kinect also returns to the tracking status to indicate if a 3D joint is tracked robustly, or inferred according to the neighbouring joints, or not-tracked when the joint is completely invisible. The 3D joints, which are occluded, belong to the inferred or not-tracked part. In our experiments, we found out that both inferred and not-tracked joints are unusually unreliable and will cause misclassifications. Thus, to solve the occlusion problem and increase the reliability, we only perform recognition when the tracking status of the essential parts related to our algorithm are tracked to avoid misclassifications.

*3) Transforming Type-1 Membership Functions to Interval Type-2 Membership Functions*

In this subsection, we present the initial design process of the IT2FLS which will be further optimized by the proposed BB-BC algorithm presented in the next subsection. Fig. 7. shows the type-1 fuzzy sets which were extracted via FCM as explained in subsection *C* in section III.

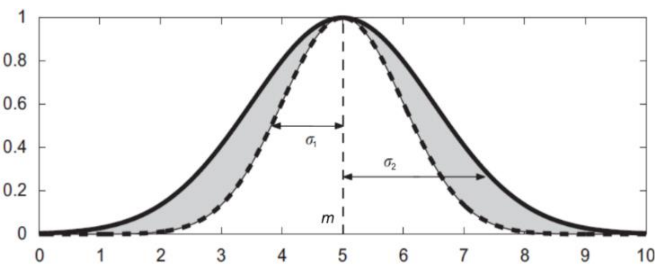



Fig. 7. Example of the type-2 fuzzy membership function of the Gaussian membership function with uncertain standard deviation σ. The shaded region is the Footprint of Uncertainty(FOU). The thick solid and dashed lines denote the lower and upper membership functions [28], [29]

In order to construct the initial type-2 MFs modelling the FOU, we transform the type-1 fuzzy sets to the interval type-2 fuzzy sets with certain mean (m) and uncertain standard deviation $\sigma$ $[\sigma_{k1}^l, \sigma_{k2}^l]$ [28], [29], i.e.,

$$\mu_k^l(x_k) = \exp[-\frac{1}{2}\left(\frac{x_k - m_k^l}{\sigma_k^l}\right)], \ \sigma_k^l \in [\sigma_{k1}^l, \sigma_{k2}^l] \qquad (6)$$

where *k = 1,…, p*; *p* is the number of antecedents; *l = 1,…, R*; *R* is the number of rules. The upper membership function of the type-2 fuzzy set can be written as follows:

$$\bar{\mu}_k^l(x_k) = N(m_k^l, \sigma_{k2}^l, x_k) \qquad (7)$$

The lower membership function can be written as follows:

$$\underline{\mu}_k^l(x_k) = N(m_k^l, \sigma_{k1}^l, x_k) \qquad (8)$$

where

$$N(m_k^l, \sigma_k^l, x_k) = \exp\left(-\frac{1}{2}\left(\frac{x_k - m_k^l}{\sigma_k^l}\right)\right) \qquad (9)$$

In order to construct the type-2 MFs in our IT2FLS, we use the standard deviation of the given type-1 fuzzy set (extracted by FCM clustering) to represent the $\sigma_{k1}^l$. $\sigma_{k2}^l$ is obtained by blurring $\sigma_{k1}^l$ with a certain $\alpha$% ($\alpha$ = *10, 20, 30, 40…*) such that

$$\sigma_{k2}^l = (1 + \alpha\%)\, \sigma_{k1}^l \qquad (10)$$

where $m_k^l$ is the same as the given type-1 fuzzy set. In order to allow for a fair comparison between the type-2 fuzzy logic system and type-1 fuzzy logic system, we used the same input features for the IT2FLS and the T1FLS.

*4) Initial Rule base construction from the raw data*

In this paper, we use an enhanced type-2 version of extended Wang-Mendel approach [28], [29], [39] to construct the initial rule base of the fuzzy system which will be further optimized by the proposed BB-BC algorithm presented in the next subsection.

*5) Optimizing the IT2FLS via BB-BC*

The main purpose of using FCM to generate the membership functions and using the Wang-Mendel method to construct the initial rule base before our BB-BC optimization is to obtain a good starting point in the search space, since the BB-BC quality of the optimization highly relies on the starting state to converge fast to the optimal position. If we started from random fuzzy sets and rules, the BB-BC will take very long time to converge to optimal values.

- *Big Bang-Big Crunch (BB-BC) Optimization*

The BB-BC optimization is an evolutionary approach which was presented by Erol and Eksin [24]. It is derived from one of the theories of the evolution of the universe in physics and astronomy, namely the BB-BC theory. The key advantages of BB-BC are its low computational cost, ease of implementation, and fast convergence. In [24], comparisons between BB-BC against Genetic Algorithm (GA) were performed. According to their comparison results, the performance of BB-BC exhibits superiority over an improved and enhanced genetic search algorithm. Furthermore, it was shown that the BB-BC outperforms the GA for many benchmark test functions and comparison experiments with a much faster convergence speed [24]. In [40], the BB-BC demonstrated better performance and outperformed the other optimisation algorithms such as genetic algorithms, evolution strategies algorithm, simulated annealing, tabu search, ant colony optimization, and harmony search. Similar comparison can be found in [41] which shows that BB-BC outperforms GA in their experiments. The reason for this fact is that, according to [24], GA suffers from premature convergence, convergence speed and execution time problems in global optimum searching as they are generally sluggish in reaching the global optimum accurately and reliably in a short period of time. By contrast, BB-BC avoids these drawbacks and finds an optimum point within the maximum number of allowed iterations. The steps followed in a BB-BC algorithm are as follows [25]:

Step 1: (Big Bang Phase): An initial generation of *N* candidates is randomly generated in the search space, similar to the other evolutionary search algorithms.

Step 2: The cost function values of all the candidate solutions are computed.

Step 3 (Big Crunch Phase): The Big Crunch phase comes as a convergence operator. Either the best fit individual or the centre of mass is computed. The centre of mass is calculated as:

$$x_c = \frac{\sum_{i=1}^{N} \frac{x_i}{f^i}}{\sum_{i=1}^{N} \frac{1}{f^i}} \qquad (11)$$

where $x_c$ is the position of the center of mass, $x_i$ is the position of the candidate, $f^i$ is the cost function value of the ith candidate, and *N* is the population size.

Step 4: New candidates are calculated around the new point calculated in Step 3 by adding or subtracting a random number whose value decreases as the iterations elapse:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



$$x^{new} = x_c + \frac{\gamma\rho(x_{max} - x_{min})}{k} \quad (12)$$

where $\gamma$ is a random number, $\rho$ is a parameter limiting search space, $x_{min}$ and $x_{max}$ are lower and upper limits, and $k$ is the iteration step.

Step 5: Return to Step 2 until stopping criteria have been met. Examples stopping criteria are: (1) current iteration number equals to the maximum iteration number; (2) the error (the difference between the current actual output and the expected output) is lower than a threshold value; (3) the accumulated running (consumed) time of the entire optimization procedure is larger than the given time for example five hours.

More details and examples regarding the application of the BB-BC technique can be found in [24], [40], [41]. To apply BB-BC, the first step is to determine what the input parameters are (i.e. which parameters in the system need to be tuned and optimized) and how to evaluate the quality and fitness of the achieved set of parameters. This allows determining the parameters in Equation (11). After that, the configuration of the BB-BC optimization can be determined by setting the search parameters such as the population size *N* and iteration count *k* to determine the search space. There is no restriction regarding these two parameters as the effect of *N* and *k* is only to determine the running time and search space. $x_{min}$ and $x_{max}$ can be determined by the application scenario. The parameter *r* is determined by a random number generated in each iteration. The parameter $\rho$ is a positive value which could be set to 1 in the first iteration and then decreased while the iteration count increases so that the search space can be narrowed down. Once the parameters are determined, the BB-BC can operate continuously until the stopping criteria are satisfied.

- *Optimizing the rule base of the IT2FLS with BB-BC*

To optimize the rule base of the IT2FLS, the parameters of the rule base are encoded into a form of a population. The IT2FLS rule base can be represented as shown in Fig. 8.

As showed in Fig. 8, $m_j^r$ are the antecedents and $o_k^r$ is the consequents of each rule respectively, where *j = 1,…, p*, *p* is the number of antecedents; *k = 1,…, q*, *q* is the number of behaviours; *r = 1,…, R*, and *R* is the number of the rules to be tuned. However, the values describing the rule base are discrete integers while the original BB-BC supports continuous values. Thus, instead of Equation (12), the following equation is used in the BB-BC paradigm to round off the continuous values to the nearest discrete integer values modelling the indexes of the fuzzy set of the antecedents or consequents.

$$D^{new} = D_c + round\left[\frac{\gamma\rho(D_{max} - D_{min})}{k}\right] \quad (13)$$

where $D_c$ is the fittest individual, r is a random number, $\rho$ is a parameter limiting search space, $D_{min}$ and $D_{max}$ are lower and upper bounds, and k is the iteration step.

In this study, the rule base constructed by the Wang-Mendel approach [28], [29], [39] is used as the initial generation of candidates. After that, the rule base can be tuned by BB-BC using the cost function depicted in Equation (14).

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m_1^1$ | … | $m_p^1$ | $o_1^1$ | … | $o_q^1$ | … | $m_1^R$ | … | $m_p^R$ | $o_1^R$ | … | $o_q^R$ |

Fig. 8. The population representation for the parameters of the rule base

- *Optimizing the Type-2 membership functions with BB-BC*

In order to apply BB-BC, the feature parameters of the type-2 membership function have to be encoded into a form of a population. As depicted in Equation (10), in order to construct the type-2 MFs, the parameter $\alpha$ has to be determined to obtain $\sigma_{k2}^l$ while $\sigma_{k1}^l$ is provided by FCM. To be more accurate, the uncertainty factors $\alpha_k^j$ for each fuzzy set of the MFs are computed, where *k = 1,…, p*, *p* is the number of antecedents; *j = 1,…, q*, *q* is the number of input features. For illustration purposes, as in the MFs of the proposed system, three type-2 fuzzy sets including LOW, MEDIUM and HIGH are utilized for modelling each of the 7 features, therefore, the total number of the parameters for the input type-2 MFs is 3×7=21. In a similar manner, parameters for the output MFs are also encoded; these are $\alpha_L^{Out}$ for the linguistic variable LOW and $\alpha_H^{Out}$ for the linguistic variable HIGH of the output MF. Therefore, the structure of the population is built as displayed in Fig. 9.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $\alpha_1^1$ | $\alpha_2^1$ | $\alpha_3^1$ | … | $\alpha_1^7$ | $\alpha_2^7$ | $\alpha_3^7$ | $\alpha_L^{Out}$ | $\alpha_H^{Out}$ |

Fig. 9. The population representation for the parameters of type-2 MFs

The optimization problem is a minimization task, and with the parameters of the MFs encoded as showed in Fig. 9 and the constructed rule base, the recognition error in our solutions space can be minimized by using the following function as the cost function.

$$f^i = (1 - Accuracy^i) \quad (14)$$

where $f^i$ is the cost function value of the *i*th candidate and $Accuracy^i$ is the scaled recognition accuracy of the *i*th candidate. The new candidates are generated using Equation (12).

## D. Recognition Phase

In our fuzzy system, the antecedents are $m_1$, $m_2$, $m_3$, $m_4$, $m_5$, $m_6$, $m_7$ and each of these antecedents is modelled by three fuzzy sets: LOW, MEDIUM, and HIGH. The output of the fuzzy system is the behaviour possibility which is modelled by two fuzzy sets: LOW and HIGH. The type-1 fuzzy sets shown in Fig. 6 have been obtained via FCM and the rules are the same as the IT2FLS.

When the system operates in real time, we measure {$m_1$, $m_2$, ..., $m_7$} on the current frame and the IT2FLC is supposed to provide the possibilities of the candidate behaviour classes: drinking/eating, sitting, standing, walking, running, and lying/falling down. In our system, each activity category utilizes the same output membership function as depicted in Fig. 6h, and product t-norm is employed while the centre of sets type-reduction for IT2FLS is used (for the compared type-1 FLS the centre of sets defuzzification is used). To recognize the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



current behaviour, our system works in the following pattern:

The Kinect v2 is continuously capturing the raw 3D skeleton data from the subjects in the real-world intelligent environment.

Then the raw real-time 3D Kinect data will be analysed by our feature extraction module to get the feature vector M = ($m_1$, $m_2$, $m_3$, $m_4$, $m_5$, $m_6$, $m_7$) modelling the behaviour characteristics in the current frame.

For the crisp input vector *M*, a type-2 singleton fuzzifier will be used to fuzzify the crisp input and obtain the upper $\overline{\mu}_{\tilde{F}_1^i}(x')$and lower ($\underline{\mu}_{\tilde{F}_1^i}(x')$) membership values.

After that, we compute the firing strength $\underline{f}^i$and $\overline{f}^i$ of each rule, where $i = 1, \dots, R$, and $R$ is the number of rules, where $\underline{f}^i(x') = \underline{\mu}_{\tilde{F}_1^i}(x'_1) * \dots * \underline{\mu}_{\tilde{F}_p^i}(x'_p)$ and $\overline{f}^i(x') = \overline{\mu}_{\tilde{F}_1^i}(x'_1) * \dots * \overline{\mu}_{\tilde{F}_1^i}(x'_p)$.

The type reduction is carried out by using the Karnik-Mendel (KM) approach [29] to compute the type reduced set defined by the interval [$y_{lk}$, $y_{rk}$]. The reason for using KM algorithm is that KM approach is the standard method and most accurate algorithm for type-reduction in in interval type-2 fuzzy systems [29].

In the end, defuzzification is computed as $\frac{y_{lk}+y_{rk}}{2}$ to calculate the output degree of the target behaviour class. For one input feature vector analysed by our fuzzy system, we will have one output degree per candidate activity class, which models the possibility of the candidate activity class occurring in the current frame.

In our application scenario within AAL spaces, we aim at recognizing the daily regular activities. However, the subject's activity sequence happening in the actual environment is not a continuous time-series due to the occlusion problems, capturing angle, and the casualness of the subject which could lead to untargeted and unknown behaviours out of our concern range. To solve this problem, we are not using shoulder functions in our membership functions since the target behaviours are only modelled by the feature values ranging in the sections returned by FCM learned from the feature data of the concerned activities. Additionally, we will check if the behaviour candidate is confident in the current frame by checking if its associated output degree is higher than a confidence threshold t, where in our experiment we set *t* = *0.62*. The confident behaviour candidates will be further considered to get the final recognition output.

In our application scenario, some of the target behaviour categories are conflicting as it is impossible for them to be happening at the same moment. Therefore, in our experiment, we divide the target behaviour categories into several conflicting groups, i.e. sitting, standing, walking, running, and lying/falling down as a group while drinking/eating is another group.

In the final step, the behaviour recognition is performed by choosing the confident candidate behaviour category with the highest output degree as the recognized behaviour class in its behaviour group. For example, if the outputs of sitting, standing, walking, running, and lying/falling down are *0.25*, *0.75*, *0.64*, *0.0*, *0.0* and the output of drinking/eating is *0.25*, then the final recognition result would be standing since its output degree is the highest among the confident candidates (which are standing and walking in this case) in the its group and the output degree of drinking/eating in the other group is lower than a confident level. However, in a very rare situation, if two confident candidate categories in a conflicting group are allocated with a same output degree, this demonstrates that the two candidates have extremely high behavioural similarity and cannot be distinguished in the current frame, our system ignores these two candidate categories in the behaviour recognition of the current frame.

## IV. Experiments and Results

In our application scenarios, we aim at recognizing the following behaviours: drinking/eating, sitting, standing, walking, running, and lying/falling down. Our experiments were performed in different places such as the intelligent apartment (iSpace) [42] and intelligent Classroom iClassroom [22] [43] at the University of Essex. We tested our methods including Type-1 Fuzzy Logic System (T1FLS) and Type-2 Fuzzy Logic System (T2FLS) compared against the non-fuzzy traditional methods including Hidden Markov Models (HMM) and Dynamic Time Warping (DTW) on *22* subjects ensuring high-levels of intra- and inter- subject variation and ambiguity in behavioural characteristics.

In the training stage, the training data were captured from different subjects where the subjects were asked to perform each target behaviour on average two to three times. This resulted in around *220* activity samples for training. In the real-world recognition stage, we divided the subjects into different groups and we performed the experiments with different subject numbers in a scene to model different uncertainty complexity.

The experiments were conducted on average with five repetitions per target behaviour by each subject in the group analysed by the real-time behaviour recognition system. This resulted in around *1,740* activity samples for testing. To perform a fair comparison, all the methods share the same input features. As in real-world environments, occlusion problems exist in our test cases leading to behavioural uncertainty caused by the occlusions of the subjects. The experiments were conducted with different subjects and different scenes in various circumstances including different illumination strength, partial occlusions, daytime and night time, moving camera, fixed camera, different monitoring angles, etc. The experiment results demonstrate that our algorithm is robust and effective in handling the high levels of uncertainties associated with real-world environments including occlusion problems, behaviour uncertainty, activity ambiguity, and uncertain factors such as position, orientation and speed, etc. The type-2 membership functions used in our system, which are constructed and optimized by BB-BC, are shown in Fig. 10.

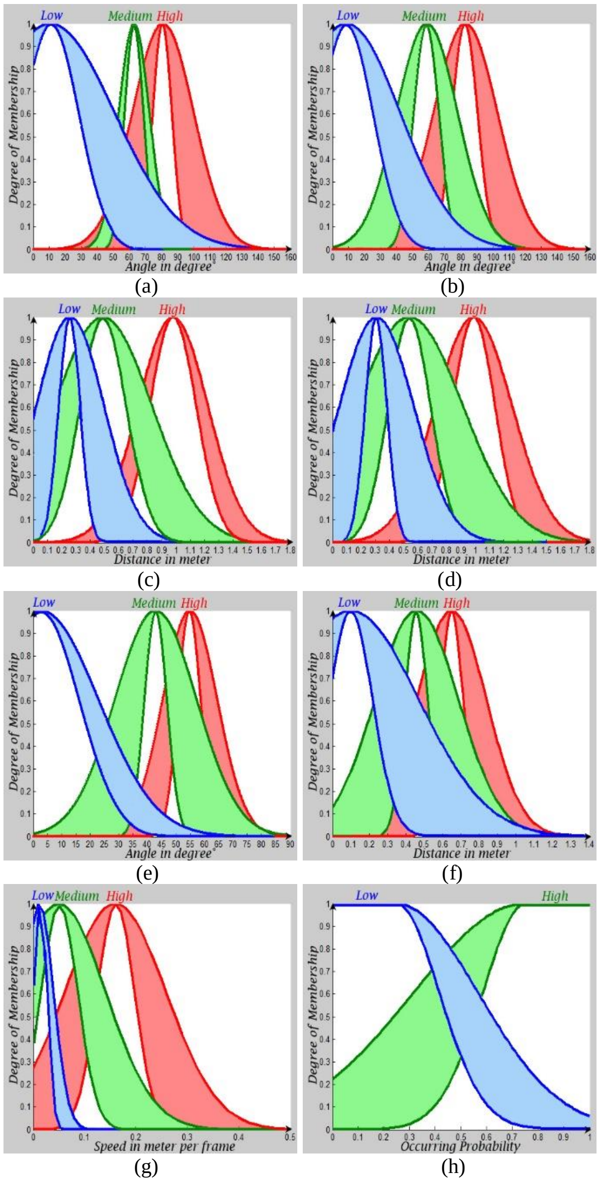



Fig. 10. Type-2 membership functions optimized by using BB-BC, (a) Type-2 MF for $m_1$ (b) Type-2 MF for $m_2$ (c) Type-2 MF for $m_3$ (d) Type-2 MF for $m_4$ (e) Type-2 MF for $m_5$ (f) Type-2 MF for $m_6$ (g) Type-2 MF for $m_7$ (h) Type-2 MF for Output

Our experiment result demonstrates that the BB-BC optimization improves the performance of our type-2 fuzzy logic system. In the BB-BC optimization procedure of the type-2 membership functions, we set $x_{min}$ and $x_{max}$ to *50*% and *300*%, which influences the FOU blurring factor $\alpha$ in type-2 MFs construction. In order to achieve robust recognition performance, in our experiment, the population size N of BB-BC is set to *200,000*.

Based on the optimized type-2 fuzzy sets and rule base by utilizing BB-BC, our IT2FLSs-based system outperforms the counterpart T1FLSs-based recognition system, as shown in Table I, where the type-2 system achieves *5.29*% higher average per-frame accuracy over the test data in the recognition phrase than the type-1 system. Our type-2 fuzzy logic system also outperforms the traditional non-fuzzy based recognition methods based on Hidden Markov Models (HMM) [14] and Dynamic Time Warping (DTW) [15]. In order to conduct a fair comparison with the traditional HMM-based and DTW-based methods, all the methods share the same input features. As shown in Table I, our IT2FLSs-based method with BB-BC optimization achieves *15.65*% higher recognition average accuracy than the HMM-based algorithm, and *11.62*% higher recognition average accuracy than the DTW-based algorithm. For the standard deviation of each subject's recognition accuracy, the T2FLS-based method is the lowest, demonstrating the stableness and robustness of our method when testing on different subjects.

When the number of subjects increases which leads to a higher possibility of occlusion problems with a higher-level of behaviours uncertainty, the difference between our method compared to the T1FLS-based method and the traditional non-fuzzy methods is even higher, as shown in Table II, Table III and Table IV. Our T2FLS-based method remains the most robust algorithm with the highest recognition accuracy which remains roughly the same with adding more users to the scene.

Due to the limitations of the field-of-view (70°×60°) and sensing distance (*5* meters) of the hardware platform of the Kinect v2, according to our experiments, we found out that the reliability of the sensing data 3D skeleton will degrade if the user is around the boundary of the field-of-view or is around the sensing distance range. Furthermore, the reliability and quality of skeleton data will degrade if there are occlusion problems caused by the crowded users. This problem generates high-levels of uncertainties in the real-world application scenario since the status of some of the 3D joints of the skeleton emain "tracked" rather than "inferred" when the human subject is not within the effective and robust sensing range. This problem is caused by the limitations of the hardware and the software package of the Kinect v2. Moreover, higher-level of uncertainties occur if the crowdedness of the users increases and the human subjects are acting freely. In our real-world experiments, the human subjects acted freely which caused occlusion problems resulting from the crowdedness of the users and objects such as tables, chairs, sofas, TV, etc. Therefore, the high-level of uncertainties significantly increase the difficulty of recognising behaviours in the real-world AAL environment. In order to test the system ability to handle the occlusion problems, we have performed particular occlusion experiments in which the human subjects were heavily occluded by the obstacles such as chairs, tables, sofas, or other human subjects. These noise factors have decreased the quality of the extracted 3D skeleton and have increased the uncertainties in behaviour representation and recognition. As shown in Table IV, our IT2FLSs-based system remains robust and maintains its accuracy (which shows the ability to handle uncertainties) while outperforming the other methods and achieving *8.94*%, *24.02*%, and *27.12*% higher accuracy than the T1FLSs, DTW, and HMM based methods respectively.

Based on the recognition results of our optimized IT2FLS, higher-level applications including video linguistic summarizations, event searching and retrieval, event playback, and human-machine interactions have been developed and deployed in iSpace and iClassroom.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



TABLE I
COMPARISON OF FUZZY-BASED METHODS AGAINST TRADITIONAL METHODS WITH ONE SUBJECT PER GROUP IN A SCENE (FIFTEEN GROUPS)

| Method | Average Accuracy | Standard Deviation |
|---|---|---|
| HMM | 70.9266% | 0.175258 |
| DTW | 74.9614% | 0.129266 |
| T1FLS | 81.2903% | 0.110410 |
| T2FLS | 86.5798% | 0.086551 |

TABLE II.
COMPARISON OF FUZZY-BASED METHODS AGAINST TRADITIONAL METHODS WITH TWO SUBJECTS PER GROUP IN A SCENE (SIX GROUPS)

| Method | Average Accuracy | Standard Deviation |
|---|---|---|
| HMM | 72.4134% | 0.078800 |
| DTW | 71.6549% | 0.051693 |
| T1FLS | 79.0394% | 0.157738 |
| T2FLS | 85.8864% | 0.092471 |

TABLE III.
COMPARISON OF FUZZY-BASED METHODS AGAINST TRADITIONAL METHODS WITH THREE SUBJECTS PER GROUP IN A SCENE (FIVE GROUPS)

| Method | Average Accuracy | Standard Deviation |
|---|---|---|
| HMM | 70.1782% | 0.042738 |
| DTW | 73.7452% | 0.103744 |
| T1FLC | 78.3855% | 0.128380 |
| T2FLC | 86.1305% | 0.082625 |

TABLE IV.
COMPARISON OF FUZZY-BASED METHODS AGAINST TRADITIONAL METHODS WITH FOUR SUBJECTS PER GROUP IN A SCENE (THREE GROUPS)

| Method | Average Accuracy | Standard Deviation |
|---|---|---|
| HMM | 69.5274% | 0.083920 |
| DTW | 70.1220% | 0.112780 |
| T1FLC | 76.6017% | 0.080618 |
| T2FLC | 84.7253% | 0.072113 |

TABLE V.
COMPARISON OF FUZZY-BASED METHODS AGAINST TRADITIONAL METHODS IN OCCLUSION PROBLEMS HANDLING

| Method | Average Accuracy | Standard Deviation |
|---|---|---|
| HMM | 60.5395% | 0.220914 |
| DTW | 63.6414% | 0.142105 |
| T1FLC | 78.7150% | 0.105513 |
| T2FLC | 87.6638% | 0.113804 |

The results of detected events and the associated video data are stored in the SQL Event database server so that further data mining can be performed by using our event summarization and retrieval software. Also, the user can easily summarize the event of interest at the given time frame and play them back. Fig. 11 provides the detection results of the real-time event detection system deployed in different real-world intelligent environments such as iClassroom and iSpace at the University of Essex.

TABLE VI
REULTS FROM THE TUKEY TEST AS COMPUTED BY THE SPSS TOOL.

| | (I) VAR00004 | (J) VAR00004 | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|---|
| Tukey HSD | 1 | 2 | -.02108 | .02136 | .759 | -.0822 | .0400 |
| | | 3 | -.10089* | .02136 | .001 | -.1620 | -.0398 |
| | | 4 | -.17480* | .02136 | .000 | -.2359 | -.1137 |
| | 2 | 1 | .02108 | .02136 | .759 | -.0400 | .0822 |
| | | 3 | -.07981* | .02136 | .009 | -.1409 | -.0187 |
| | | 4 | -.15372* | .02136 | .000 | -.2148 | -.0926 |
| | 3 | 1 | .10089* | .02136 | .001 | .0398 | .1620 |
| | | 2 | .07981* | .02136 | .009 | .0187 | .1409 |
| | | 4 | -.07391* | .02136 | .015 | -.1350 | -.0128 |
| | 4 | 1 | .17480* | .02136 | .000 | .1137 | .2359 |
| | | 2 | .15372* | .02136 | .000 | .0926 | .2148 |
| | | 3 | .07391* | .02136 | .015 | .0128 | .1350 |

*The mean difference is significant at the 0.05 level

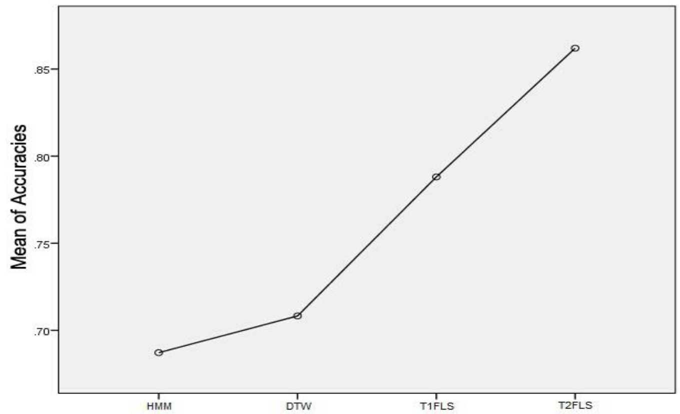



Fig.11 Plot for group means comparison as extracted from the Tukey test within SPSS.

We have conducted statistical analysis using the Analysis of Variance (ANOVA) methodology inorder to statistically verify our results and detect if there is statistically relevant difference among the compared techniques. The Four Groups involved in the statistical analysis were Group1 (HMM technique), Group 2 (FTW technique), Group 3 (T1FLS technique) and Group 4 (T2FLS technique). The resulting *p-value* was 0.00015 which is lower than the *level of significance* $\alpha$= 0.05, which means that we can reject the null hypothesis and affirm that there exist statistical differences between the multiple distributions associated with each of the compared techniques. We have then applied the post-hoc Tukey test (using the SPSS software tool) where we conducted first the Levene's test which is needed for the Tukey test to establish the equality of variances. After passing the Levene's test, we conducted the Tukey test whose results are reported in Table VI. Fig.11 displays the means of the accuracies of the four compared approaches as reported via SPSS. As can be seen through Table VI, there is no statistically significant difference between Group 1 (HMM technique) and Group 2 (DTW technique) as the *sig* value (which is 0.759) was bigger than 0.05. However, it can be seen that Group 4 (T2FLS) is significantly statistically different from the other compared techniques. This statistical difference can be compounded with the T2FLS based technique achieving the best mean accuracy reported in Fig.11 to confirm the superiority of the T2FLS based technique as reported in the abovementioned comparisons.

The results of detected events and the associated video data are stored in the SQL Event database server so that further data mining can be performed by using our event summarization and retrieval software. Also, the user can easily summarize the event of interest at the given time frame and play them back

Fig. 12 provides the detection results of the real-time event detection system deployed in different real-world intelligent environments such as iClassroom and iSpace at the University of Essex. The number of subjects changes according to the application scenario. In Fig 12a, two students are using our immersive learning platform [43] in iClassroom with one Kinect v2. In Fig 12b, the system analysed the activity of the three subjects in the scene in the iClassroom. In Fig 12c, behaviours recognition is performed in the iSpace with four

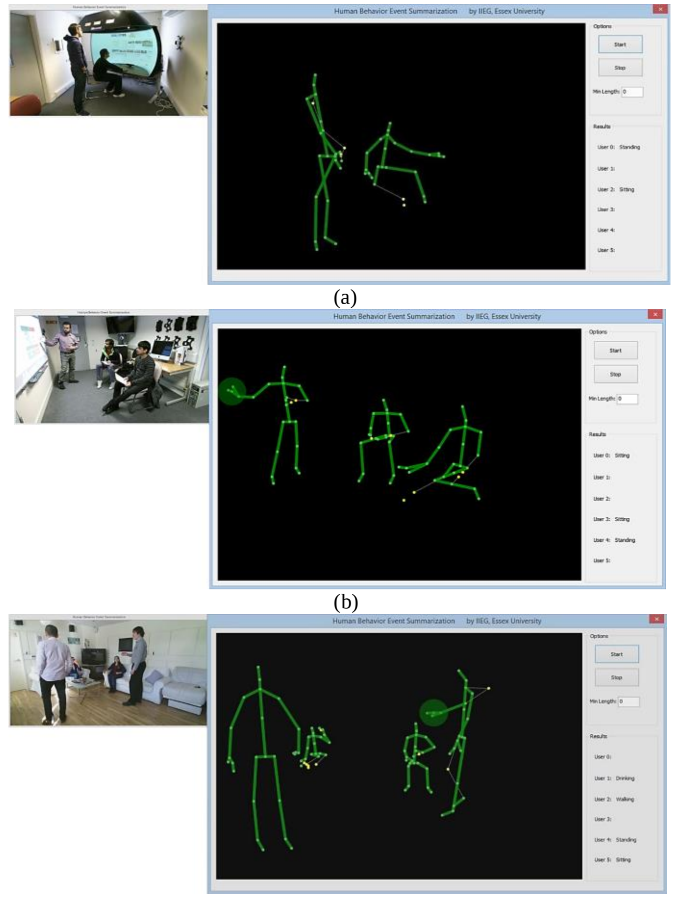
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



subjects. As the scenario is in the living environment, the users have more freedom to act casually and the occlusion problems are more likely to happen with a large crowd of subjects, these factors lead to higher-levels of uncertainty.

As can be seen, the user *1* who is drinking coffee is heavily occluded by the table in front, as well as the user 2 who is walking towards the door. Our IT2FLS-based recognition system handles the high-levels of uncertainty robustly and returns the correct results.

As shown in Fig. 13a, to retrieve the interesting events of a certain subject conducted during a fixed time period, we inputted a subject number and time duration, and performed event retrieval via the front-end GUI. After that, the relevant retrieved events were shown in the result list, from where we selected the retrieved event and played back the HD video. Similarly, in Fig 13b, we were interested in the drinking activities that happened in the iSpace. Therefore, we selected the “Drinking” activity from the event category and also provided a certain time period. Then, the events associated with “Drinking” during the given time period were retrieved and shown in the result list for the user to play back.



(a) (b) (c)

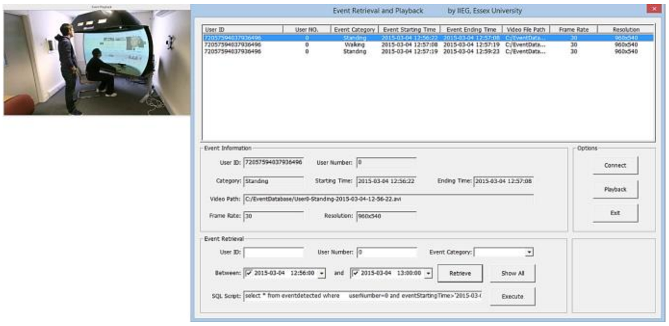
Fig. 12. Detection results from our real-time IT2FLS-based recognition system, (a) recognition results in the iClassroom with two subjects in the scene (b) recognition results in the iClassroom with three subjects in the scene (c) recognition results in the iSpace with four subjects in the scene leading to occlusion problems and high-levels of uncertainty

## V. CONCLUSIONS

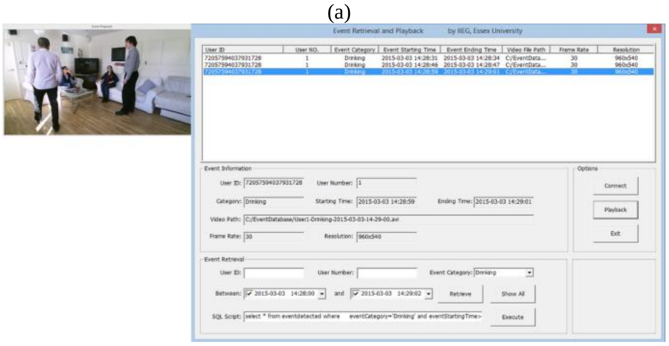
To construct real world AAL environments, there is a need to develop intelligent systems which are capable of realising context awareness regarding the activities and behaviours of the human users in AAL such that particular assisted or healthcare services can be provided to the users. 3D vision techniques can provide vital and accurate context awareness information for AAL by modelling the behaviour characteristic with massive data. However, high-levels of uncertainties caused by the behaviour uncertainty, activity ambiguity and noise factors associated with the real-world environments exist in the captured 3D data. In this paper, we introduced a framework for behaviour recognition and event linguistic summarization utilizing a RGB-D sensor Kinect v2 based on BB-BC optimised Interval Type-2 Fuzzy Logic Systems (IT2FLSs) for AAL real world environments. We have shown that the proposed system is capable of handling high-levels of uncertainties caused occlusions, behaviour ambiguity and environmental factors. Our proposed system has been successfully deployed in real world environments occupied with various users ensuring high-levels of intra- and inter- subject behavioural uncertainty. Our results demonstrated that the BB-BC based optimization paradigm is effective in tuning and optimizing the parameters of our fuzzy system. In addition, our experiment results with single users show that the proposed IT2FLS handles the high-levels of uncertainties well and achieves robust recognition of *86.57*% and outperformed the T1FLS counterpart by an enhancement of *5.28*% as well as other traditional non-fuzzy systems including the HMM-based system and DTW-based method by *15.65*% and *11.61*%, respectively. Moreover, it was shown that the proposed IT2FLS delivers consistent and robust recognition accuracy while the T1FLS and other conventional methods based on HMM and DTW shows degradations in recognition accuracy when increasing the number of users.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems





(a)



(b)

Fig. 13. Event retrieval and playback, (a) Event retrieval and playback with a given subject number and time period in iClassroom (b) Event retrieval and playback with a given event category and time period in iSpace

## REFERENCES

[1] World Health Organization, “10 facts on ageing and the life course,” http://www.who.int/features/factfiles/ageing/en/index.html, [Online], reviewed October 2014.

[2]H. Hagras, “Embedding Computational Intelligence in Pervasive Spaces,” IEEE Pervasive Computing, vol.6, no.3, pp. 85-89, May/June 2007.

[3] H. Hagras, “Employing Computational Intelligence to Generate More Intelligent and Energy Efficient Living Spaces,” The International Journal of Automation and Computing, Volume 5, No. 1, pp. 1-9, January 2008.

[4] V. Callaghan, G. Clarke, M. Colley and H. Hagras, “Embedding Intelligence: Research Issues for Ubiquitous Computing,” Proceedings of the Ubiquitous Computing in Domestic Environments Conference, pp. 13-14, Nottingham, UK, September 2001.

[5]T. Vos, “The Video Surveillance Industry – Market size,” online available at Cameramanager.com

[6] M. Nambu, K. Nakajima, M. Noshiro, and T. Tamura, “An algorithm for the automatic detection of health conditions,” Engineering in Medicine and Biology Magazine, IEEE, vol. 24, no. 4, pp. 38–42, 2005.

[7] W. Wu, L. Au, B. Jordan, T. Stathopoulos, M. Batalin, W. Kaiser, A. Vahdatpour, M. Sarrafzadeh, M. Fang, and J. Chodosh, “The smartcane system: An assistive device for geriatrics”, Proceedings of 3rd international conference on Body area networks, 2008.

[8] J. Wan, C. Byrne, G. O’Hare, and M. O’Grady, “Orange alerts: Lessons from an outdoor case study,” Proceedings of 5th International Conference on Pervasive Computing Technologies for Healthcare, IEEE, pp. 446–451, 2011.

[9] C. Lin, M. Chiu, C. Hsiao, R. Lee, and Y. Tsai, “Wireless health care service system for elderly with dementia,” IEEE Transactions on Information Technology in Biomedicine, vol. 10, no. 4, pp. 696–704, 2006.

[10] N. Barnes, N. Edwards, D. Rose, and P. Garner, “Lifestyle monitoring technology for supported independence,” Computing & Control Engineering Journal, vol. 9, pp. 169–174, Aug. 1998.

[11] J. Hoey, K. Zutis, V. Leuty amd A. Mihailidis, “A tool to promote prolonged engagement in art therapy: design and development from arts therapist requirements,” Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility, pp. 211-218, 2010.

[12] R. Levinson, “The Planning and Execution Assistant and Trainer (PEAT),” The Journal of head trauma rehabilitation, vol. 12, no. 3–4, pp. 769–775, 1997.

[13] X. Ji, H. Liu, Y. Li, “Viewpoint insensitive actions recognition using hidden conditional random fields,” Knowledge-Based and Intelligent Information and Engineering Systems, Springer Berlin Heidelberg, pp.369-378, 2010.

[14] E. Kim, H. Sumi, and C. Diane, “Human activity recognition and pattern discovery,” Pervasive Computing, vol.9, no.1, pp.48-53, 2010.

[15] M. Reyes, D. Gabriel, and E. Sergio, “Feature weighting in dynamic time warping for gesture recognition in depth data,” Proceedings of IEEE International Conference on Computer Vision Workshops, pp. 1182-1188, 2011.

[16] D. Anderson, R. Luke, J. Keller, M. Skubic, M. Rantz, and M. Aud, “Linguistic summarization of video for fall detection using voxel person and fuzzy logic,” Computer Vision Image Understanding, vol. 113, pp. 80–89, 2008.

[17] D. Anderson, R. Luke, J. Keller, M. Skubic, M. Rantz, and M. Aud, “Modeling human activity from voxel person using fuzzy logic,” IEEE Transactions on Fuzzy Systems, vol. 17, no. 1, pp. 39-49, 2009.

[18] O. Lara, and A. Miguel, “A survey on human activity recognition using wearable sensors,” Communications Surveys and Tutorials, IEEE, vol.15, no.3, pp.1192-1209, 2013.

[19] G. Trivino, and A. van der Heide, “Linguistic summarization of the human activity using skin conductivity and accelerometers,” Proceedings of 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU), pp. 1583-1589, 2008.

[20] B. Yao, H. Hagras, M. Alhaddad, D. Alghazzawi, “A fuzzy logic-based system for the automation of human behavior recognition using machine vision in intelligent environments,” Soft Computing, pp.1–8, 2014.

[21] J. Yaouanc, J. Poli, “A fuzzy spatio-temporal-based approach for activity recognition,” Advances in Conceptual Modeling, pp. 314-323, 2012.

[22] K. Almohammadi, B. Yao, and H. Hagras, “An interval type-2 fuzzy logic based system with user engagement feedback for customized knowledge delivery within intelligent E-learning platforms,” Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 808-817, 2014.

[23] B. Yao, H. Hagras, D. Ghazzawi, and M. Alhaddad, “An interval type-2 fuzzy logic system for human silhouette extraction in dynamic environments,” Autonomous and Intelligent Systems, Springer, 2012, pp. 126–134.

[24] O. Erol and I. Eksin, “A new optimization method: big bang–big crunch,” Advances in Engineering Software, vol.37, no. 2, pp. 106-111, 2006.

[25] T. Kumbasar, I. Eksin, M. Guzelkaya and E. Yesil, “Adaptive fuzzy model based inverse controller design using BB-BC optimization algorithm,” Expert Systems with Applications, vol. 38, no. 10, pp.12356-12364, 2011.

[26] J. Han, L. Shao, D. Xu and J. Shotton, “Enhanced computer vision with microsoft kinect sensor: A review,” IEEE Transactions on Cybernetics, vol. 43, no. 5, pp. 1318-1334, 2013.

[27] B. Yao, H. Hagras, D. Ghazzawi, and M. Alhaddad, “A Type-2 Fuzzy Logic based system for linguistic summarization of video monitoring in indoor intelligent environments,” Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 825-833, 2014.

[28] H. Hagras, “A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots,” IEEE Transactions on Fuzzy Systems, vol. 12, no. 4, pp.524-539, 2004.

[29] J. Mendel, “Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions,” Upper Saddle River, NJ: Prentice-Hall, 2001.

[30] C. Lynch. H. Hagras, V. Callaghan, “Embedded Interval Type-2 Neuro-Fuzzy Speed Controller for Marine Diesel Engines”, Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2006), pp. 1340-1347, Paris, France, July 2006.

[31] N. Pal and J. Bezdek, “On cluster validity for the fuzzy c-means model,” IEEE Transaction on Fuzzy Systems, vol. 3, pp. 370-379, 1995.

[32] J. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact, well separated cluster,” Cybernetics, vol. 3, no. 3, pp. 32–57, 1973.

[33] OP. Popoola, and K Wang, “Video-based abnormal human behavior recognition—A review” IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 42, no. 6, pp. 865-878, 2012.

[34] H. Nait-Charif, SJ. McKenna, “Activity summarisation and fall detection in a supportive home environment,” Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 4, IEEE, 2004.

[35] JS. Tham, YC. Chang, A. Fauzi, and M. Faizal, “Automatic identification of drinking activities at home using depth data from RGB-D camera,” 2014 International Conference on Control, Automation and Information Sciences (ICCAIS), 2014.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2016.2514366, IEEE Transactions on Fuzzy Systems



[36] M. Maierdan, K. Watanabe, and S. Maeyama, “Human behavior recognition system based on 3-dimensional clustering methods,” 13th International Conference on Control, Automation and Systems (ICCAS), pp. 1133-1137, 2013.
[37] G. Acampora, P. Foggia, A. Saggese, M. Vento (2015): A Hierarchical Neuro-Fuzzy Architecture for Human Behavior Analysis. In: Information Sciences, 2015.
[38] G. Acampora, P. Foggia, A. Saggese, M. Vento (2012): Combining Neural Networks and Fuzzy Systems for Human Behavior Understanding. In: Proceedings of the "IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)", pp. 88-93, IEEE, 2012, ISBN: 978-1-4673-2499-1.
[39] L. Wang, “The MW method completed: A flexible system approach to data mining,” IEEE Transactions on Fuzzy Systems, vol. 11, no. 6, pp. 768–782, 2003.
[40] A. Kaveh and S. Talatahari, “A discrete big bang-big crunch algorithm for optimal design of skeletal structures,” Asian Journal of Civil Engineering, vol. 11, no. 1, pp. 103-122, 2010.
[41] M. Afshar, and I. Motaei, “Constrained big bang-big crunch algorithm for optimal solution of large scale reservoir operation problem,” Int. J. Optim. Civil Eng., vol. 2, pp. 357-375, 2011.
[42] H. Hagras, F. Doctor, A. Lopez, V. Callaghan, “An Incremental Adaptive Life Long Learning Approach for Type-2 Fuzzy Embedded Agents in Ambient Intelligent Environments,” IEEE Transactions on Fuzzy Systems, vol. 15, no. 1, pp. 41-55, February 2007.
[43] A. Peña-Rios, V. Callaghan, M. Gardner, and M. Alhaddad, “Developing xReality objects for mixed-reality environments,” Intelligent Environments, 2013.

**Bo Yao** obtained his Bachelor's degree in Computer Science and Master’s degree in Signal and Information Processing from Guilin University of Electronic Technology, Guilin, China, and the Ph.D. degree in Computer Science from the University of Essex, Colchester, U.K. He is now a research associate at the University of Essex and works in a collaborative research project with Selex ES. His major research interests are computational intelligence, machine learning, and computer vision. During his Ph.D, he was awarded the best student paper award in FUZZ-IEEE 2014. He began his career of computer vision researcher/developer in 2004 by joining a computer vision research group with whom he started up a technology company manufacturing smart CCTV cameras for the industries of surveillance and security.

**Hani Hagras** (M’03–SM’05, F’13) received the B.Sc. and M.Sc. degrees in electrical engineering from Alexandria University, Alexandria, Egypt, and the Ph.D. degree in computer science from the University of Essex, Colchester, U.K. He is a Professor in the School of Computer Science and Electronic Engineering, Director of the Computational Intelligence Centre and the Head of the Fuzzy Systems Research Group in the University of Essex, UK. His major research interests are in computational intelligence, notably type-2 fuzzy systems, fuzzy logic, neural networks, genetic algorithms, and evolutionary computation. His research interests also include ambient intelligence, pervasive computing and intelligent buildings. He is also interested in embedded agents, robotics and intelligent control. He has authored more than 200 papers in international journals, conferences and books. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) and he is also a Fellow of the Institution of Engineering and Technology (IET (IEE). He was the Chair of IEEE Computational Intelligence Society (CIS) Senior Members Sub-Committee. His research has won numerous prestigious international awards where most recently he was awarded by the IEEE Computational Intelligence Society (CIS), the 2013 Outstanding Paper Award in the IEEE Transactions on Fuzzy Systems and he was also awarded the 2006 Outstanding Paper Award in the IEEE Transactions on Fuzzy Systems. He is an Associate Editor of the IEEE Transactions on Fuzzy Systems. He is also an Associate Editor of the International Journal of Robotics and Automation, the Journal of Cognitive Computation and the Journal of Ambient Computing and Intelligence. He is a member of the IEEE Computational Intelligence Society (CIS) Fuzzy Systems Technical Committee and IEEE CIS conference committee. Prof. Hagras chaired several international conferences where he will act as the Programme Chair of the 2017 IEEE International Conference on Fuzzy Systems, Naples, Italy, July 2017 and he served as the General Co-Chair of the 2007 IEEE International Conference on Fuzzy systems London.

**Daniyal Alghazzawi** (M’12–SM’15) obtained his Bachelor's degree with honor in Computer Science from King Abdulaziz University (KAU) in 1999. He completed his master's degree and doctorate in the field of Computer Science at the University of Kansas at the United States in 2007. He also obtained another master's degree in Teaching and Leadership from University of Kansas in 2004. He also obtained the certificate of Management International Leadership (LMI) and has been the Head of the Information Systems department, Faculty of Computing and Information Technology at King Abdulaziz University for over five years during which he organized many workshops, and international and domestic conferences. He is currently an Associate Professor in the Department of Information Systems, Faculty of Computing & Information Technology at King Abdulaziz University. He is also the head of the Information Security Research Group at King Abdulaziz. He has published 60 papers in various international journals, conferences and books in the field of Intelligent Systems and Information Security. His research interests include intelligent environments, computational intelligence, Smart e-Learning and Information Security.

**Mohammed J. Alhaddad** (M’11) received his Master from Essex University in 2001, and he obtained PhD from Essex University 2006, UK. He became Chairman of Information Technology Department at King Abdul-Aziz University. He is a member of IEEE, ACM and Saudi Computer Society. His research interests are :Network Security, Artificial Intelligence, Robots, fuzzy logic, neural networks, genetic algorithms, Brain Computer Interfaces, Data Mining, Semantic Query Optimization, co-operative query answering, distributed databases and Deductive databases. He is an Associate Editor of many international journals.