IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 14, 2017, accepted September 12, 2017, date of publication September 19, 2017, date of current version October 12, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2754291

Engagement Evaluation for Autism Intervention by Robots Based on Dynamic Bayesian Network and Expert Elicitation

YONGLI FENG, QINGXUAN JIA, MING CHU, AND WEI WEI
School of Automation, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Ming Chu (buptchuming@163.com)

This work was supported by the National Natural Science Foundation of China under Grant 61573066 and Grant 61327806.

ABSTRACT Robots as therapy tools have been researched in intervention for children with autism. During the interaction between robots and autistic children, engagement is an important metric which can be used to express whether robot's behavior is suited to the current context. The evaluation of engagement is a key prerequisite to improve the autonomous ability of robots in intervention. In this paper, we propose a new model to evaluate the engagement of children with autism. The proposed model is developed based on the dynamic Bayesian network, and the parameters of the model are obtained by fuzzy logic and expert elicitation. After determining the input features and the classification of engagement, the topology of the model is established. Afterward, experts' opinions are collected based on linguistic variables. Based on triangular fuzzy number, the parameterization of the model is realized by fuzzification, aggregation, and defuzzification. Finally, the model is validated by experiment. The result demonstrates that proposed model satisfies the actual demands and the result of engagement classification can provide the input condition for the decision making of the robot.

INDEX TERMS Engagement evaluation, autism, expert elicitation, dynamic Bayesian network, fuzzy logic.

I. INTRODUCTION

Autism, or Autism Spectrum Disorder, refers to a broad range of developmental disorders [1]. The main symptoms of autism often include two types, one is impairment in social communication and social interaction, and the other is restricted or repetitive patterns of behavior, interests or activities. Previous studies show that autism is a lifelong developmental disorder and the pathogenesis of autism is unclear for which no effective medication treatment has been used yet [2]. On the other hand, early and individualized interventions are considered to be crucial to encourage children with autism to improve their level of social communication in daily life [3]–[5]. However, the severity and clinical symptoms are varied from one individual to another and there is not an intervention method suitable for every patient with autism [6]. During the intervention, it is necessary to make appropriate strategies for individuals with autism.

The fact of the current situation is that the incidence of autism has reached a higher level (e.g., about 1 in every 88 children in the United States), and there is a tendency to further increase [1]. However, the force of rehabilitation is insufficient. In addition, training process of autism rehabilitation is boring and could cause great physical and mental pressure on practitioners. Consequently, it is necessary to study more scientific and efficient intervention methods of autism.

In recent years, researchers have shown increasing interest in innovative technologies of interaction for the interventions of autistic children [7]. In those studies, the socially assistive robots (SAR) are considered as support tools for autism therapy to provoke and encourage social behaviors of autistic children [8]. Up to now, more than 10 different robot platforms (e.g., Nao, KASPAR, Keepon, and etc.) have been applied to encourage children with autism to develop their social interaction skills [12]. During the autism therapy, robots are used in intervention tasks for joint attention [9], imitation [10], turn-talking [11], and so on.

In order to ensure the robot becomes an effective social entity, its behavior must adapt to the current content. In the previous work, robots were controlled generally by the operators through the method of tele-operation [13]. Although this approach meets the needs of social interactions between robots and children, it cannot be extended on a large scale. Because this approach requires the operators not only have the knowledge of autism intervention, but also have the ability to operate the robot. Therefore, in the long term, robots should have a certain degree of ''intelligence'' and can interact with children autonomously. However, in the existing researches, little effort has been spent on the study of these skills about robot autonomy for autism intervention. Autonomous interaction between robots and autistic children is the main sign of robot autonomy. However, in order to achieve the autonomous interaction between robots and autistic children, there are some challenges to be overcome. For example, robots should perceive and understand user's behaviors, and select the appropriate behavior to respond to the user. In the process of social interaction, engagement is an important metric to measure the effect of the interaction. The engagement can be assessed and classified and the result of classification can be used as the input condition for the robot's decision making.

Aiming at the deficits of autism in social communication and social interaction, we carried out the applied researches of robots in the intervention of autism. In this paper, the main purpose is to present a method to evaluate the engagement of autistic children during social interaction with robots. The main contributions are as follows: (1) We build a new model based on Dynamic Bayesian Network (DBN) to evaluate the engagement for children with autism. This model can fulfil the task of engagement assessment efficiently. (2) Sample data of autistic children are difficult to collect, the parameterization of the model is restricted by data driven. In our study, we obtain the parameters of the model by expert elicitation and deal with opinions of experts through the fuzzy logic. In this way, experts' knowledge and experience are fully applied in the modeling process, and the problem about lack of sample data can be solved.

We remark that our study in this paper is focused on building engagement assessment model, and does not involve in collecting and processing about features of autistic children (e.g., voice and image processing). The remainder of this paper is organized as follows: section II provides an overview of related work; section III induces correlation elementary knowledge of DBN; section IV presents details of model building process; section V introduces the experiment to validate our model; in section VI, we conclude this study and look forward to the future.

II. RELATED WORK

In recent years, researchers have shown great interest in the evaluation of engagement for human-robot interaction. They classified the engagement state of users by their behaviors. In some way, engagement evaluation is one of the emotional recognition.

Several methods which did not consider the particularity of users have been proposed for engagement classification. Hall *et al.* [14] handled nonverbal gestures to classify the engagement of the participants. Rich *et al.* [15] developed a model for classifying the engagement based on gesture and speech. Ivaldi *et al.*[16] proposed a model to classify engagement of children by analyzing their postures and body motion, when they were playing chess with the robot. In addition to these, much more content can be found in [17]. Koo *et al.* [18] used the relative motion between the interactive object and the robot as the input characteristic, and identified the intent of human by hidden Markov model. Ooko *et al.* [19] analyzed the association of head posture and human engagement in the interaction, and classified their engagement through decision tree. Vaufreydaz *et al.* [20] set up the training sets and test sets based on the acquisition of distance, posture, voice and so on, and realized the intention classification by artificial neural network and support vector machine.

For the children with autism, little of efforts have been spent on the study of engagement assessment during robots-children interaction. Feil-Seifer and Mataric [21] presented a recognition system to distinguish positive reactions and negative reactions of autistic children. They collected the features of distance through an overhead camera and classified based on Gaussian Mixture Models and a naive-Bayes classifier. Krupa *et al.* [22] proposed a wearable wristband for acquiring physiological signals which include galvanic skin response and heart rate variability. Then they used support vector machine to predict the state of children such as happy, neutral, and involvement. Liu *et al.* [23] presented a physiology-based method to detect the state of autistic children. In this study, they used wearable sensors to extract the physiology signals of children and used support vector machine to classify.

To sum up, the previous researches have achieved certain results, but there are some problems to be solved. Firstly, discriminating models of engagement for ordinary usually involve strict rules for their use. But children with autism always lack understanding for these rules. So, it is always unsuitable to use these models to evaluate engagement for them. Secondly, some features were obtained by intrusive devices, which could make the autistic children averse. Finally, the classification algorithms which have been used, requiring a large amount of sample data for model training. However, for children with autism, it is difficult to acquire the relevant data. Therefore, the robustness of these models is limited. Different from these work, we present a new method to solve these problems based on DBN and obtain the parameters of the model from experts' opinions by linguistic variables and fuzzy logic.

III. DYNAMIC BAYESIAN NETWORK

Bayesian Network(BN) is a directed acyclic graph which is composed of nodes and arrows [24]. The arrow, which points to the child node from the parent node, expresses the relationship between these two nodes. In BN, each node is labeled with a quantitative probability information. In addition, the relationship between a parent node and its child node is measured by conditional probability. There are some advantages to choose BN for modeling, which are listed as follows [25]: it can explicitly express the causal relationship between each node; it can fuse different sources of information; it can deal with uncertainty powerfully; it can use prior knowledge and subjective probability to infer.

The classification based on BN is a kind of process of probabilistic inference. The basic task of probabilistic reasoning is to calculate the posterior probability of query variables based on evidence variables which can be observed. In some cases, hidden variable (in this paper, it denotes the variable which is neither query variable nor evidence variable) is used in the probabilistic inference. The typical reasoning process can be express as [26]:

$$P(S|\boldsymbol{e}) = kP(S, \boldsymbol{e}) = k\sum_{\boldsymbol{h}} P(S, \boldsymbol{h}, \boldsymbol{e}) \tag{1}$$

Where S represents a class of query variable, k is a normalization factor. Set $\boldsymbol{e}$ and set $\boldsymbol{h}$ represent evidence variables and hidden variables respectively. The joint probability in (1) can be calculated by [26]:

$$P(n_1, \cdots, n_n) = \prod_{i=1}^{n} P(n_i|parents(n_i)) \tag{2}$$

Where $parents(n_i)$ is the parent node of n_i. Combined with (1) and (2), the goal of probabilistic inference is realized.

BN can be applied to uncertain environments. However, the real world is dynamic and the interaction between robots and autistic children is a continuous process. DBN inherits the properties of BN and extends standard BN with the sequences of time. It can describe the change of variables with time through the transition model. Therefore, we develop the evaluation model based on DBN.

IV. THE DESIGN OF ENGAGEMENT EVALUATION MODEL

In practical application, the process of model building based on DBN usually includes two parts: development of the topological structure and parameterization. In the following subsection, we introduce both parts respectively.

A. THE STRUCTURE DESIGN OF THE MODEL

The main task of structural design based on DBN is to determine the nodes of the model and the dependencies between them. In our study, the evidence variables of the model are composed of behavioral features of children with autism. The query variables are expressed by the classes of engagement. In addition, in order to calculate conveniently, the hidden variable is used in the model.

1) FEATURE SELECTION

As discussed in the related work, many features have been selected to classify the engagement. In this study, we selected features which can reflect the intention of autistic children and can be easily obtained during the interaction.

In psychological research, body language is an important communication channel [27]. It plays an crucial role in the expression of emotions, attitudes, and intentions. Among them, facial orientation is a kind of body language, it can indicate the interaction object and the focus of attention in the interpersonal interaction. For instance, in the process of interaction, when a person faces to the interactive object, it shows that he is interested in the content. Otherwise, he is not interested in this interaction. In addition, the duration of facial orientation is also an important feature which can express the intention of people.

Interpersonal distance is a kind of special language in interpersonal interaction which can tell people psychological state of other sides. Hall, an anthropology professor, categorized interpersonal distance into four types: intimate, personal, social and public distance [28]. These four kinds of distance often change dynamically and these changes are synonymous with psychological changes of people.

Acoustic analysis has always been an important research area in the field of emotion recognition [29]. Although language development of children with autism is slow, we can identify their emotions by sound which they make. They can express their negative emotions by cry and scream, and using the sound of laughter to express their positive emotions.

These three kinds of features listed above can affect people's judgment on the engagement of children during autism intervention. Besides, these features can be detected easily by sensors which are integrated into the robot and processed by the existing algorithm. Therefore, in our study, we choose the features of face orientation, interpersonal distance and sound to evaluate the engagement for children in autism intervention.

2) CLASSIFICATION OF ENGAGEMENT

A purpose of engagement evaluation is to classify, then based on the result of classification, the robot can adjust its behavior. There are a variety of classified methods in previous studies. In [13], the engagement was classified into six ratings which include intense engagement, engagement, slight interest, neutral, noncompliance, intense noncompliance. In [21], the engagement was divided into positive reaction and negative reaction. The authors used happiness, neutral and involvement to express the engagement in [22].

The classification of engagement is closely related to the task of the robot. So, in this study, we combine task design and behavior planning of the robot to classify the engagement of autistic children into interest, neutral and aversion from the high level to the low level. For this method of classification, when the result of evaluation is interest, the robot will continue and carry out the interactive tasks. When the result of evaluation is neutral, the robot will strengthen current interaction and try to attract the attention of the children. When the result of evaluation is aversion, the robot will adjust its behavior to reassure the children. To summarize, this classification method can satisfy our requirements.

IEEE Access

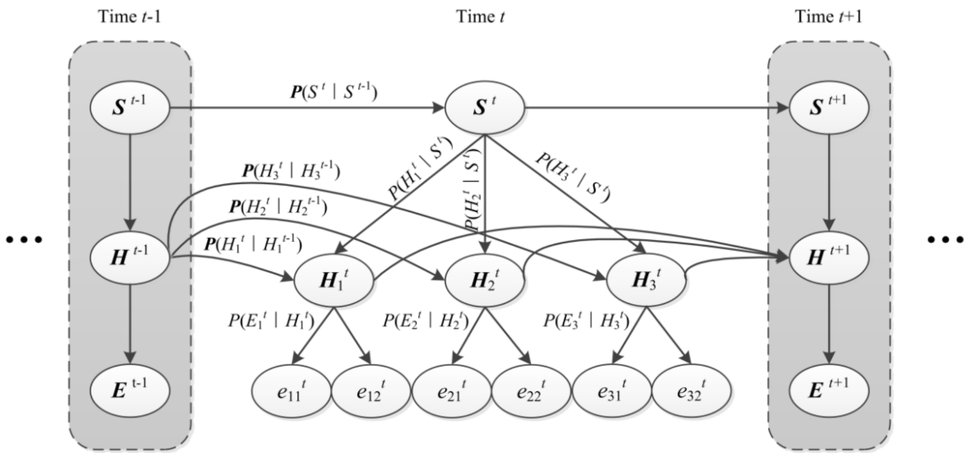


FIGURE 1. The proposed model for engagement evaluation.

TABLE 1. The description of the variables of proposed model.

Query variables	Hidden variables	evidence variables: Basic behavior	evidence variables: Value and explanation
	$\boldsymbol{H}_1$: the results of engagement evaluation by face orientation $\boldsymbol{H}_1$={H_{11},H_{12},H_{13}} H_{11}: interest H_{12}: neutral H_{13}: aversion	$\boldsymbol{E}_1$: Face orientation $\boldsymbol{E}_1$={E_{11}, E_{12}} E_{11}: face orientation (E_{11}-value: e_{11}) E_{12}: duration (E_{12}-value: e_{12})	e_{11}=0: autistic children face to the robot e_{11}=1: autistic children do not face the robot e_{12}=0: duration is less than 3 seconds e_{12}=1: duration is between 3 and 5 seconds e_{12}=2: duration is more than 5 seconds
$\boldsymbol{S}$: the results of engagement evaluation $\boldsymbol{S}$={S_1, S_2, S_3} S_1: interest S_2: neutral S_3: aversion	$\boldsymbol{H}_2$: the results of engagement evaluation by interpersonal distance $\boldsymbol{H}_2$={H_{21},H_{22},H_{23}} H_{21}: interest H_{22}: neutral H_{23}: aversion	$\boldsymbol{E}_2$: Interpersonal distance $\boldsymbol{E}_2$={E_{21}, E_{22}} E_{21}: interpersonal distance (E_{21}-value: e_{21}) E_{22}:change of state (E_{22}-value: e_{22})	e_{21}=0: the distance is less than 1.2 meters e_{21}=1: the distance is between 1.2 and 3 meters e_{21}=2: the distance is more than 3 meters e_{22}=0: autistic children close to the robot e_{22}=1: autistic children keep still e_{22}=2: autistic children leave the robot
	$\boldsymbol{H}_3$: the results of engagement evaluation by acoustic state $\boldsymbol{H}_3$={H_{31},H_{32},H_{33}} H_{31}: interest H_{32}: neutral H_{33}: aversion	$\boldsymbol{E}_3$:Acoustic state $\boldsymbol{E}_3$={E_{31}, E_{32}} E_{31}: acoustics state (E_{31}-value: e_{31}) E_{32}: duration (E_{32}-value: e_{32})	e_{31}=0: the sound of positive emotion e_{31}=1: the sound of neutral emotion (e.g., silence, the sounds do not belong to the positive or negative mood) e_{31}=2: the sound of negative mood (e.g., cry, scream, and et al.) e_{32}=0: duration is less than 3 seconds e_{32}=1: duration is between 3 and 5 seconds e_{32}=2: duration is more than 5 seconds

3) THE STRUCTURE OF MODEL

Through the analysis of the front sections in this paper, we propose an engagement evaluation model based on DBN and the first-order Markov assumption. The structure of proposed model is shown in Fig.1.

In this model, the query variables are expressed by one node, which represents the result of engagement evaluation. There are three hidden variables which express the evaluation results based on the features of face orientation, interpersonal distance and sound respectively. The evidence variables consist of six nodes and these nodes describe the behavioral features which we have chosen. In the real word, the variables of behavior features are always continuous, which has the unlimited number of possible values. It is difficult to calculate the conditional probability for each value explicitly. So the variables are discretized based on psychological research results and experts' advice. The descriptions of the variables are shown in Table 1.

Therefore, combined with (1) and (2), the results of engagement evaluation can be obtained in time $t(t \geq 1)$ by (3), as shown at the bottom of the next page. Where, k_1 and k_2 are normalization factors. $\boldsymbol{P}(S^t|\boldsymbol{e}^t) = [s_1^t, s_2^t, s_3^t]$

and $\boldsymbol{P}(H_i^t|\boldsymbol{e}^t) = [h_{i1}^t, h_{i2}^t, h_{i3}^t]$ represent the calculated values of query variables and hidden variables in time t respectively. $\boldsymbol{P}\left(S^t\left|S^{t-1}\right.\right)$ and $\boldsymbol{P}\left(H_i^t\left|H_i^{t-1}\right.\right)$ are 3-order confused matrices which are used to denote transition relations between adjacent time slices. The symbol '$\circ$' indicates that the corresponding elements between the two matrices are multiplied.

In this section, the structure of engagement evaluation model has been proposed. In the next step, we will parameterize the model and determine the model's conditional probability tables and transition models.

B. THE PARAMETERIZATION OF MODEL

Usually, the parameters of DBN can be obtained by the algorithms of data driven learning, such as Expectation Maximization, which needs a certain amount of sample data. However, there are some drawbacks to this approach. For instance, the cost of data collection and label of samples is high, the size and quality of training set affect learning effect, and the generalization ability of model in different data sets is limited. In addition, in the study of engagement evaluation for autistic children, it is difficult to collect the sample data for using. The reasons are as follows: the related researches are not systematic and there is no dedicated database; the parents of autistic children do not want to put the relevant data released or used in the study because of the privacy.

Different from the learning algorithms based on data, another method based on domain knowledge can be used to determine the parameters [25]. The basic idea of this method is that the domain experts give quantitative data to make decisions when exact data are not available. Generally speaking, people always use qualitative language to evaluate behaviors of others but quantitative analysis. So, in the study of engagement assessment, we introduce linguistic variables based on domain knowledge to elicit the parameters of the model.

1) THE EVALUATION OF DOMAIN EXPERTS

Domain experts refer to the persons who have high attainments within a certain range. They generally have a good command of knowledge, work in related fields, or have the deep understanding of the field. In real world application, the domain experts can give their opinions which can be carried out in linguistic variables based on their experiences and knowledge. The linguistic variable is a kind of conventional quantitative expression which can evaluate the uncertainty [30]. In the field of engineering psychology, the probability of event occurrence is expressed with seven linguistic variables include very low, low, fairly low, medium, fairly high, high and very high [31]. Reference to this kind of expression method, we use the set $\boldsymbol{L} = \{$Very Low, Low, Fairly Low, Medium, Fairly High, High, Very High$\}$ [VL, L, FL, M, FH, H, VH] to express the engagement of autistic children in interaction with robots from low degree to high degree. In this way, experts can choose appropriate linguistic variables to give their opinions.

TABLE 2. Weighting scores of experts.

Constitution	Classification	score
Professional position	Professor, researcher	4
	Therapist	3
	Parent	2
	Others (e.g., students who study in this domain)	1
Time of experience (in years)	Greater than 10	4
	5 to 10	3
	2 to 5	2
	<2	1
Education qualification	Doctor's degree	4
	Master's degree	3
	Bachelor's degree	2
	Others	1

The opinions of domain experts are subjective. In order to ensure the objectivity of the model parameters, the multiple experts should be invited. The experts' opinion can be elicited individually and then these opinions are combined. In this process, to reflect the differences of experts' assessment, different weights from 0 to 1 should be assigned to every expert. In this study, $W(E_i)$ is the weight factor of i-th expert, which can be obtained according to Table 2.

Although the linguistic variables are convenient for experts to give their professional assessment, the parameters of the model are crisp numbers and cannot apply the linguistic variables to express directly. Therefore, we need to convert the these variables into specific values.

2) TRANSFORMING LINGUISTIC VARIABLES TO PRECISE VALUES

Triangular fuzzy number is a useful method to transform the fuzzy linguistic variable into certain numerical. The triangular fuzzy number can be denoted by $\boldsymbol{A} = (l, m, u; 1)$ and the membership function is written as:

$$\mu_A(x) = \begin{cases} \dfrac{x-l}{m-l}, & l < x \le m \\ \dfrac{u-x}{u-m}, & m < x \le u \\ 0, & otherwise \end{cases} \tag{4}$$

$$\begin{cases} \boldsymbol{P}\left(S^t\left|\boldsymbol{e}^t\right.\right) = k_1 \boldsymbol{P}\left(S^{t-1}\left|\boldsymbol{e}^{t-1}\right.\right)\boldsymbol{P}\left(S^t\left|S^{t-1}\right.\right) \circ \sum_{i=1}^{3}\left(\boldsymbol{P}\left(H_i^t\left|S^t\right.\right)\boldsymbol{P}\left(H_i^t\left|\boldsymbol{e}^t\right.\right)\right) \\ \boldsymbol{P}\left(H_i^t\left|\boldsymbol{e}^t\right.\right) = k_2 \boldsymbol{P}\left(H_i^{t-1}\left|\boldsymbol{e}^{t-1}\right.\right)\boldsymbol{P}\left(H_i^t\left|H_i^{t-1}\right.\right) \circ \left(\boldsymbol{P}\left(e_{i1}^t\left|H_i^t\right.\right) \circ \boldsymbol{P}\left(e_{i2}^t\left|H_i^t\right.\right)\right. \quad i = 1, 2, 3 \end{cases} \tag{3}$$

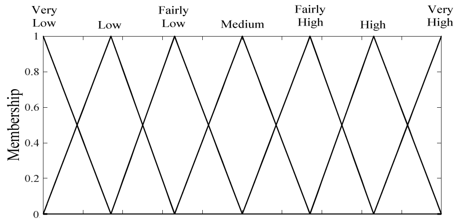


FIGURE 2. Fuzzy numbers representing linguistic variables.

TABLE 3. Fuzzy representation of linguistic variables.

Linguistic Variables	Fuzzy Numbers
Very Low (VL)	(0.00, 0.00, 0.17; 1)
Low (L)	(0.00, 0.17, 0.33; 1)
Fairly Low (FL)	(0.17, 0.33, 0.50; 1)
Medium (M)	(0.33, 0.50, 0.67; 1)
Fairly High (FH)	(0.50, 0.67, 0.83; 1)
High (H)	(0.67, 0.83, 1.00; 1)
Very High (VH)	(0.83, 1.00, 1.00; 1)

When evaluation object cannot be accurately measured, the triangular fuzzy number can be used to deal with this situation by fuzzification, aggregation, and defuzzification. In this way, important quantitative information of the model can be acquired even if the sample data are not known.

Fuzzification: In order to get the certain numerical to express the experts' assessment which use linguistic variables in set $\mathbf{L}$, the triangular fuzzy number $\boldsymbol{A}_i = (l_i, m_i, u_i; 1)$ corresponding can be obtained by (5).

$$\begin{cases} l_i = (i-1)/n \\ m_i = i/n \\ u_i = (i+1)/n & \qquad i = 0, 1, \cdots, n \\ u_i = 1, \qquad u_i \geq 1 \\ l_i = 0, \qquad l_i \leq 0 \end{cases} \tag{5}$$

In this paper, we can get $n = 6$ from set $\mathbf{L}$, then the fuzzy set is calculated and the membership function as shown in Fig.2.

In addition, we can obtain the fuzzy numbers which corresponding to set $\mathbf{L}$ as shown in Table 3.

Aggregation: In the parameterization of the model, the opinions from experts need to be aggregated. In this way, every parameter of the model can be obtained. Assume that the number of experts is n, they give their assessment results for a problem respectively based on the linguistic variables set $\mathbf{L}$. Let $\boldsymbol{A}_i = (l_i, m_i, u_i; 1)$ and $\boldsymbol{A}_j = (l_j, m_j, u_j; 1)$ be two fuzzy numbers which represent the opinions of expert E_i $(i = 1, 2, \ldots, n)$ and expert $E_j (j = 1, 2, \ldots, n)$. Where the values of $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ can be obtained from Table 3. The method of aggregation as follows [32]:

Step 1: Determine the similarity of expert E_i's opinion to expert E_j's opinion as in (6).

$$S\left(E_i, E_j\right) = 1 - \frac{1}{3}\left(\left(\left|l_i - l_j\right|\right) + \left(\left|m_i - m_j\right|\right) + \left(\left|u_i - u_j\right|\right)\right) \tag{6}$$

Step 2: Calculate the average degree of similarity of expert E_i by (7), where n stands for the number of experts.

$$AD(E_i) = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} S\left(E_i, E_j\right) \tag{7}$$

Step 3: Determine the relative degree of opinion of expert E_i by (8).

$$RD(E_i) = \frac{AD\left(E_i\right)}{\sum_{i=1}^{n} AD\left(E_i\right)} \tag{8}$$

Step 4: The consensus weight of expertE_i can be obtained by (9). Where $W(E_i)$ represents a given weight of expert E_i based on Table 4. $\beta \in [0, 1]$ is a relaxation factor which measures the importance of $W(E_i)$ and $RE(E_i)$ in the equation.

$$CW\left(E_i\right) = \beta \cdot W\left(E_i\right) + (1-\beta) \cdot RD\left(E_i\right) \tag{9}$$

Step 5: Finally, the comprehensive assessment of experts for an event can be obtained by (10).

$$\boldsymbol{A}_{\text{comp}} = \sum_{i=1}^{n} CW\left(E_i\right) \cdot \boldsymbol{A}_i \tag{10}$$

Defuzzification: The fuzzy numbers cannot be directly applied to the proposed model, and they need to be converted to specific values. In this paper, the conversion can be achieved by the method of integral value [33]. The expression of integral value can be defined as (11).

Where I_{T}^{α} is the result of transformation of fuzzy number, and $\alpha \in [0, 1]$ is an index of optimism which is weighted the objectivity of evaluation. For $\alpha = 0$ and $\alpha = 1$, I_{T}^{α} get the extreme value respectively, and when $\alpha = 0.5$ we can obtain a moderate value. $I_L(\boldsymbol{A})$ and $I_R(\boldsymbol{A})$ are the integral value of the inverse function of the left and right membership functions. $g_A^{\mathrm{L}}(y)$ and $g_A^{\mathrm{R}}(y)$ represent the inverse function of the left and right membership functions.

$$\begin{aligned} I_{\mathrm{T}}^{\alpha} &= \alpha I_{\mathrm{R}}(\boldsymbol{A}) + (1-\alpha) I_{\mathrm{L}}(\boldsymbol{A}) \\ &= \alpha \int_0^1 g_A^{\mathrm{R}}(y)\,\mathrm{d}y + (1-\alpha) \int_0^1 g_A^{\mathrm{L}}(y)\,\mathrm{d}y \end{aligned} \tag{11}$$

Based on (4) and (11), the crisp value, which related to triangular fuzzy number, can be expressed by (12).

$$I_{\mathrm{T}}^{\alpha} = \frac{1}{2}\left((1-\alpha)\, l + m + \alpha u\right) \tag{12}$$

3) DETERMINE THE PARAMETERS

In our study, we selected five experts to give their opinions to evaluate the engagement of children with autism when they interact with the robot. They are composed of the parent of an autistic child, an autistic therapist, two researchers of autistic education and a professor of psychology for special children. We use weighting factors, which can be obtained from Table 2, to represent the relative importance of experts. The weight factor of different expert is shown in Table 4. In order to obtain the opinions of experts, a questionnaire is

TABLE 4. Weighting factors of experts.

Expert	Classification	Experience	Education qualification	Score	Weighting factor
E_1	Parent	4 years	Bachelor's degree	6	0.133
E_2	Therapist	4 years	Bachelor's degree	7	0.155
E_3	Researcher	12 years	Doctor's degree	12	0.267
E_4	Researcher	3 years	Bachelor's degree	8	0.178
E_5	Professor	13 years	Doctor's degree	12	0.267

TABLE 5. The processing of an example.

Expert	Opinion	Fuzzy number	Average degree $AD(E_i)$	Relative degree $RD(E_i)$	Consensus weight $CW(E_i)$	Comprehensive assessment A_{comp}	Conditional probability
E_1	H	(0.67, 0.83, 1.00; 1)	0.875	0.194	0.164		
E_2	H	(0.67, 0.83, 1.00; 1)	0.875	0.194	0.175		
E_3	FH	(0.50, 0.67, 0.83; 1)	0.917	0.204	0.235	(0.558, 0.724, 0.887; 1)	0.723
E_4	FH	(0.50, 0.67, 0.83; 1)	0.917	0.204	0.191		
E_5	FH	(0.50, 0.67, 0.83; 1)	0.917	0.204	0.235		

TABLE 6. Conditional probability table 1.

H_1	$P(e_{11}=0)$	$P(e_{11}=1)$	$P(e_{12}=0)$	$P(e_{12}=1)$	$P(e_{12}=2)$
H_{11}=Interest	0.624	0.376	0.207	0.310	0.483
H_{12}=Neutral	0.309	0.691	0.488	0.353	0.159
H_{13}=Aversion	0.164	0.836	0.107	0.270	0.623

TABLE 7. Conditional probability table 2.

H_2	$P(e_{21}=0)$	$P(e_{21}=1)$	$P(e_{21}=2)$	$P(e_{22}=0)$	$P(e_{22}=1)$	$P(e_{22}=2)$
H_{21}=Interest	0.431	0.382	0.187	0.433	0.385	0.182
H_{22}=Neutral	0.287	0.407	0.306	0.320	0.401	0.279
H_{23}=Aversion	0.150	0.258	0.592	0.121	0.392	0.487

TABLE 8. Conditional probability table 3.

H_3	$P(e_{31}=0)$	$P(e_{31}=1)$	$P(e_{31}=2)$	$P(e_{32}=0)$	$P(e_{32}=1)$	$P(e_{32}=2)$
H_{31}=Interest	0.482	0.426	0.092	0.150	0.363	0.487
H_{32}=Neutral	0.268	0.467	0.265	0.500	0.384	0.116
H_{33}=Aversion	0.121	0.314	0.565	0.141	0.280	0.579

designed. The content of the questionnaire is closely related to the evaluation model and the set of linguistic variables.

In the questionnaire, some questions are described and experts answer these questions by means of linguistic variables. We gathered the questionnaires from each expert and dealt with experts' opinions by the method which has been described in the previous section. For example, there is an event in the questionnaire: if a child is interested in the engagement, which degree do you think he/she faces to the robot ($e_{11} = 0$)? Based on this example, experts give their opinions, and we can obtain the conditional probability without normalization $P(e_{11} = 0|H_{11})$ by (6)-(12). The detail of this process is shown in Table 5. In this process, we set the relaxation factor $\beta = 0.5$ and the optimism index $\alpha = 0.5$.

Repeating the above process, we get the parameters of the model. The conditional probability tables between hidden variables and evidence variables are shown in Table 6, Table 7, and Table 8. In the model, the conditional probabilities between the query variables and the hidden variables can be considered as the weight factors. We use these values to measure the role of each feature in the evaluation engagement. In practice, the laughter of autistic children is insignificance in most of the time. So, we asked the experts to give the appropriate answers for the different value based on acoustic state of autistic children. The conditional probability table between query variables and evidence variables is shown in Table 9. The confusion matrix is used to represent the state of transition relationship from moment $t-1$ to t. In the study, transition matrixes of the model are equal ($\boldsymbol{P}(H_1^t|H_1^{t-1}) = \boldsymbol{P}(H_2^t|H_2^{t-1}) = \boldsymbol{P}(H_3^t|H_3^{t-1}) = \boldsymbol{P}(S^t|S^{t-1})$), then the transition matrix is shown in Table 10.

At this point, the model of engagement evaluation is built completely.

V. EXPERIMENTAL RESULTS

In this section, we verify the performance of the proposed model by experiment.

TABLE 9. Conditional probability table 4.

e_{31}	$P(H_1\|S)$	$P(H_2\|S)$	$P(H_3\|S)$
0	0.358	0.326	0.316
1	0.429	0.326	0.245
2	0.186	0.227	0.587

TABLE 10. The transition matrix.

	S_1^t	S_2^t	S_3^t
S_1^{t-1}	0.852	0.095	0.053
S_2^{t-1}	0.107	0.811	0.082
S_3^{t-1}	0.033	0.096	0.871

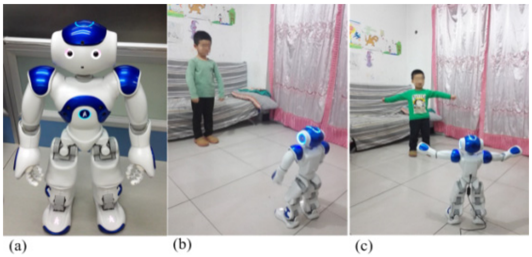


FIGURE 3. (a) The robot platform NAO, (b) the robot is interacting with child A, (c) the robot is interacting with child B.

A. THE ROBOT PLATFORM

At present, a variety of robotic platforms have been applied to studies of interventional therapy for children with autism. These robots are different from structure to function and each robot can offer help in intervention through their unique skills. In our study, we use the robot Nao to support our study which is shown in Fig.3. Robot Nao has a lovely appearance and it can be accepted by children with autism easily. Besides, the robot Nao integrates a variety of sensors and actuators, so that it can collect environment information and express its behavior conveniently.

In our experiments, we use Python 2.7 to develop the application for robot NAO. Before using the engagement evaluation model, the robot needs to collect and process the features of the child. The robot uses the video camera which is located in the forehead to capture the image of the interactive environment (image resolution is 320∗240). Then the robot detects and locates the child by HOG-Linear SVM, and detects the face of the child based on Haar feature. On this basis, the robot can estimate the child's head pose through the random forest. Then, the estimated results are divided into two types: face to the robot and do not face to the robot (when robot cannot detects the face of the child we believe that the child do not face to the robot). The robot estimates the distance to the child by combining the monocular vision with the ultrasonic sensors. In the process of engagement evaluation, we mainly recognize autistic children's laughter, crying and screaming. The robot acquires the sound signal through the microphones which installed in its head, and then MFCC feature of the sound is extracted and using SVM for classification.

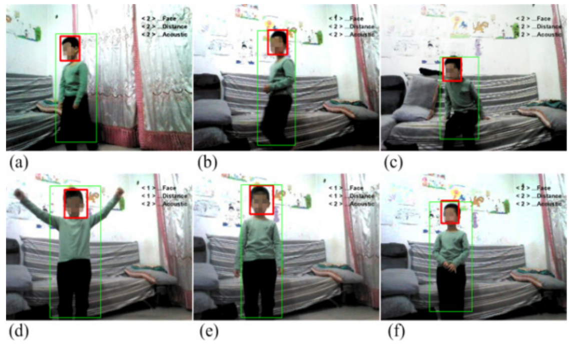


FIGURE 4. Images (a) to (f) were acquired by the robot's video camera, and illustrate the sequence of changes of the child A when the robot interacted with him.

B. INTERACTION SCENARIO DESIGN

During the interaction, the robot guided the autistic child to complete the interaction task. This experiment is mainly used to verify the performance of the engagement evaluation model, so the interaction scenario can be described as follows: 1) When the child is detected, the robot adjusts its position towards the child, moves to the appropriate distance; 2) The robot motivate the child to interact through language and action; 3) If the result of engagement evaluation is interest, the robot will guide the child to imitate its action; 4) If the result of engagement evaluation is neutral, the robot will strengthen current interaction and try to attract the attention of the children; 5) If the result of evaluation is aversion, the robot will attempt to arouse the interest of the child by talking and acting. If the state of aversion continues for a certain period, the robot will adjust its behavior to reassure the child.

In our study, we took the task of action imitation (the robot guides the child to imitate its action) as an example to experiment. The child A (who is a boy and 5 years old) and the child B (who is a boy and 4 years old) were invited to participate in the study, and the study was supported by their parents.

C. THE RESULTS OF EVALUATION

In the experiment, the robot evaluated children's engagement by proposed engagement evaluation model in two scenarios of interaction.

Interaction 1: Fig.4 illustrates the image sequence when the robot interacted with child A. Besides, we use some typical events to briefly describe the interaction process, as shown in Table 11.

Fig.5 illustrates the results of engagement evaluation when the robot interacted with the child A. By comparison, the results of the engagement evaluation by proposed model are consistent with the actual state of the child A in the interaction.

Interaction 2: Fig.6 illustrates the image sequence when the robot interacted with child B. Besides, we use some typical events to briefly describe the interaction process, as shown in Table 12.

TABLE 11. The description of process of interaction 1.

Time	Interactive content	The state of the child
3s	The robot detected the child and adjusted its position.	The child did not pay attention to the robot
10s	The robot initiated an interactive request and said "hello, I am Qiqi， can you play with me?".	The child was not interested in the robot, he went to sofa, and glanced the robot.
25s	The robot begin to change the color of its eyes and attempted to attract the child.	The child began to face the robot and observed it.
50s	The robot began to guide the child to imitate its action.	The child imitated the action of the upper limbs of the robot
110s	The robot suggested the child take a break.	The child began to look around and did not know what he should do.
140s	The robot suggested that the child imitated its action once again.	The child imitated the action of the robot once again.
200s	The robot finished the interaction and said "Bye-bye!".	The child turned and left the robot.

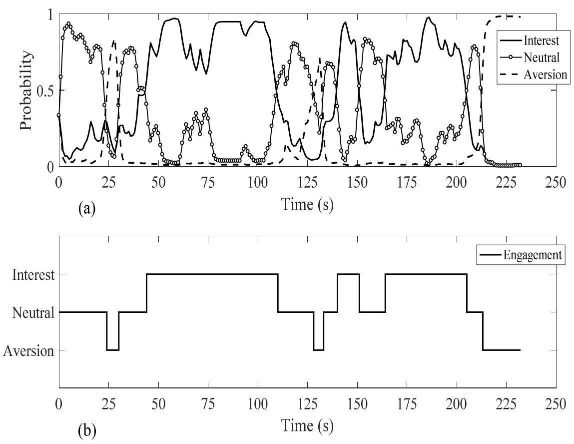


FIGURE 5. The results of engagement evaluation when the robot interacts with the child A (acquisition period is 1s). (a) describes the trend of engagement by probability values. (b) describes the evaluation result of child's engagement.

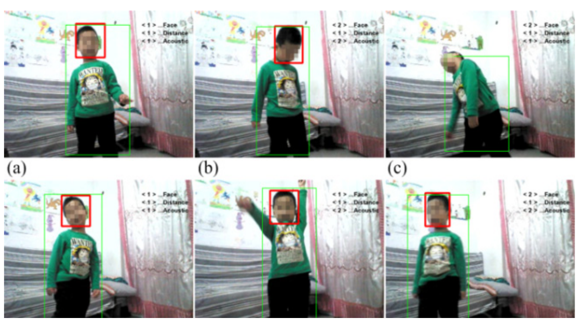


FIGURE 6. Images (a) to (f) were acquired by the robot's video camera, and illustrate the sequence of changes of the child B when the robot interacted with him.

Fig.7 illustrates the results of engagement evaluation when the robot interacted with the child B. By comparison, the results of the engagement evaluation by proposed model

TABLE 12. The description of process of interaction 2.

Time	Interactive content	The state of the child
3s	The robot detected the child	The child faced the robot and observed it.
10s	The robot initiated an interactive request and said "hello, I am Qiqi， can you play with me?".	The child began to laugh and twist his body.
30s	The robot began to guide the child to imitate its action.	The child imitated the action of the robot. However, he sometimes did not pay attention to the robot.
85s	The robot finished the interaction and said "Bye-bye!".	The child stood for a moment and then walked away.

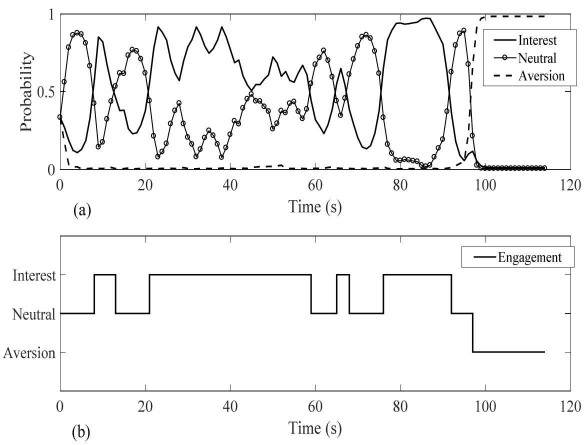


FIGURE 7. The result of engagement evaluation when the robot interacts with the child B (acquisition period is 1s). (a) describes the trend of engagement by probability values. (b) describes the evaluation result of child's engagement.

TABLE 13. The results of comparison for interaction 1.

	Interest	Neutral	Aversion
Interest	93.22%	6.78%	0%
Neutral	4.24%	94.07%	1.69%
Aversion	0%	6.45%	93.55%

are consistent with the actual state of the child B in the interaction.

In the study, we asked the experts to rate the level of engagement in videos of the interaction (the period is 1s), and then we compared that with the results which were evaluated by proposed model. The results of comparisons are shown in Table 13 and Table 14.

Table 13 and Table 14 show the confusion matrixes of evaluation accuracy by the proposed model when compared with the experts' opinions. The overall performance of our model is 93.6% correct. Classification errors are centered between adjacent categories. In this case, the results of engagement evaluation meet the needs

TABLE 14. The results of comparison for interaction 2.

	Interest	Neutral	Aversion
Interest	91.94%	8.06%	0%
Neutral	5.71%	94.29%	0%
Aversion	0%	5.56%	94.44%

of the robot to develop the appropriate strategy of the interaction.

VI. CONCLUSION

Classifying whether an autistic child is engaged with the robot is the first step for using robots as therapy tool. Engagement classification can provide the input condition for the decision making of the robot. For example, when the decision-making strategy of the robot is developed based on reinforcement learning, the degree of engagement which has been evaluated can be used to measure the state of the interactive object. In this paper, an effective method is proposed for evaluating engagement when the robot interacted with autistic children. According to the characteristics of interaction task, an evaluation model based on DBN was designed, and the parameterization of the model was accomplished by transforming the qualitative evaluation from experts as parameters of the model by fuzzy logic. The key advantage of this method is that the experts' opinions are expressed in the model and some restrictions based on data-driven methods are avoided. It effectively reduces the difficulty of model building and has certain practical value. The experimental results illustrate that the model performs excellent in practical applications and satisfies the actual demands.

The model of engagement evaluation is developed based on Dynamic Bayesian Network and expert elicitation. Nodes and relationships between nodes of the model can be given by experts' experience. So, the proposed method has greater potential in practical application, and it can be used in other companion robots for different types of interactions. The steps as follows: 1) select the features according to the interaction task; 2) build the topological structure of the model; 3) design the language variable set; 4) design the questionnaire based on the language variable set and the structural relationship of the model to collect advice of experts; 5) obtain the parameters of the model by fuzzification, aggregation, and defuzzification based on experts' opinions; 6) correct the parameters of the model by experiment.

While the study has offered an effective approach of modeling for assessing engagement, there are still a number of interesting avenues to explore. In the proposed method, subject to the experimental platform and our existing perceptual techniques, we choose some features as input to the model, including time, face orientation, interpersonal distance, and acoustic state. In future research, we will consider the application of other features (such as eye tracking/pupil tracking) in the process of evaluation, and update the detection methods to obtain these features. In addition, it is a dynamic process during the interaction between robots and autistic children (interaction task or interaction object is changing). The model with fixed parameters has some limitations in dealing with these situations. In the future study, we will expand the model and give the learning capacity to the model and design experiments for different interactive objects to verify our model, such as different genders, various kinds of level in autism, different age groups and so on. At the same time, the appropriate decision-making strategy based on the evaluation results of engagement will be studied. In this way, the estimation accuracy will be further improved and the robustness will be enhanced for different users and different interactive environments.

REFERENCES

[1] J. J. Cabibihan, H. Javed, M. Ang, Jr., and S. M. Aljunied, "Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism," *Int. J. Social Robot.*, vol. 5, no. 4, pp. 593–618, 2013.

[2] C. Lord and S. L. Bishop, "Autism spectrum disorders: Diagnosis, prevalence, and services for children and families," *Soc. Res. Child Develop.*, vol. 27, no. 2, pp. 1–26, 2010.

[3] P. Srinivasan, "A review of dietary interventions in autism," *Ann. Clin. Psychiatry*, vol. 21, no. 4, pp. 237–247, 2009.

[4] L. Escobedo, C. Ibarra, J. Hernandez, M. Alvelais, and M. Tentori, "Smart objects to support the discrimination training of children with autism," *Pers. Ubiquitous Comput.*, vol. 18, no. 6, pp. 1485–1497, 2014.

[5] C. A. G. J. Huijnen, M. A. S. Lexis, and L. P. de Witte, "Matching robot KASPAR to autism spectrum disorder (ASD) therapy and educational goals," *Int. J. Social Robot.*, vol. 8, no. 4, pp. 1–11, 2016.

[6] S. Thill, C. A. Pop, T. Belpaeme, T. Ziemke, and B. Vanderborght, "Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook," *Paladyn*, vol. 3, no. 4, pp. 209–217, 2012.

[7] S. Boucenna *et al.*, "Interactive technologies for autistic children: A review," *Cognit. Comput.*, vol. 6, no. 4, pp. 1–19, 2014.

[8] A. Peca, "Robot enhanced therapy for children with autism disorders: Measuring ethical acceptability," *IEEE Technol. Soc. Mag.*, vol. 35, no. 2, pp. 54–66, Jun. 2016.

[9] B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn, "Robot-mediated joint attention in children with autism: A case study in robot-human interaction," *Interact. Stud.*, vol. 5, no. 5, pp. 161–198, 2004.

[10] S. M. Srinivasan, K. A. Lynch, D. J. Bubela, T. D. Gifford, and A. N. Bhat, "Effect of interactions between a child and a robot on the imitation and praxis performance of typically devloping children and a child with autism: A preliminary study," *Perceptual Motor Skills*, vol. 116, no. 3, pp. 885–904, 2013.

[11] K. Dautenhahn, "Socially intelligent robots: Dimensions of human-robot interaction," *Philos. Trans. Roy. Soc. London A, Math. Phys. Sci.*, vol. 362, no. 1480, pp. 679–704, 2007.

[12] P. Pennisi, A. Tonacci, and G. Tartarisco, "Autism and social robotics: A systematic review," *Autism Res.*, vol. 9, no. 2, pp. 58–67, 2016.

[13] B. Scassellati, H. Admoni, and M. Metric, "Robots for use in autism research," *Annu. Rev. Biomed. Eng.*, vol. 14, no. 1, pp. 275–294, 2012.

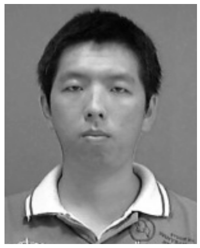
[14] J. Hall, T. Tritton, A. Rowe, C. Melhuish, U. Leonards, and A. Pipe, "Perception of *own* and robot engagement in human–robot interactions and their dependence on robotics knowledge," *Robot. Auto. Syst.*, vol. 62, no. 3, pp. 392–399, 2014.

[15] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner, "Recognizing engagement in human-robot interaction," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact. (HRI)*, Mar. 2010, pp. 375–382.

[16] S. Ivaldi, S. M. Anzalone, W. Rousseau, O. Sigaud, and M. Chetouani, "Robot initiative in a team learning task increases the rhythm of interaction but not the perceived engagement," *Frontiers Neurorobot.*, vol. 8, no. 5, pp. 1–23, 2014.

[17] S. M. Anzalone, S. Boucenna, S. Ivaldi, and M. Chetouani, "Evaluating the engagement with social robots," *Int. J. Social Robot.*, vol. 7, no. 4, pp. 75–82, 2015.

[18] S. Koo and D. S. Kwon, "Recognizing human intentional actions from the relative movements between human and robot," in *Proc. 18th Int. Symp. Robot Human Interact. Commun.*, Toyama, Japan, 2009, pp. 939–944.
[19] R. Ooko, R. Ishii, and Y. I. Nakano, "Estimating a user's conversational engagement based on head pose information," in *Proc. Int. Conf. Intell. Virtual Agents*, Berlin, Germany, 2011, pp. 262–268.
[20] D. Vaufreydaz, W. Johal, and C. Combe, "Starting engagement detection toward a companion robot using multimodal features," *Robot. Auto. Syst.*, vol. 75, pp. 4–16, Sep. 2015.
[21] D. Feil-Seifer and M. J. Mataric, "Distance-based computational models for facilitating robot interaction with children," *J. Human-Robot Interact.*, vol. 1, no. 1, pp. 1–23, 2012.
[22] N. Krupa, K. Anantharam, M. Sanker, S. Datta, and J. V. Sagar, "Recognition of emotions in autistic children using physiological signals," *Health Technol.*, vol. 6, no. 2, pp. 1–11, 2016.
[23] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Online affect detection and robot behavior adaptation for intervention of children with autism," *IEEE Trans. Robot.*, vol. 24, no. 4, pp. 883–896, Aug. 2008.
[24] L. Uusitalo, "Advantages and challenges of Bayesian networks in environmental modelling," *Ecol. Model.*, vol. 203, no. 3, pp. 312–318, 2007.
[25] G. Zhang and V. V. Thai, "Expert elicitation and Bayesian network modeling for shipping accidents: A literature review," *Safety Sci.*, vol. 87, pp. 53–62, Aug. 2016.
[26] J. Stuart Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. London, U.K.: Pearson Education, 2012.
[27] A. Beck, B. Stevens, K. A. Bard, and L. Cañamero, "Emotional body language displayed by artificial agents," *ACM Trans. Interact. Intell. Syst.*, vol. 2, no. 1, pp. 1–29, 2012.
[28] S. Fotios, B. Yang, and J. Uttley, "Observing other pedestrians: Investigating the typical distance and duration of fixation," *Lighting Res. Technol.*, vol. 47, no. 5, pp. 548–564, 2015.
[29] L. Kessous, G. Castellano, and G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *J. Multimodal User Interfaces*, vol. 3, no. 1, pp. 33–48, 2010.
[30] S. Kabir, M. Walker, E. Rüde, P. Securius, and Y. Papadopoulos, "Fuzzy temporal fault tree analysis of dynamic systems," *Int. J. Approx. Reason.*, vol. 77, pp. 20–37, Oct. 2016.
[31] S. Rajakarunakaran, A. M. Kumar, and V. A. Prabhu, "Applications of fuzzy faulty tree analysis and expert elicitation for evaluation of risks in LPG refuelling station," *J. Loss Prevention Process Ind.*, vol. 33, pp. 109–123, Jan. 2015.
[32] H. M. Hsu and C. T. Chen, "Aggregation of fuzzy opinions under group decision making," *Fuzzy Sets Syst.*, vol. 79, no. 3, pp. 279–285, 1996.
[33] T. S. Liou and M. J. J. Wang, "Ranking fuzzy numbers with integral value," *Fuzzy Sets Syst.*, vol. 50, no. 3, pp. 247–255, 1992.



YONGLI FENG received the master's degree from the North China University of Science and Technology. He is currently pursuing the Ph.D. degree with the School of Automation, Beijing University of Posts and Telecommunications. His main research interests include human–robot interaction and artificial intelligence.

QINGXUAN JIA received the bachelor's degree from the Shandong University of Technology, and the master's and Ph.D. degrees from the Beijing University of Aeronautics and Astronautics. He is currently a Professor with the Beijing University of Posts and Telecommunications. His specific interests include robotics, artificial intelligence, and virtual reality.

MING CHU received the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2010. He is currently an Associate Professor with the Automation School, BUPT. His recent research areas are flexible dynamics, nonlinear control, and robust control.

WEI WEI received the bachelor's degree from the China University of Mining and Technology, Beijing, and the master's degree from the Beijing University of Posts and Telecommunications, where she is currently pursuing the Ph.D. degree. Her research interests include computer vision, affective computing, and human–computer interaction.