The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Interpretation of Fuzzy Voice Commands for Robots Based on Vocal Cues Guided by User's Willingness

A. G. Buddhika P. Jayasekara, Keigo Watanabe, Kazuo Kiguchi, and Kiyotaka Izumi

Abstract—**This paper proposes a method for interpretation of fuzzy voice commands based on the vocal cues. The fuzzy voice commands include fuzzy linguistic terms like "little" and quantitative meaning of such terms depends on the environmental conditions. Therefore the robot's perception of the corresponding environment is modified by acquiring the user's perception through a series of vocal cues. The user's willingness to change the robot's perception is identified based on the vocal cues to improve the adaptation process. The primitive behaviors related to the end-effector movements of a robot manipulator are considered and evaluated by a behavior evaluation network (BEN). A vocal cue evaluation system (VCES) is used to evaluate the vocal cues to adapt the robot's perception by modifying the BEN. The user's satisfactory level for the robot's movements is utilized to track the satisfaction of the user. A situation of cooperative rearrangement of the user's working space is used to illustrate the proposed system by a PA-10 robot manipulator.**

I. INTRODUCTION

Human-friendly robotic systems are being emerged in the human working environments. Since the humans prefer to use voice instructions, responses, and suggestions including uncertain information involuntarily in their daily interactions, the ability of the robot companion to understand the uncertain information is crucial. A human-friendly robot equipped with human-like voice communication capabilities will be able to help disabled people, to increase the capabilities of the aged people, to support daily activities, etc. [1]–[4]. Normally, humans possess a natural ability of adapting to other humans and artifacts. The mutual adaptation is important in acquiring the information from the voice commands in human-human communications. Such adaptation can improve the usability of the robot companions in human-robot interaction [5].

The ability of understanding the fuzzy linguistic information plays a key role in human-robot interaction. In Pulasinghe et al. [6], robot control by using information-rich voice commands such as "move a very little forward" has been studied. Such voice commands inherit uncertain information due to the inclusion of fuzzy linguistic terms like "very little." Generally, the voice commands, which include fuzzy linguistic information, are referred as fuzzy voice commands (FVCs) [7]–[9]. In [10], a method has been proposed to evaluate the FVCs by considering a set of previous movements of the robot as internal rehearsals. In addition, the FVCs have been utilized to teach composite behaviors based on the primitives for posture control of a robot manipulator [11]. However, the quantitative meaning of such uncertain information depends on the environmental conditions and the user's expectations. The main limitation of the above-mentioned methods is that the system of understanding and quantifying the fuzzy linguistic information in voice commands is predetermined. However, the ability to adapt the system for understanding the fuzzy linguistic information based on the user's desire and the environmental conditions is important.

The number of interactive steps required to adapt the robotic system toward the user's desire is vital and lower number of interactions is preferred for human-like interaction. In [12], a method has been proposed to adapt the robot's perception of the fuzzy linguistic information based on user feedback. The main limitation is the requirement of exceptional large number of adaptation steps to achieve the user's satisfaction. The above-mentioned limitation is occurred due to the predetermined nature of the vocal cue evaluation process and inability to acquire the user's willingness. Therefore, this paper proposes a method for interpreting the fuzzy voice commands based on vocal cues guided by the user's willingness to change the robot's perception. Here, the robot's perception on the fuzzy linguistic information is modified by acquiring the user's perception with his willingness to change the robot's perception based on a series of vocal cues. The system overview is discussed in section II. Next section III discusses the evaluation process of the fuzzy linguistic information. Finally, results and conclusions are presented.

II. SYSTEM OVERVIEW

This paper proposes a method to understand the FVCs by evaluating the fuzzy linguistic information based on the user's guidance as a series of vocal cues. The proposed system consists of an interaction manager, a behavior evaluation network (BEN), and a vocal cue evaluation system (VCES). The end-effector movements of a robot manipulator are considered as the primitive behaviors for the proposed system. The BEN is utilized to evaluate the primitive behaviors and implemented by using a fuzzy-neural network. The VCES is introduced to evaluate the vocal cues. Functional overview of the proposed method is shown in Fig.1. The interaction manger is utilized to manage the interaction between the human user and the robot. The voice commands are categorized into two groups: a FVC

A.G.B.P. Jayasekara, Kazuo Kiguchi, and Kiyotaka Izumi are with the Department of Advanced Systems Control Engineering, Saga University, 1-Honjomachi, Saga 840-8502, Japan (e-mail: buddhika@ieee.org, {kiguchi, izumi}@me.saga-u.ac.jp).

Keigo Watanabe is with the Department of Intelligent Mechanical Systems, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan (e-mail: watanabe@sys.okayama-u.ac.jp).

978-1-4244-6676-4/10/$25.00 ©2010 IEEE

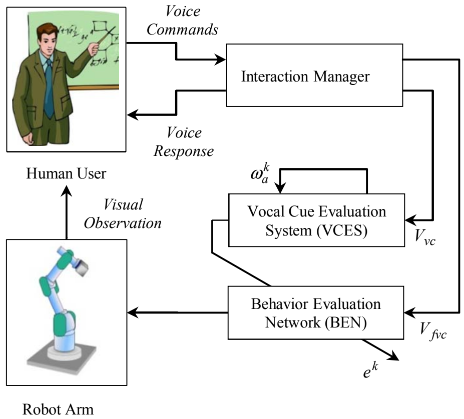

Fig. 1. Functional overview of the proposed system.

(V_{fvc}) and a vocal cue (V_{vc}). The vocal cues and the FVCs are fed into the VCES and the BEN respectively.

The fuzzy linguistic terms are assessed by modifying the robot's perception toward the user's perception of the environmental conditions. In the adaptation phase, it is fulfilled by navigating the end-effector of the robot manipulator for a selected set of tasks and acquiring the user's perception on the robot movements by vocal cues. After each movement, the BEN is adapted based on the evaluated value of the vocal cue and the user's willingness to change the robot's perception that is acquired by considering a series of vocal cues. Here, the modified value of the movement error e^k is obtained by modifying the evaluated value of the vocal cue based on the user's willingness to change the robot perception ω_a^k for ath fuzzy linguistic term for kth action group. Likewise, the adaptation process is continued by performing the tasks until the user is satisfied on the robot's movements for the FVCs. After the adaptation phase, the system is capable of understanding the fuzzy implications in the FVCs based on the current environmental conditions as expected by the user.

III. Evaluation of Fuzzy Linguistic Information

The fuzzy linguistic information is evaluated by adapting the BEN based on the evaluated value of the movement error from the VCES. The end-effector movements of the robot manipulator are considered as the primitive behaviors. Fuzzy voice motion commands that affect only the end-effector are used to activate the primitive behaviors [7]. The detailed descriptions of the possible set of fuzzy voice motion commands are shown in Table I. The user has been given the freedom to select any combination of action and action modification according to the expertise knowledge on the situation. Here, the lexical symbol "medium" is omitted in the user commands.

A. Behavior Evaluation Network

The fuzzy linguistic terms are evaluated and quantified based on the previous output of the corresponding action. The proposed BEN is implemented by adapting the system into a fuzzy-neural network architecture and the proposed

TABLE I
Fuzzy Voice Motion Commands

Action Groups	Actions	Modification
G1	Move forward Move backward	very little little (medium) far
G2	Move left Move right	
G3	Move up Move down	

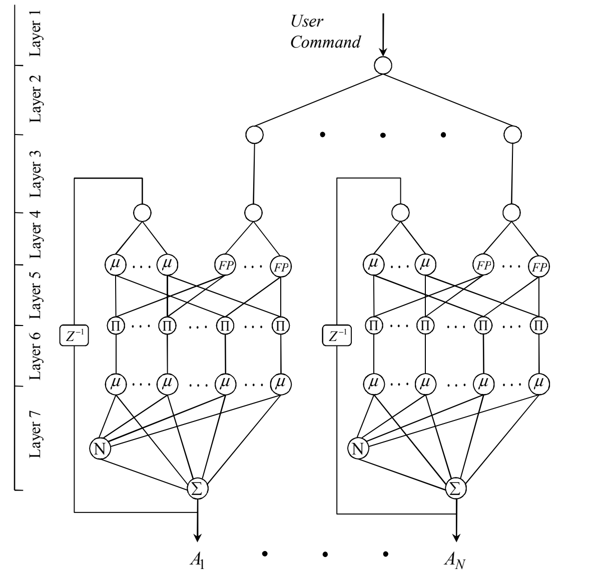

Fig. 2. Fuzzy-neural network for BEN. Primitive behaviors are evaluated and quantified based on a fuzzy predicate and the previous output of the corresponding action A_i where i = 1,2,3.

structure is shown in Fig. 2. The available actions are grouped into three action groups by considering the similarity of movements as shown in Table I. Separate behavior evaluation sections are proposed for each action group in the fuzzy-neural network and included from layer 3 to layer 7.

Layer 1 transmits the user commands directly to the next layer. Layer 2 acts as an action selection layer. Any node i activates the output depending on the existence of the corresponding action A_i in the user command where i = 1,2,3. Then, the corresponding fuzzy predicate is transferred to the next layer. Layer 3 consists of two types of nodes; one is a command node to pass the fuzzy predicate included in the user command and the other is a node to acquire the previous value of the corresponding action. Layer 4 acts as the fuzzification layer of the fuzzy-neural network. In layer 5, the nodes labeled as $\prod$ compute the T-norm by taking the algebraic product between the incoming signals. The output of any node s represents the firing strength of the sth rule, i.e., u_s^k. Layer 6 links the fuzzy antecedent part to the consequent part. Any node t of kth action represents a triangular membership function with center $a_t^k \in [(a_t^k)_L, (a_t^k)_H]$ and width $b_t^k \in [(b_t^k)_L, (b_t^k)_H]$ where t = 1,…,T. The nodes in the final layer generate the output and act as the defuzzification layer. Then the following function can be used to simulate the center of area method of defuzzification. It can be

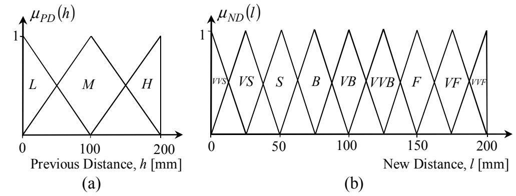

Fig. 3. (a) and (b) represent the initial membership functions for the previous distance and the new distance. The fuzzy labels are defined by, *L*: Low, *M*: Medium, *H*: High for *h* and *VVS*: Very Very Small, *VS*: Very Small, *S*: Small, *B*: Big, *VB*: Very Big, *VVB*: Very Very Big, *F*: Far, *VF*: Very Far, *VVF*: Very Very Far, for *l*.

formulated based on the sum-product composition for Mamdani fuzzy system [13], [14]:

$$A_k = \frac{\sum_{t=1}^{t=T} u_t^k a_t^k b_t^k}{\sum_{t=1}^{t=T} u_t^k b_t^k} \tag{1}$$

The initial membership functions for the previous movement ($\mu_{PD}(h)$) and the new movement ($\mu_{ND}(l)$) are defined by assuming the uniform distribution over the universe of discourse. They are illustrated in Fig.3. Here, the membership function for the previous distance is used to initialize the corresponding parameters of the layer 4. Likewise, the parameters of the layer 7 are also initialized based on the parameters of the $\mu_{ND}(l)$. The connection weights of the layer 7 are adjusted based on the user's guidance.

B. Vocal Cue Evaluation System

The VCES is introduced to evaluate the vocal cues and assess the movement error. A vocal cue includes user's directives to modify the robot behaviors. As an example, the vocal cue "too large" includes the meaning that the robot movement for the corresponding user command is too large according to the environmental arrangement. Furthermore, user asks to adapt the system to obtain a lower output. Therefore, robot can extract the required degree of adjustments for more suitable movement.

The VCES is realized by using a fuzzy inference system. The fuzzy linguistic information in the vocal cue is interpreted by assuming that the behavior change depends on the robot movement observed by the user. Therefore, the observed robot movement and the vocal cue are considered as the inputs of the fuzzy inference system. The evaluated error is the output. The vocal cue consists of a set of command components (i.e. "too large," "slightly large," "good," "slightly small," and "too small") that are considered as singleton membership functions. The membership functions for the robot movement and the corresponding evaluated error and the rule base are used as in [12]. The center of area method of defuzzification is used to obtain the output [13].

C. Adaptation of Behavior Evaluation Network Guided by User's Willingness

The BEN is adapted toward the environmental conditions according to the user's guidance based on the vocal cues and the user's willingness to change the robot's perception. The movement error is modified based on the user's willingness to change the robot's perception. In the training phase, a selected set of tasks are performed and the corresponding vocal cues are also acquired. After each movement, the parameters of the BEN are modified based on the modified movement error. Likewise, the training process is continued until the user feels a satisfactory level of the robot's movements for the FVCs.

The state of satisfaction for the robot's movements is identified by considering the percentage of the vocal cue "good" in a sequence of previous movements. The user's comment "good" implies that the user has satisfied about the robot's movement and the assessed value of the corresponding fuzzy predicate. Therefore, the number of vocal cue "good" in a certain sequence of movements represents a measure about the overall satisfaction of the user. However, the user's satisfaction for robot's movements in a certain instant depends on the action group. Therefore, satisfactory levels of the user towards the robot's movements are considered separately. The satisfactory level for *k*th action group $\xi_k \in [0,1]$ is defined by

$$\xi_k = \frac{N_G^k}{N_T^k} \tag{2}$$

where N_G^k is the number of vocal cue "good" in the robot movements of action group *k* from the number of total vocal cues N_T^k in the recent previous movements.

The user's willingness to change the robot's perception is identified as a parameter that can extract the user's motivation to change an assessment for a particular fuzzy linguistic term by considering a series of vocal cues. Therefore, the rate of adaptation is improved by acquiring the information related to the repetition of the vocal cues. According to the user's guidance, the required adaptation is different for each fuzzy linguistic term. A separate parameter is selected for each fuzzy linguistic term in each action group. As an example, two vocal cues, i.e., "too small," "too small," for the FVCs with fuzzy linguistic term "far" in action group G1 is considered. Here, the second vocal cue gives more emphasis on the required modification than the first. Therefore, the information included in the repetition of the vocal cues is utilized to find the user's willingness. The user's willingness to change the robot's perception for *a*th fuzzy predicate in *k*th action group is defined by

$$\omega_a^k = (1-\xi_k)\sum_{m=1}^{N_T^k} \beta_m \tag{3}$$

where

$$\beta_m = \begin{cases} 0.6 & \text{if } V_{vc} \text{ is "too small"} \\ 0.2 & \text{if } V_{vc} \text{ is "slightly small"} \\ 0.0 & \text{if } V_{vc} \text{ is "good"} \\ -0.2 & \text{if } V_{vc} \text{ is "slightly large"} \\ -0.6 & \text{if } V_{vc} \text{ is "too large"} \end{cases} \tag{4}$$

Here, the component for the user's willingness to change the robot's perception from the vocal cue related to *m*th previous movement is β_m and the corresponding values are selected experimentally. The user's willingness is defined as a function of the satisfactory level to consider the interrelation between them.

Adaptation of the BEN is implemented by training the parameters of the membership functions for the new distance ($\mu_{ND}(l)$) corresponding to the parameters of the layer 7. The membership parameter trainings corresponding to the network weight training for the tth node of kth action at the ($\ell+1$)th time step are given by

$$a_t^k(\ell+1)=\begin{cases} a_t^k(\ell)+\eta e_1^k u_t^k & \text{if } a_t^k(\ell+1)\in\left[\left(a_t^k\right)_L,\left(a_t^k\right)_H\right] \\ a_t^k(\ell) & \text{if otherwise}\end{cases} \tag{5}$$

$$b_t^k(\ell+1)=\begin{cases} b_t^k(\ell)+\eta e_2^k u_t^k & \text{if } b_t^k(\ell+1)\in\left[\left(b_t^k\right)_L,\left(b_t^k\right)_H\right] \\ b_t^k(\ell) & \text{if otherwise}\end{cases} \tag{6}$$

where η represents the learning rate. e_1^k and e_2^k are the modified values of the movement error e, which are decided by the VCES based on the vocal cue and calculated by

$$e_i^k=\delta_i\left(1+\left|\omega_a^k\right|\right)\left(1-\xi_k\right)e,\ i=1,2 \tag{7}$$

in which δ_1 and δ_2, where $0<\delta_1,\ \delta_2\le 5$, are defined to match the corresponding ranges. ξ_k is the level of user's satisfaction for robot's movements in action group k. ω_a^k is the user's willingness to change the robot's perception for ath fuzzy predicate in kth action group.

In addition, we have assumed that the vocal cue includes certain guidance to adapt the neighboring membership functions also from a reduced amount. An excitatory function is introduced to improve and accelerate the adaptation process. Therefore, the parameters for the nth neighboring node of tth node, where $n = t–2,\ t–1,\ t+1,\ t+2$, are considered. The membership parameter trainings for the neighboring nodes of tth node for kth action at the ($\ell+1$)th time step are given by

$$a_n^k(\ell+1)=\begin{cases} a_n^k(\ell)+\eta e_1^k u_t^k & \text{if } a_n^k(\ell+1)\in\left[\left(a_n^k\right)_L,\left(a_n^k\right)_H\right] \\ a_n^k(\ell) & \text{if otherwise}\end{cases} \tag{8}$$

$$b_n^k(\ell+1)=\begin{cases} b_n^k(\ell)+\eta e_2^k u_t^k & \text{if } b_n^k(\ell+1)\in\left[\left(b_n^k\right)_L,\left(b_n^k\right)_H\right] \\ b_n^k(\ell) & \text{if otherwise}\end{cases} \tag{9}$$

where the modified values of the movement error are calculated by

$$e_i^k=\delta_i\left(1+\left|\omega_a^k\right|\right)f_k e,\ i=1,2 \tag{10}$$

in which the excitatory function f_k for kth action group is defined by

$$f_k=\begin{cases} 0.2\left(1-\xi_k\right) & \text{if } n=t-2 \\ 0.6\left(1-\xi_k\right) & \text{if } n=t-1 \\ 0.6\left(1-\xi_k\right) & \text{if } n=t+1 \\ 0.2\left(1-\xi_k\right) & \text{if } n=t+2 \end{cases} \tag{11}$$

where the excitatory function f_k for kth action group is defined as a function of the satisfactory level to avoid the unnecessary distortions in the final steps of the training phase.

IV. Results and Discussions

A situation of cooperative rearrangement of the user's working space is considered to illustrate the system. The working space consists of a working table, a storing table, and a rack. The length, the width, and the height of the working table, the storing table, and the rack were selected as $0.5\times0.8\times0.17\,m^3$, $0.45\times0.3\times0.34\ m^3$, and $0.12\times0.33\times0.3\ m^3$ respectively. The proposed system was implemented based on a PA-10 robot manipulator with 7-DOFs [15]. In addition, the system was equipped with the robot manipulator and the controller, a personal computer running on Windows operating system, and a microphone. The parameters related to the adaptation of the BEN were chosen as $\eta = 0.1$, $\delta_1 = 1.5$, and $\delta_2 = 4.5$ experimentally.

In the adaptation phase, the user commented on the robot's movements by visually observing the demonstration. Likewise, a series of tasks (i.e. sequences of primitive movements) were performed under the guidance of the user. The adaptation process was continued until the user satisfied with the robot's movements. The variation of the satisfactory levels by considering the user's willingness to change the robot's behavior and one without considering it are shown in Fig. 4. The user's satisfaction was identified based on a satisfactory limit of 90% (i.e. for all $\xi_k = 0.9$ where $k = 1,2,3$). The number of steps required in the adaptation phase to achieve the user's satisfaction is drastically reduced with the consideration of the user's willingness.

The final set of parameters for the membership functions of the new distance $\mu_{ND}(l)$ corresponding to the parameters of the Layer 7 of the BEN after the adaptation is shown in Table II. Here, the center $a_t^k\in\left[\left(a_t^k\right)_L,\left(a_t^k\right)_H\right]$ and the width $b_t^k\in\left[\left(b_t^k\right)_L,\left(b_t^k\right)_H\right]$ of each membership value of the fuzzy variable for the new distance are presented. The lower and upper bounds of the membership parameters were selected experimentally [12]. The universe of discourse of the membership functions for the previous distance $\mu_{PD}(h)$ in the BEN and the membership functions for the observed robot movement in the VCES were also modified in each adaptation step. They were modified based on the universe of discourse of the membership function for the new distance $\mu_{ND}(l)$. Initially, the universe of discourse was selected as [0, 200] for all the action groups. After the adaptation phase, each universe of discourse for membership functions in action groups G1, G2, and G3 was changed to [0.00, 247.86], [0.00, 349.25], and [0.00, 179.72] respectively.

Few analogous tasks were performed before and after the adaptation process and analyzed to elaborate the proposed method. Here new set of tasks that was not used in the adaptation phase is considered. The selected sequences of movements for the above-mentioned tasks are illustrated in Fig. 5. The detailed descriptions of the user commands and the corresponding end-effector positions of the tasks EF and $\overline{EF}$ are shown in Table III and Table IV respectively. The fuzzy implications "medium" and "far" were mainly used for controlling the end-effector of the robot manipulator

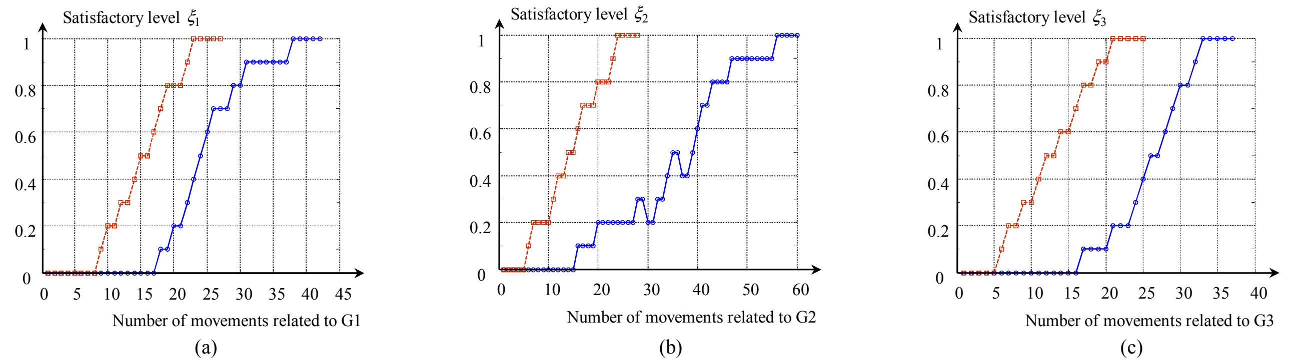

Fig. 4. (a), (b), and (c) represents the variation of the satisfactory level with the number of movements for action groups G1, G2, and G3 respectively. The variation without considering the user's willingness to change the robot's perception is represented by the blue continuous line. The red dotted line shows variation of the satisfactory level after considering the user's willingness. N_T^k is selected as 10 previous movements.

TABLE II
THE FINAL SET OF PARAMETERS FOR MEMBERSHIP FUNCTIONS OF NEW DISTANCE AFTER THE ADAPTATION

Action Groups	Parameters of the membership function for new distance [mm]								
	VVS	*VS*	*S*	*B*	*VB*	*VVB*	*F*	*VF*	*VVF*
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
(Initial value)	0.00	25.00	50.00	75.00	100.00	125.00	150.00	175.00	200.00
G1	0.00	11.41	25.36	70.62	132.42	177.99	208.07	227.01	247.86
G2	0.00	11.23	29.80	72.61	128.76	197.49	272.46	308.52	349.25
G3	0.00	10.81	19.37	34.88	62.99	101.62	135.74	156.92	179.72
	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
(Initial value)	25.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	25.00
G1	16.88	28.34	34.07	47.08	71.11	84.32	87.21	82.67	39.70
G2	19.06	36.05	37.03	48.41	68.67	97.33	130.14	115.12	73.16
G3	17.08	33.12	29.25	23.26	25.66	35.08	41.49	45.81	23.97

in coarse movements. As an example, the robot movement for the user command "move far left" is considered from the above-mentioned example and it is marked by † in Table III and Table IV. Here, the adaptation has resulted an increment of the robot movement for the above-mentioned command from 165.88 mm to 272.35 mm. It means that the user has expected a larger movement for that command according to the environmental conditions. The fine movements were handled by using the fuzzy predicates like "little" and "very little." The reduction in the evaluated value for the user command "move very little down" (i.e., marked by ‡ in Table III and Table IV) can be explained similarly. According to the results, the manipulator movements for the user commands containing the fuzzy linguistic terms "very little" and "little" were reduced in all the action groups. The evaluated values of the user commands containing fuzzy linguistic terms "medium" and "far" were increased in the movements of action groups G1 and G2 and reduced in the movements of action group G3. Therefore, the capability to use the FVCs for navigating the end-effector of the robot manipulator for coarse and fine movements was improved. This particular result implies that the robot movements for the FVCs were performed as expected by the user according to the contextual information. Thereby, the number of voice commands required to complete a particular task was reduced.

Before the adaptation phase, the robot's movements for the user commands are different from the expected values of the user. Therefore, there is a possibility to occur overshoots in the movements. Such unexpected movements were occurred in the above-mentioned example and the voice commands, which were used to correct the movement, are pointed by § in Table III. As an example, in task EF, voice command "move very little up" has been issued after the voice command "move very little down" to correct such an unexpected movement. The results verified that the adaptation of the system for understanding the fuzzy linguistic information toward the environmental conditions reduced the possibility to occur such overshoots in the navigation. Finally, the adaptation of the system based on the environment improved the overall usability of the system.

V. CONCLUSION

This paper has proposed a method for understanding the FVCs by evaluating the fuzzy linguistic information based on the user's guidance. The robot's perception on the fuzzy linguistic information was adapted based on a series of vocal cues. The adaptation process was significantly improved due to the consideration of the user's willingness to change the robot's perception based on a series of vocal cues. Thereby the number of steps required in the adaptation phase to

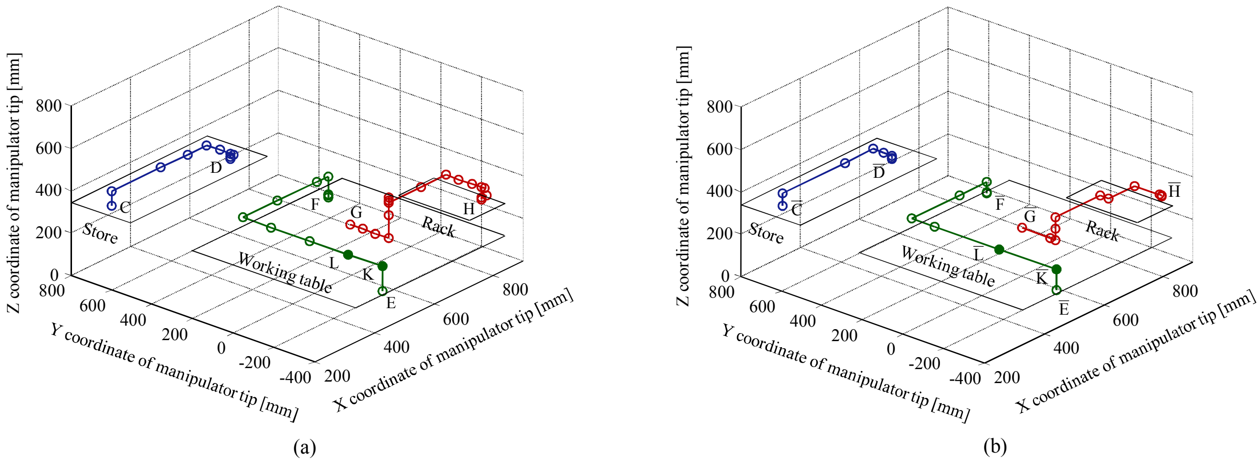

Fig. 5. (a) represents manipulator tip movements for the tasks CD, EF, and GH, which were performed before the adaptation. (b) represents manipulator tip movements for the tasks $\overline{CD}$, $\overline{EF}$, and $\overline{GH}$, which were performed after the adaptation.

TABLE III
MOVEMENTS FOR TASK EF BEFORE THE ADAPTATION

User command	Tip position [mm]			l[mm]
	x	y	z	
(Point E)	463.44	–368.69	152.02	
Move up	463.37	–368.69	267.86	115.84
Move far left †	463.38	–202.81	267.82	165.88
Move far left	463.30	–13.99	267.69	188.82
Move far left	463.30	176.73	267.65	190.72
Move left	463.29	317.62	267.62	140.89
Move forward	572.08	317.76	267.73	108.79
Move forward	699.79	317.62	267.77	127.71
Move very little forward	735.99	317.64	267.79	36.20
Move little down	736.08	317.64	183.21	–84.58
Move very little down ‡	736.08	317.59	167.81	–15.40
Move very little up §	736.06	317.62	177.39	9.58

TABLE IV
MOVEMENTS FOR TASK $\overline{EF}$ AFTER THE ADAPTATION

User command	Tip position [mm]			l[mm]
	x	y	z	
(Point $\overline{E}$)	463.04	–369.06	153.58	
Move up	463.01	–369.05	224.94	71.36
Move far left †	462.97	–96.70	224.83	272.35
Move far left	462.96	209.53	224.73	306.23
Move little left	462.93	319.93	224.73	110.40
Move forward	642.00	320.05	224.79	179.07
Move little forward	735.35	319.87	224.79	93.35
Move little down	735.35	319.94	183.93	–40.86
Move very little down ‡	735.34	319.96	175.25	–5.68

achieve the user's satisfaction was drastically reduced. The proposed system is capable of evaluating the fuzzy implications in a contextual dependent manner. It enabled the capability to adapt the robotic supporter toward the environmental conditions. Therefore, the user's capability of using voice commands including the fuzzy linguistic information for coarse and fine movements was also enhanced.

REFERENCES

[1] P. Menzel and F. D'Aluisio, *Robo sapiens: evolution of a new species*, The MIT Press, England, 2000.

[2] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura, "The intelligent ASIMO: system overview and integration," in *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, Switzerland, Oct. 2002, pp. 2478–2483.

[3] M. Fujita , "On activating human communications with pet-type robot AIBO," *Proc. IEEE*, vol. 92, no. 11, pp. 1804–1813, Nov. 2004.

[4] R. Alqasemi and R. Dubey, "Maximizing manipulation capabilities for people with disabilities using a 9-DoF wheelchair-mounted robotic arm system," in *Proc. of the IEEE Int. Conf. on Rehabilitation Robotics*, Noordwijk, Jun. 2007, pp. 212–221.

[5] Y. Mohammad and T. Nishida, "Human adaptation to a miniature robot: Precursors of mutual adaptation," in *Proc. of the IEEE Intl. Symp. on Robot and Human Interactive Communication*, Munich, Germany, Aug. 2008, pp. 124–129.

[6] K. Pulasinghe, K. Watanabe, K. Izumi, and K. Kiguchi, "A modular fuzzy-neuro controller driven by spoken language commands," *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 34, pp. 293–302, 2004.

[7] C. Jayawardena, K. Watanabe, and K. Izumi, "Controlling a robot manipulator with fuzzy voice commands using a probabilistic neural network," *Neural Computing and Applications*, vol. 16, pp. 155–166, 2007.

[8] C. Jayawardena, K. Watanabe, and K. Izumi, "Posture control of robot manipulators with fuzzy voice commands using a fuzzy coach-player system," *Advanced Robotics*, vol. 21, pp. 293–328, 2007.

[9] A. Chatterjee, K. Pulasinghe, K. Watanabe, and K. Izumi, "A Particle-Swarm-Optimized Fuzzy-Neural Network for Voice-Controlled Robot Systems," *IEEE Trans. on Industrial Electronics*, vol. 52, pp. 1478–1489, 2005.

[10] A.G.B.P. Jayasekara, K. Watanabe and K. Izumi, "Evaluating fuzzy voice commands by internal rehearsal for controlling a robot manipulator," in *Proc. of the ICROS-SICE International Joint Conference*, Fukuoka, Japan, Aug. 18–21, 2009, pp. 3130–3135.

[11] A.G.B.P. Jayasekara, K. Watanabe and K. Izumi, "Interactive dialogue for behavior teaching to robots based on primitive behaviors with fuzzy voice commands," in *Proc. of the Int. Conf. on Information and Automation for Sustainability*, Colombo, Sri Lanka, Dec. 12–14, 2008, pp. 1–6.

[12] A.G.B.P. Jayasekara, K. Watanabe, K. Kiguchi, and K. Izumi, "Adaptation of robot behaviors toward user perception on fuzzy linguistic information by fuzzy voice feedback," in *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication*, Toyama, Japan, Sep. 27–Oct. 2, 2009, pp. 395–400.

[13] J.S.R. Jang, C.T. Sun, and E. Mizutani, *Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, India, 1997.

[14] C.T. Lin and C.S.G. Lee, "Reinforcement structure/parameter learning for neural-network-based fuzzy logic control systems," *IEEE Trans. on Fuzzy Systems*, vol. 2, pp. 46–63, 1994.

[15] Mitubishi Heavy Industries Ltd., *PA-10 Portable General Purpose Intelligent Arm Programming Manual–rev.1*.