The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Adaptation of Robot Perception on Fuzzy Linguistic Information by Evaluating Vocal Cues for Controlling a Robot Manipulator

A.G.B.P. Jayasekara
Dept of Advanced Systems Control Engineering
Saga University
1 Honjo machi, Saga 840-8502, Japan.
buddhika@ieee.org

Keigo Watanabe
Dept of Intelligent Mechanical Systems
Okayama University
3-1-1 Tsushima-naka, Kita-ku
Okayama 700-8530, Japan.
watanabe@sys.okayama-u.ac.jp

Kazuo Kiguchi and Kiyotaka Izumi
Dept of Advanced Systems Control Engineering
Saga University
1 Honjo machi, Saga 840-8502, Japan.
{kiguchi,izumi}@me.saga-u.ac.jp

Abstract

This paper proposes a method for adapting robot's perception on fuzzy linguistic information by evaluating vocal cues. Robot's perception on fuzzy linguistic information such as "very little" depends on the environmental arrangement and the user's expectations. Therefore robot's perception on the corresponding environment is modified by acquiring user's perception through vocal cues. Fuzzy linguistic information related to primitive movements are evaluated by a behavior evaluation network (BEN). Vocal cue evaluation system (VCES) is utilized to evaluate the vocal cues for modifying the BEN. The proposed system is implemented by a PA-10 robot manipulator.

1 Introduction

Voice communication is a better option among the other alternatives for human-robot interaction. Ability of the robot companion to understand the uncertain information is crucial for effective human-robot interaction. A successful human-friendly robot equipped with human-like voice communication capabilities will be able to help disabled people, to assist the aged people, to help in complex tasks such as surgery, etc. [1].

In Pulasinghe et al. [2], robot control by using information-rich voice commands such as "move a very little forward" has been studied. Generally, the voice commands, which include fuzzy linguistic information are referred as fuzzy voice commands (FVCs) [3]. The quantitative assessment for such 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. Normally, humans possess a natural ability of adapting to other humans and artifacts. The mutual adaptation is important in acquiring the information that included in the voice commands in human-human communications [4].

Therefore this paper proposes a method for interpreting fuzzy linguistic information by adapting the robot's perception based on user's perception on corresponding environment. Here, the user's perception is acquired based on vocal cues. The system overview is discussed in section 2. Next section 3 discusses the adaptation process of the robot's perception based on vocal cues. Finally, summary is presented.

2 System Overview

We have proposed a method to understand FVCs by evaluating fuzzy linguistic information based on user's guidance. Functional overview of the proposed method is shown in Fig. 1. Vocal cues and FVCs are fed into the VCES and the BEN respectively. Here, end-effector movements of a 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. If the user command is a vocal cue, the BEN is adapted based on the assessment of the movement error e. If the user command is an FVC, the corresponding primitive behavior is activated and fuzzy

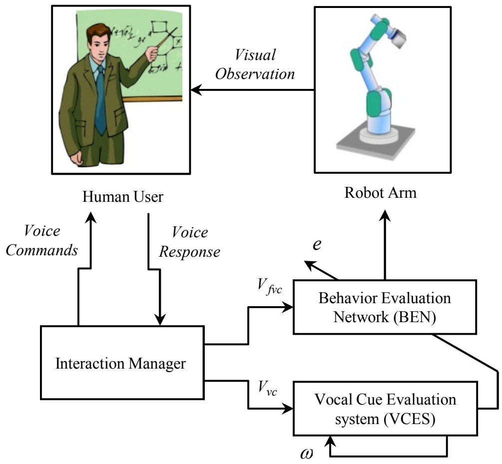

Figure 1: Functional overview of the proposed system.

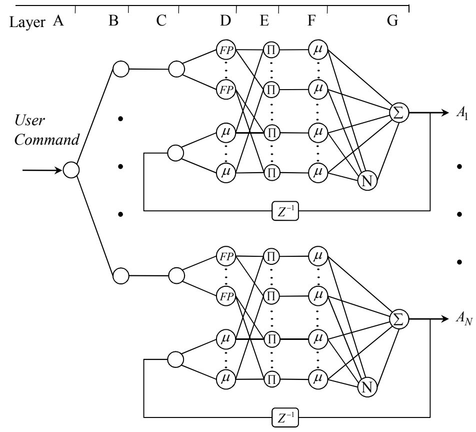

Figure 2: Fuzzy-neural network for the BEN.

Table 1: 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	

linguistic information is evaluated by the BEN. In addition, the user's willingness to change the robot's perception is identified as a parameter that is capable of improving the adaptation process by acquiring the information related to repetition of vocal cues. After the adaptation phase, the system can be used to navigate the tip of the robot manipulator by using FVCs.

3 Adaptation of Robot Perception

A quantitative assessment for a fuzzy linguistic term such as " little" highly depends on the environmental conditions. Therefore, it is proposed to acquire the user's desire through vocal cues for adapting the robot's perception toward the user's perception.

3.1 Behavior Evaluation Network

Fuzzy implications in FVCs are interpreted by the BEN. The corresponding primitive movement is identified based on the action and the action modification phrases. Fuzzy linguistic terms are evaluated and quantified based on the previous output of the corresponding action. The proposed structure is shown in Fig. 2. The available actions are grouped into three action groups by considering the similarity of movements and shown in Table 1. Separate behavior evaluation sections are proposed for each action group in the BEN [5].

Layer A transmits the user commands directly to the next layer. Layer B acts as an action selection layer. Layer C 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 D acts as the fuzzification layer of the fuzzy-neural network. Layer F 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 output can be formulated based on the sum-product composition for Mamdani fuzzy system [6]:

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

The initial membership functions for the previous movement and the new movement are defined by assuming the uniform distribution over the universe of discourse. Here, the membership function for the previous distance and the new distance are used to initialize the corresponding parameters of the layer D and layer G respectively. The connection weights of the layer G are adjusted by applying the backpropergation algorithm in the training phase based on the movement error.

3.2 Vocal Cue Evaluation System

The VCES is introduced to evaluate the vocal cues for assessing the movement error e. A vocal cue includes user's directives to modify the robot behaviors. 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 user feedback 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.

3.3 Adaptation of Behavior Evaluation Network

The BEN is adapted toward the environmental conditions based on the user's guidance. In the training phase, the BEN is adapted based on the movement error that is calculated by the VCES. The training process is continued by considering a selected set of tasks until the user feels a satisfactory level of robot's movements. Satisfactory level of the user toward the robot's movements for kth 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 user feedback "Good" in robot's movements of action group k from number of total vocal cues N_T^k, which is from recent previous movements.

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 user's willingness to change the robot's perception for ath fuzzy predicate in kth action group is defined by

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

Here, the component for user's willingness to change the robot's perception from the vocal cue related to mth previous movement is β_m.

Aadaptation of the BEN is implemented by training the parameters of the membership functions for new distance corresponding to the parameters of the layer G. The membership parameter training corresponding to the network weight training for the tth node and its neighboring nodes of kth action at the $(\ell+1)$th time step are given by

$$a_t^k(\ell+1) = \begin{cases} a_t^k(\ell) + \eta e_1^k u_t^k & \text{if } a_t^k(\ell+1) \\ & \quad \in [(a_t^k)_L, (a_t^k)_H] \\ a_t^k(\ell) & \text{otherwise} \end{cases} \tag{4}$$

$$b_t^k(\ell+1) = \begin{cases} b_t^k(\ell) + \eta e_2^k u_t^k & \text{if } b_t^k(\ell+1) \\ & \quad \in [(b_t^k)_L, (b_t^k)_H] \\ b_t^k(\ell) & \text{otherwise} \end{cases} \tag{5}$$

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

$$e_i^k = \delta_i(1+|\omega_a^k|)f_k e, \quad i=1,2 \tag{6}$$

δ_1 and δ_2, where $0 < \delta_1, \delta_2 \leq 5$, are defined to match the corresponding ranges. ω_a^k is the user's willingness to change the robot's perception for ath fuzzy predicate in kth action group. The excitatory function f_k for kth action group is defined accordingly.

4 Summary

A set of tasks, which consists of sequences of primitive movements were used in the adaptation phase. The proposed system was implemented based on 7-DOF PA-10 robot manipulator. The end-effector movements of the robot manipulator were used as the primitive behaviors. 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.

The adaptation process was continued until the user satisfied with the robot's movements. The user's satisfaction was identified based on a satisfactory limit of 90% (i.e. for all $\xi_k = 0.9$ where $k = 1,2,3$). Variation of the satisfactory levels of action group G2 by considering the user's willingness to change the robot's behavior and without considering it are shown in Fig. 3. The final set of parameters for the membership functions of new distance are shown in Table 2. The universe of discourse of the membership functions for previous distance is also adjusted accordingly.

According to the results, the manipulator movements for the user commands containing 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

Table 2: The final set of parameters for membership functions of new distance after the adaptation

Action group	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)	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	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
	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
(Initial)	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

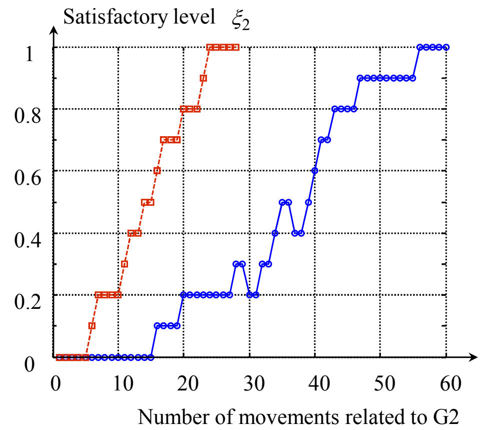

Figure 3: Variation of satisfactory level with number of movements for action group G2. The Variation without considering the user's willingness and with considering the user's willingness are represented by blue continuous line and Red dotted line respectively.

of action group G1 and G2 and reduced in the movements of action group G3. This particular result implies that the robot movements for FVCs were performed as expected by the user according to the contextual information. The user's capability of using voice commands including fuzzy linguistic information for coarse and fine movements is also enhanced. The number of voice commands required to complete a particular task was reduced. In addition, the possibility of occurring overshoots also minimized. Finally, the adaptation of the system based on the environment improved the overall usability of the system.

References

[1] P. Menzel and F. D'Aluisio, *Robosapiens*, The MIT Press, England, 2000.

[2] 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.

[3] 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.

[4] 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.

[5] 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. 2009, pp. 3130-3135.

[6] 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.