Asian Journal of Information Technology 4 (12):1201-1207, 2005

A Voice Guiding System for Autonomous Robots

Brahim Boulebtateche, Mohamed Fezari and Mouldi Bedda
Laboratory of Automatic and Signal, Faculty of Engineering, Department of Electronics
Badji Mokhtar University, Annaba BP.12, Annaba, 23000, Algeria

Abstract: In this study , a voice guidance system for autonomous robots is designed based on microcontroller. The proposed system consists of a microcontroller and a voice recognition processor that can recognize a limited number of voice patterns. The commands of autonomous robots are classified and organized such that one voice recognition processor can distinguish robot commands under each class. Thus, the proposed system can distinguish more voice commands than one voice recognition processor can. A voice command system for three autonomous robots is implemented with a microcontroller from Microchip PIC16F876, a voice recognition processor RSC364 from Sensory and a set of Infra-red emitters- receivers. A proposal is also outlined for integrating the voice command system into a reinforcement learning scheme in order to enhance the performance of learning task by autonomous robots. This work and design have taken 10 months, starting on january 2004.

Key words: Autonomous robot, microcontroller, voice recognition, reinforcement learning

INTRODUCTION

This new component which is the microcontroller, integrating within one chip all the necessary elements to design a mother board and being used in many different industrial, biomedical and robotics applications, is very important and essential in the design of small robots. Many techniques have been developed in order to enhance robot behaviour based on smart sensors and microcontrollers efficiently [1-4].

It has been known that the design of a mother board based on microprocessor is not an easy way, it takes time to design the board and cost more money to provide the necessary components. Therefore, the intelligent robots or the intelligent systems implemented by microcontrollers can reduce time of development.

There have been many research projects dealing with robot control, among these projects, there are some projects that build intelligent systems[5-7]. Since we have seen human-like robots in science fiction movies such as in I ROBOT movie, making intelligent robots or intelligent systems became an obsession within the research group. A micromouse[5] is a computer controlled autonomous mobile robot that can find a predetermined destination when placed in an unknown maze. A vision-guided autonomous vehicle is proposed as an alternative of the micromouse competition[6]. The track of a vision-guided autonomous vehicle can be easily constructed. Most of components are off-the-shelf products and the level of challenge on software can vary to suit the designer. However, since image processing and control of the car are accomplished by a personal computer, a vision-guided autonomous vehicle project is more appropriate in courses related to computer vision, artificial intelligence and fuzzy logic rather than those related to microcontrollers. In the MIROSOT(Micro -Robot World Cup Soccer Tournament), intelligent autonomous robots compete with each other in the soccer game[7]. But the soccer robot system requires a colour vision system that consists of a personal computer, a colour camera and a colour vision board. In this study , a voice command system for autonomous robots, based on the use of low-cost microcontrollers, is proposed. Compared with previous projects by[5-7], the cost of the proposed voice command system is much lower. Also, the variation of the proposed voice command system and other applications besides autonomous robots is possible depending on characteristics of the microcontroller and different sensors.

In order to command autonomous robots by voices as in (Fig.1), a voice recognition system has to be designed. There is a lot of research about designing voice recognition systems. Most of them have been built in workstations, personal computers, or DSP (Digital Signal Processing) chips by developing voice recognition algorithms. Depending on which algorithm is applied, voice recognition systems of different characteristics can be made. Therefore, the voice recognition system built in workstations, personal computers, or DSP chips can achieve various characteristics. Also, they can achieve

Corresponding Author: Brahim Boulebtateche, Laboratory of Automatic and Signal, Department of Electronics, Faculty of Engineering,Badji Mokhtar University, Annaba P.12, Annaba, 23000, Algeria

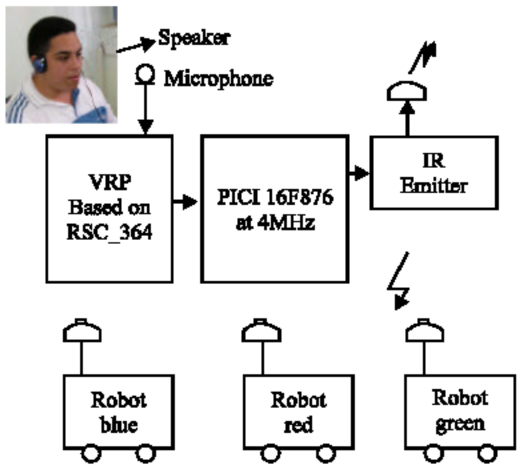


Fig.1: Voice command system for autonomous robots

high voice recognition rate under a general environment.

However, because techniques about voice recognition built in workstations, personal computers, or DSP chips are complicated and needs more time for development, a voice recognition processor is proposed in this study . Which is a dedicated processor that can recognize voices and therefore if the voice recognition processor is used, a voice recognition system can be built without knowing any voice recognition algorithms.

The proposed voice recognition system is used as a voice command system for autonomous robots. The effectiveness of the voice input compared with other input devices in robot systems was discussed in[8,11]. In order to command the tasks of autonomous robots by voices, the same number of words as tasks has to be recognized. Since one voice recognition processor can only recognize a limited number of voices (words) , the tasks of autonomous robots more than the limited number cannot be distinguished if one voice recognition processor is used alone. In this study, a RISC architecture microcontroller is combined with one voice recognition processor for controlling different robots[12,13].

Given the tasks of autonomous robots, each robot task is classified and is organized such that the number of robot tasks under each class of commands is not more than the maximum recognition number of one voice recognition processor. Then one voice recognition processor can distinguish robot tasks under each class. In addition to the hardware details on both voice command system and autonomous robots design, a section in voice command integration in autonomous robot learning is discussed[14-18],

THE PROPOSED VOICE COMMAND SYSTEM

The proposed voice command system consists of a special low power microcontroller and a voice recognition processor VRP that can recognize a limited number of words. Let the maximum number of recognizable words by this VRP be $N > 2$, the number of commands to be given to a robot are $p > 2$, then the maximum number of robots to be controlled by this VRP is $n=N/p$ (Fig. 2). So the orders can be grouped into sets, each set of order is assigned to a selected robot. In this case each robot can be affected different tasks. The VRP works in two phases: The training phase and the recognition phase or verification phase. In the training phase, the operator will be asked to pronounce command words one by one. The VRP extracts patterns from acoustic signal and store them in its internal memory. During this phase, the operator might be asked to repeat a word many times, specially if the word pronunciation is quite different from the first time.

In the recognition phase, given a voice command, the VRP recognizes the voice command by matching the most similar pattern in its memory unit and then sends the recognition result to the microcontroller. According to the recognition result, the microcontroller selects a group and provides another memory unit to the VRP. This memory unit contains the voice commands under the selected group and therefore the VRP waits for the next voice command under the selected group. If a voice command *is* recognized by the VRP, the microcontroller makes a command to a robot to perform the corresponding task and the VRP waits for the following voice command.

VOICE COMMAND FOR THREE AUTONOMOUS ROBOTS

A voice command system and three autonomous robots that receive voice commands are implemented. The voice commands of robots are selected from a set of actions used to control a vehicle (forward, backward, left,

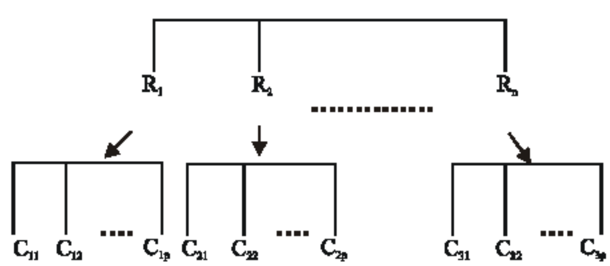


Fig.2: The structure of voice command for n robots

A 3	A 2	A 1	A 0	C 7	C 6	C 5	C 4	C 3	C 2	C 1	C 0

Fig.3: The frame format

Table 1: The meaning of voice commands

Red	Choose the first robot, red colour
Blue	Choose the second robot, blue colour
Green	Choose the third robot, green colour
Forward	Go forward, action on M1
Backward	Change direction back, action on M1
Right	Turn right, action on M2
Left	Turn left, action on M2
Stop	Stop the movement, stops M1 and M2

right and stop) and their meanings are listed as in Table 1. However, voice commands for the three robots may be made different from each other as mentioned early.

Voice command system description: The voice command system as shown in Fig.5 consists of a voice recognition processor RSC-364[19], a microcontroller from Microchip the PIC16F876[20], a Infra-red (IR) emitter module and a microphone. The RSC-364 is a speaker-dependent isolate-word recognition processor that can recognize up to 15 words in stand-alone mode , up to 60 words or phrases can be recognized in slave mode, the response time is less than 200 ms to a given voice input. The RSC family of microcontrollers are low-cost 8-bit designed for use in consumer electronics. All products of the RSC family are fully integrated and include a speech processor, A/D Converter, D/A converter, ROM as program memory and RAM circuitry on chip.

The RSC-364 also include on chip pre-amplification. To train and to recognize 15 words, an external 8K–byte EEPROM (Electrical Erasable Programmable Read Only Memory) needs to be connected to the RSC-364. No software programming of the RSC-364 is needed. The Microchip PIC16F876 is an 8 bit microcontroller designed to handle high-speed calculations and fast input/output operations. Since a PIC16F876 includes a five channel 10bit A/D converter, a serial port, Three 8 bit high-speed I/O and two PWM (pulse-width-modulated) modules, this microcontroller can be used to design a simple controller. Typical applications using the PIC16F876 include data acquisition system, robot motion control and domestic appliance control.

To transmit orders to different robots, an Infra-red LED (Light Emitting Diode) is used. In the implemented voice system, the PIC16F876 plays an important role as an interface between the VD364 and IR transmitter in sending commands to three autonomous robots.

Training: In order to train the recogniser in the manual mode of the VD-364, a push button is used to initialise the training mode, the system asks the user to enter the word, which should be entered into the RSC-364 through a microphone connected to the RSC-364. During training, the VRP stores patterns of each word in its memory.

Voice commands RED, BLUE and GREEN are used to select between the three robots, respectively. RED, BLUE and GREEN are the words representing the three colours of the robots. Since voice commands of each robot are stored in different memory areas, different voice commands may be trained depending on each robot.

Recognition: In the recognition mode, when the voice input through the microphone is detected, the VD-364 tries to find the most similar voice pattern to this input and sends a byte representing the recognition result to the PIC16f876. This byte is sent to the PIC16F876 through a TTL buffer 74LS245. Depending on this byte, the PIC16F876 sends an appropriate frame format of 12 bits as is shown in figure 3, where the four first bits represent the address of the robot chosen A3-A0 and the remaining bits represent the command to be executed by that robot C7-C0. This format allows the control up to 16 robots with 256 commands. Finally the frame is modulated and transmitted by the IR emitter.

AUTONOMOUS ROBOTS AND LEARNING

Hardware description of autonomous robots: The structures of the mechanical hardware and the computer board of autonomous robots in this study are similar to those[5,7] (Fig. 4). However, since our autonomous robots need to perform simpler tasks than those in [5][10] do, they can be more easily constructed. The computer board of each autonomous robot consists of a PIC16F876, with 8K-instruction EEPROM (Electrically Programmable Read Only Memory) and 256 bytes of DATA EEPROM memory, two H bridges drivers using BD134 and BD133 transistors for DC motors, an IR receiver module which is TSOP1736 and a four bit micro-switch to fix the address of each robot. Each autonomous robot performs the corresponding task in response to a received command as given in Table 1. Commands and their corresponding tasks recognized by related robots may be changed in order to enhance or modify the application.

Voice command integration in autonomous robot learning: Recently, a fully autonomous agent such as mobile robots has become one of the central issues in Artificial Intelligence. To reach this ultimate goal, many techniques have been used to deal with the problem of learning for autonomous robots. Among them, we may cite Reinforcement Learning scheme which plays an important role in attempting to perform such a difficult task. Reinforcement Learning is a mean of learning from

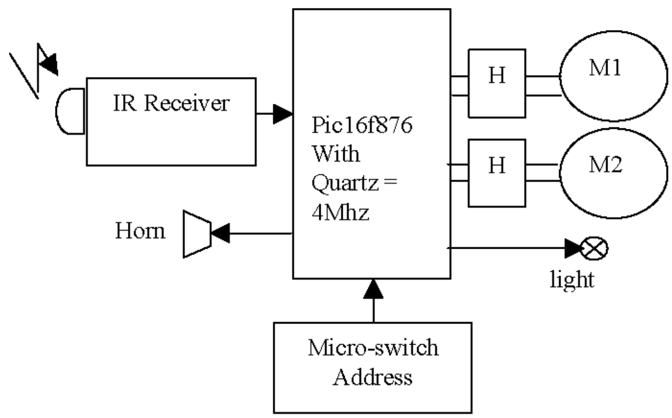


Fig. 4: Autonomous robot block diagram

Fig. 5: Overview of the voice command system

experiences that involve trial and errors only. The performance of the autonomous robot are then evaluated in terms of a scalar function (called reward function) which is available from the interaction of the robot with its environment.

The objective of a reinforcement learning algorithm is to maximise the expected cumulative reward after following a given policy of actions by updating some value function or utility function over a time horizon[14]. Those methods based on Q-learning , for

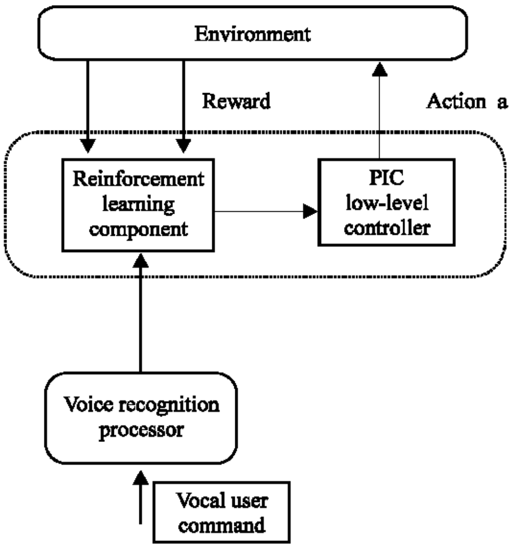


Fig. 6: Reinforcement learning process

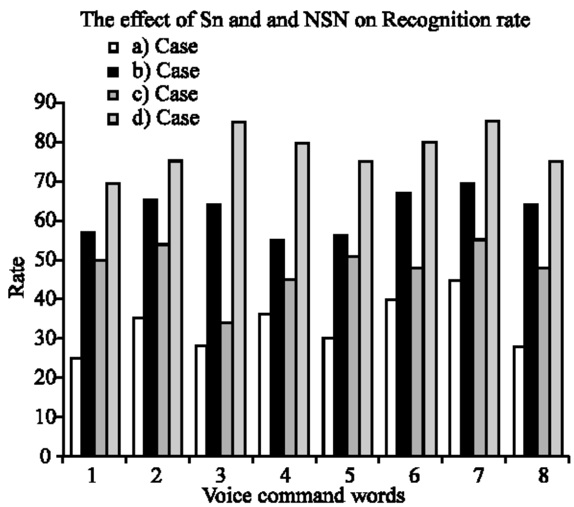


Fig. 7: The effect of SN or NSN in and out the laboratory.

instance, try to adjust an utility function that is used to establish a preference order within the action space available to each state of the robot[15]. Usually, the learned policy is then represented as the action with highest utility value for a given state. Thus, the goal of learning is to obtain an optimal policy which maps states to actions while optimising the expected utility function for each state.

Unfortunately, some difficulties have been encountered when attempting to apply reinforcement learning to real practical problems with a huge or infinite number of states, for example the problem of robots navigating in a complex real environment. For this kind of issues, learning can require a very large if not an infinite number of trials in order to establish an acceptable policy for autonomous robots. In addition, robots may behave in an unpredictable way that can lead to a disaster. To overcome some of these difficulties, many researchers have proposed to introduce some sort of supervisory into the learning loop in order simplify the learning problem to some extent[16,18]. Following this trend in our work, we intend to integrate the Voice Recognition Processor as a part of a reinforcement learning component within the overall system of the autonomous robots (Fig. 6).

The figure shows a schematic of our learning system where the user commands play the role of training input to the reinforcement learner. Q-learning, a well established technique in reinforcement learning theory, can be used here to estimate the utility function Q(s, a) by adjusting its value each time an action a is executed from a given state s to reach a new state s' according to the following formula:

$$Q_{new}(s, a) = Q_{old}(s, a) + \alpha(r + \gamma \max_{a'} Q(s', a') - Q_{old}(s, a))$$

Where r is the obtained reward at time t, α represents a step size parameter and γ is a discount factor between 0 and 1. Higher-level user commands in the form of intermittent input from the voice recognition processor are entered into the learning component serving as temporary guidance to choose more plausible actions for the autonomous robot. We believe that implementing reinforcement learning within this framework will greatly help the robot to acquire autonomy in shorter time. Details of the implementation of this proposed structure will be described during a future work.

TESTS ON THE VOICE COMMAND SYSTEM

The developed system has been tested within the laboratory of L.A.S.A, there were two different conditions to be tested: The distance of the microphone from the speaker and the rate of recognition in stationary noise and non stationary noise (NSN). The system, first, had been tested in the laboratory and outside in order to detect the environment effect on the recognition rate. After testing the recognition of each word 25 times in the following conditions: a) outside the Laboratory (LASA) with NSN, b) outside the LASA with stationary noise (SN), c) inside the LASA with NSN and d) inside the LASA with SN. The results are shown in figure 7 where the numbers in the horizontal axis correspond to the order of voice command words as they appear in Table 1.

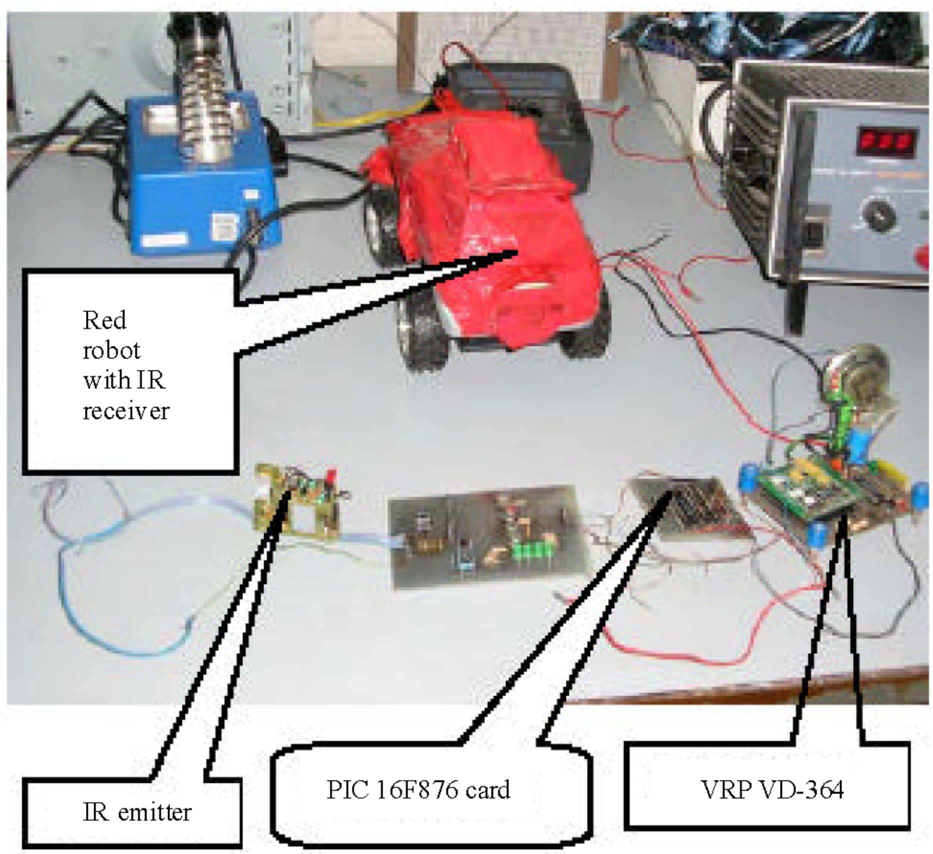
DISCUSSION

A voice command system for autonomous robots is proposed and is implemented based on microcontrollers. Since the proposed system consists of a VRP, a microcontroller and other low-cost components namely IR transmitters, the hardware design can be easily carried out. The results of the tests shows that a better recognition rate can be achieved inside the laboratory and specially if the phonemes of the selected word for voice command are quite different. However, a good position of the microphone and additional filtering may enhance the recognition rate. In order to recognize voice commands more than the maximum recognition number provided by one voice recognition processor, all voice commands can be grouped and organized such that the VRP always tries to find the most similar voice pattern from one memory unit. The number of the voice commands in one memory unit is not more than the maximum recognition number of one voice recognition processor. Several more interesting applications of the proposed system different from previous ones are also possible. Variation of the proposed voice command system to fit other applications besides autonomous robots is under study. The use of Radio Frequency transmission would give more hardware autonomy to the designed system. Our future work will focus also on developing the software part in order to implement the interaction between the supervisory aspect of the user interface through the voice recognition processor and the reinforcement learning component of the autonomous robot. In addition, we intend to assess the overall performance of the proposed structure in the context of multi-robots environment as concerning the response of each robot to vocal user commands issued at higher level of abstraction.

REFERENCES

1. Strangio, C.E., 1988, Microcontroller instruction in the engineering laboratory. IEEE Transactions on Education, 31: 172-176.
2. Simpson, R.C., D. Poirot and F. Baxter, 2002. The Hephaestus smart wheelchair system, IEEE Trans. on Neural Systems and Rehabilitation Engineering,10: 118-122,.
3. Beritelli, F., S. Casale and A. Cavallaro, 1998. A robust voice activity detector for wireless communications using soft computing. IEEE J. on Selected Areas in Communications (JSAC), special Issue on Signal Processing for Wireless Communications, 16: 435-439.
4. Hara, J.M.O. and R.E. Olsen, 1988, Control device effects on telerobotics manipulation operations, Robotics and Autonomous Systems, 4: 1-4.
5. Chen, N., H. Chung and Y.K. Kwon, 1995, Integration of micromouse project with undergraduate curriculum: A Large-Scale Student Participation approach, IEEE Transactions on Education, 38: 136-144.
6. Chen, N., 1997. A Vision-Guided autonomous vehicle: An alternative micromouse competition, IEEE Transactions on Education, 40: 253-258.
7. Kim, J.H., *et al,* 1997. Cooperative Multi-Agent robotic systems: From the Robot-Soccer perspective, 1997 Micro-Robot World Cup Soccer Tournament Proceedings, Taejon, Korea, pp: 3-14.
8. Suhm, B., B. Myers and A. Waibel, 2001. Multimodal error correction for speech user interfaces. ACM Transactions on Computer-Human Interaction, 8: 60-98.
9. Nicoud, J., 1996. Microrobots for research and applications. Micro-Robot World Cup Soccer Tournament Proceedings, Taejon, Korea, pp: 11-16.
10. Renevey, P., R. Vetter and J. Kraus, 2001, Robust speech recognition using missing feature theory and vector quantization. In: Proc. of 7th European Conference on Speech Communication Technology (EUROSPEECH-01), ESCA, 2: 1107–1110.
11. Nishimoto , T.*et al.,* 1993. Improving human interface in drawing tool using speech, mouse and Key-board. Proceedings of the 4th IEEE International Workshop on Robot and Human Communication, Tokyo, Japan, pp: 107-112.
12. Sheridan, T.B., 1993. Space teleoperation through time delay. Review and Prognosis, IEEE Transactions on Robotics and Automation, 9: 592-606.
13. Kim, W.J., *et al.,* 1998. Development of a voice remote control system, Proceedings of the 1998 Korea Automatic Control Conference, Pusan, Korea, pp: 1401-1404.
14. Schaal, S., 1999, Is Imitation Learning the route to humanoid robots?, Trends in Cognitive Science, 3:233-242.
15. Smart,W.D. and L. Kaebling. 2000. Practical reinforcement learning in continuous spaces. In: Proceedings of the International Conference on Machine Learning.
16. Sutton, R.S. and A.G. Barto, 1998. Reinforcement Learning : An Introduction , MIT Press, Cambridge, MA.

17. Watkins, C.J.C.H., 1989. Learning from delayed rewards, Ph.D. Dissertation, Cambridge University, Cambridge, England.
18. Dorigo, D. and M. Colombetti, 1994, Robot shaping: Developing Autonomous Agents Through Learning, Artificial Intelligence, 71 : 321-370.
19. RSC-364 Manual, Sensory Company.
20. Data sheet PIC16F876 from Microchip inc. User's Manual, 2001.


Red robot with IR receiver
IR emitter
PIC 16F876 card
VRP VD-364