ICIC Express Letters
Part B: Applications
Volume **14**, Number **2**, February **2023**
 ISSN 2185-2766
pp. **117–123**

LOW-COST CAT ROBOT USING SPEECH RECOGNITION SYSTEMS FOR PROMOTION

Widodo Budiharto

Computer Science Department
School of Computer Science
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
wbudiharto@binus.edu

Received July 2022; accepted October 2022

Abstract. *Intelligent robot becomes a trend for giving education and entertainment around the world, whereas humans are used to interact with Natural Language (NL) in the social context. This paper presents research on the development of speech recognition for Cat robots named as ANITA V4 focused on promotion using minimum storage. As part of social robot, Cat robot should have the features for giving information about majors in the faculties and schools with natural interaction with prospective students. Speech recognition is still difficult because of sources of variation such as number of types of words in vocabulary, intonation and only tuned on specific speaker. To solve those problems, we use hardware-based speech recognition trained with different speaker and general text to speech systems module for giving response to user using Indonesian language. The experimental results and analysis were presented, and we are sure this model of robot can be used widely on marketing department and education at the university. The comparative experiments with 3 persons show that the system is able to respond question from user.*
Keywords: Cat robot, Education, Speech recognition systems, Indonesian language

1. **Introduction.** A social robot is a special kind of autonomous robot that interacts and communicates with humans by following social behaviors and rules attached to its role. In recent years, it has been suggested that social robots have potential as entertainment, tutors, and educators. Cat robot for education should focus the interactions to be pleasant, challenging, and pedagogically sound for student [1]. The development of Cat robot as part of social robot for huge variety of tasks such as teaching, and learning has become a popular and challenging research topic. Many problems must be solved in intelligent robot such as natural interaction, speech, and face recognition systems.

The need of Cat robots with natural interaction, knowledge and excellence in recognition in the future is very high. For example, in our previous work, we can successfully imitate behavior of teacher to the robot using common words and sentences from teachers (behavior-based Cat robot) [2]. There has been an increased focus on how social robots may help engage children in learning activities. Robots have been shown to help increase interaction levels in larger classrooms, correlating with an improvement in children's learning ability [3]. For example, [4] has shown that a group introduction in the kindergarten prior to one-on-one interactions with the robot influenced the subsequent interactions positively and employed Cat robots for teaching English language [5]. Robot also should have imitation (reinforcing behavior) and social learning [6]. The three-year research project of Edurob [7] aims to explore their impact on teaching, learning and the curriculum. It is collaboration between the Association of Independent Schools South Australia (AISSA), which owns robots, and academics from Swinburne University, the University of Queensland, and the Queensland University of Technology. Teachers and

DOI: 10.24507/icicelb.14.02.117

students can access the technology at different entry points, depending on their skills and experience.

This paper aims to propose a contribution in an architecture of Cat robot and present experimental results of the development of intelligent Cat robot, and we limit our focus on how robot is able to accept questions and correctly respond to prospective students. Build the real Cat robot which has capability to imitate real marketing or teacher and handle the uncertainty of the real world based on cognitive psychology [8]. The organization of the paper consists of Section 1 for introduction, Section 2 for related works, Section 3 for the proposed method, and experimental result and discussion in Sections 4 and 5. The Cat robot with Arduino as a main controller is shown in Figure 1.

FIGURE 1. Our ANITA V4, which can be used at university for promoting and explaining about the majors, registration procedures, information about university, and giving entertatinment such as singing [2,8]

2. **Related Works.**

2.1. **Speed recognition systems.** Sound waves are one-dimensional. At every moment in time, they have a single value based on the height of the wave. To turn this sound wave into numbers, we just record the height of the wave at equally spaced points called sampling. EasyVR is a multi-purpose speech recognition module designed to add versatile, robust, and cost-effective speech and voice recognition capabilities to virtually any application. EasyVR is the second-generation version of the successful VRbot module and builds on the features and functionality of its predecessor. Along with features like 32 user-defined Speaker Dependent (SD) triggers and a host of built-in Speaker Independent (SI) commands, the EasyVR 3 plus adds convenient features such as firmware update capability, 8ohm speaker output and additional SI languages. A simple and robust serial protocol (9600 8-N-1 default) can be used to access these functions from the user's microcontroller boards. The EasyVR 3 can be powered by anywhere between 3.3 V and 5.5 V, and typically consumes 12 mA of current in operation. Example of speech recognition module is shown in Figure 2.

2.2. **Text to speech systems.** May research develop a robot for promotion such as AQuRo [10], this quadruped Cat robot has novel bio-inspired capabilities. RoboPrax is a humanoid robot as tools for the promotion of STEM-education in German schools [11]. The problem with those robots is so complex and has limitation in speech recognition. Text to speech is a system for translating text into speech. The Emic 2 text to speech module

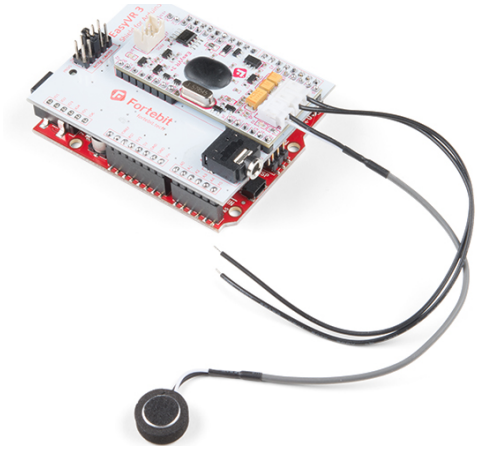

FIGURE 2. EasyVR 3 shield for speech recognition module [9]

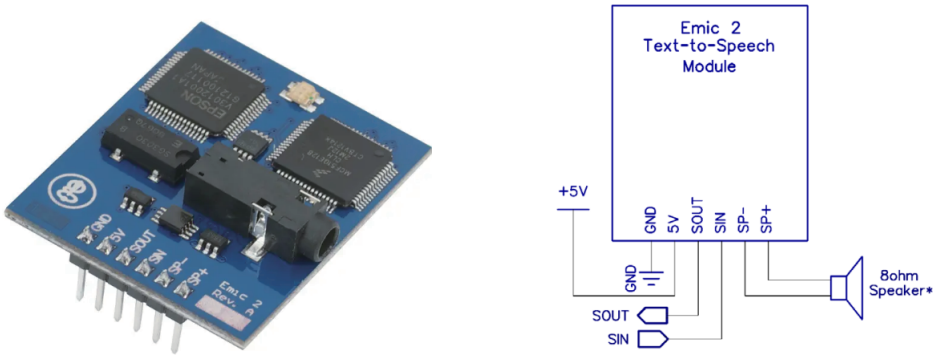

FIGURE 3. Emic 2 text to speech module and schematic [12]

is an unconstrained, multi-language voice synthesizer that converts a stream of digital text into natural sounding speech output. Using the universally recognized DECtalk text to speech synthesizer engine, Emic 2 provides full speech synthesis capabilities for any embedded system via a simple command-based interface [12].

3. Proposed Method.

3.1. **Architecture and flowchart of the Cat robot.** Cat robot has been shown to be a powerful tool of Science, Technology, Engineering, and Math (STEM) education. A central aspect of STEM education is problem-solving, and robots serve as excellent means for teaching problem-solving skills. Robots become important tools for teaching computer science and college engineering introductory. Based on the mounting success of robotics courses world-wide, there is an active movement to develop robot hardware and software in service of education, starting from the youngest elementary school ages and up. Robots become important tools for teaching computer science and college engineering introductory [14-16]. Natural interaction is the common denomination used by designers and developers of human-computer interfaces to refer to a user interface that is effectively invisible and remains invisible as the user continuously learns increasingly complex interactions [13].

To communicate naturally, the robot should have speech recognition capability. Thus, our systems are designed with input both from camera and microphone from the tablet. We improve our previous research [7] such as the software and Tablet PC replaced with Arduino and modules shown in Figure 4.

The microphone provided with the module is an omnidirectional electret condenser microphone with almost flat frequency response in the range 100 Hz – 20 kHz. For text to

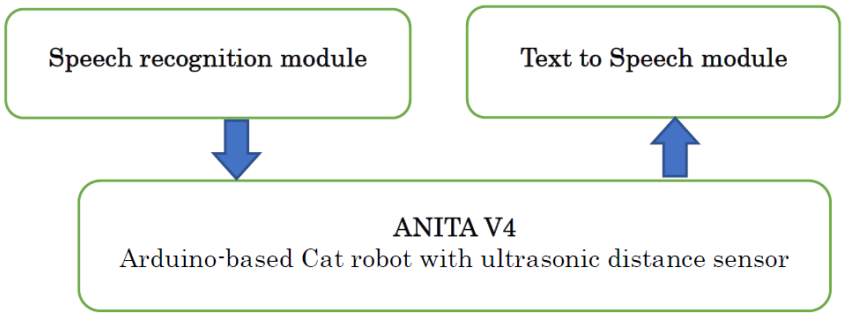

FIGURE 4. Architecture of our intelligent Cat robot with Arduino controller, speech recognition and text to speech systems

speech system, we use Emic text to speech module from Parallax. It is high-quality speech synthesis for English and Spanish languages and contains nine pre-defined voice styles comprising male, female, and child. It provides dynamic control of speech and voice characteristics, including pitch, speaking rate, and word emphasis [9]. For speech recognition systems, we use EasyVR 3 module. It provides multi-purpose speech recognition module designed to add versatile, robust, and cost-effective speech recognition capabilities to almost any application. The EasyVR 3 Plus module can be used with any host with a UART interface powered at 3.3 V – 5 V, such as PIC and Arduino boards. Some application examples include home automation, such as voice-controlled light switches, locks, curtains, or kitchen appliances, or adding "hearing" to the most popular robots on the market [9].

Based on the literature, there is significant positive correlation between lecturer with friendly relationship and good report for students [17], so soft sentences given to the student, welcoming message and giving entertainment when learning are very important. Examples of soft sentences and lecturer's behavior in the robot are as follows.

- Good morning students, I am a robot, my name is Robo-Marketing. I will help marketing and lecturer to explain about our majors. Would you please choose A. Majors, B. Registration, C. About BINUS and others such as Learn Math and Entertainment?
- Your answer is correct, you are very smart. Keep learning and respect with your parent and teacher.
- Don't forget to pray before sleep tonight, honey.

3.2. **Flowchart of the system.** The flowchart of the systems is shown in Figure 5.

Based on the flowchart in Figure 5, explanation of the flow of the system is detailed.

1) The program will detect a user using distance sensor, if user is detected then robot will say welcoming message to user and ask him/her whether he/she is ready to ask something/learning or not. The formal sentences are imitated from the marketing.
2) If user is ready for asking, there are some options for user such as asking question and learning mathematics, doing entertainment (singing and storytelling) and asking questions to robot.
3) The program will loop if the user wants to listen the majors and information about the university and answer the question from robot.
4) If user wants to continue asking, then robot will stop asking the user and say "goodbye".

4. **Experimental Result.** The experiment was conducted with 3 students (user) as volunteer at our robotics laboratory. The robot is able to accept command such as information about majors, teaching basic mathematics, singing and answer questions. The comparative experiment with 3 students is shown in Table 1; we see that without noise and distance 50 cm has the best accuracy. If the machine did not get the correct answer compared with the vocabulary, it will ask again, so the average accuracy with the training

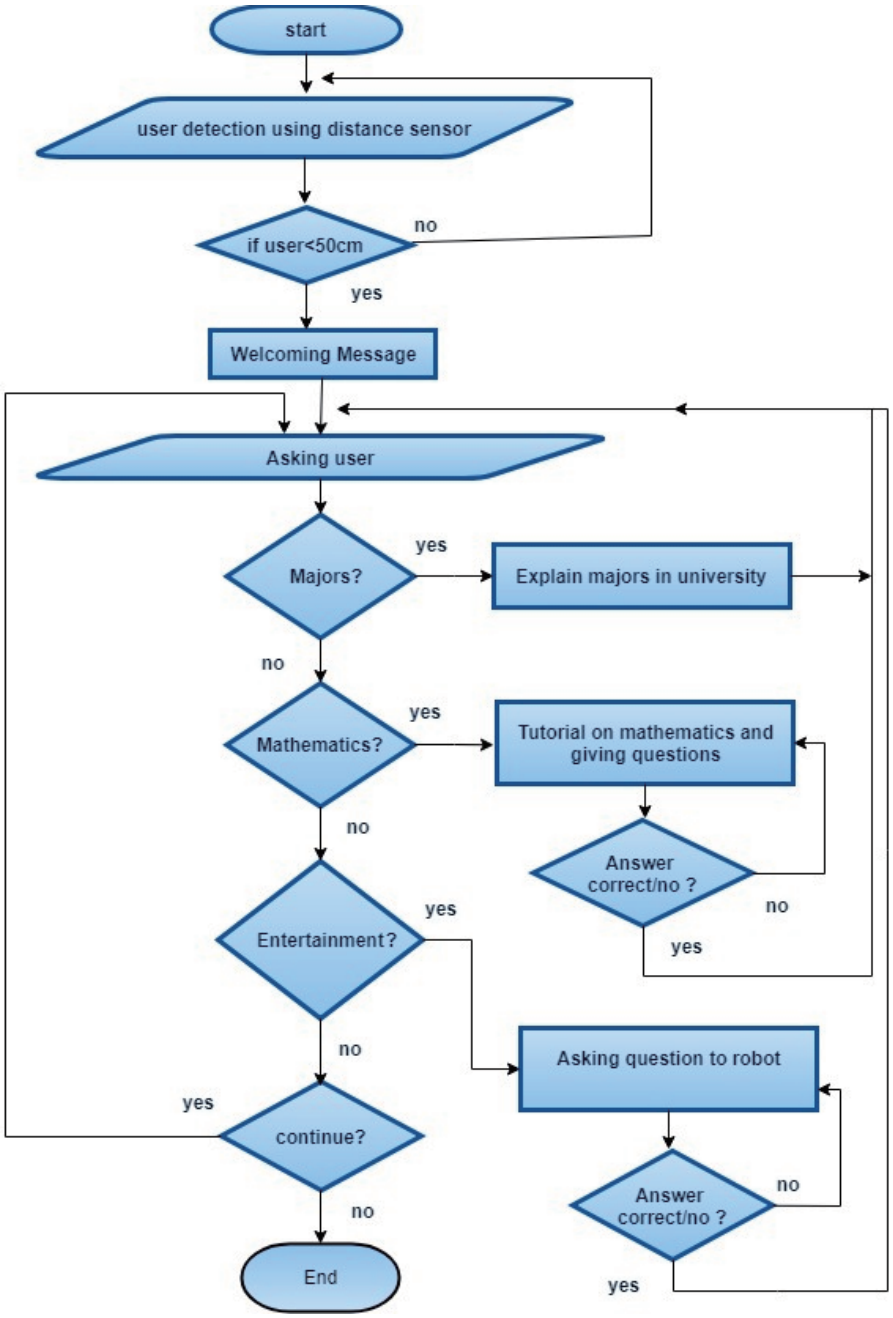

FIGURE 5. Flowchart of the system

TABLE 1. Accuracy with 3 persons and varying the distance of user

Subjects	**Quite**		**Noisy**	
	50 cm	120 cm	50 cm	120 cm
User 1	90%	86.7%	86.7%	80%
User 2	80%	73.3%	66.7%	46.7%
User 3	100%	93.3%	86.7%	66.7%
Average accuracy	**90%**	**84.4%**	**80%**	**64.5%**

set combined with 1 from male's voice and 1 from female's voice reaches 90%. However, with noise, it may be because the environment is very noisy, and the system asks twice but still does not get the correct answer. The reduced signal will cause error in words translation. Overall, the robot is able to hear speech of students and give the good response.

TABLE 2. Preferred sentences by students/user

Subjects	Soft sentences	Formal sentences
User 1	Like	Dislike
User 2	Like	Dislike
User 3	Like	Dislike

User also said that they like clear answer and soft sentences from robot, because they feel like welcoming by marketing. Table 2 shows that users like soft sentences rather than formal sentences.

However, some of the users felt the robot's voice was still unnatural and did not adapt to situations by changing tone or pitch because of the limitation of the hardware. The children still preferred more natural behavior from the robot as possible. Of course, these problems should be solved for future work. We also have a solution for improving the correct reception for the robot, by giving answer by multiple choices. So, students just answer multiple choice from A-C rather than answer using words and sentences. Human-robot cooperation based on visual communication should be learned further [18].

5. **Conclusions.** The development of intelligent Cat robot for marketing and education is quite a challenging task. Natural interaction and imitating good behavior of teacher to robot are important aspect in developing intelligent Cat robot. We successfully propose the system of speech recognition for Cat robot. Fun aspect is given as well because students learn best when they are relaxed and focused. The experimental results show that the robot hears answer from users at the distance 50-120 cm using general purpose speech recognition systems and electret microphone and soft sentences needed by the students. For the future work, our robot will be challenged with more complex spoken interaction due to more variation of actions to do and the ability for recognizing emotions, more natural and long speech should be proposed.

Acknowledgment. This work was supported by BINUS University, Jakarta, Indonesia.

REFERENCES

[1] T. Belpaeme, P. Vogt, R. van den Berghe et al., Guidelines for designing social robots as second language tutors, *International Journal of Social Robotics*, vol.10, pp.325-341, 2018.

[2] W. Budiharto and A. D. Cahyani, Behavior-based humanoid robot for teaching basic mathematics, *Internetworking Indonesia Journal*, vol.9, no.1, pp.33-37, 2017.

[3] M. Fridin, Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education, *Comput. Educ.*, vol.70, pp.53-64, 2014.

[4] M. Fridin, Kindergarten social assistive robot: First meeting and ethical issues, *Comput. Hum. Behav.*, vol.30, pp.262-272, 2014.

[5] M. Alemi, A. Meghdari and M. Ghazisaedy, Employing Cat robots for teaching English language in Iranian junior high-schools, *International Journal of Humanoid Robotics*, vol.11, no.3, 2014.

[6] *Robot-Based Pedagogy Requirements: Results from Stakeholder Interviews*, http://www.edurob.eu/assets/edurob_pedagogyrequirementsreport2.pdf, Accessed on 8 February, 2022.

[7] *Teaching in Cat Robot*, https://www.teachermagazine.com.au/articles/teaching-and-learning-with-Cat-robots, Accessed on 3 August, 2022.

[8] W. Budiharto et al., EduRobot: Intelligent cat robot with natural interaction for education and entertainment, *Procedia Computer Science*, vol.116, pp.564-570, 2017.

[9] *EasyVR3 Plus Shield for Arduino*, https://www.sparkfun.com/products/15453, Accessed on 10 January, 2020.

[10] A. A. Saputra, N. Takesue, K. Wada, A. J. Ijspeert and N. Kubota, AQuRo: A cat-like adaptive quadruped robot with novel bio-inspired capabilities, *Front. Robot. AI*, vol.8, 562524, DOI: 10.3389/frobt.2021.562524, 2021.

[11] S. Zeaiter, Using humanoid robots as tools for the promotion of STEM-education in schools (RoboPraX), in *Pioniergeist, Ausdauer, Leidenschaft. Festschrift zu Ehren von Prof. Dr. Jürgen Handke*, Philipps-Universität Marburg, 2020.

[12] *Emic 2 Text to Speech Module*, https://www.parallax.com/package/emic-2-text-to-speech-module-downloads/, Accessed on 4 February, 2022.

[13] G. Bugmann, Effective spoken interfaces to service robots: Open problems, *AISB'05: Social Intelligence and Interaction in Animal, Robots and Agents-SSAISB 2005 Convention*, Hatfield, UK, pp.18-22, 2005.

[14] J. B. Weinberg and X. Yu, Robotics in education: Low cost platforms for teaching integrated systems, *IEEE Robotics & Automation Magazine*, vol.10, no.2, pp.4-6, 2003.

[15] D. Feil-Seifer and M. J. Matarić, Human robot interaction, *Encyclopedia of Complexity and Systems Science*, pp.4643-4659, 2009.

[16] T. Hsiu, S. Richards, A. Bhave, A. Perez-Bergquist and I. Nourbakhsh, Designing a low-cost, expressive educational robot, *Proc. of the Conference on Intelligent Robots and Systems*, vol.3, pp.2404-2409, 2003.

[17] S. Nayereh, Review on the impact of teachers' behaviour on students' self-regulation, *Procedia Social and Behavioral Sciences*, vol.114, pp.130-135, 2014.

[18] J. K. Tan and A. Sato, Human-robot cooperation based on visual communication, *International Journal of Innovative Computing, Information and Control*, vol.16, no.2, pp.543-554, DOI: 10.24507/ijic ic.16.02.543, 2020.