

論文 / 著書情報
Article / Book Information

Title	An Overview of Speech Recognition Research for Indonesian Language
Author	Dessi Puji Lestari, Sadaoki Furui
Journal/Book name	Proceedings of the 18th Indonesian Scientific Conference in Japan, , , pp. 89-90
発行日 / Issue date	2010, 8

An Overview of Speech Recognition Research for Indonesian Language

Dessi Puji Lestari¹, Sadaoki Furui²

Department of Computer Science, Graduate School of Information Science and Engineering,
Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552 Japan

¹Email address: dessi@furui.cs.titech.ac.jp

²Email address: furui@cs.titech.ac.jp

Abstract. *A speech recognition system (ASR) is generally constructed by a set of technologies that allows a computer to transform sound input recorded through a microphone to a sequence of words. State of the art speech recognition technologies use statistical methods. It requires a large scale speech and text corpus to train their acoustic and language models. Researches on ASR have been extensively conducted for many languages, including English, Japanese, and Chinese. It also has already been integrated to many kinds of applications, such as voice-control applications, dictation, and telephone solutions. For Indonesian languages, there have been significant research activities on ASR in recent years. This paper presents some of the speech recognition-related works for Indonesian languages. Due to difficulties of finding a large scale corpus for both speech and text, in early works on Indonesian ASR which started in 2003, preliminary researches focused on phone recognition, isolated word recognition, and small vocabulary ASR [1]. Other researchers also tried to build a large vocabulary continuous speech recognition system (LVCSR) for Indonesian language using a rapid approach; they used a cross-language approach to build phoneme-based speech recognition [2]. Under the A-STAR Project, PT. Telkom Indonesia, in the collaboration with the ATR Japan, collected a large scale speech corpus and developed an Indonesian LVCSR system [3]. Their target applications were speech recognizers for hearing and speaking impaired people and automatic machine translation for Asian Languages using a spoken dialogue system. Different from their research, our researches are focusing on how to improve Indonesian LVCSR performance affected by the speech variation problem. It has been revealed that proper nouns and English words are the main sources of recognition errors. By using some model adaptation techniques, we have successfully improved the performance of the ASR [4]. Our target application is spoken query-based information retrieval. While previously mentioned researches are mainly conducted abroad (Japan), a research has also been conducted in an Indonesia University (UI) to develop an Indonesian LVCSR system [5]. Although there have been some good reports on the development of Indonesian ASR in recent years, many works still need to be conducted to make this technology applicable to our daily life. This includes LVCSR that is robust to the dialect problem. Since Indonesia has hundreds of local languages and dialects, the performance of the ASR might be decreased due to the dialect variations. Another research area is spontaneous speech recognition. Indonesian spontaneous speech has characteristics different from formal and read speech. Thus, this needs to be further investigated.*

Keywords: Indonesian Language, Automatic Speech Recognition (ASR), Large Vocabulary Continuous Speech Recognition System (LVCSR).

Category: Applied Sciences/Engineering

Presentation Media: Oral Presentation

References

- [1] Indrayanti, L., Dutono, T., Nakagawa, A., Chisaki, Y., and Usagawa, T. (2003) "Preliminary Development of Indonesian LVCSR: Database Preparation and Acoustic Modelling for Isolated Word Recognition", *Proc. Joint Conference of Electrical and Electronics Engineers*, Kyushu, 12-1P-02.

- [2] Sakti, A., Markov, K., and Nakamura, S. (2005). "Rapid Development of Initial Indonesia Phoneme-Based Speech Recognition Using Cross Language Approach", *O-COCOSDA*, Jakarta, Indonesia.
- [3] Sakti, S., Kelana, E., Riza, H., Markov, K., and Nakamura, S. (2008). "Development of Indonesian Large Vocabulary Continuous Speech Recognition System within A-STAR Project", *Workshop on Technologies and Corpora for Asia-Pacific Speech Translation*, Hyderabad, India.
- [4] Lestari, D.P., and Furui, S. (2010), "Adapatation to Pronunciation Variations in Indonesian Spoken Query-Based Information Retrieval", *IEICE Trans. Information and System*, Vol.E93-D, No.9. (To be published)
- [5] Baskoro, S., and Adriani, M. (2008). "Developing an Indonesian Speech Recognition System". *Proc. 2nd International MALINDO Workshop*, Cyberjaya, Malaysia.