

研 究 主 論 文 抄 録

論文題目

Framework of Keyword Spotting for Bahasa Indonesia Speech under Multiple Speaker Condition
(複数話者環境におけるインドネシア語音声のキーワードスポッティングのためのフレームワーク)

熊本大学大学院自然科学教育部 工学専攻 人間環境情報教育プログラム
(主任指導 宇佐川 毅 教授)

論文提出者 Muhammad Bagus Andra (ムハンマド バグス アンラ)

主論文要旨

Bahasa Indonesia is one of the most prominent low-resource Languages that still lack development in regards to communication-assisting technology and linguistic research. In this research, we have developed a communicating-assisting technology that encompasses the keyword spotting and transcription generation as a specialized function of speech recognition. The proposed model is designed by considering the linguistic properties of Bahasa Indonesia to optimize its performance. Unlike many other previous research, our proposed framework is able to adapt with situation such as noisy environment and multiple speaker. In addition, our proposed model is also able to automatically generate transcript for each speaker and assign it to its respective owner. Portability and performance is also important aspects that we constantly improve in our design and our design were able to achieve satisfying performance in mini computer such as raspberry-pi and mini microphone array. We have shown that the system is applicable in many fields such as UAV and mobile robot auditory.

The system combines Reinforced Learning (RL) Model with pitch-aware speech separation to identify the speakers in a concurrent speech. A Recurrent Neural Network (RNN) is utilized to generate the text transcript which is later improved by an external language model and spelling correction model. The proposed system was able to identify up to 5 speakers with a variable degree of confidence and generate a transcript for each of them with better quality compared to other methods when evaluated with several metrics.