Detection of Speaker Direction Based on the On-and-Off Microphone Combination for Entertainment Robots

Takeshi Kawabata, Masashi Fujiwara, and Takanori Shibutani

Abstract: An important function of entertainment robots is voice communication with humans. For realizing them, accurate speech recognition and a speaker-direction detection mechanism are necessary. The direct-noise problem is serious in such speech processing. The microphone attached to the robot body receives not only human voice but also motor and mechanical noises directly. The direct noises are often larger than distance voices and fatally degrade the speech recognition rate. Even if the microphone close to the user ("on-mic") is used for speech recognition, the body microphones ("off-mic") are still necessary for detecting the speaker direction under the severe condition with direct noises. This paper describes a new method for detecting the speaker direction based on the on-and-off microphone combination. The system searches for the spectral elements of "on-mic" voice in the other "off-mic" channels. The segregated power ratio or the time delay between the "off-mic" channels is used for detecting the speakers direction. Experiments show that the proposed method effectively improves the direction detection accuracy during the robot moves.