

SPOKESPERSON DETECTION METHOD FOR AUTONOMOUS ROBOT IN COMPLEX COMMUNICATION ENVIRONMENT, BASED ON IMAGE PROCESSING

JIA MA¹, MOTOYUKI SUZUKI² AND FUJI REN^{2,3}

¹Graduate School of Advanced Science Technology Education

²Institute of Technology and Science

The University of Tokushima

2-1 Minamijosanjima-cho, Tokushima 770-8506, Japan

{ majia; suzuki_m; ren }@is.tokushima-u.ac.jp

³Beijing University of Posts and Telecommunications

Beijing, 100876, P. R. China

Received January 2009; revised July 2009

ABSTRACT. *For providing a man-machine communication system, the computer or the robot may be enabled to understand and simulate the command from human. As a prerequisite, the system should be enabled to recognize the spokesperson, who is speaking to it. However, in the realistic scene, the noise from the complex environment may cause some difficulties for the system to realize this function. By imitating the performance of human, we try to use the eyes of the robot to solve the problem. In this research, we propose a new spokesperson detection method based on image processing. The system might find and identified all the possible spokesperson around it, by detecting and recognizing the faces in the video from the cameras on the robot. It might also find out the spokesperson who is speaking to it, by judging the mouth action of possible spokesperson. Using eye cameras and an omni-directional camera, we tried to realize the system and some experiments were conducted to verify the new method.*

Keywords: Spokesperson detection, Complex communication environment, Face detection, Face verification, Mouth action judgment

1. Introduction. Recently, it turns to be the question which was interesting to many researchers, that practicing the freely communication between human and robot. AI technology can cause the current robot showing more and more complex and reasonable response to peoples' instruction. Furthermore, people hope that the computer or robot can actively talk with human, and may even understand the human sentiment. In order to develop this kind of system, researchers attempt to use the method of natural language processing, sound processing and image processing. As a premise, this kind of system must be able to find the target people naturally and actively. However, as we know, there is not any man-machine communication system, which is possible to satisfy this requirement [1-4].

The current research about robot, especially the humanoid robot, is mainly concentrated in the domain of action control [5]. The researchers try to make robot's movement becoming more precisely and naturally. However, as we know, they are not so cared about the man-machine communication system. Otherwise, the researchers of man-machine conversation system pay more attention to the recognition of language and speech. Therefore, when the user was determined, the existing man-machine communication system displayed much better. In such a system, the user was determined with some certain order. But in the reality, the user is possibly not familiar with the system instruction. Therefore the