

Title	人 - ロボット相互作用における人の性格特性推定に向けた選択的マルチモーダル融合アプローチ
Author(s)	申, 志豪
Citation	
Issue Date	2021-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/17477
Rights	
Description	Supervisor: 丁 洛榮, 先端科学技術研究科, 博士

Doctoral Dissertation

A selective multi-modal fusion approach to
inferring human personality traits in
human-robot interaction

SHEN Zhihao (1720413)

Chong Laboratory

Japan Advanced Institute of Science and Technology

Information Science

A selective multi-modal fusion approach to inferring human personality traits in human-robot interaction

SHEN Zhihao (1720413)

March 15, 2021

With the population aging and sub-replacement fertility problems increasingly prominent, many countries have started promoting robotic technology for assisting people toward a better life. The robot was designed with the appearances that are similar to human's. And more importantly, many robots also were endowed with many capabilities such as synchronized verbal and nonverbal behaviors, emotion recognition, and many others for acquiring a high quality interaction between robot and its users.

It has been found that the personality traits are playing very important roles in human-human interactions. With an increasing number of research on personality traits, their relationship to many important aspects of life, such as job performance, health-related behaviors, emotion, and many others have been revealed. Therefore, understanding personality traits is useful for predicting human behaviors, and understanding human's mind and how personality traits affect the attitude and behaviors towards other people. Once the robots are endowed with the capability of recognizing human personality traits, the robots then will be able to adjust their behaviors such as voice volume, speech rate, and body gestures to enhance the degree of user engagement.

For achieving this goal, a pilot experiment for personality traits recognition was conducted for testing the feasibility of inferring personality traits from nonverbal behavior features, and finding more practical problem in human-robot interaction. Some features which are head motion, gaze, body motion, voice pitch, voice energy, and Mel-Frequency

Cepstral Coefficient were extracted to describe human's nonverbal behaviors. Each feature showed its advantage in a different aspect. However, different nonverbal features can provide different personality traits classification results. It is not a standard way of drawing the conclusion for declaring the user's personality traits. On the other hand, the camera was fixed to make sure that the background did not change, same strategy also was applied in many related studies for the same purpose. However, this conflict with the idea that robot that was enabled to understand human personality traits aims to behave more properly.

Therefore, a new paradigm of human-robot interaction as close to the real situation as possible was designed, the following three main problems were also addressed: (1) fusion of visual and audio features of human interaction modalities, (2) integration of variable length feature vectors, and (3) compensation of shaky camera motion caused by movements of the robot's communicative gesture. Same nonverbal features including head motion, gaze, and body motion, voice pitch, voice energy, and Mel-Frequency Cepstral Coefficient were extracted from a camera mounted on the robot performing verbal and body gestures during the interaction. Then, the system was geared to fuse these feature and deal with variable length multiple feature vectors. Lastly, considering unknown patterns and sequential characteristics of human communicative behavior, a multi-layer Hidden Markov Model that improved the classification accuracy of personality traits and offered notable advantages of fusing the multiple features was proposed. The results were thoroughly analyzed and supported by psychological studies. The proposed multimodal fusion approach is expected to deepen the communicative competence of social robots interacting with humans from different cultures and backgrounds.

Keywords: Human-Robot Interaction; Personality Traits Recognition; Multimodal Feature Fusion; Nonverbal Features; Multi-layer Hidden Markov Model; Machine Learning Model.

PUBLICATION

JOURNAL

1. **Zhihao Shen**, A. Elibol, N. Y. Chong, Multimodal feature fusion for better understanding of human personality traits in social human robot interaction. *Robotics and Autonomous Systems*.²
2. **Zhihao Shen**, A. Elibol, N. Y. Chong, (2020). Understanding nonverbal communication cues of human personality traits in human-robot interaction. in *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 6, pp. 1465-1477, November 2020, URL: <https://doi.org/10.1109/JAS.2020.1003201>, doi: 10.1109/JAS.2020.1003201.

CONFERENCE PAPER

1. **Zhihao Shen**, A. Elibol, N. Y. Chong, (2020). Multimodal Feature Fusion for Human Personality Traits Classification. *Proceedings of the 2020 17th International Conference on Ubiquitous Robots (UR)*, 565–566. URL: <https://ci.nii.ac.jp/naid/120006874934>
2. **Zhihao Shen**, A. Elibol, N. Y. Chong, (2019). Nonverbal behavior cue for recognizing human personality traits in human-robot social interaction. *2019 4th IEEE International Conference on Advanced Robotics and Mechatronics ICARM*, July 2019, 402–407. URL: <https://doi.org/10.1109/ICARM.2019.8834279> (**Best Paper Award in Advanced Robotics**)
3. **Zhihao Shen**, A. Elibol, N. Y. Chong, (2019). Inferring Human Personality Traits in Human-Robot Social Interaction. *ACM/IEEE International Conference on Human-Robot Interaction*, March 2019, 578–579. URL: <https://doi.org/10.1109/HRI.2019.8673124>
4. **Zhihao Shen**, H. Lee, S. Jeong and N. Y. Chong, (2017). Informative sequential selection of variable-sized patches for image retrieval, *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, Kanazawa, November 2017, pp. 1169-1172, URL: <https://doi.org/10.23919/SICE.2017.8105625>, doi: 10.23919/ SICE.2017.8105625.
5. **Zhihao Shen**, S. Jeong, H. Lee and N. Y. Chong, (2017). Informative sequential patch selection for image retrieval, *2017 IEEE International Conference on Information and Automation (ICIA)*, Macau, October 2017, pp. 213-218, doi: 10.1109/ICInfA.2017.8078908.

²This journal paper has been revised and is waiting for the second review.