2019-08-19

Multi-modal Personalisation in Long-Term Human-Robot Interaction

Irfan, B

Multi-modal Personalisation in Long-Term Human-Robot Interaction

Bahar Irfan
CRNS, University of Plymouth, UK
bahar.irfan@plymouth.ac.uk

I. Introduction

Long-term human-robot interaction will be an integral part of domestic applications, rehabilitation and education in the future. However, long-term interactions based on a fixed set of behaviours, bring about certain challenges due to the repeated interactions with users, such as decreased user engagement. Personalisation can improve user engagement and facilitate building rapport and trust [1]. User recognition is an essential step towards achieving personalisation in the interaction for multiple users. Contrary to the common problems in the computer vision field, users might not be known in advance in a human-robot interaction scenario, which is known as an *open world recognition* problem. Moreover, users are encountered continuously and incrementally. Hence, deep learning approaches may fall short for meeting those requirements due to the *catastrophic forgetting* problem defined as the severe loss of performance on previously learned classes after the introduction of a new class [2]. Moreover, these approaches may take a vast amount of time for recognition and re-training in low computational power systems.

Incremental learning is not sufficient to adapt to the changes in the environment, such as those in the appearance of a user. *Online learning* allows for updating previous beliefs and user models to overcome this problem. Furthermore, inaccurate or noisy data, such as a blurry image or illumination changes, can result in low performance if only face recognition (FR) is used to identify users [3]. Ancillary physical or behavioural characteristics called *soft biometrics*, such as age and gender can be used to improve the recognition performance [4] and overcome issues related to similarities between users.

Consequently, we designed a multi-modal incremental Bayesian network (MMIBN) [5], which is the first in combining multi-modal biometric information for sequential and incremental learning for open world recognition.

II. Methodology

The proposed multi-modal incremental Bayesian network combines weighted soft biometrics, namely, gender, age, height and time of interaction, with a primary biometric, FR, through a naive Bayes classifier model.

We introduced a two-step ad hoc mechanism for MMIBN to identify if a user is known: (1) FR threshold is used to establish an "unknown" state, (2) the highest resulting posterior probability is compared to the second highest to evaluate the quality of estimation. When a new user is encountered, the Bayesian network is expanded to allow incremental learning by scaling the conditional probabilities of the modalities. Online learning is achieved through Expectation Maximization with an adaptive learning rate based on Maximum Likelihood.

III. Results

We initially evaluated our model with 14 participants over four weeks period [5]. We used the proprietary algorithms of the Pepper robot (SoftBank Robotics Europe) to obtain the multi-modal biometric information. Our results suggested that the proposed models increase the identification rate up to 1.4% in open-set and 4.4% in closed-set recognition compared to 90.3% of FR. Moreover, MMIBN performed worse with online learning. However, due to the low number of participants and the limited age range, we concluded that a larger dataset was necessary to evaluate the capabilities of the system.

Accordingly, we created a multi-modal long-term user recognition dataset with 200 users of varying characteristics based on the IMDB-Wiki dataset [6]. We compared our proposed approach to a state-of-the-art open world recognition approach, Extreme Value Machine [7] (EVM). The results showed that the proposed MMIBN models improved the identification rate significantly and substantially compared to soft biometrics, FR and EVM. Our results concluded that online learning either decreases recognition performance compared to fixed likelihoods or performs at the same significance level, which might result from accumulating noise of the identifiers.

IV. Ongoing Work

We applied the proposed user recognition model for a barista robot that learns the preferences of customers. We are currently analysing the results of a recent real-world study with 17 participants with repeated interactions over a week.

Moreover, we are using the proposed model for recognising users in cardiac rehabilitation therapy to personalise the interaction based on the information about the patients' previous sessions and their progress during the therapy [8]–[10].

References

[1] B. Irfan, A. Ramachandran, S. Spaulding, D. F. Glas, I. Leite, and K. L. Koay, "Personalization in long-term human-robot interaction," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019.

[2] J. L. Mcclelland, B. L. Mcnaughton, and R. C. O'Reilly, "Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory," *Psychological Review*, vol. 102, no. 3, 1995.

[3] W. Wójcik, K. Gromaszek, and M. Junisbekov, "Face recognition: Issues, methods and alternative applications," in *Face Recognition - Semisupervised Classification, Subspace Projection and Evaluation Methods*, S. Ramakrishnan, Ed. InTech, 2016, ch. 02.

[4] A. K. Jain, A. A. Ross, and K. Nandakumar, *Introduction to Biometrics*. Springer Publishing Company, Incorporated, 2011, ch. Multibiometrics.

[5] B. Irfan, N. Lyubova, M. Garcia Ortiz, and T. Belpaeme, "Multi-modal open-set person identification in HRI," in *2018 ACM/IEEE International Conference on Human-Robot Interaction Social Robots in the Wild workshop*, 2018.

[6] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision (IJCV)*, 2016.

[7] E. M. Rudd, L. P. Jain, W. J. Scheirer, and T. E. Boult, "The extreme value machine," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 3, 2018.

[8] J. S. Lara, J. Casas, A. Aguirre, M. Munera, M. Rincon-Roncancio, B. Irfan, E. Senft, T. Belpaeme, and C. A. Cifuentes, "Human-robot sensor interface for cardiac rehabilitation," in *2017 International Conference on Rehabilitation Robotics (ICORR)*, 2017, pp. 1013–1018.

[9] J. Casas, B. Irfan, E. Senft, L. Gutiérrez, M. Rincon-Roncancio, M. Munera, T. Belpaeme, and C. A. Cifuentes, "Social assistive robot for cardiac rehabilitation: A pilot study with patients with angioplasty," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018.

[10] J. Casas, B. Irfan, E. Senft, L. Gutierrez, M. Rincon-Roncancio, M. Munera, T. Belpaeme, and C. A. Cifuentes, "Towards a SAR system for personalized cardiac rehabilitation: A patient with PCI," in *2018 ACM/IEEE International Conference on Human-Robot Interaction Personal Robots for Exercising and Coaching workshop*, 2018.