·**Editorial**·

Emotion recognition for human-computer interaction

Jianhua TAO[1,2,3]

1. *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing* 100081, *China*

2. *CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing* 100081, *China*

3. *School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing* 100081, *China*

* **Corresponding author,** jhtao@nlpr.ia.ac.cn

Emotion recognition is to quantify, describe and recognize different emotional states through the behavioral and physiological responses generated from emotional expressions. Emotion recognition is an important field due to its wide applications in many tasks, such as dialogue generation, social media analysis and intelligent system. It builds a harmonious human-computer environment by enabling the computer systems and devices to recognize and interpret human affects. Emotion recognition models are built using multimodal information such as audio, video, text and so on. It is important to consider emotion characteristics of humans in the design and presentation of intelligent interaction. We have selected seven papers that provide the latest updates on the development of emotion recognition technology covering micro-expression spotting and recognition, speech emotion recognition, physiological signal emotion recognition, emotional dialog generation and so on.

The automatic detection of emotional states from human speech has long been regarded as a challenging task. Zheng et al. proposed a novel multi-scale discrepancy adversarial (MSDA) network for conducting multiple timescales domain adaptation for cross-corpus speech emotion recognition (SER). The proposed model is effective for solving cross-corpus SER. Huang et al. widened a larger view of the feature frames with frame skipping and temporal pooling on the feature level and employed a skip RNN on the model level. These promising results suggest that skip LSTM can lead to important improvements in continuous emotion recognition. Zhao et al. proposed attention-based model, called SATN, combines self-attention with knowledge transfer for SER tasks. This structure enables the model to learn long-term data with complex spatio-temporal patterns and capture the dependencies between these patterns.

Human-machine dialog generation is an essential topic of research in the field of natural language processing. Generating high-quality, diverse, fluent, and emotional conversation is a challenging task. Based on generative adversarial networks (GANs), Chen et al. proposed a new emotional dialog generation framework called EMC-GAN to address the task of emotional dialog generation. The proposed EMC-GAN model is capable of generating consistent, smooth, and fluent dialog that conveys pre-specified emotions.

Facial micro-expressions could reveal the true emotions that a person may be attempting to suppress, hide, disguise, or conceal. Such expressions have a wide range of application in public safety and clinical diagnosis. In this research, Pan et al. discuss challenges that remain unresolved alongside future work to be completed in the field of micro-expression analysis. A comprehensive review on the topic of spotting and recognition used in micro-expression analysis databases and methods, is conducted, and advanced technologies in this area are summarized.

In anticipation of its great potential application to natural human-computer interaction and health

monitoring, heart-rate (HR) estimation based on remote photoplethysmography has recently attracted increasing research attention. Lu et al propose a neural architecture search approach for HR estimation to automatically search a lightweight network that can achieve even higher accuracy than a complex network while reducing the computational cost.

The automatic detection of frustration can enable the development of adaptive systems that can adapt a game to a user's specific needs through real-time difficulty adjustment, thereby optimizing the player's experience and guaranteeing game success. Song et al present a speech-based approach for the automatic detection of frustration during game interactions, a specific task that remains underexplored in research. This is an interesting application for emotion recognition.

This special issue covers the latest advances in the theories, technologies, and applications of human computer interactions that have been achieved by researchers and engineers in this field. We hope that through the publication of this special issue, it will provide useful reference for readers engaged in the related technology research and applications. We would like to thank Virtual Reality & Intelligent Hardware for their guidance and assistance in the publication of this special issue and also thank all reviewers for their timely, patient, and detailed reviews.

Jianhua Tao
19 January 2021