

Using speech data to recognize emotion in human gait

Angelica Lim and Hiroshi G. Okuno

Graduate School of Informatics, Kyoto University
Sakyo-ku, Kyoto, Japan 606-8501
{angelica, okuno}@kuis.kyoto-u.ac.jp

Abstract. Robots that can recognize emotions can improve humans' mental health by providing empathy and social communication. Emotion recognition by robots is challenging because unlike in human-computer environments, facial information is not always available. Instead, our method proposes using speech and gait analysis to recognize human emotion. Previous research suggests that the dynamics of emotional human speech also underlie emotional gait (walking). We investigate the possibility of combining these two modalities via perceptually common parameters: Speed, Intensity, irRegularity, and Extent (SIRE). We map low-level features to this 4D cross-modal emotion space and train a Gaussian Mixture Model using independent samples from both voice and gait. Our results show that a single, modality-mixed trained model can perform emotion recognition for both modalities. Most interestingly, recognition of emotion in gait using a model trained uniquely on speech data gives comparable results to a model trained on gait data alone, providing evidence for a common underlying model for emotion across modalities.

Keywords: robot emotions, emotional gait, emotional voice, affect recognition