

Accuracy of automatic emotion recognition from voice

Damien Dupré^{*,a}, Gary McKeown^b

^a*The Insight Centre for Data Analytics, University College Dublin, Republic of Ireland*

^b*School of Psychology, Queen's University Belfast, United Kingdom*

Abstract

Voice is one of the main communication channels to express emotions (Poria, Cambria, Bajpai, & Hussain, 2017). Because Affective Computing is aiming to create systems and algorithms to automatically analyse people's emotion, several companies such as Affectiva, Beyond Verbal and Audeering have already developed automatic systems to analyse vocal expression of emotions. However little is known about the accuracy of such systems.

In order to evaluate the accuracy of automatic emotion recognition from voice, we have processed vocal expressions from the GEMEP database with "SensAI Emotion" developed by Audeering. The GEMEP database is made of audio-video recordings from 10 actors performing 17 different emotional scenarios in which they have to vocalize the same non-words with an emotional tonality (Bänziger & Scherer, 2010). SensAI Emotion analyses emotions from speech and render a value for 23 affective states as well as for valence and arousal dimensions (Eyben, Scherer, & Schuller, 2018).

In term of category recognition, the accuracy of SensAI to label GEMEP vocal expression of emotion is 6.67%. However this low result is partly due to the high number of different affective state labels recognized. In order to bypass the bias of label matching we compared the accuracy to recognize valence and arousal dimensions. The results show an accuracy of 0.56 (CI95%[0.46,0.65]) for valence and 0.73 (CI95%[0.64,0.81]) for arousal recognition.

Automatic emotion recognition from voice is a growing topic of research in Affective Computing. Even if the recognition of emotion categories remains a challenge given the diversity of affective states, the accuracy of a system like SensAI Emotion provides promising results in the recognition of valence and arousal.

References

Bänziger, T., & Scherer, K. R. (2010). Introducing the geneva multimodal emotion portrayal (gemep) corpus. *Blueprint for Affective Computing: A Source-*

*Corresponding Author

Email addresses: damien.dupre@ucd.ie (Damien Dupré), g.mckeown@qub.ac.uk (Gary McKeown)

book, 271–294.

Eyben, F., Scherer, K. R., & Schuller, B. W. (2018). Method for automatic affective state inference and an automated affective state inference system. Google Patents.

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37, 98–125.