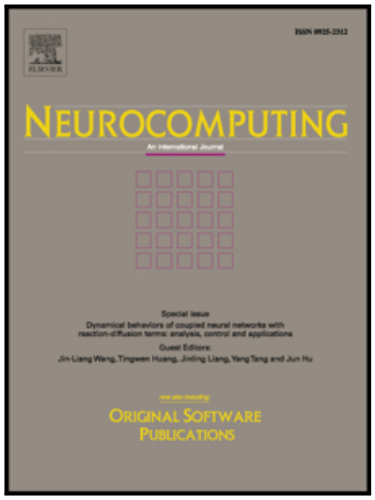
Communicated by Dr. Nianyin Zeng

Accepted Manuscript



Speech emotion recognition based on feature selection and extreme learning machine decision tree

Zhen-Tao Liu, Min Wu, Wei-Hua Cao, Jun-Wei Mao, Jian-Ping Xu, Guan-Zheng Tan

PII: S0925-2312(17)31356-5
DOI: 10.1016/j.neucom.2017.07.050
Reference: NEUCOM 18746

To appear in: *Neurocomputing*

Received date: 15 April 2017
Revised date: 3 July 2017
Accepted date: 31 July 2017

Please cite this article as: Zhen-Tao Liu, Min Wu, Wei-Hua Cao, Jun-Wei Mao, Jian-Ping Xu, Guan-Zheng Tan, Speech emotion recognition based on feature selection and extreme learning machine decision tree, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.07.050

Speech emotion recognition based on feature selection and extreme learning machine decision tree☆

Zhen-Tao Liu[a,b], Min Wu[a,b], Wei-Hua Cao[a,b,*], Jun-Wei Mao[a,b], Jian-Ping Xu[a,b], Guan-Zheng Tan[c]

[a]*School of Automation, China University of Geosciences, Wuhan 430074, China*
[b]*Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China*
[c]*School of Information Science and Engineering, Central South University, Changsha 410083, China*

Abstract

Feature selection is a crucial step in the development of a system for identifying emotions in speech. Recently, the interaction between features generated from the same audio source was rarely considered, which may produce redundant features and increase the computational costs. To solve this problem, feature selection method based on correlation analysis and Fisher is proposed, which can remove the redundant features that have close correlations with each other. To improve the recognition performance of the feature subset after proposal feature selection further, an emotion recognition method based on extreme learning machine (ELM) decision tree is proposed according to the confusion degree among different basic emotions. A framework of speech emotion recognition is proposed and the classification experiments based on proposed classification method by using Chinese speech database from institute of automation of Chinese academy of sciences (CASIA) are performed. And the experimental results show that the proposal achieved 89.6% recognition rate on average. By proposal, it would be fast and efficient to discriminate emotional states of different speakers from speech, and it would make it possible to realize the interaction between speaker-independent and

☆This work was supported by the National Natural Science Foundation of China under Grants 61403422 and 61273102, the Hubei Provincial Natural Science Foundation of China under Grant 2015CFA010, the 111 project under Grant B17040, and the Fundamental Research Funds for National University, China University of Geosciences (Wuhan).

*Corresponding author: Wei-Hua Cao, email: weihuacao@cug.edu.cn