# Speaker-independent speech emotion recognition by fusion of functional and accompanying paralanguage features

## 融合功能性副语言的与人无关语音情感识别方法

JZUSC (Comput. & Electron.)

**Citation:** Qi-rong Mao, Xiao-lei Zhao, Zheng-wei Huang, Yong-zhao Zhan, 2013. Speaker-independent speech emotion recognition by fusion of functional and accompanying paralanguage features. *Journal of Zhejiang University-Science C (Computers & Electronics)*, 14(7):573-582. [doi:10. 1631/jzus.CIDE1310]

- Since functional paralanguage includes considerable emotion information and is insensitive to speaker changes, this paper deals with the association of functional paralinguistic information with the accompanying paralanguage, in order to achieve higher and more robust speaker-independent recognition accuracy.

- The paralanguage fusion recognition algorithm (PFRA) combining the functional paralanguages and the accompanying paralanguages is proposed for speech emotion recognition.

- The speaker-independent speech emotion recognition experiment results on the emotional paralanguages contained speech emotion database recorded by our research group show that, compared with the single channel speech emotion recognition method not fusing functional paralanguages, linear weighted method, Dempster-Shafer (DS) evidence theory, and Bayes fusion method, the proposed fusion algorithm is superior in accuracy and stability.

JZUS-C (Comput. & Electron.)

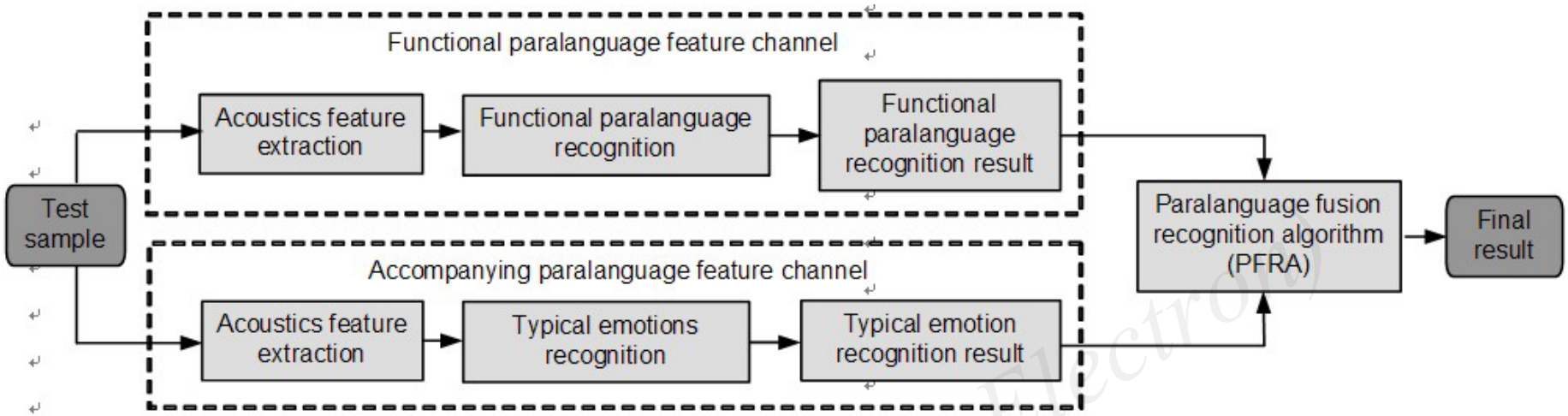



**Fig. 1 Recognition process fusing the functional paralanguages and the accompanying paralanguages**

**Table 1 Comparison results of different fusion methods**

| Fusion method | Precision (%) | Recall (%) | |
|---|---|---|---|
| | | Accuracy | F1-measurement |
| LW | 67.66±2.09 | 65.31±2.15 | 66.49±2.17 |
| DS | 66.93±2.99 | 64.67±2.65 | 65.80±2.87 |
| Bayes | 67.57±2.01 | 65.92±1.73 | 66.75±1.89 |
| SERBF | 65.29±1.95 | 64.17±1.85 | 64.73±1.90 |
| PFRA | 68.94±1.99 | 67.32±1.65 | 68.13±1.87 |

Note: LW stands for the linear weighted fusion method, DS stands for DS theory, Bayes stands for the Bayes fusion method, SERBF stands for the single-channel based speech emotion recognition method, which uses only the accompanying paralanguages, and PFRA stands for our proposed method