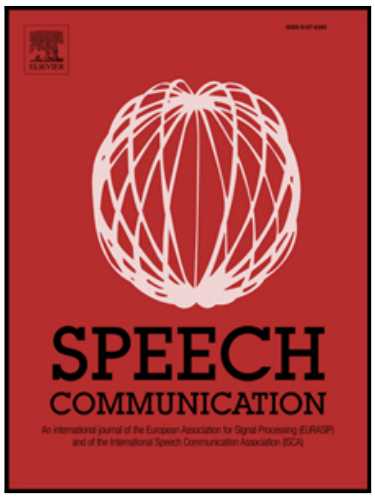

Fusion of Bottleneck, Spectral and Modulation Spectral Features for Improved Speaker Verification of Neutral and Whispered Speech

Milton Sarria-Paja, Tiago H. Falk

PII: S0167-6393(17)30470-3
DOI: 10.1016/j.specom.2018.07.005
Reference: SPECOM 2581

To appear in: *Speech Communication*

Received date: 22 December 2017
Revised date: 11 July 2018
Accepted date: 26 July 2018

Please cite this article as: Milton Sarria-Paja, Tiago H. Falk, Fusion of Bottleneck, Spectral and Modulation Spectral Features for Improved Speaker Verification of Neutral and Whispered Speech, *Speech Communication* (2018), doi: 10.1016/j.specom.2018.07.005

Highlights

- We have addressed the problem of speaker verification (SV) based on whispered speech.
- Two fusion schemes (score-level and i-vector level) were implemented to overcome existing challenges observed for the task at hand.
- The use of bottleneck features in the context of speaker verification using whispered speech.
- Addressed challenges: Short duration utterances (4.5 seconds average), no whispered speech data during enrollment from target speakers, and negative effects seen when adding whispered speech recordings during enrollment
- Dedicated systems per vocal effort offer a promising solution. To this end, a neutral/whispered speech classification system was implemented