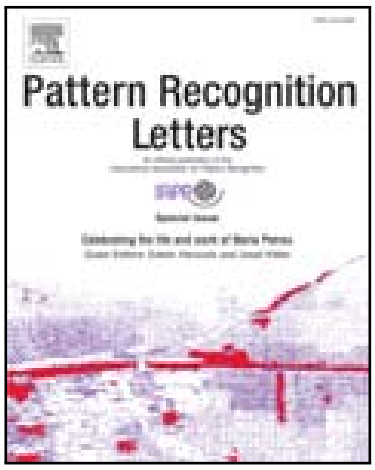

Exploiting Alternative Acoustic Sensors for Improved Noise Robustness in Speech Communication

Panikos Heracleous, Jani Even, Fumiaki Sugaya,
Masayuki Hashimoto, Akio Yoneyama

PII: S0167-8655(18)30304-0
DOI: 10.1016/j.patrec.2018.07.014
Reference: PATREC 7242

To appear in: *Pattern Recognition Letters*

Received date: 28 June 2017
Revised date: 11 March 2018
Accepted date: 6 July 2018

Please cite this article as: Panikos Heracleous, Jani Even, Fumiaki Sugaya, Masayuki Hashimoto, Akio Yoneyama, Exploiting Alternative Acoustic Sensors for Improved Noise Robustness in Speech Communication, *Pattern Recognition Letters* (2018), doi: 10.1016/j.patrec.2018.07.014

Authorship Confirmation

Please save a copy of this file, complete and upload as the "Confirmation of Authorship" file.

As corresponding author I, Panikos Heracleous, hereby confirm on behalf of all authors that:

1. This manuscript, or a large part of it, has not been published, was not, and is not being submitted to any other journal.
2. If presented at or submitted to or published at a conference(s), the conference(s) is (are) identified and substantial justification for re-publication is presented below. A copy of conference paper(s) is(are) uploaded with the manuscript.
3. If the manuscript appears as a preprint anywhere on the web, e.g. arXiv, etc., it is identified below. The preprint should include a statement that the paper is under consideration at Pattern Recognition Letters.
4. All text and graphics, except for those marked with sources, are original works of the authors, and all necessary permissions for publication were secured prior to submission of the manuscript.
5. All authors each made a significant contribution to the research reported and have read and approved the submitted manuscript.

Signature Panikos Heracleous Date July 14, 2017

List any pre-prints:

Relevant Conference publication(s) (submitted, accepted, or published):
[1] P. Heracleous, J. Even, C. T. Ishi, T. Miyashita, N. Hagita, "Fusion of standard and alternative acoustic sensors for robust automatic speech recognition", in Proc. of IEEE ICASSP 2012. DOI: 10.1109/ICASSP.2012.6289002

Justification for re-publication:
In [1], preliminary experimental results on automatic speech recognition when using body-conducted acoustic sensors are reported. In addition, a simple concatenative feature fusion and a late fusion method are also introduced.
The current comprehensive study on human-human and human-machine communication includes also some of the results reported in [1]. However, the current study significantly differs from the [1], and the differences are as follows:

- Instead of the very simple, and often problematic, concatenative feature fusion, the current study focuses on the more advanced and effective multi-stream HMM decision fusion. Using this fusion method, higher word accuracies were obtained compared to the concatenative feature fusion.
- In [1], the acoustic models used were trained using clean training speech data. For testing, clean and real noisy data were used. In the current study, acoustic models trained using artificial noisy training data were tested using both artificial noisy data (matched case) and real noisy data. This experiment aims to show the noise robustness of the body-conducted sensors.
- In the current study, ANOVA and t-test significance statistical tests are introduced and applied in the comparison of several fusion methods and comparison of different acoustic sensors.
- The current study contains the results of subjective evaluation of three acoustic sensors in noisy and clean environments, and when using male and female stimuli. Specifically, the authors were interested in analysing and measuring the behaviour of the three acoustic sensors in clean and noisy environment due to their importance in human-human communication (e.g., soldiers, pilots, etc.).
- An important phenomenon, which occurs in speech communication in noisy environments is also addressed. Specifically, the Lombard effect when using body-conducted sensors is also analysed. The analysis shows the appearance of the Lombard effect when using body-conducted sensors.