

## Accepted Manuscript

Front-end Speech Enhancement for Commercial Speaker Verification Systems

Sefik Emre Eskimez, Peter Soufleris, Zhiyao Duan,  
Wendi Heinzelman

PII: S0167-6393(17)30248-0  
DOI: [10.1016/j.specom.2018.03.008](https://doi.org/10.1016/j.specom.2018.03.008)  
Reference: SPECOM 2548



To appear in: *Speech Communication*

Received date: 30 June 2017  
Revised date: 26 February 2018  
Accepted date: 15 March 2018

Please cite this article as: Sefik Emre Eskimez, Peter Soufleris, Zhiyao Duan, Wendi Heinzelman, Front-end Speech Enhancement for Commercial Speaker Verification Systems, *Speech Communication* (2018), doi: [10.1016/j.specom.2018.03.008](https://doi.org/10.1016/j.specom.2018.03.008)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Front-end Speech Enhancement for Commercial Speaker Verification Systems

Sefik Emre Eskimez<sup>a</sup>, Peter Soufleris<sup>b</sup>, Zhiyao Duan<sup>a</sup>, Wendi Heinzelman<sup>a</sup>

<sup>a</sup>University of Rochester, 500 Wilson Blvd, Rochester, NY 14627, USA

<sup>b</sup>Voice Biometrics Group, 12 Penns Trail, Newtown, PA 18966, USA

---

## Abstract

Commercial speaker verification systems are an important component in security services for various domains, such as law enforcement, government, and finance. These systems are sensitive to noise present in the input signal, which leads to inaccurate verification results and hence security breaches. Traditional speech enhancement (SE) methods have been employed to improve the performance of speaker verification systems. However, to the best of our knowledge, the impact of state-of-the-art speech enhancement techniques has not been analyzed for text-independent automatic speaker verification (ASV) systems using real-world utterances. In this work, our contribution is twofold. First, we propose two deep neural network (DNN) architectures for SE, and we compare the performance of the proposed networks with the existing work. We evaluate the resulting SE networks using the objective measures of perceptual evaluation of speech quality (PESQ) and short-time objective intelligibility (STOI). Second, we analyze the performance of ASV systems when SE methods are used as front-end processing to remove the non-stationary background noise. We compare the resulting equal error rate (EER) using our DNN based SE approaches, as well as existing SE approaches, with real customer data and the freely available *RedDots* dataset. Our results show that our DNN based SE approaches provide benefits for speaker verification performance.

---

*Email addresses:* [eeskimez@ur.rochester.edu](mailto:eeskimez@ur.rochester.edu) (Sefik Emre Eskimez),  
[peter.soufleris@voicebiogroup.com](mailto:peter.soufleris@voicebiogroup.com) (Peter Soufleris), [zhiyao.duan@rochester.edu](mailto:zhiyao.duan@rochester.edu)  
(Zhiyao Duan), [wendi.heinzelman@rochester.edu](mailto:wendi.heinzelman@rochester.edu) (Wendi Heinzelman)

Download English Version:

<https://daneshyari.com/en/article/6960619>

Download Persian Version:

<https://daneshyari.com/article/6960619>

[Daneshyari.com](https://daneshyari.com)