

Accepted Manuscript

UPDATING THE SILENT SPEECH CHALLENGE BENCHMARK  
WITH DEEP LEARNING

Yan Ji , Licheng Liu , Hongcui Wang , Zhilei Liu , Zhibin Niu ,  
Bruce Denby

PII: S0167-6393(17)30096-1  
DOI: [10.1016/j.specom.2018.02.002](https://doi.org/10.1016/j.specom.2018.02.002)  
Reference: SPECOM 2533



To appear in: *Speech Communication*

Received date: 9 March 2017  
Revised date: 19 January 2018  
Accepted date: 20 February 2018

Please cite this article as: Yan Ji , Licheng Liu , Hongcui Wang , Zhilei Liu , Zhibin Niu , Bruce Denby , UPDATING THE SILENT SPEECH CHALLENGE BENCHMARK WITH DEEP LEARNING , *Speech Communication* (2018), doi: [10.1016/j.specom.2018.02.002](https://doi.org/10.1016/j.specom.2018.02.002)

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.

# UPDATING THE SILENT SPEECH CHALLENGE BENCHMARK WITH DEEP LEARNING

Yan Ji<sup>1</sup>, Licheng Liu<sup>1</sup>, Hongcui Wang<sup>2</sup>, Zhilei Liu<sup>1</sup>, Zhibin Niu<sup>1</sup>, Bruce Denby<sup>1</sup>

<sup>1</sup>Tianjin University, Tianjin, China; <sup>2</sup>Zhejiang W.E. University, Hangzhou, China

## ABSTRACT

The term “Silent Speech Interface” was introduced almost a decade ago to describe speech communication systems using only non-acoustic sensors, such as electromyography, ultrasound tongue imaging, or electromagnetic articulography. Although the use of specialized sensors in speech processing is challenging, silent speech research remains an active field that can often profit from new developments in traditional acoustic speech processing – for example recent advances in Deep Learning. After an overview of Silent Speech Interfaces and their special challenges, the article presents new results in which a 2010 benchmark study, called the Silent Speech Challenge, is updated with a Deep Learning strategy, using the same input features and decoding strategy as in the original Challenge article. A Word Error Rate of 6.4% is obtained with the new method, compared to the published benchmark value of 17.4%. Additional results comparing new auto-encoder-based features with the original features at reduced dimensionality, as well as decoding scenarios on two different language models, are also presented. The Silent Speech Challenge archive has furthermore been updated to contain both the original and the new auto-encoder features, in addition to the original raw data.

*Index Terms*—*silent speech interface, multimodal speech recognition, deep learning, language model*

## 1. INTRODUCTION

### 1.1. Silent speech interfaces and challenges

A Silent Speech Interface [1], or SSI, is defined as a device enabling speech processing in the absence of an exploitable audio signal – for example, speech recognition obtained exclusively from video images of the mouth, or from electromagnetic articulography sensors (EMA) glued to the tongue. Classic applications targeted by SSIs include:

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات