

Distant Speech Recognition Using Laser Doppler Vibrometer

Authors : Yunbin Deng

Abstract : Most existing applications of automatic speech recognition relies on cooperative subjects at a short distance to a microphone. Standoff speech recognition using microphone arrays can extend the subject to sensor distance somewhat, but it is still limited to only a few feet. As such, most deployed applications of standoff speech recognitions are limited to indoor use at short range. Moreover, these applications require air passway between the subject and the sensor to achieve reasonable signal to noise ratio. This study reports long range (50 feet) automatic speech recognition experiments using a Laser Doppler Vibrometer (LDV) sensor. This study shows that the LDV sensor modality can extend the speech acquisition standoff distance far beyond microphone arrays to hundreds of feet. In addition, LDV enables 'listening' through the windows for uncooperative subjects. This enables new capabilities in automatic audio and speech intelligence, surveillance, and reconnaissance (ISR) for law enforcement, homeland security and counter terrorism applications. The Polytec LDV model OFV-505 is used in this study. To investigate the impact of different vibrating materials, five parallel LDV speech corpora, each consisting of 630 speakers, are collected from the vibrations of a glass window, a metal plate, a plastic box, a wood slate, and a concrete wall. These are the common materials the application could encounter in a daily life. These data were compared with the microphone counterpart to manifest the impact of various materials on the spectrum of the LDV speech signal. State of the art deep neural network modeling approaches is used to conduct continuous speaker independent speech recognition on these LDV speech datasets. Preliminary phoneme recognition results using time-delay neural network, bi-directional long short term memory, and model fusion shows great promise of using LDV for long range speech recognition. To author's best knowledge, this is the first time an LDV is reported for long distance speech recognition application.

Keywords : covert speech acquisition, distant speech recognition, DSR, laser Doppler vibrometer, LDV, speech intelligence surveillance and reconnaissance, ISR

Conference Title : ICSLP 2019 : International Conference on Spoken Language Processing

Conference Location : Bali, Indonesia

Conference Dates : January 14-15, 2019