

Speech signal processing technology for smart devices to achieve multilingual speech translation service

November 17 2015

Hitachi today announced that it has developed a speech signal processing technology for smart devices to achieve a better multilingual speech translation service on the market. By removing background noise excluding speaker's voice, this innovative technology offers a speech recognition capability in noisy urban street environments in which its noise level is 70 dB. In addition, its automatic detection of speech intervals enhances usability with an accurate recognition of speech timing without requiring user to press a button for determining the intervals. This technology will contribute to the commercialization of the multilingual speech translation service at service counters in various stores or at information center in public transportation systems.

As the growing popularity of visiting Japan, the number of foreign tourists has been increasing every year. Consequently, a demand of multilingual [speech translation](#) services is rising from the practical needs of performing effective communications between foreign tourists and local service counter clerks without feeling language barrier in public transportation services or shopping centers.

However, in a crowded and noisy environment such as public transportation or shopping center, to specifically recognize speaker's voice for translation service is quite challenging due to the background [noise](#) that is recorded by microphone. In order to enhance [noise reduction](#), Hitachi has been developing the innovative noise reduction

[technology](#) on special purpose device using multiple microphones. Furthermore, an issue of conventional multilingual speech translation service is that users must press a button for translating each phrase of their conversations. This is very inconvenient for users when they often carry many bags in a situation of visiting service counter for information or services.

Based on the speech signal processing technology that has been cultivated by Hitachi for many years, Hitachi has developed a speech signal technology for general purpose [smart devices](#) instead of special purpose device. This newly developed technology has achieved the multilingual speech translation using smart device under a crowded environment such as public transportation area or shopping center. It is also capable of automatically recognizing speech intervals accurately without pressing any button to determining speech timing for translation.

The following are the features of the developed speech signal processing technology.

1. Noise reduction utilizing microphone inputs of multiple smart devices

In the conventional multi-microphone-based noise reduction technology on special purpose devices, noise is reduced by using the time difference among the microphones. Specifically, its process is to collect speaker's voice that is closest to one microphone first, then to collect other voices from other microphones. The voice processing is to identify the direction of the targeted speech source and remove any noise from other directions. This technology is not easy to apply to the smart devices available on the market due to the slight differences among the devices that cause small gap in recording timing. To solve this problem, the developed technology separates target's voice and background noise

using the differences of sound energy that is less easily to be influenced by timing gap of noise signals. Then, by correcting the time differences from timing gap of noise signals while comparing sound sources, the high-accuracy noise reduction using the time-difference-based approach as same as special purpose devices has been achieved.

2. Decreasing the time for speech input

The newly developed speech signal processing technology is capable of reducing noise and enhancing user's voice that offers accurate automatic recognition of speech intervals. As a result, there is no need to press any button for determining speech intervals. Furthermore, it is capable of decreasing the input time, and responding to continuous input for simultaneous translation for each phrases as live chat due to the accurate speech intervals.

The newly developed technology performs its speech processing and translation on the cloud system. Therefore, users can use this system easily by installing the dedicated application into the existing smart devices.

To confirm the performance of this innovative technology, we constructed a prototype system using a multilingual speech translation engine developed by National Institute of Information and Communications Technology and two general purpose smart devices, and carried out a validation experiment. As a result, we confirmed that the developed technology is capable of translating [speech](#) in a noisy urban street environment in which the noise level is 70 dB.

Hitachi will promote the development of this technology for practical applications, and contribute to provide high satisfactory hospitality services to Japan where many foreigners will visit.

Provided by Hitachi

Citation: Speech signal processing technology for smart devices to achieve multilingual speech translation service (2015, November 17) retrieved 25 April 2024 from

<https://phys.org/news/2015-11-speech-technology-smart-devices-multilingual.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.