TYPE OF INDUSTRY

科学技術・大学

情報通信研究機構

NICT 先端研究 ⑯

雑音除去 深層学習と効果的に併用

情報通信研究機構（NICT）が音声翻訳技術の実証実験として公開している多言語音声翻訳アプリ「VoiceTra」は、スマートフォンアプリであるため受音部（マイク）と口の距離が近く、屋内外で利用しても周囲の雑音は大きな問題ではないという認識であった。しかし、「VoiceTra」の音声ログを精査したところ、幾つかの国においては、極めて騒々しい環境での利用機会が多いという実態が判明した。

また、NICTからライセンス提供を受けた幾つかの企業は、タブレット端末などを用いたハンズフリー音声翻訳サービスを展開しつつあり、音声翻訳の入り口に当たる音声認識の耐雑音性が大きな課題となった。

この課題に対して筆者は、粒子フィルターという物体追跡手法を応用した雑音除去技術は実環境下での音声認識性能改善に大きく寄与し、幾つかの商用システムに実装された。

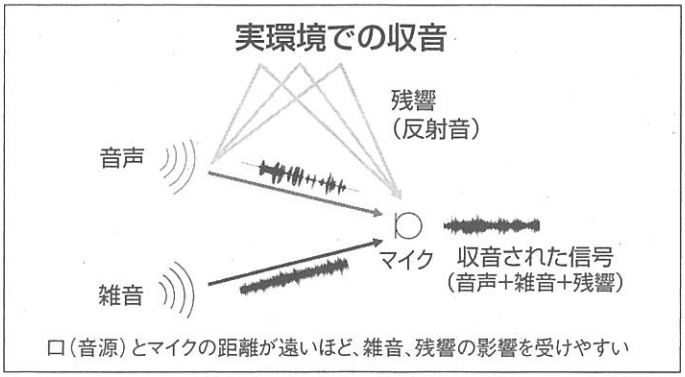
その後、収音信号から音声の時間区間を推定する音声区間検出技術の研究を行ったのち、並列非線形カルマンフィルターという手法を軸として音声区間検出と雑音除去の統合技術を考案した。この技術はさまざまなBツーBツーC事業を発掘し、商用音声認識エンジンのみならず、民生の収音機器など、多くの場面で活用されている。

その後、深層学習の台頭により情勢が大きく変わった。学習機構そのものが強力であるため、データさえ大量に集めてしまえば、雑音を除去せずとも十分な性能が得られるようになった。

雑音除去の有用性が薄れ、研究に対するアイデンティティーが崩壊しそうになったが、先ごろ雑音除去と深層学習を効果的に併用する方法を考案した。結果は良好であり、捲土重来の第一歩を踏み出した。

深層学習は強力な技術ではあるが、それだけでは解決できない課題も多々存在する。これまで培ってきた経験をもとに、今もなお、さまざまな角度から課題に挑戦している。

（火曜日に掲載）



先進的音声翻訳研究開発推進センター・先進的音声技術研究室主任研究員　藤本　雅清

04年龍谷大博士課程単位取得退学、05年学位取得。ATR、NTT研究所を経て16年より情報通信研究機構主任研究員。実環境音声認識に関する研究に従事。博士（工学）。