

「よく聴こえる」拡声システムのための音声合成システムの構築

室蘭工業大学 大学院工学研究科 しくみ情報系領域 助教 小林 洋介

写真1の**拡声器**は、簡便に利用できることから、音声情報伝達のためにあらゆる場所に設置されている。しかし、直感的に利用できるが故に最適な使い方がなされていなく、不明瞭で聴こえにくい。一般に聴き取りを改善する意図で、単純な音量調整（いわゆる「ボリュームを上げる」操作）が行われる。しかし、放送地点での単純な音量増加では、拡声される空間での反響・残響が増加し、不明瞭になってしまう。このような音声の改善について、電気音響・音声科学・聴覚心理分野の先行研究から以下の6点がすでに明らかとなっている。

- ① マイクロホンやアンプなどの機器設定を自分の発話に最適化
- ② 騒音の周波数特性を考慮したイコライゼーションにより音声の音色を調整する
- ③ ゆっくりと発話する・文節間の間（ポーズ）を長めにする
- ④ ボソボソと話さず、声を張り上げる（ロンバート発話と呼ばれる）
- ⑤ 発話中に「えー」や「あー」などのフィラーや間投詞を入れない
- ⑥ 発話する文章をより簡易な単語で簡潔な文章とする

これらのうち、①と②に関しては**機器の操作に習熟**していれば可能だが、現実的には難しい。さらに③から⑥の発話そのものの工夫には、声優やアナウンサーなどの**専門家のスキル**となっており、業務として拡声器を扱う人であってもほとんど意識されていない。特に災害時には、落ち着いた放送は精神的にも困難である。

この問題を解決するため本研究では、**機械学習**（≡いわゆる**人工知能**）を用いて自動的に聴き取りやすい音声に変換するシステムを実現する。具体的には図1のように発話そのままではなく、発話したい内容だけを簡潔に確実に放送できるように**音声認識**による発話内容の文字起こしと、絶対に伝えなければならない**情報の取得**と非常に明瞭に聴取できる**音声合成**の3点が必要である。これらの技術はどれも機械学習の応用分野として個々に研究されており、本研究では特に音声合成部の最適化を行い、**本人同様の声質**を目指す。

【将来実用化が期待される分野】

機械が能動的にサポートする音支援システムの実現



写真1 屋外拡声器の例

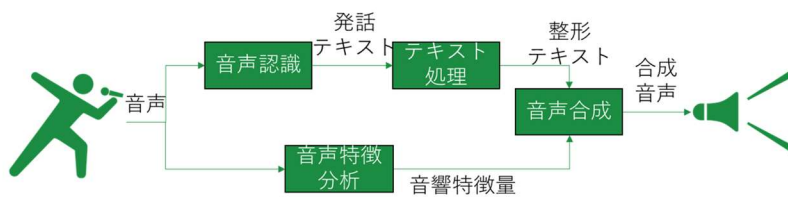


図1：提案システムの処理フロー