



## 主役登場

### “機械による声”が当たり前になる未来

## 井島 勇祐

NTTメディアインテリジェンス研究所  
特別研究員

「最近の音声合成では、こんな音声がつくれるようになったんだ」。NTTグループ内外で、一緒にお仕事をさせていただいている方々から、このようなお言葉をいただく機会が増えてきたと感じています。私は2009年の入社以来、一貫して音声合成に関する研究、プロダクト開発に従事しており、特集記事で紹介している深層学習に基づく最新の音声合成エンジン「Saxe」も成果の1つです。この音声合成エンジンでは、例えばバーチャルアバターがアニメの1シーンのようなツンデレ風の音声で返答をしてくれるといった、これまでの音声合成技術では実現が難しかった多種多様な表現が可能になっています。そして、この先の技術進展によっては、SF映画などで見かける、コミュニケーションロボット等が自分の身近な人の声で応対してくれる、あたかも人間のような感情表現が可能なパーソナルエージェントといった「機械による声」が当たり前になる未来の実現は夢ではなくなりつつあります。

こうした未来の実現に向けた課題はまだ多くありますが、その1つは声による表現だと考えています。私は業務で、発声のプロであるアナウンサー、声優とお仕事をさせていただく機会が多くあります。そのたびに、プロによる表現力にただただ驚かされるのと同時に、現在の音声合成技術には至っていないことが数多くあるのだと痛感させられます。例えば、台本や小説からキャラクターの心情や人間関係を汲み取って、それを声によって表現することができる表現力の多様性、私たちやディレクターからの表現

に対する指示を即座に理解して細やかに表現を修正することができる表現力の柔軟性等です。一方、現在の音声合成技術ではそういったことはできず、常に同じ表現の合成音声しかつくることができません。今後はこのような表現力の強化に関する研究を推進することで、現在は音声合成技術の適用が難しいサービス領域でも、音声合成技術を使っただけのようにしていきたいと考えています。

また、仮に素晴らしい研究成果が完成したとしても、それを広く世に使っていただくことができなければ意味はなく、研究成果を高いレベルでプロダクトとして開発することも非常に重要です。プロダクト開発のためには、サービスに応じて異なる処理速度、メモリ量、運用コスト等さまざまな要件をクリアする必要があるため、サービスを主管する研究所以外の方々との協力が必要不可欠です。厳しい要件にこたえるためには、アルゴリズムや実装における工夫等の研究とは異なる大きな困難が待ち受けています。しかし、新しいサービスを立ち上げようとする熱意のある方々と一緒に働くことができるのは、研究開発を進めるうえで大きなモチベーションとなると感じています。

今後も私の目標とする「機械による声」が当たり前になる未来の実現に向けて、NTTグループ内外の方々と協力しながら、研究とプロダクト開発の両面での活動を続けていきたいと思えます。