

一覧表で読み解く 音声生成 AI ワールド

佐藤 聖

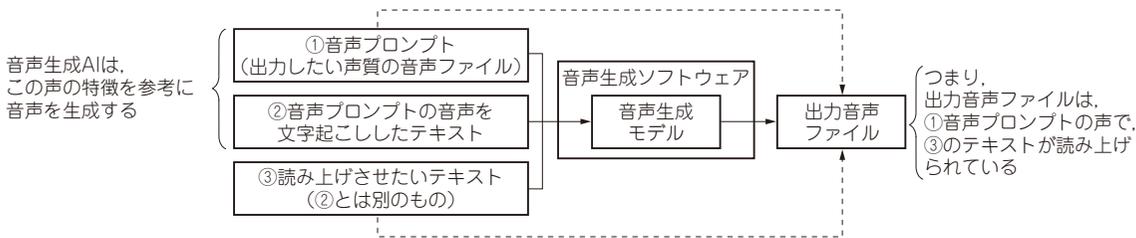


図1 音声生成AIで音を生成するイメージ

音声ファイルの中で話されている文章(②のテキスト)と違う内容を、テキストで入力することで、声は同じなのに話される内容が変わる

言葉や発音のニュアンスには感情や情緒が込められ、これがコミュニケーションの深みを構成します。人間同士の対話や情報の伝達において、音声は豊かなコミュニケーションを実現するための鍵となっています。最近の音声生成AIは進歩が著しく、人間同士のよう自然な対話、多言語対応、アクセシビリティ向上、カスタマイズ性など、多岐にわたる利点を提供しています。これが日常生活やビジネス環境において、新しい可能性を広げる技術として注目されています。

本稿では音声生成AIのおさえておきたい基礎知識を解説します。

音声生成のイメージ

音声生成AIは、テキストを自然な音声に変換する技術です。音声生成AIを活用した音声生成ソフトウェアは、入力されたテキストを解析し、音声生成モデルから生成された人間の声のような音声波形を出力します。音声として生成された音は、厳密には人間の声ではありませんが、人間の声とAIが作り出した音声を判別することが難しくなっています。現在、テキストだけでなく音声も入力できるものがあります。これを音声プロンプトと言います。つまり、音声やテキストから別の音声を生成することができます(図1)。

音声生成AI注目の理由

音声は、テキスト情報よりも情報量が多く感情やニュアンスをより豊かに伝えることができます。従来は、コンピュータや装置から一方的な情報提供という使い方が多かったのですが、図2、図3のようなAI連携により、人間とコンピュータ、人間と人間のコミュニケーションが変化すると考えられます。次に想定される音声生成AIの主な活用例を示します。

● 音声アシスタントとの対話

従来の音声アシスタントは、ストーリー・ベースの会話しかできませんでしたが、近年では音声合成技術から音声生成AIに置き換わり始めています。音声生成AIに置き換わることで、自然な会話だけでなく、さまざまな言語にも対応し、ユーザの使用言語に合わせた感情やニュアンスも再現できるようになるでしょう。

● 異なる言語間でのコミュニケーション

音声生成AIは、異なる言語間のコミュニケーションの促進に一役買うと考えられます。音声生成AIは異なる言語やアクセントに柔軟に対応できます。これにより、国際的なビジネスや交流が円滑に進み、異なる文化や言語圏の人々が効果的にコミュニケーションをとることが可能となっています。