

音声認識技術

コンピュータと人、あるいは人と人とのコミュニケーションを支援するメディア処理技術として、NTTでは幅広いサービス分野に適用可能な音声認識技術（開発名：VoiceRex）の研究開発を実施しています。VoiceRexは、事前の学習を要せずに、マイクや電話・スマートフォンからの音声入力に対して、高い認識精度を発揮できる特長があります。また、小学生から高齢者までの幅広い年齢層の音声にも対応可能です。また、従来水準をはるかに上回る、1,000万単語を登録した音声認識用辞書を用いたリアルタイム連続音声認識も実現可能です。また、人間の脳の動作を模した高度な情報処理方式を音声認識にも応用し、認識精度を飛躍的に向上させることに成功しました。

応用サービス例としては、自治体などにおける議会録作成支援があります。議会の審議音声を音声認識により高精度に文字化。誤り部分のみを手で修正することによって、効率的に議事録を作成することが可能です。また、利用者がシステムに話しかける自然な発話の認識にも対応しているので、音声対話によって情報検索を行うサービスなどにも適用可能です。さらに、人と人が自由に会話する話し言葉音声の認識技術の開発を進め、今後も、快適、便利な音声コミュニケーションの実現をめざして研究開発を進めます。

●自由発話音声認識技術(VoiceRex)とその応用例

