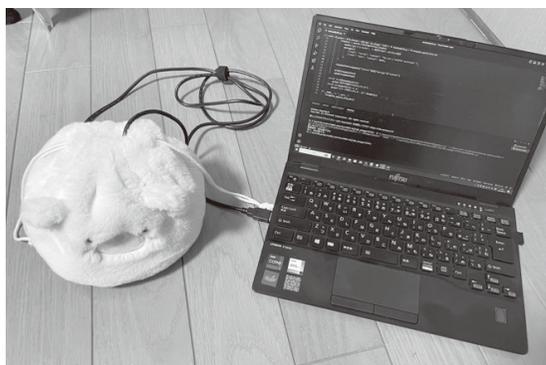


ChatGPT と会話する装置を作る

古川 玲



(a) 装置全体



(b) 内部にはマイクとスピーカ

写真1 ChatGPT と会話できる装置を作る

トライすること

ChatGPTには対話型のウェブUIが用意されています。テキストを打ち込むと、内容に沿った回答を返してくれます。質問に対して的を射た回答や、ちょっと知ったかぶりをした回答など、多岐にわたる反応を返してくれ、まるで人が回答しているかのように思えます。これがテキストでのやりとりではなく、音声による会話でやりとりができれば、実際に人とコミュニケーションを取っているように感じるのではないかと思います(図1)。最近のPCにはマイクやスピーカが付いていることが多く、音声認識用のAPIも充実してきています。そこでChatGPTと会話できる装置を、ChatGPTに聞きながら作ってみましょう(写真1)。

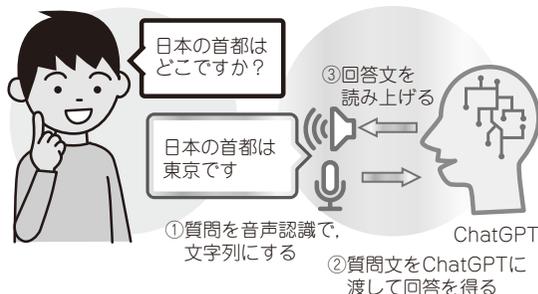


図1 ChatGPT と会話する装置に必要な3つの処理

ステップ1：マイク入力を音声認識で文字列にする

まず、PCに接続したUSBマイクより入力した音声を認識し、文字列にする処理を実装する方法をリスト1のようにChatGPTに聞いてみます。

リスト1はspeech_recognitionライブラリを使って、USB接続したマイクから音声を取得し、グーグルの音声認識APIを使って音声をテキストに変換します。日本語の音声認識を行うためにlanguage='ja-JP'を設定しています。

プログラムを実行すると、コマンドラインに「話しかけてください...」と表示されるので、音声を入力します。音声認識された場合は、その結果がテキストで表示されます。グーグルの音声認識APIを利用しているため、APIキーが必要です。

得られたコードをVisual Studio CodeとPython 3.11環境で試します。

ライブラリのインストールで提示されたコマンドを実行すると、SpeechRecognitionがインストールされます。

USB接続のマイクをPCに接続し、サンプル・コードを実行し、マイクに話しかけると、音声認識結果が出力されます。これで、声に出した内容をPCのプログラムに渡すことができそうです。