

第2章 手ぶらで俺的AIライフ・ロガーを作る

サンプルで初体験! ラズパイ×カメラで人工知能

ご購入はこちら

金田 卓士

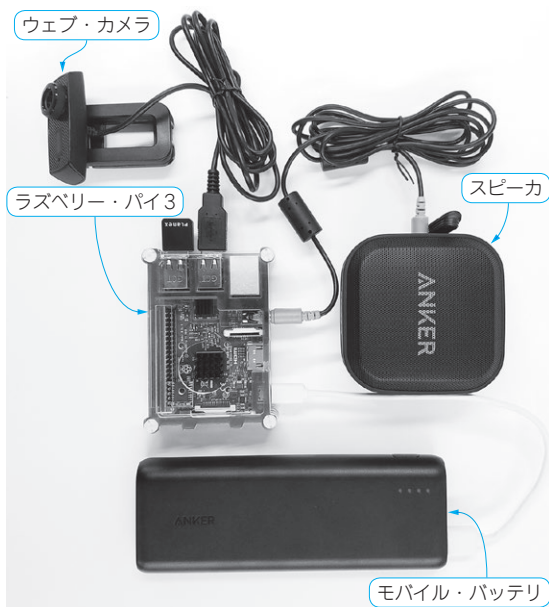


写真1 ラズベリー・パイ3とカメラを使って行動を人工知能解析してテキストに残す「俺的AI日記コンピュータ」

本稿では、カメラ画像を10秒ごとに記録して、画像から自動生成したキャプション(説明文)を音声で発話するという「俺的AI日記コンピュータ」を作成します(写真1)。ラズベリー・パイ3とクラウドAI「Microsoft Cognitive Services(以下MCS)」を使っています。

ラズパイ×カメラ×クラウドAPIで作る「俺的AI日記コンピュータ」

● ハードは5000円/ソフトは無料で試せる

俺日記を生成するには、記録装置を常に携帯する必要があります。スマホを使ってもよかったのですが、乱暴に扱って壊したくありません。そこで5000円で購入できるLinuxコンピュータであるラズベリー・パイを利用します。

マイクロソフトのAPIを利用した理由は、第1章でも述べましたが、同社が数々の画像認識コンテストで優勝しており、画像認識では他社よりも優れていると

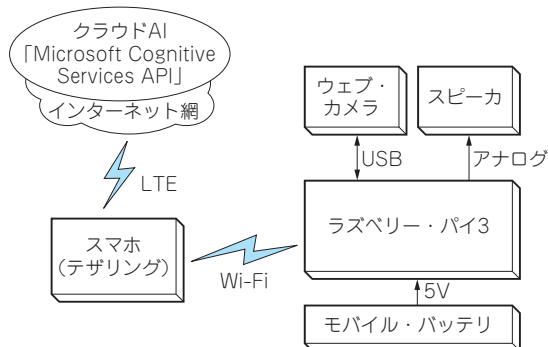


図1 俺的AI日記コンピュータの人工知能はマイクロソフトのクラウドAPI & サンプルで実現する

判断したからです。また、今回の「画像から文章を自動生成」してくれるAPIを持っているのも同社だけです。このAPIは5000トランザクション/月まで無料で使えるので、およそ5000回、画像から文章を生成できます。

● システム構成

システム構成は図1の通りです。ラズベリー・パイ3にUSB接続のウェブカメラとスピーカーが接続されており、Wi-Fiを経由してインターネットに接続できるようになっています。そして、ウェブカメラから取得した画像を、インターネットを経由して、クラウドAIの各種APIに接続できるようになっています。Wi-Fiについてはスマートフォンのテザリングを利用しました。

● 鍵となる「日記」の自動生成にはクラウドのAPIを利用

この装置を作成するにあたり、肝になってくるのが画像からのキャプションの自動生成です。本装置ではMCSのComputer Vision API(以下CV API)を利用し、これを実現しています。CV APIは、RESTで画像を送信すると、JSON形式で画像の解析結果を返してくれるサービスです。表1に解析結果としてどのようなレスポンスが返ってくるかまとめました。図2に画像を音声として再生するまでの流れを示します。