

My アルゴリズムで聞き分け率アップ! ボイス制御マシンにも使える!?

# ラズベリー・パイでトライ! 音声認識

お試しプログラムも用意!



第5回

データベースの精度UPに欠かせない  
番外編：認識精度の評価方法

安藤 章悟

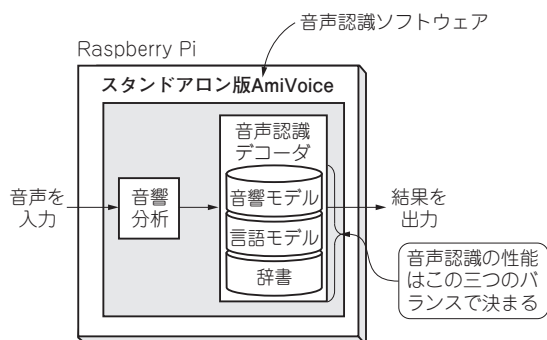


図1 音声認識の3種の神器データベース：音響モデル/言語モデル/辞書は評価をしながら性能を上げていく

## ● データベースの精度が認識のキモ!

音声認識処理で使う図1の音響モデル・言語モデル・辞書は、一度作ったら完成というわけではありません。学習データを増やしたり構造を見直すなどの改善を繰り返して、より良いエンジンを構築していくことになります。しかし、多少改善しても体感できないくらいの小さな差しか出ないことが多いため、常に性能を定量的に評価しながら、どのような工夫をするとどれだけ良くなるのかを見極めることが重要です。

この定量評価の指標によく使うものは、認識精度 (accuracy) と認識率 (correct) です。

## 認識性能の測定手順

認識精度と認識率は図2のようにして測定します。

### ▶ステップ1：測定したい音声と正解のテキストを用意する

まず、音声データと、その音声は何としゃべっているかを書いたテキスト・データ (正解文) の組を用意します。この音声とテキストの組を評価用データやテスト・セットと呼びます。

### ▶ステップ2：音声を認識させて結果を出力する

この評価用データの音声を、性能測定したい音響モデル・言語モデル・辞書を読み込んだ音声認識デコーダに処理させ、音声認識結果を出力します。

### ▶ステップ3：認識結果と正解のテキストを比較する

そして、この音声認識結果と正解文がどれだけ一致しているかを計算して、認識率を求めます。

## ● 性能の指標：認識精度と認識率

図3に認識精度と認識率の定義式を示します。

正解文と認識結果が完全に一致したときに認識率は100%になり、挿入誤りや削除誤り、置換誤りがあればその分だけ認識率が下がります。認識精度と認識率の違いは、挿入誤りを考慮するかどうかです。認識率は正解の単語がどれだけ出力されたかを表しているので、挿入誤りが入っていても数値は下がりにません。それに対して認識精度は認識結果がどれだけ正しいかを

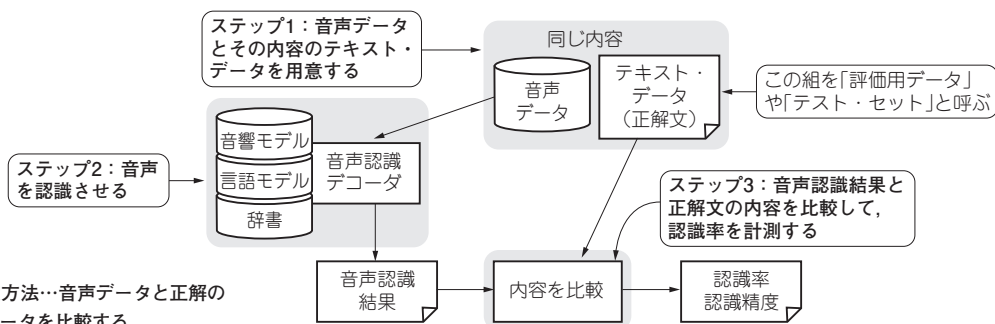


図2  
認識率の測定方法…音声データと正解の  
テキスト・データを比較する