

DP マッチングによる単語音声の認識

麻生英寿
21C1005

June 29, 2023

Contents

1 目的

DP マッチングのアルゴリズムを利用して、小語彙の音声認識の実験を行う。

2 実験方法

DP マッチングのアルゴリズムによって音声データの単語間距離を計算するプログラムを作成する。そのプログラムに、100 単語の音声データのテンプレートに対して、同じ発声内容の 100 単語を未知入力音声として、順に入力していく。入力された音声データの発声内容を判定し、その正答率を計算する。

2.1 使用したデータセットについて

この実験で使用した音声データセットには、2 人の話者がそれぞれ同じ 100 種類の単語を発話したものが含まれている。話者は 100 種類の単語の発話を 2 回行っているため、合計 400 個の音声データが含まれている。1 つの音声データにはファイル名、発声内容、フレーム数とフレーム数分の 15 次のメルケプストラム特徴量である。

2.2 単語間距離の計算

実験では DP マッチングのアルゴリズムを用いて 2 つの音声データの単語間距離を計算する。単語間距離の計算は以下の手順で行う。

入力として与えられた 2 つの音声データのフレーム長をそれぞれ I と J とする。 $(i, j), 0 < i \leq I, 0 < j \leq J$ で表せられる 2 つの音声データの各ノードのメルケプストラム特徴量の距離を計算し、 (i, j) でのフレームの距離を局所距離 $d(i, j)$ と表す。音声データの最初のフレームから任意のフレームまでの局所距離の総和を累積距離 $g(i, j)$ とする。最終フレームでの累積距離 $g(I, J)$ が最小となるような経路を探索することで、単語間距離を計算する。

2.3 最終フレームでの累積距離の計算

はじめに、初期条件を以下のように設定する。

$$g(0, 0) = d(0, 0)$$

次に境界条件を以下のように設定する。

$$j > 0 \rightarrow g(0, j) = g(0, j - 1) + d(0, j)$$

$$i > 0 \rightarrow g(i, 0) = g(i - 1, 0) + d(i, 0)$$

最後に、再帰的に以下の式を用いて、最終フレームでの累積距離 $g(I, J)$ を計算する。

$$g(i, j) = \min \begin{bmatrix} g(i, j - 1) & + & d(i, j) \\ g(i - 1, j - 1) & + & 2d(i, j) \\ g(i - 1, j) & + & d(i, j) \end{bmatrix}$$

3 実験結果

音声認識の正答率は次のようになった。

3.1 音声認識の正答率

モデル/認識対象	city011	city012	city021	city022
city011		99%	90%	84%
city012	100%		92%	86%
city021	83%	91%		99%
city022	86%	94%	100%	

Table 1: 音声認識の正答率

4 考察

A 付録

付録に関する内容をここに書きます。

A.1 付録 A

付録 A の内容をここに書きます。

A.2 付録 B

付録 B の内容をここに書きます。

参考文献

ここに参考文献の引用を書きます。

References

- [1] 著者名, タイトル, 出版社, 年.
- [2] 著者名, タイトル, 出版社, 年.