

知的システム構成論課題

矢入先生担当分

航空宇宙工学専攻修士一年
37-186305 荒居秀尚

2018 年 8 月 9 日

1 選択したテーマ

今回の課題に向けて選択したテーマは「隠れマルコフモデルと EM アルゴリズムを用いた音声データの潜在状態の推定」としました。音声データとしては、研究室の研究会の様子を録音したもの、約 160 分程度のものの一部を使用し、潜在状態として話者を分類することを目的として学習を行いました。今回のデータはサイズが大きかったため、前半 5 分を学習に使用し、その後の 5 分を評価用に用いました。前半に含まれているのは 5 人分の音声だったため、それに無音区間と雑音分のクラスを加えた 7 クラスへのクラスタリングのタスクとして HMM を用いました。

このテーマを選択した理由として、個人的に以前から研究会の文字起こしをしたいと考えており、特に話者を識別してアノテーションを加えたいという目的がありました。今回はそれを HMM を用いて出来ないか試してみたという例になります。

2 前処理

音声データを扱うにあたって、メル周波数ケプストラム係数に変換を行いました [1]。

参考文献

[1] 篠田浩一. 音声認識. 機械学習プロフェッショナルシリーズ. 講談社, 2017.