

音声データについての扱い

人類と人工知能

2122025 小林千沙希

今回は「音声データについての扱い」というテーマなので、音声データがどういうところに活用されているのか、また音声データ分析を Python で扱う手法を手順を示しながら述べていきたいと思う。

1. 音声データの活用事例

普段私たちは日常生活で音楽ソフトにある曲をプレイヤーに落としたいとき、mp3 変換というものをする。このように、私たちは普段音声データを用いる場面が予想より多い。だが視野を広くすれば、私たちの身の回りに応用されている。ここでは3つ紹介する。

- ・音声認識

声を発することで人物を特定、認証したり画像がなくても音のみの状況下で何が存在しているか判別したりすることができる。

- ・異常検知

動物や設備の異常検知で家畜の声の異常を検知して健康管理をしたり、機械設備の異常音を検知したりすることで故障を予知し、防ぐことができる。

- ・遠隔操作

特定の機械音を発すると街や駅のディスプレイ表示を切り替えるなどの遠隔操作が可能。

2. Python を用いて音声データ分析

- ・ライブラリ (librosa) をインストール

- ・音声データを読み込み

画像の読み込みと同様。読み込み可能なデータは wav,mp3。

- ・音声データの波形を描画

librosa.load で呼んだ音声データは Sin 波のグラフを描くことが可能。

- ・振幅の平均値を計算

データ毎に振幅の平均値を計算し、他のデータと比較することで音の特徴を確認できる。

- ・ゼロクロス数を計算

ゼロクロス数とは雑音の入り具合がわかる数値。データ毎に計算し他のデータと比較して特徴がないか見る。

- ・メルスペクトログラムを計算

音声波形を複数の sin 波に分解し周波数と振幅と時間の情報を持たせ、さらに高音低音で粗さを考慮したデータをメルスペクトログラムという。可視化して特徴を見る。