

声認証システムの精度向上についての検討

Y210122 小栗正也

指導教員 三好 力 教授

1. はじめに

近年、音声を鍵として生体認証を行う声認証システムの技術の利用が拡大してきている。声認証システムは虹彩認証や指紋認証とは異なり専用のセンサを必要とせずマイクを用意するだけで利用することができるため容易に利用でき、非接触のため電話などの通信を経由することでも認証が可能だという利点がある一方で、他の生体認証と比べて周りの環境に左右されるため識別精度の向上が困難である。そこで本研究では精度向上を図るために実際に音声を録音して雑音の処理を行い、機械学習を用いることで識別率が向上するかの検証を行う。

2. 提案手法

本研究では雑音除去の方法として音声信号に対して短時間フーリエ変換を行うことにより人間の音声の周波数帯以外に含まれる音声信号を雑音分と推定して窓関数をかけることによる手法と、発声源から近いマイクと遠いマイクの2つのマイクを使用してそれぞれのマイクが録音した音声信号の振幅差を利用することによって雑音の除去を行う。雑音の除去をした音声信号に対して特徴量の抽出を行って、QDA(二次判別分析)を用いて機械学習することによって識別を行う。また、距離の異なる2つのマイクの音声源からの到達時間差を求めることによって発生源の位置を推定し、発生源が人か機械なのかを見分けてなりすまし攻撃への対策を行う。

3. 実験

事前実験として2つのマイクを利用して音声を録音し、その音声に含まれている雑音を除去して特徴量を抽出し、機械学習で識別して発声推定の精度を測定したが、マイクを2本使用するとかえって精度が悪くなってしまったという結果になったため、実験は8人の音声を1本のマイクで録音した。

それによる実験結果の一部を図1に示す。

正答率		
ノイズ除去		
被験者	有り	無し
A	90%	80%
B	70%	70%
C	80%	60%
D	100%	100%
E	80%	80%
F	90%	80%
G	90%	90%
H	70%	80%

図1 実験結果

QDAを用いた識別ではノイズ除去有りだと正答率が平均83.75%で、除去無しだと平均80%というなり、SVMを用いた識別ではノイズ除去有りのだと63%、ノイズ除去無しだと58%という結果になった。全体的にQDAを用いた機械学習の方が正答率は高くなり、ノイズ除去を行うことでわずかに正答率が高くなるという結果が得られた。

4. 終わりに

本研究では話者照合システムの精度向上を目的として2本のマイクを使用する手法と短時間フーリエ変換によるノイズ除去とQDAでの機械学習の手法を提案したが、いずれも先行研究[1]の精度を超えることができず、有効性を示すことができなかった。原因としては2本のマイクを使うアイデアが失敗した事と、登録した音声データの数が少なかった事だと考えられる。今後の課題としては学習する音声データを増やして機械学習の精度を上げることと、本実験では有効でなかったなりすまし攻撃への対策として合成音声を見抜くシステムの導入を検討する。

参考文献

[1] NECによる音声識別

https://jpn.nec.com/press/201902/20190219_01.html