

録音位置の違いによる 音声信号のずれ検知

2020年 2月 19日

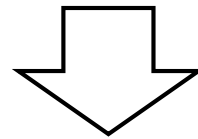
1200304 尾辻明里

ネットワーク信号処理研究室

背景

音声認識技術の発展

- スマートフォンの音声入力型情報検索システムなど広く利用
 - 目的音以外の音が混ざった場合に認識率が低くなる
- 選択的聴取ができていない



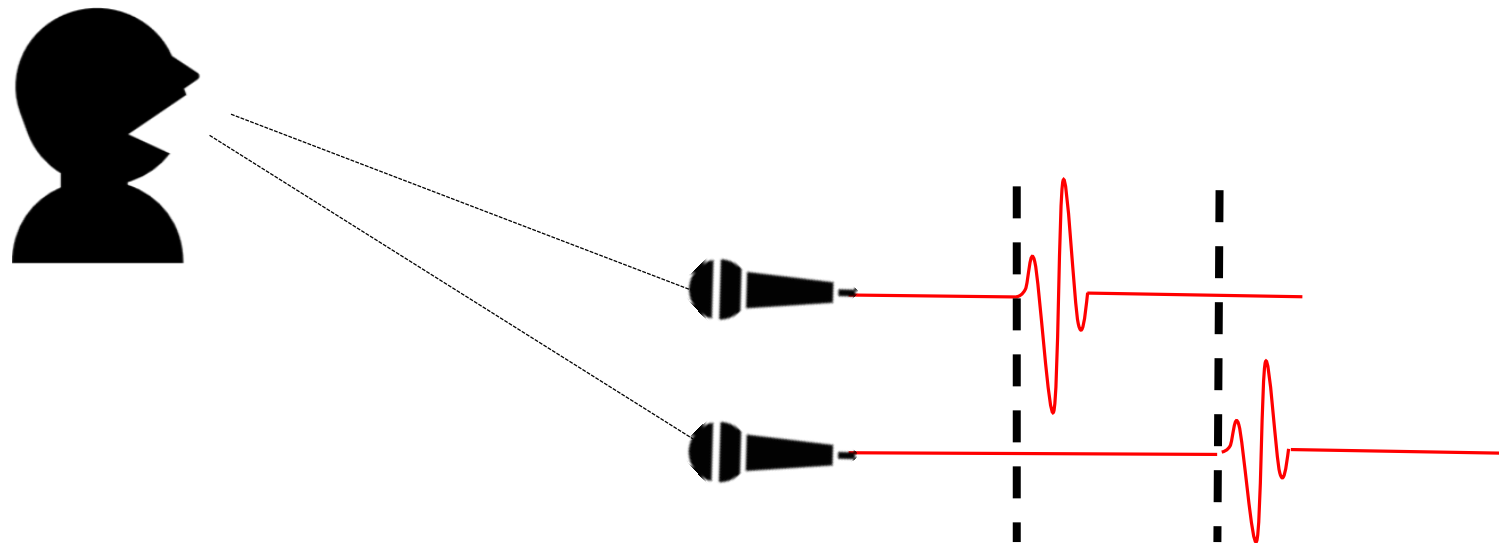
所望音のみを抽出する技術が必要

研究の目的

録音された音声の中から所望音のみを抽出する

➤複数の音が入った音から所望音以外の音を取り除く

→ (録音された音)-(所望音以外の音)=(所望音)



研究の目的

録音された音声の中から所望音のみを抽出する

➤複数の音が入った音から所望音以外の音を取り除く

→ (録音された音)-(所望音以外の音)=(所望音)



**録音位置の違いによる時刻のずれを検知し
所望信号を抽出する**

録音位置の違いによる音声信号のずれ検知

観測信号の時刻のずれはマイクと音源の距離の差による

入力時刻のずれ s は

$$s = f_s \times \frac{d}{c} \quad (f_s: \text{サンプリング周波数}, d: \text{マイク間の距離}, c: \text{音速})$$

と予測されるはず
→ 実際のずれを測定する

測定環境

- 使用した音声

- 破裂音と約1秒間の人間の声を含む8秒の録音音声

- サンプリング周波数 44.1kHz

- 場所

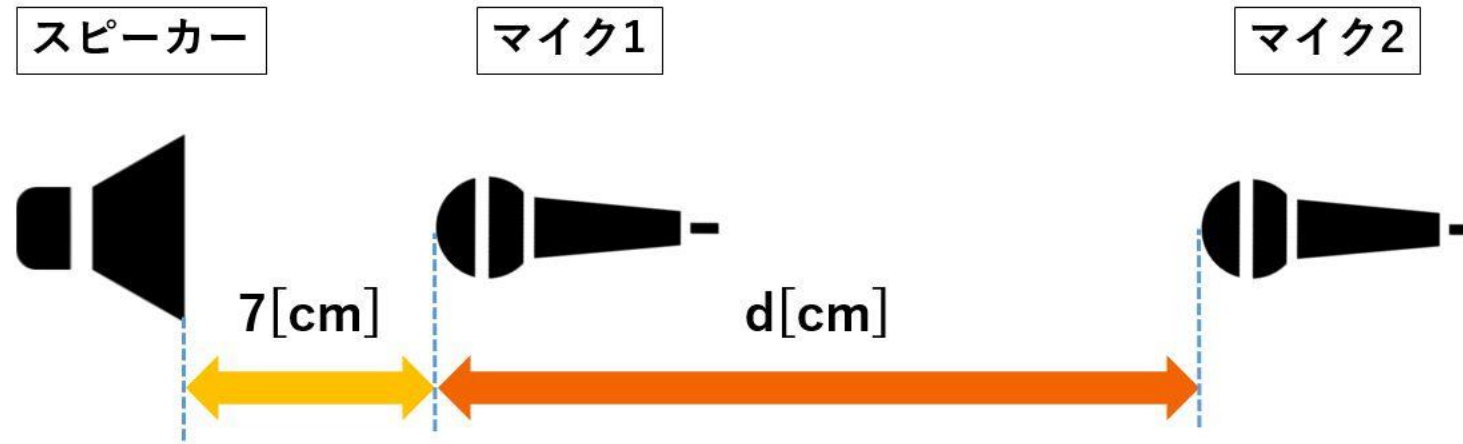
- 研究室内の共有スペース

- 気温

- 21.2度



ずれの測定



- マイク間の距離を20cm～25cmまで1cmずつ変化させる
- 信号の時刻のずれをマイク1とマイク2の信号の相互相関をとる
- 相関係数が一番高い → 信号のずれ(実際のずれ)とする

予測したずれの誤差

距離毎のずれの比較

マイク間の距離[cm]	20	21	22	23	24	25
予測されるずれ[サンプル]	25	26	28	29	30	32
実際のずれ[サンプル]	23	24	26	26	28	29
相関係数	0.901	0.897	0.915	0.924	0.920	0.917

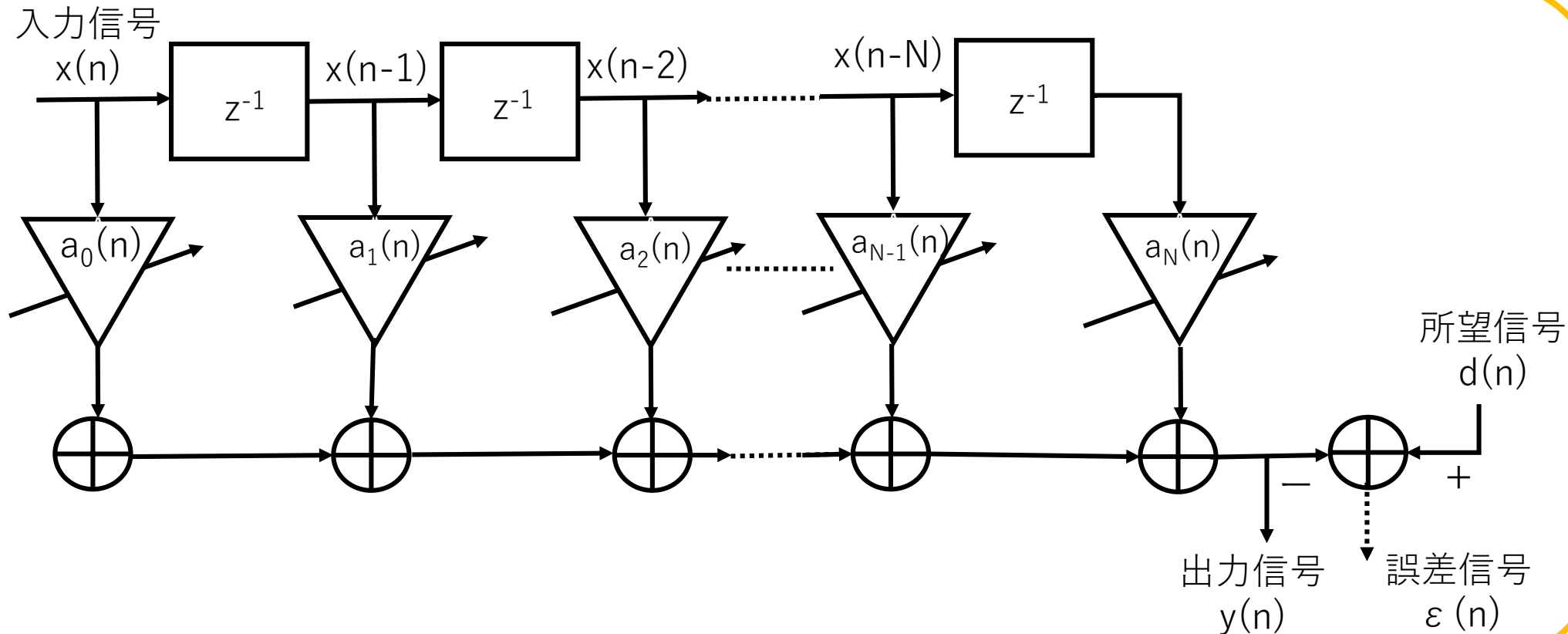
残響や減衰により信号がずれている



適応フィルタで時刻以外のずれを合わせる

適応フィルタ

$\varepsilon(n) = y(n) - d(n)$, $\|\varepsilon\|^2 \rightarrow 0$ ➡ 信号の一致を確認



まとめ

- 入力時刻のずれは相互相関をとることで検知
- 適応フィルタを使うことで信号の一致を確認

➤今後の課題

- スピーカーやマイクが一直線上にある場合のみ信号が一致
→ その他の配置でも信号の一致を可能にさせる