

概要

音源方向検出の性能向上を目指した

検討(1): 反射音の軽減

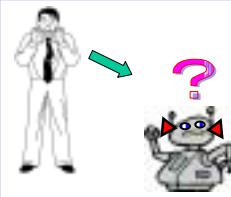
一般の室内では目的音方向とは異なった方向から来る反射音が存在する。この反射音の影響を軽減するため、ピークホールド処理を用い有効性の検討を行った。

検討(2): 音源による反射音の影響の違い

音声を用いて音源方向検出を行う上で、録音した音声を用いてスピーカから発生させる場合と、人間が実際に発声する場合で検出結果の違いが生じた。このことについて、検討を行った。

はじめに

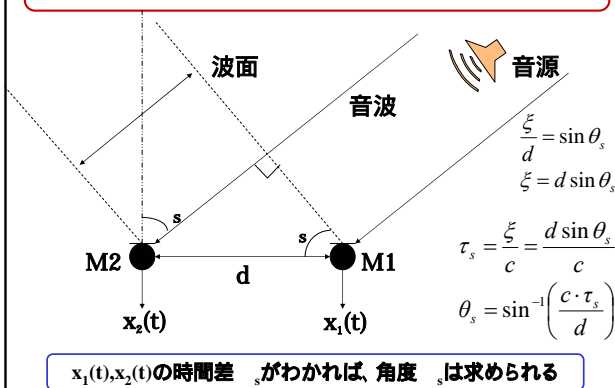
- 音源方向検出の必要性
 - ロボットの耳
 - 騒音の発生源、他
- 従来技術
 - 相関法(時間差検出)
 - 高分解能法(MUSIC、他)
- 問題点
 - 一般の室内では反射音の影響により検出性能が低下する。



本報告の流れ

- 提案法の有効性を様々な条件で確認
 - 拍手音を用いて
 - 大規模な講義室(丹羽ホール)
 - マイクホンを反射板に近づけた場合
 - 音声を用いて
 - マイクホンを2枚の反射板に近づけた場合
- スピーカと人間の特性による音源方向検出結果の違いについて検討

2つのマイクロホンによる音源方向検出

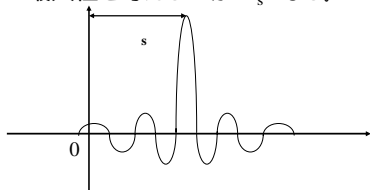


時間差 s の検出(従来法: 相互相関関数)

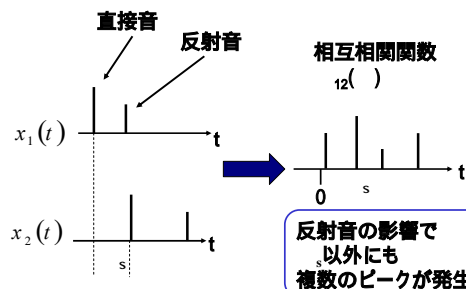
- 2つの受信信号 $x_1(t), x_2(t)$ の相互相関関数

$$\phi_{12}(\tau) = \frac{1}{N} \sum_{t=1}^N x_1(t) \cdot x_2(t+\tau) \quad \dots (1)$$

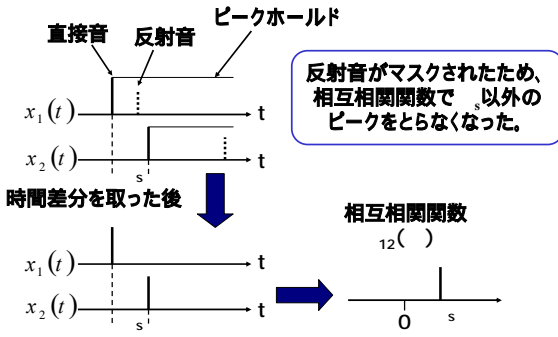
t : 離散時間
の最大値を与える τ が s となる。



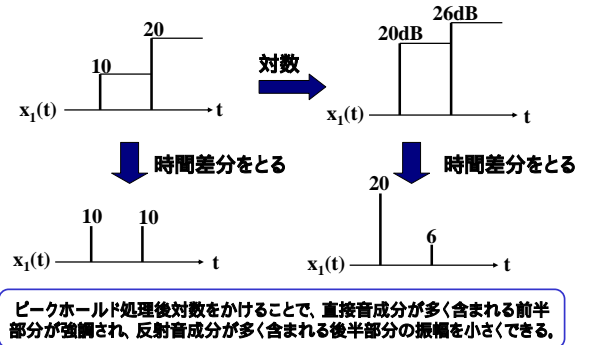
従来法に及ぼす反射音の影響



ピークホールド処理

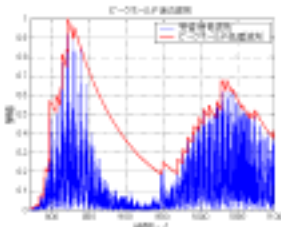


対数操作

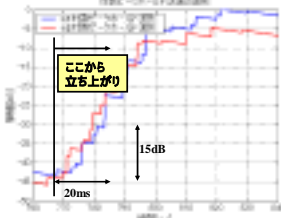


ピークホールド処理 + 対数操作 (実測例)

ピークホールド処理後の波形例



ピークホールド処理後対数操作を行った波形例



複数回の発生音に対応可能とするため、室内残響相当の減衰を持たせている。

20msの間に15dB以上の振幅変化の時発生音の立ち上がりのみなし、50msの区間で相関を取っている。

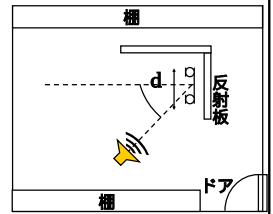
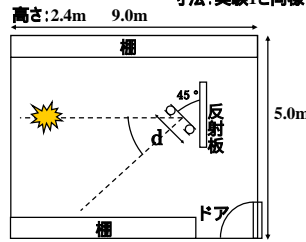
実験条件

共通
場所: 研究室
残響時間: 400[ms]

実験2 (右図)
音源: 男声3名
(スピーカ、人間)

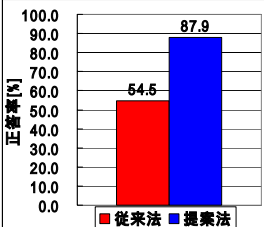
実験1 (左図)
音源信号: 拍手音
反射板: 1枚設置

音源信号: 音声
反射板: 2枚設置
寸法: 実験1と同様

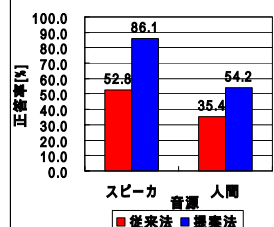


結果

実験1 (拍手音) 検出結果



実験2 (音声) 検出結果



±10°の範囲で、相互相関関数を用いた方法を従来法とし、正答率を判定

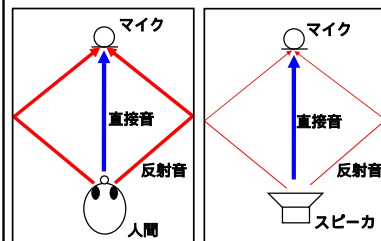
拍手音で十分な性能向上が確認できた。

1. 音声においても、従来法に比べ性能向上が確認できた。
2. 発声させた音源の違いで、音源方向検出結果にも違いが見られた。

検出結果の違い

指向性が弱い場合(人間)
発生した音が拡散する傾向にある。直接音に対して、反射音も同様な大きさでマイクに入力される。

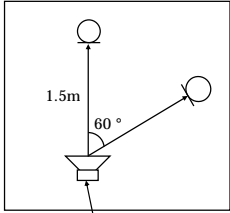
指向性が強い場合(スピーカ)
発生した音が向いている方向に大きく、それ以外は小さい。直接音に対して、反射音は小さい音でマイクに入力される。



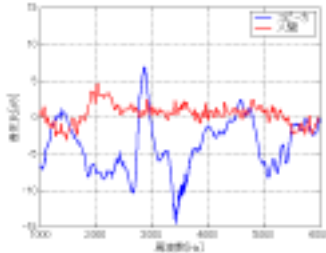
仮説1: 指向性の違い (人間は指向性が弱く、スピーカは指向性が強い)
仮説2: 人間が音源の場合受音信号には反射音が多く含まれる

スピーカと人間の指向特性

音源から正面と60°方向にマイクロホンを設置し録音



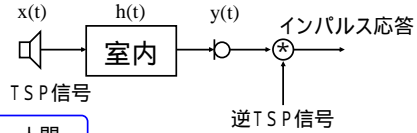
スピーカ、人間



人間: 正面方向に対して60°方向にも同様の音の大きさ **指向性弱い**
 スピーカ: 正面方向に対して60°方向は音が小さい **指向性強い**

インパルス応答測定

スピーカ



人間



インパルス応答使用理由
 直接音と反射音の分離が可能
 反射音成分の量を評価できる

音源による反射音の違い

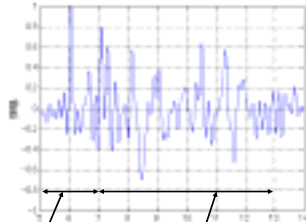
口元 マイクロホン間のインパルス応答

各音源(スピーカ、人間)のインパルス応答から、直接音と反射音のエネルギー比になるDR比を求める
 $DR比 = \frac{\text{直接音のエネルギー}}{\text{反射音のエネルギー}}$

各音源のDR比

音源	DR比[dB]
スピーカ	0.39
人間	-5.48

スピーカに対して人間の方が、約5dB反射音の影響が大きいことが確認できた。



直接音: ±1ms

反射音: 直接音以降、約6ms

結論

音源方向検出の性能向上を目指し、以下の結果を得た

結論(1): 反射音の軽減

1. 拍手音において、反射音の影響が大きい条件で、十分な性能向上が得られた。
2. 音声においても、反射音の影響が大きい条件で、性能向上した。しかし、十分な性能ではなかった。

結論(2): 音源による反射音の影響の違い

音源による検出結果の違いについては、各音源の指向特性の違いから、反射音成分の量に違いが出ていると言うことを、DR比で確認した。