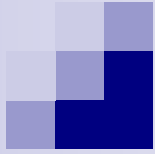


県立広島大学 ・ 経営情報学部
経営情報学科
オープンキャンパス2016 模擬講義資料

Prefectural University of Hiroshima Faculty of Management and Information Systems
Department of Management Information Systems

身近な音環境の話

環境情報処理研究室
折本寿子

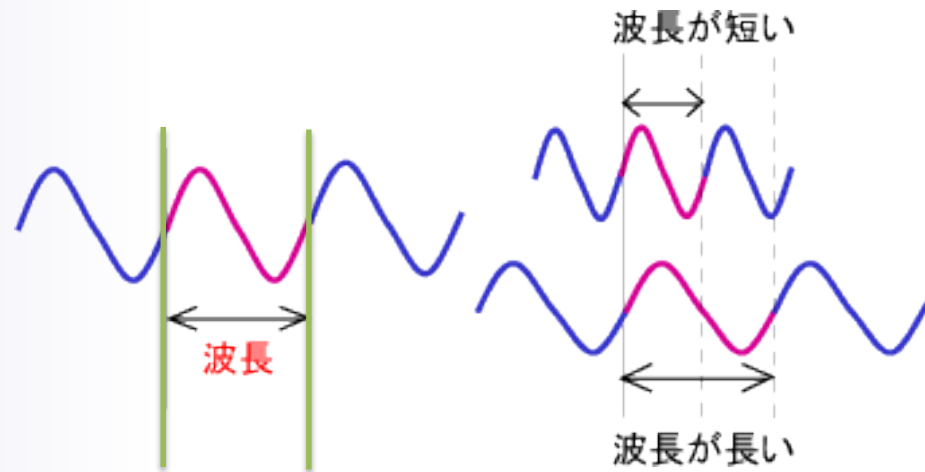


目次

1. 音の基礎知識
2. 音の性質
3. 身近な音環境
4. 音声復元のための情報処理
5. 音声復元の実験
6. まとめ

1.音の基礎知識(1)

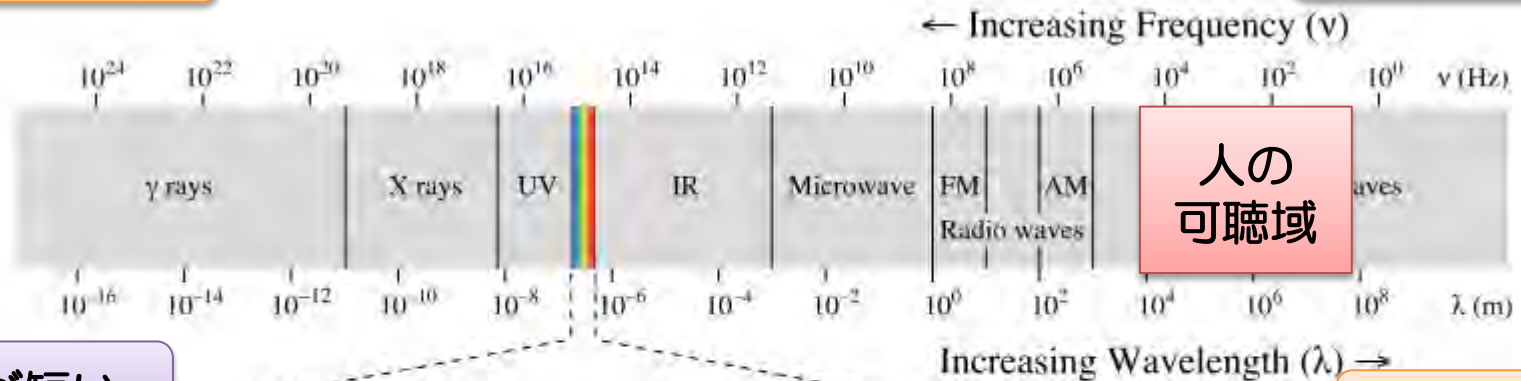
- 1オクターブ
周波数が2倍
- 周波数
1秒間に繰り返される振動の回数(単位:Hz)
- 波長
1周期に進む距離



1.音の基礎知識 (2)

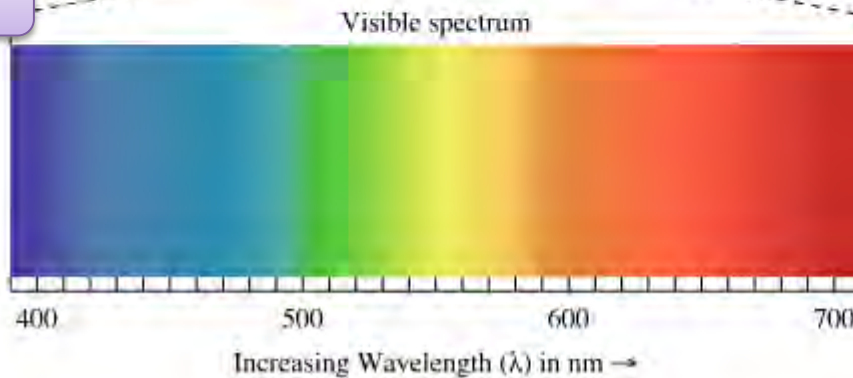
周波数が高い

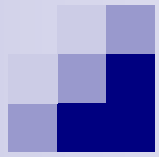
周波数が低い



波長が短い

波長が長い





モスキート音

聞くことができる年齢

★ 13KHz 

31~40歳

★ 15KHz 

25~30歳

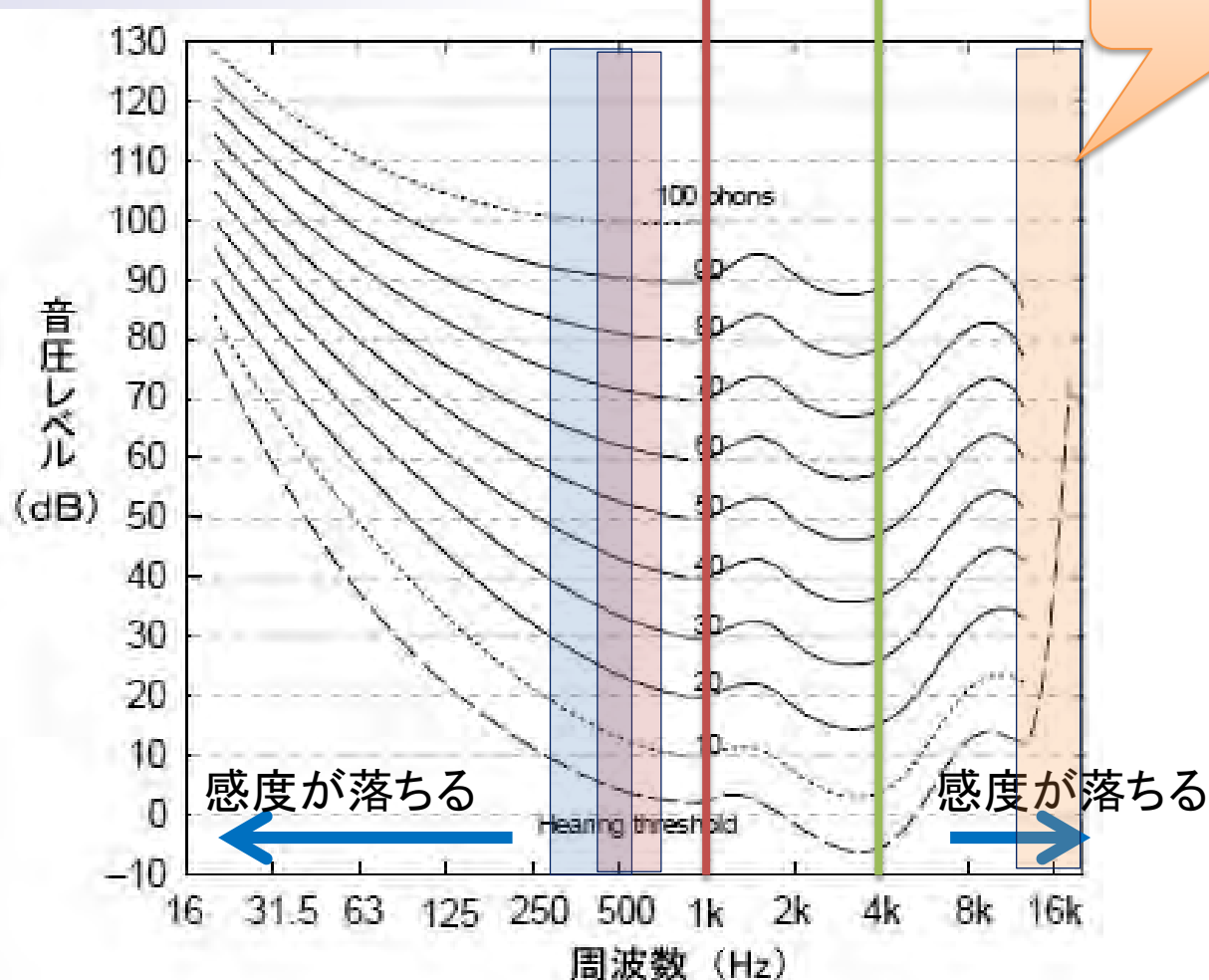
★ 17KHz 

18~24歳

★ 19KHz 

13~17歳

等ラウドネス曲線



人の声の周波数

- ・男性 ⇒ 300～550Hz
- ・女性 ⇒ 400～700Hz



1.音の基礎知識 (3)

- 音の単位 : dB(デシベル)

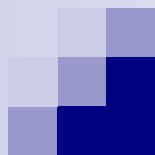
圧力変動の振幅の2乗の常用対数値を10倍したものの

例) $20\text{dB} = 10\log_{10}(10.0^2),$

$80\text{dB} = 10\log_{10}(10000.0^2)$

目的は・・・

大きな数字をわかりやすい数値で表現



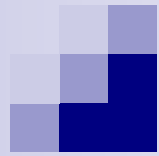
音の大きさ

項目	dB
• 皆さんの話す普段の声	60
• 図書館の管内	40
• ファミリーレストランの店内	65
• 航空機の機内	80
• バスの車内	70
• 蝉の声	75
• 静かな住宅地の昼	45
• 静かな住宅地の夜	30
• ガード下	100
• 飛行機離着陸直下	120



2.音の性質

- **快適な音, 不快な音(例:騒音)**
⇒自分の立場で感じ方が違う
(心理的問題が発生)
- **何かを知らせる音**
⇒時報, 危険
(自転車のベル, 車のクラクション等)
- **判断材料になる音**
⇒打音検査(例:トンネルが崩壊した時の検査)



3. 身近な音環境（1）

- サウンドデザイン, サウンドスケープ

音を聞いてみてください。



3. 身近な音環境 (2)

➤ マスキングの種類

- ✓ **周波数マスキング**: 周波数があまり変わらない音が連続して入ってきた場合, 後から入った音が聞こえにくい

例) 銀行の待合室のBGM

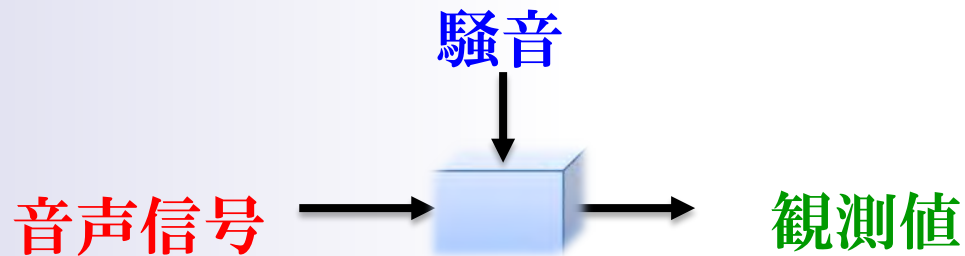
- ✓ **時間マスキング**: ある音が鳴り止んだ直後に別の音を短く(0.1秒程度)鳴らした場合, 後の音が前の音に掻き消されて聞こえなくなる現象

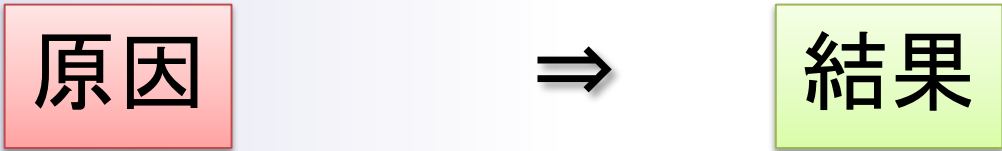
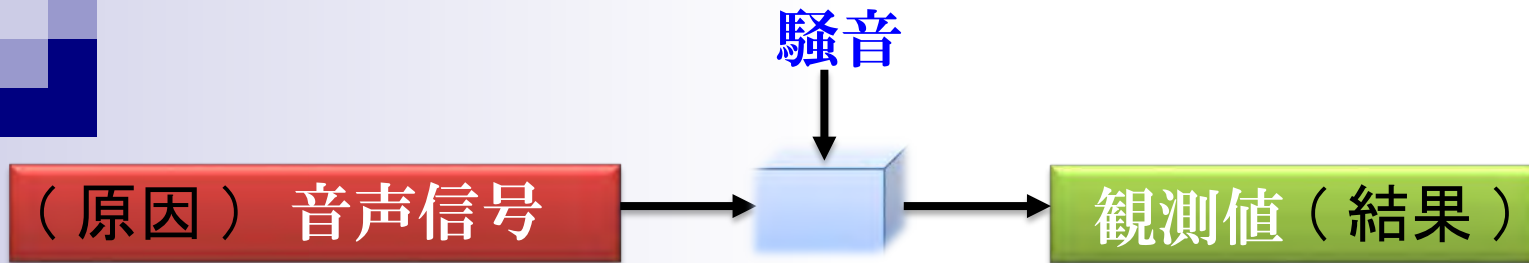
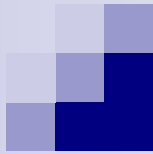
例) MP3で利用されている圧縮

4. 音声復元のための情報処理

- 耳に入ってくる声 : 観測値
- 私が話している声 : 音声信号
- エアコンの音など : 騒音

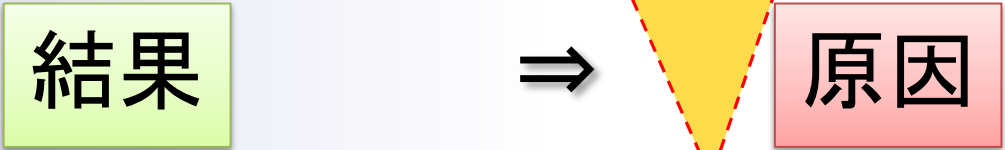
$$\text{音声信号} + \text{騒音} = \text{観測値}$$





- コンサートをweb予約 ⇒ チケットを購入

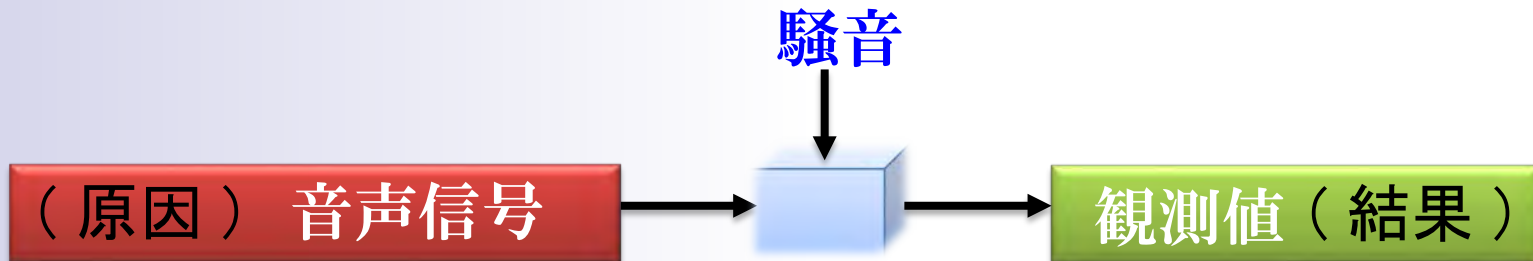
ベイズ定理を使って
出すことができる



- チケットを購入 ⇒ webで予約
電話予約
代理店で予約 など

4-2. ベイズ定理

結果から原因を推定できる



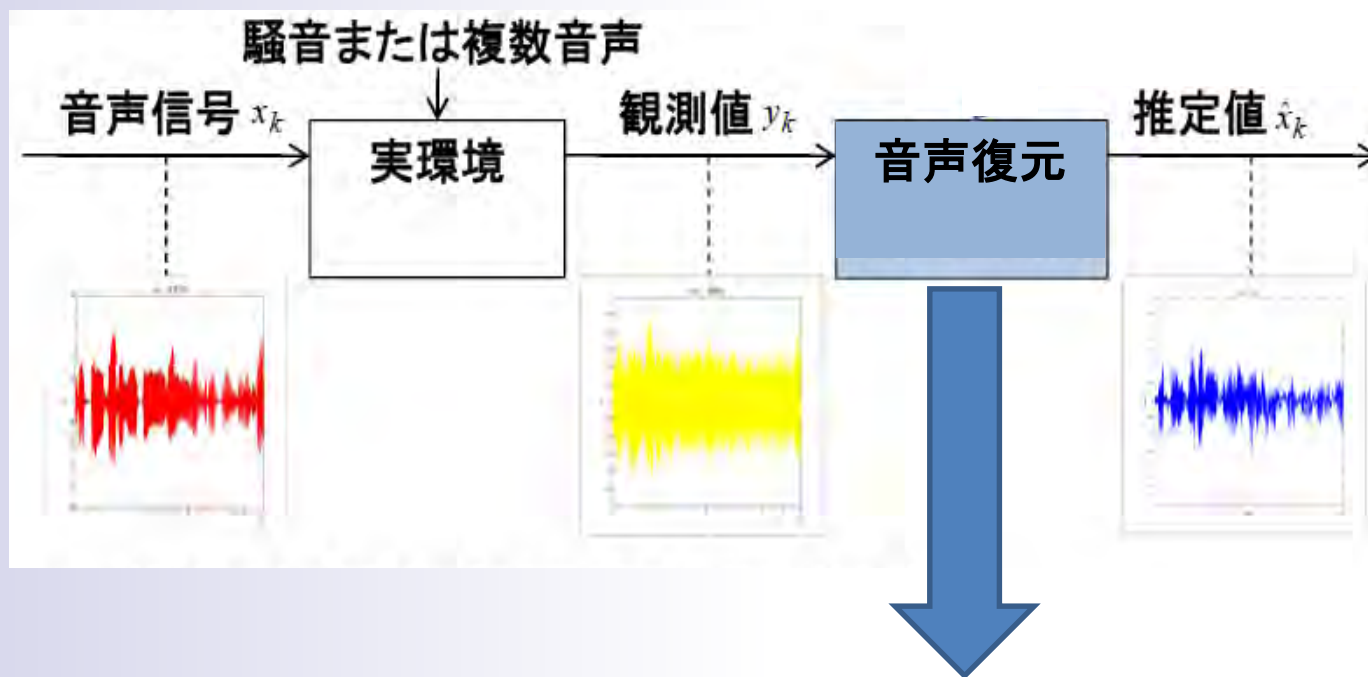
$$P(\text{原因} \mid \text{結果}) = \frac{P(\text{結果} \mid \text{原因}) \times P(\text{原因})}{P(\text{結果})}$$

$P(\text{原因} \mid \text{結果})$; 観測 (結果) に基づく事象 (原因) の確率

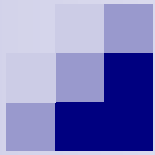
$P(\text{原因})$; 事象 (原因) が発生する確率

4-3. 音声復元のための情報処理

音声信号に対する新たなデジタルフィルタの開発



- カルマンフィルタ
- 音声復元手法(本日紹介する手法)



音声復元手法

- x_k : k 時刻における**音声信号**
- y_k : **騒音**が混入した**観測値**

(k : 離散時刻)

時間変動未知



β_k

未知統計量

未知システム特性

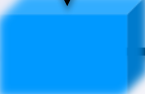


α_k

騒音 v_k



音声信号 x_k

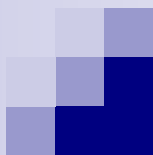


y_k 観測値



推定

フィルタを考える研究
(濾過フィルタみたいなもの)



音声信号 x_k
観測値 y_k から 音声信号と観測値の関係 α_k を推定
時間による音声変化 β_k

ベイズ定理:

$$P(x_k, \alpha_k, \beta_k | Y_k) = \frac{P(x_k, \alpha_k, \beta_k, y_k | Y_{k-1})}{P(y_k | Y_{k-1})}$$
$$= \frac{P_0(x_k | Y_{k-1}) P_0(\alpha_k | Y_{k-1}) P_0(\beta_{k-1} | Y_{k-1}) \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{r=0}^{\infty} B_{lmnr} \varphi_l^{(1)}(x_k) \varphi_m^{(2)}(\alpha_k) \varphi_n^{(3)}(\beta_k) \varphi_r^{(4)}(y_k)}{\sum_{r=0}^{\infty} B_{000r} \varphi_r^{(4)}(y_k)}$$

ただし、 $B_{lmnr} = \langle \varphi_l^{(1)}(x_k) \varphi_m^{(2)}(\alpha_k) \varphi_n^{(3)}(\beta_k) \varphi_r^{(4)}(y_k) | Y_{k-1} \rangle$, $Y_k = \{y_1, y_2, \dots, y_k\}$ (K時刻までの観測値の集合)

5. 音声復元の実験

	概要
音声信号	女声信号, 男声信号を利用
騒音	白色雑音, 有色雑音, 機械音を採用
比較対象	拡張カルマンフィルタ
評価	推定誤差と性能評価



推定誤差

- 値が小さいほど優れている

$$\text{推定誤差} = \sqrt{\frac{\sum (x_k - \hat{x}_k)^2}{N}}$$

(x_k :元の音声信号 \hat{x}_k :推定値 N :データ数)

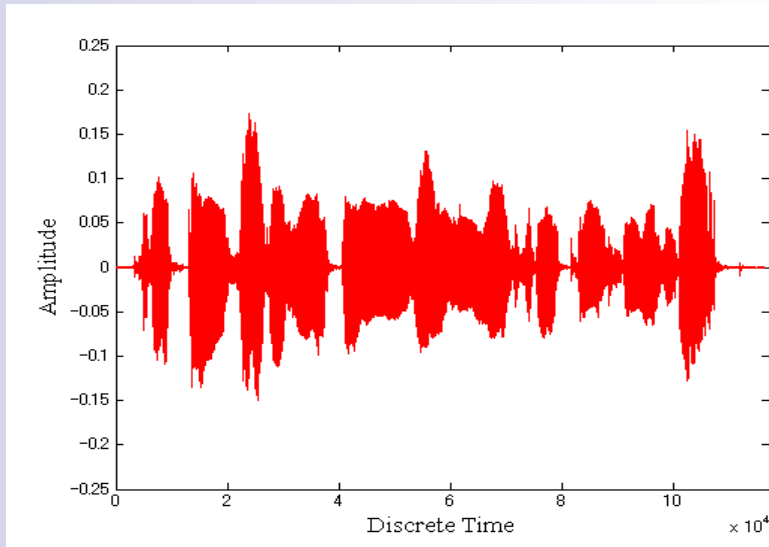


性能評価

- 値が大きいくほど優れている $\left(= 10 \log_{10} \frac{\sum x_k^2}{\sum (x_k - \hat{x}_k)^2} [dB] \right)$

5-1-a. 女声信号に適用

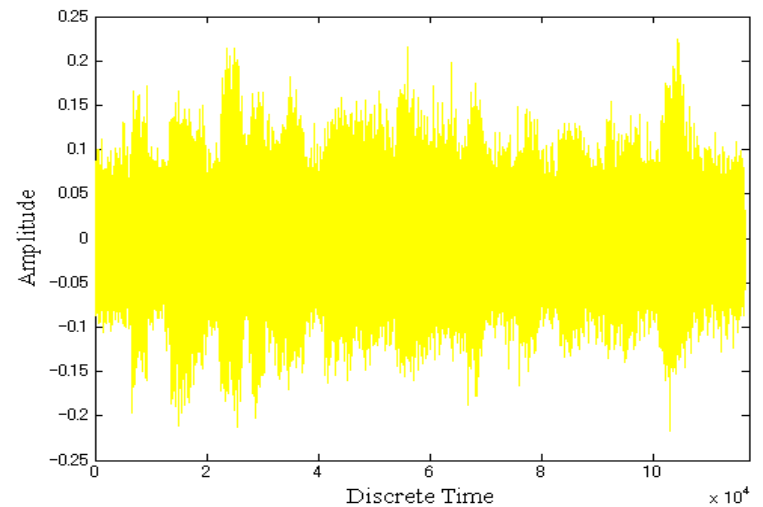
🌸 音声信号



🌸 リスニング 🌸

🌸 騒音 (白色雑音)

混入下での観測値

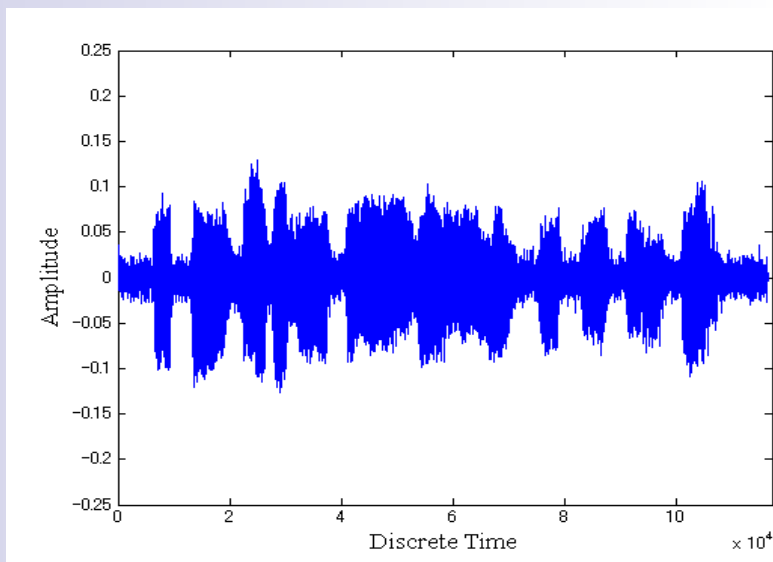


🌸 リスニング 🌸

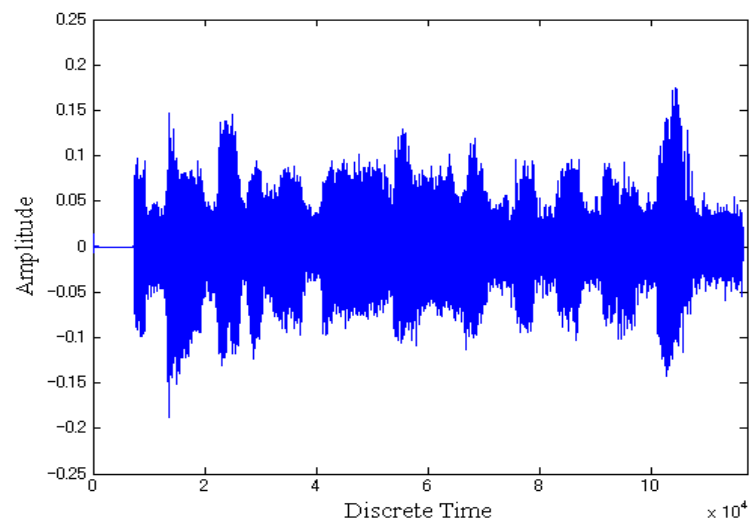
5-1-b. 女声信号に対する推定結果

■ 元の信号
■ 観測値
■ 推定値

🌸 本手法



🌸 拡張カルマンフィルタ



🌸 リスニング 🌸

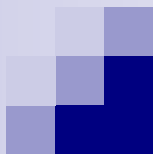
🌸 リスニング 🌸

推定誤差	0.02504	0.03028
性能評価	2.263	0.6107



6. まとめ 音声復元の応用例

- **音は快適な音にも不快な音にもなりうる性質**
 - － 生活の場において、誰でも加害者にも被害者にもなる
 - － 後に残らない性質を利用して、様々な部分に応用できる可能性を持ち合わせている
- **音声認識システム**
 - － 工場などにおける施設・設備巡回点検修理業務
 - － 工事現場の記録・状況報告などハンズフリー業務
 - － 音声入力によるカーナビや車内設置電話への応用



ご清聴ありがとうございました。

