

Web会議システムにおける エコーキャンセリング技術

2020.9.9
金沢工業大学
島内 末廣

もくじ

1. エコーキャンセリングとは？
2. エコーキャンセリングの動作原理
3. エコーキャンセリングの理想と現実
4. まとめ

1. エコークャンセリングとは？

そもそも、エコーとは？

Web会議などで発言した自分の声が、
こだまのように遅れてスピーカから聴こえる現象

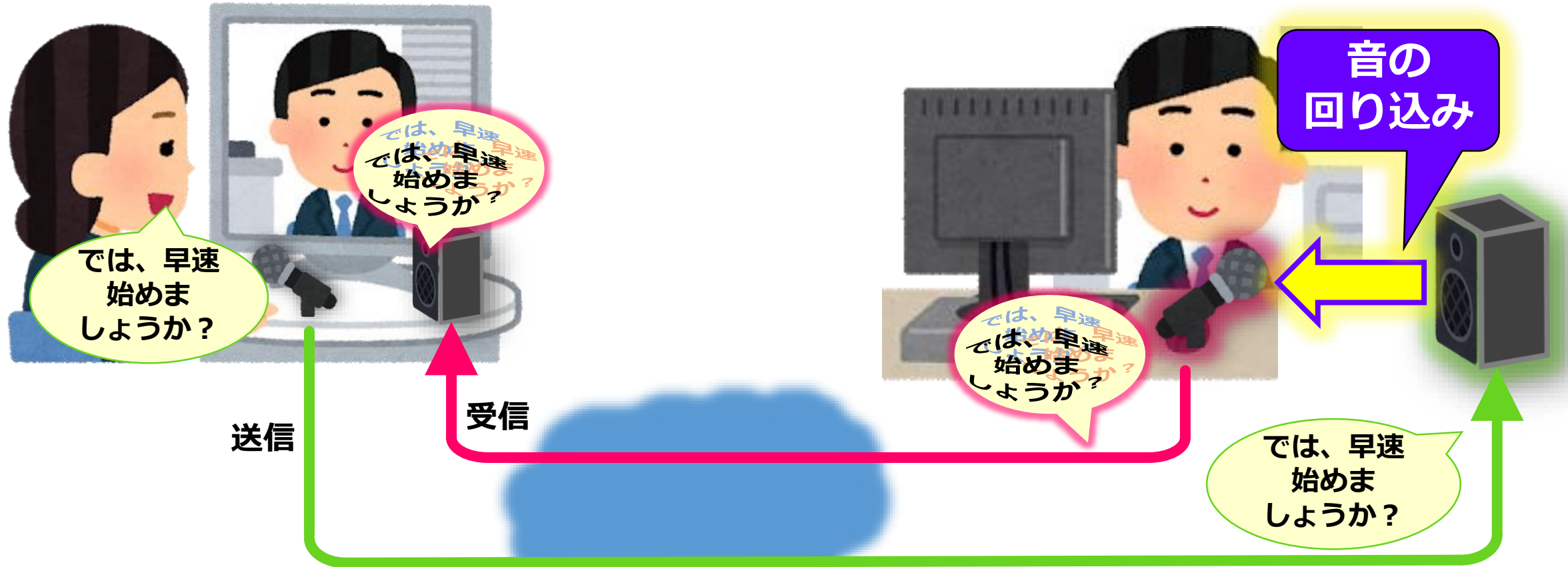


エコーが発生すると・・・

- 参加者にとって
音声が反響したり、ハウリングしたりと
非常に聴き取りにくい
- 発話者にとっても
発言中に自分の声が遅れて聞こえてくると
非常に話しにくい

エコーはなぜ生じる？

相手地点のスピーカ・マイク間の音の回り込み = "音響エコー"
相互の地点の回り込みのループによる発振現象 = "ハウリング"



エコーの発生を回避するには？

- 発話者以外のマイクをミュート(OFF)にする
大規模会議、遠隔授業、講演会など
- ヘッドホンやイヤホンで音声を聞く
Web会議に参加しない人が周囲にいる場合など
- エコーキャンセリング技術を用いる
1 地点に複数人が参加する会議、
双方向のかけ合いが多い会議、長時間の会議など

エコーキャンセリング技術

- 不要なエコーを除去し、通話相手の音声のみを受聴できるようにする技術
- 巷のノイズキャンセリング技術との違いは？
(どちらも不要な音を除去するのが目的)

不要な音を除去する技術の例

ノイズキャンセリング付
ヘッドホン



不要な音 = 音波

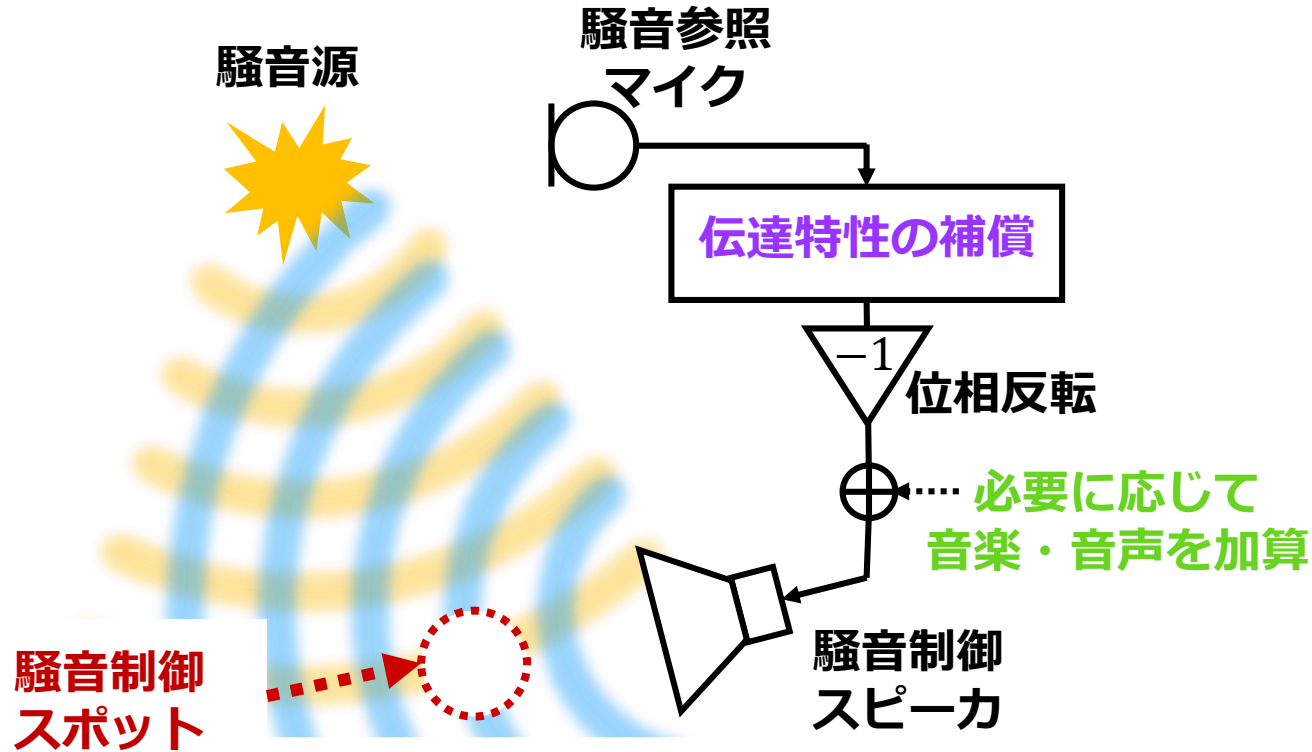
ノイズキャンセリング付
ボイスレコーダ



不要な音 = 信号

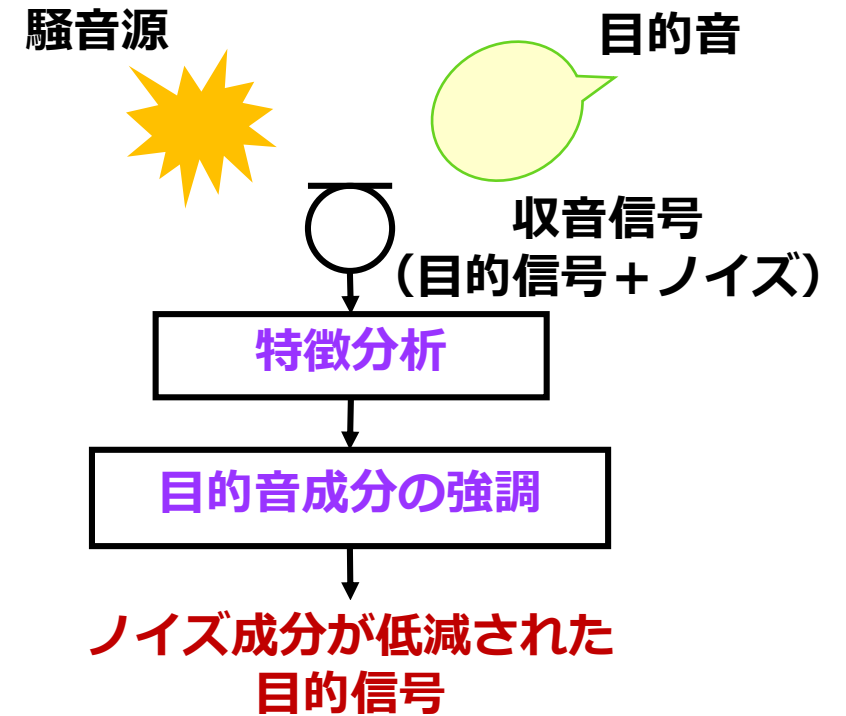
不要な音を除去する技術の例

音を音で消す技術 (音波の伝達特性を考慮)



不要な音 = 音波

信号ノイズを除去する技術 (音源の性質の違いを考慮)



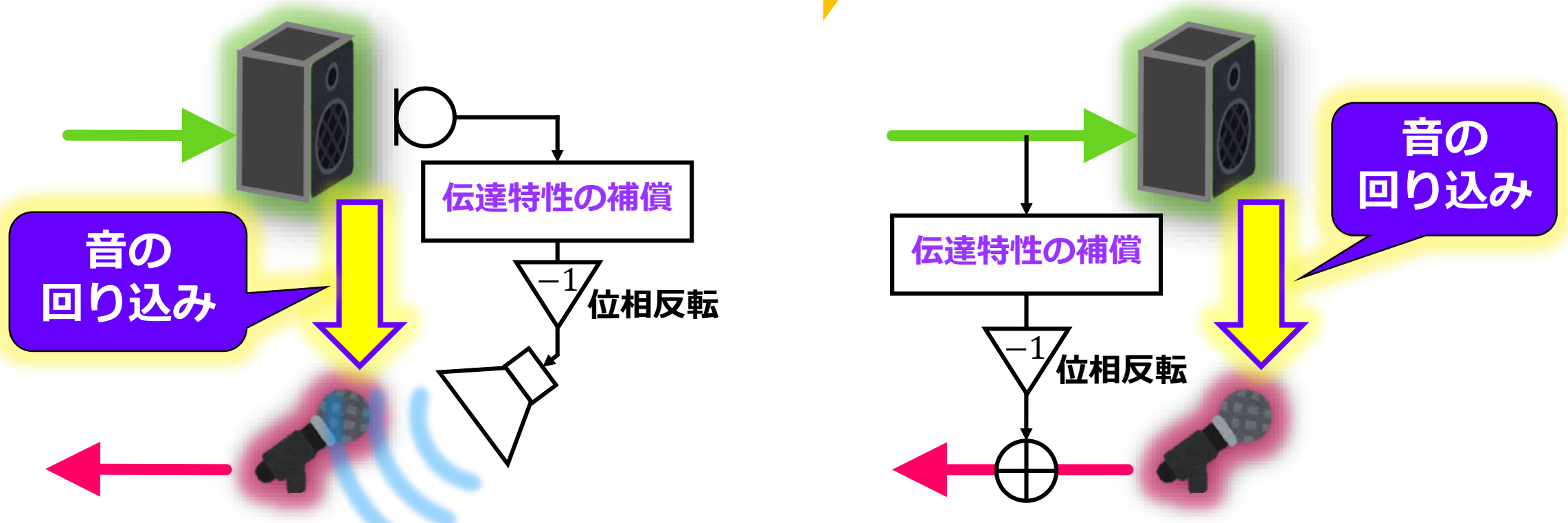
不要な音 = 信号

エコーを除去するには？

音で音を消す技術が
適用できるかも

けれども

わざわざ
音として消さなくても・・・

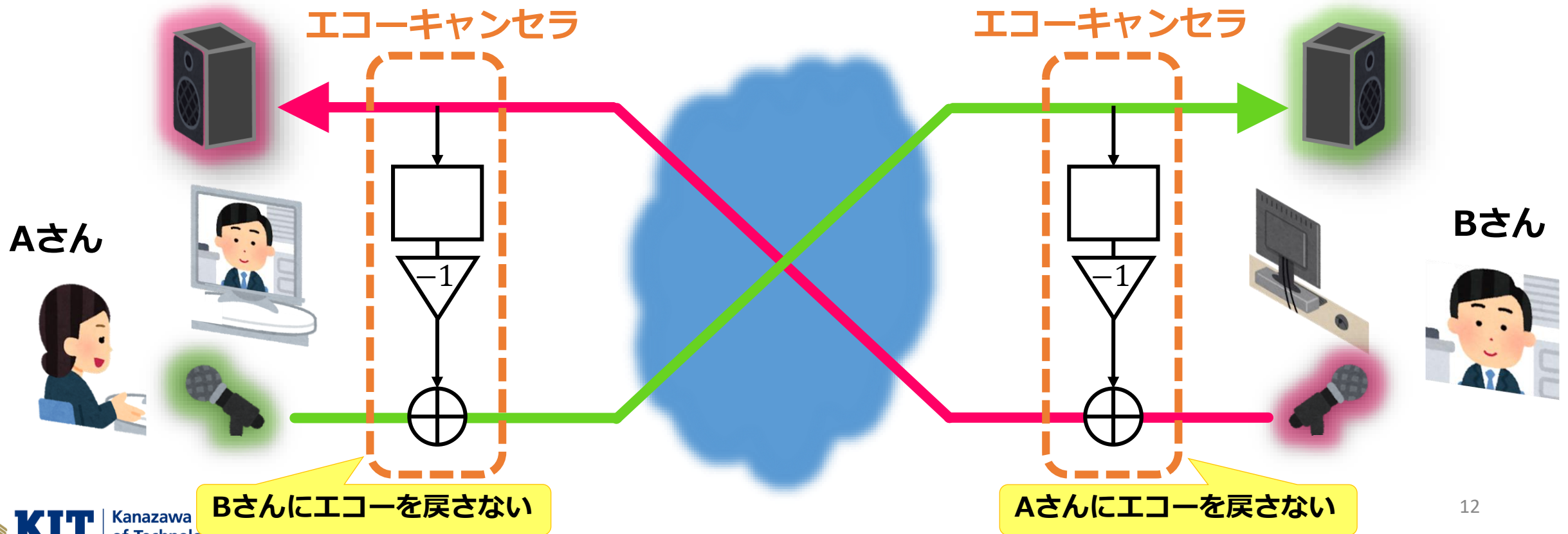


エコーキャンセリング技術：
音波の伝達特性を考慮して、信号から不要音(エコー)を除去

Web会議などに適用する場合

各地点のスピーカとマイクの間（通話相手のために）実装

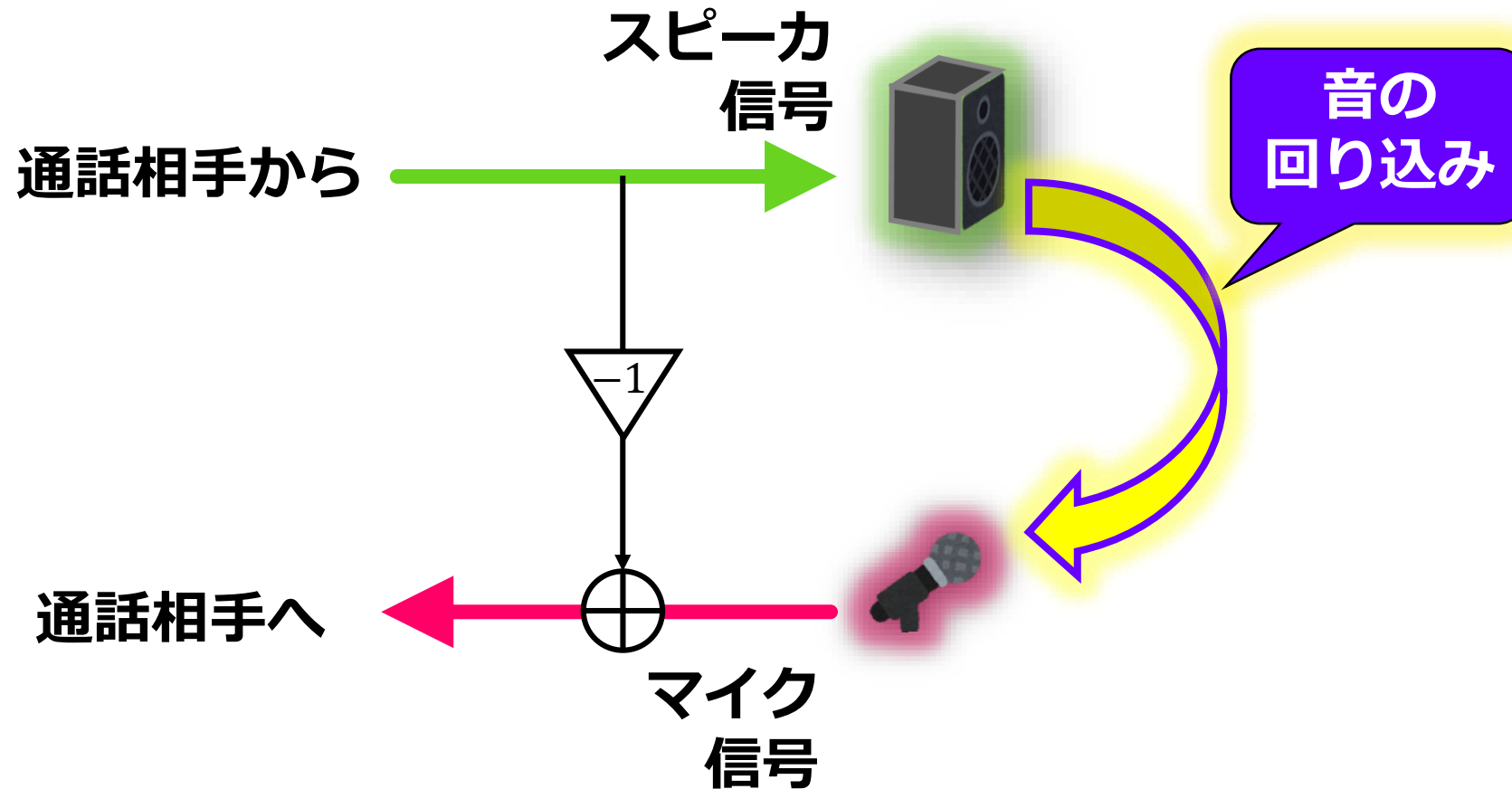
- Web会議アプリと連携したソフトウェアとして
- 専用のハードウェア端末として



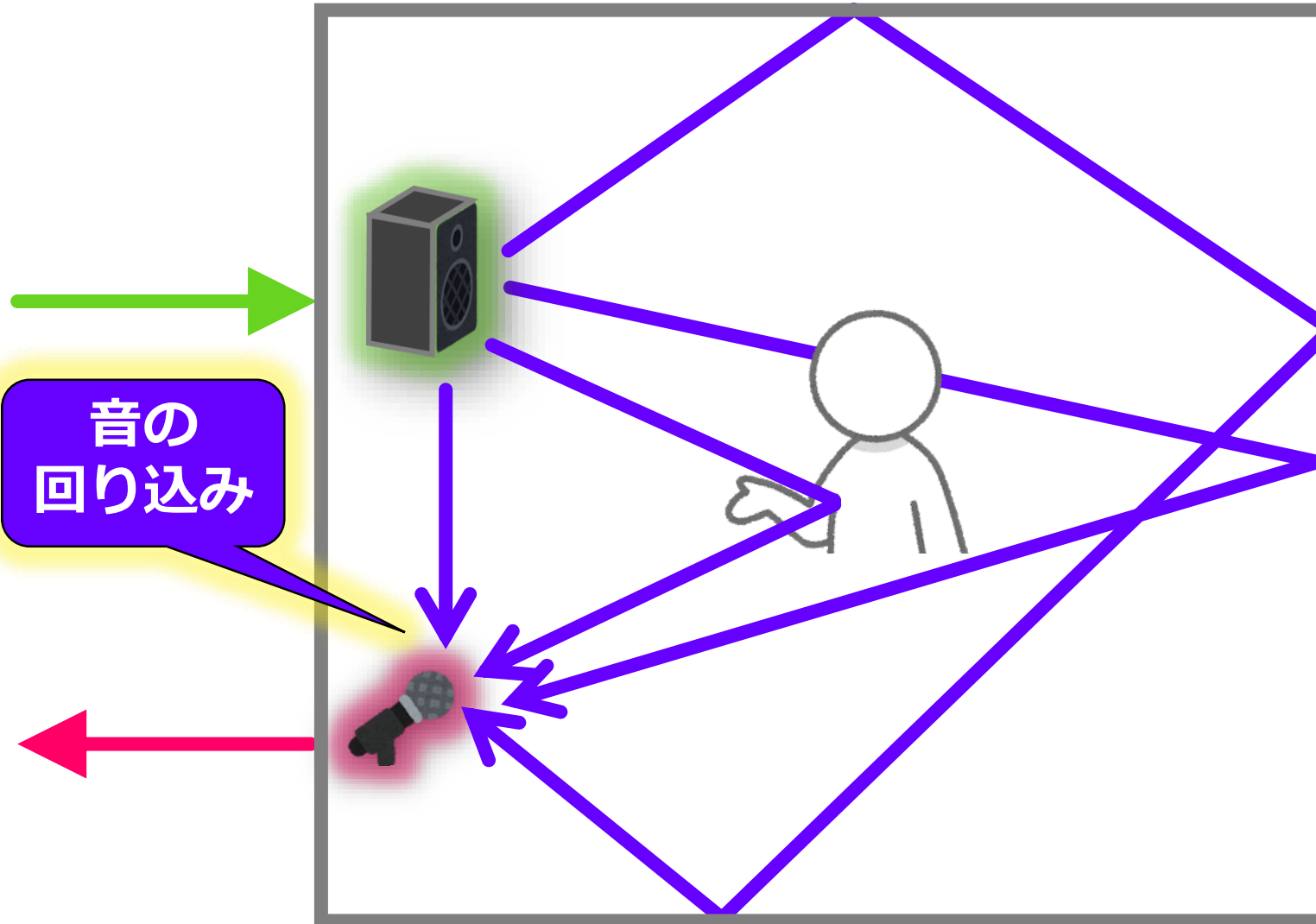
2. エコーキャンセリングの動作原理

これで解決すればよいですが . . .

スピーカ信号をマイク信号から差し引くだけでは
ほとんどエコーは消えません



複雑な反射を経て音は回り込む

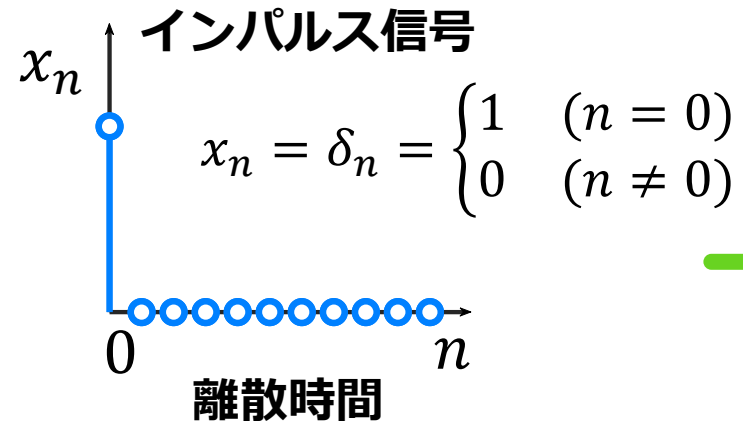


室内の壁・天井・床、
人やモノからの反射の影響を
考慮する必要がある

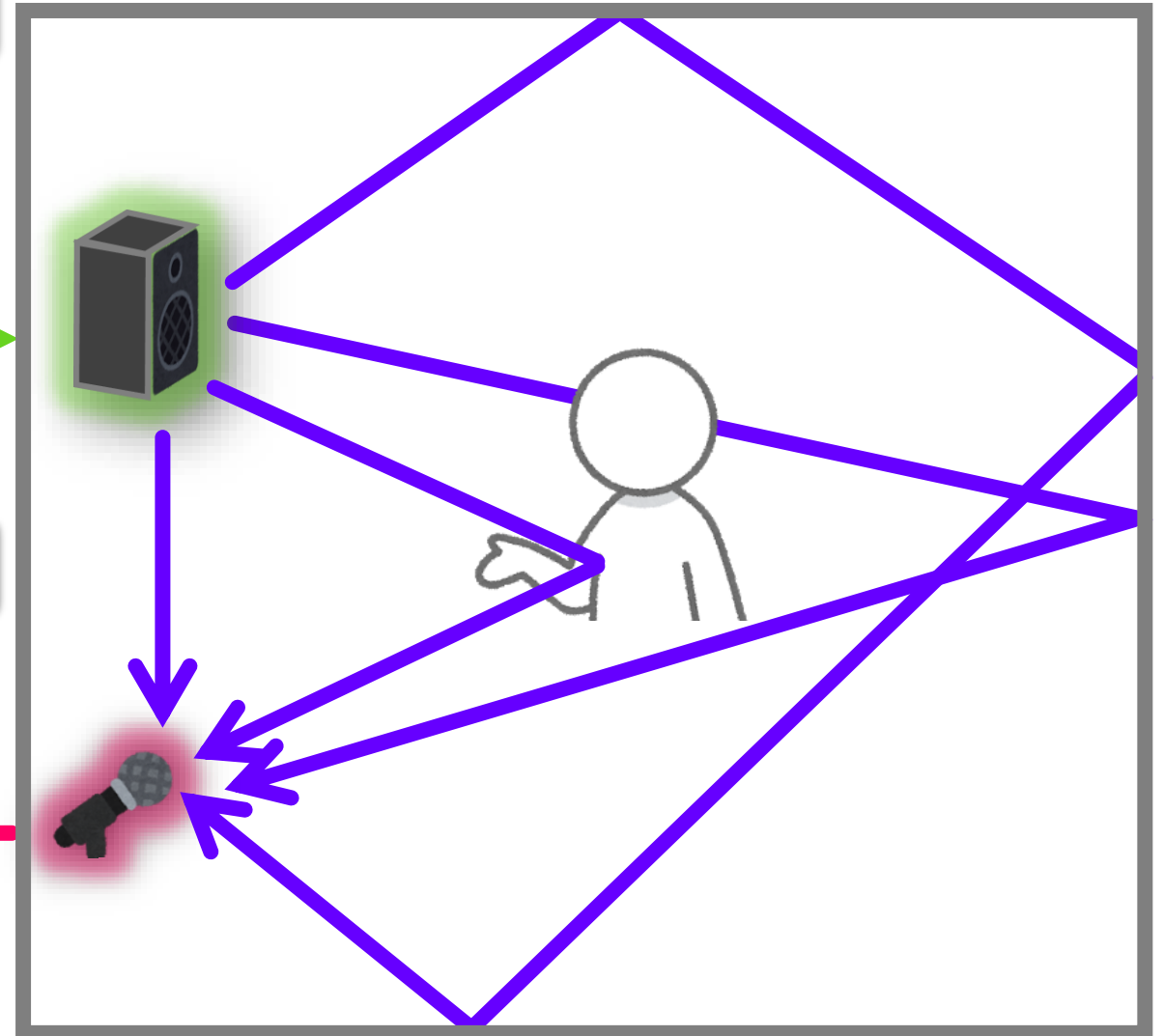
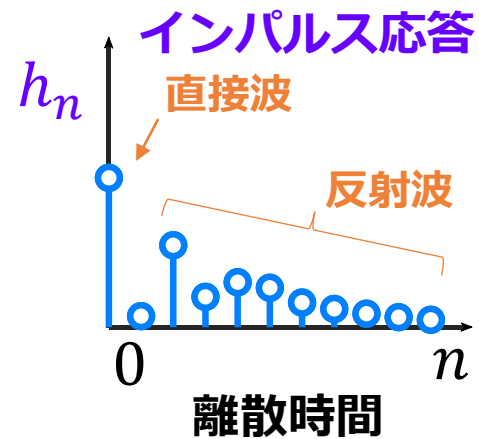
※厳密には
スピーカ、マイク、アンプ、
PCのオーディオ入出力
の特性なども影響しますが
簡単のため無視します

反射の様子 ⇒ インパルス応答

スピーカからインパルス信号を再生

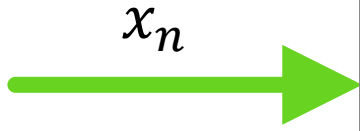
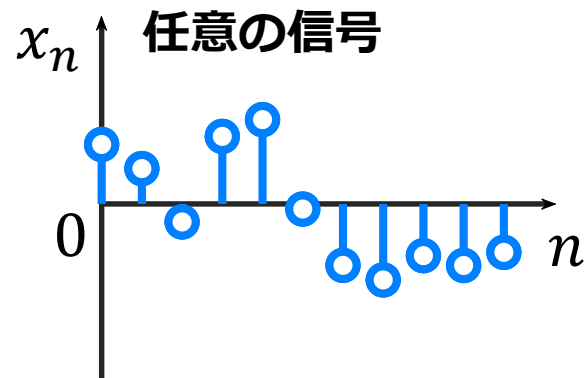


反射波の様子が時系列で観測できる

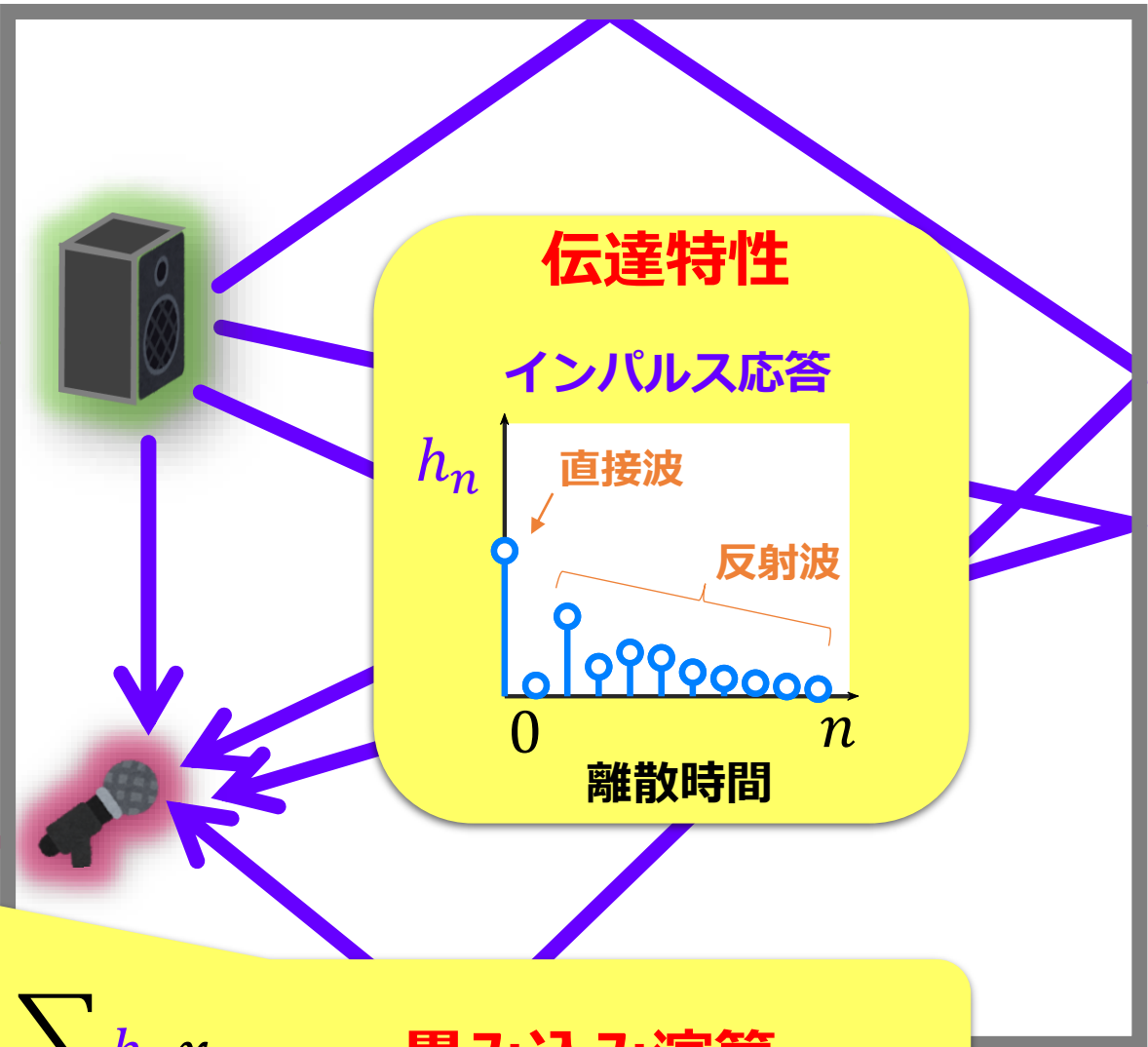
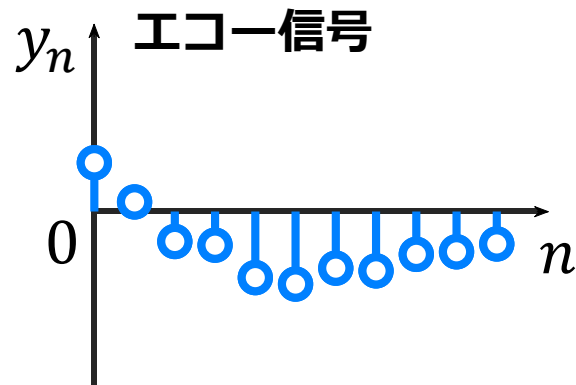


インパルス応答 ⇒ 任意のエコー信号

スピーカから任意の信号を再生



エコー信号 = インパルス応答との畳み込み



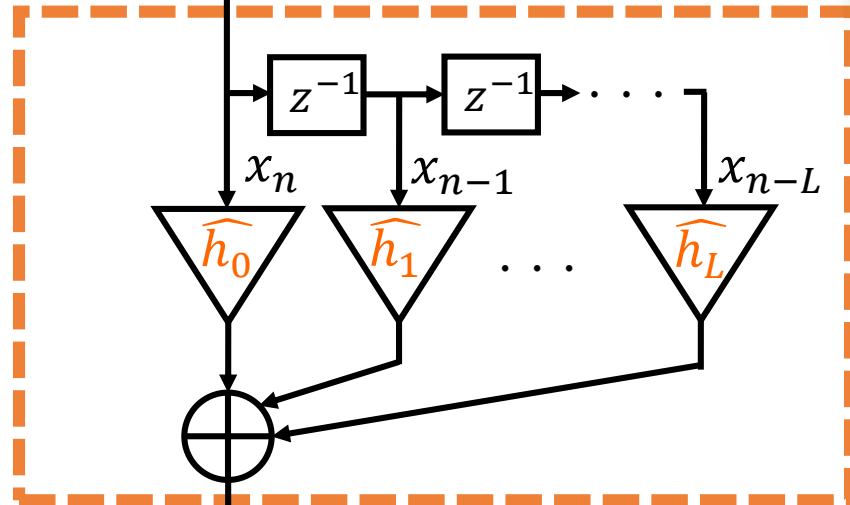
$$y_n = \sum_k h_k x_{n-k}$$

畳み込み演算

ディジタルフィルタでエコー信号を模擬

任意のスピーカ信号 x_n

ディジタルフィルタ
インパルス応答が
そのまま係数(FIR型)



エコー模擬信号

$$\hat{y}_n = \sum_k \hat{h}_k x_{n-k}$$

残差

$$e_n = y_n - \hat{y}_n$$

\hat{h}_k と h_k が一致するほど、
残差は小さくなる (エコーが消える)



音の
回り込み



マイク信号 (エコー信号)

$$y_n = \sum_k h_k x_{n-k}$$

インパルス応答を手に入れたい

インパルス応答は

- 個々のユーザ環境によって異なる
- 人の動きなどに応じて時々刻々と変化する

このため

- 会議中、リアルタイムでの観測が必要
- 会議を中断しないよう、特殊な観測信号を用いずに通話音声の情報に基づき観測（推定）したい

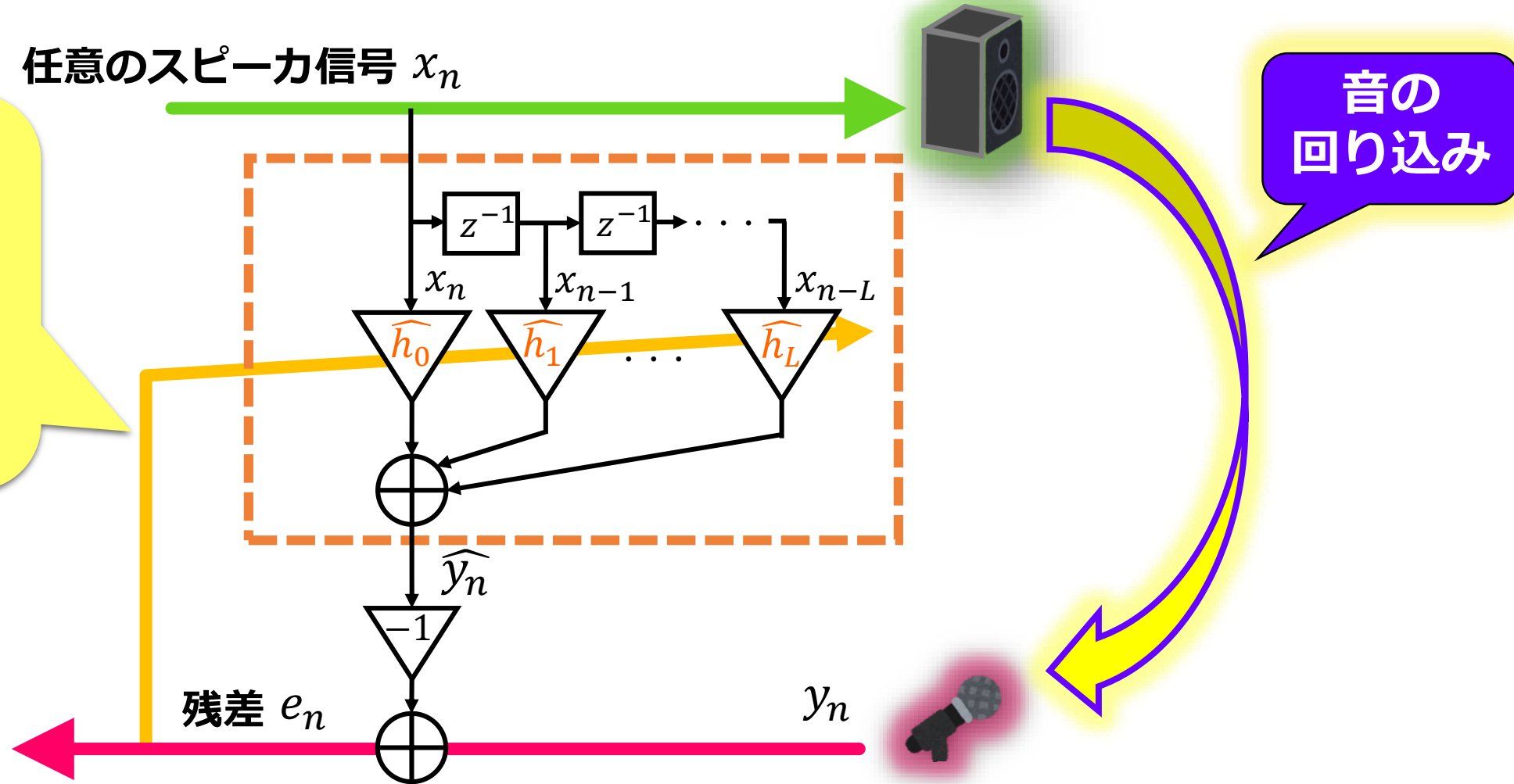
適応デジタルフィルタ

任意のスピーカ信号 x_n

残差 e_n を評価して
フィルタ係数が

$$\widehat{h}_k \rightarrow h_k$$

となるように
オンライン学習
させる



適応学習アルゴリズムの簡単な例

インパルス応答長が2サンプルの場合

実際のエコー信号

$$y_n = \sum_{k=0}^1 h_k x_{n-k} = h_0 x_n + h_1 x_{n-1} = [h_0, h_1] \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}$$

エコー模擬信号

$$\widehat{y}_n = \sum_{k=0}^1 \widehat{h}_k x_{n-k} = \widehat{h}_0 x_n + \widehat{h}_1 x_{n-1} = [\widehat{h}_0, \widehat{h}_1] \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}$$

残差

$$e_n = y_n - \widehat{y}_n = [h_0 - \widehat{h}_0, h_1 - \widehat{h}_1] \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}$$

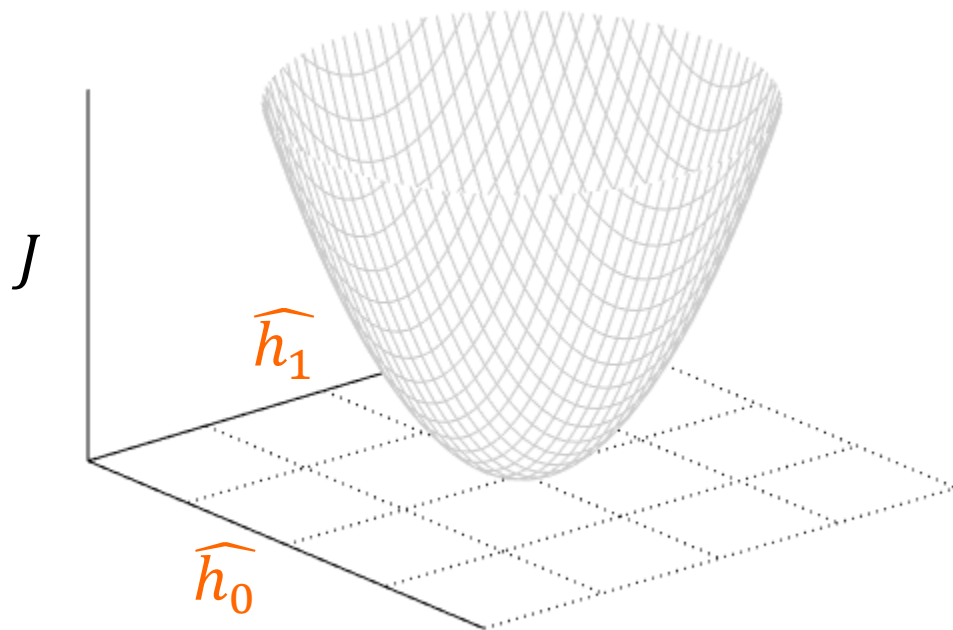
残差の2乗を評価する

残差の2乗を損失関数として評価する

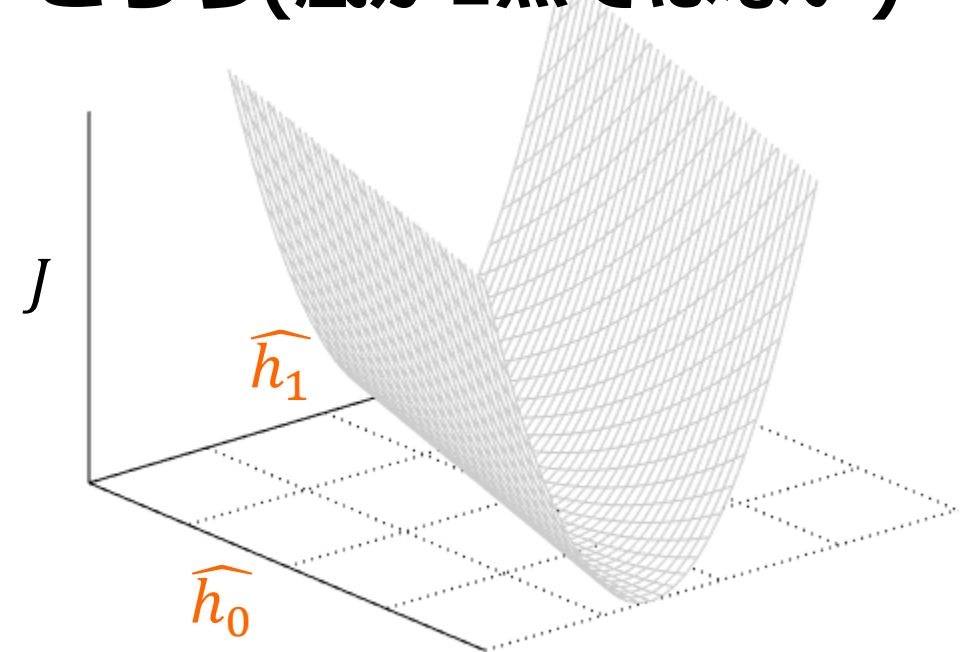
$$J = \frac{1}{2} e_n^2 = \frac{1}{2} \left([h_0 - \widehat{h}_0, h_1 - \widehat{h}_1] \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} \right)^2$$

この損失関数をグラフ化すると

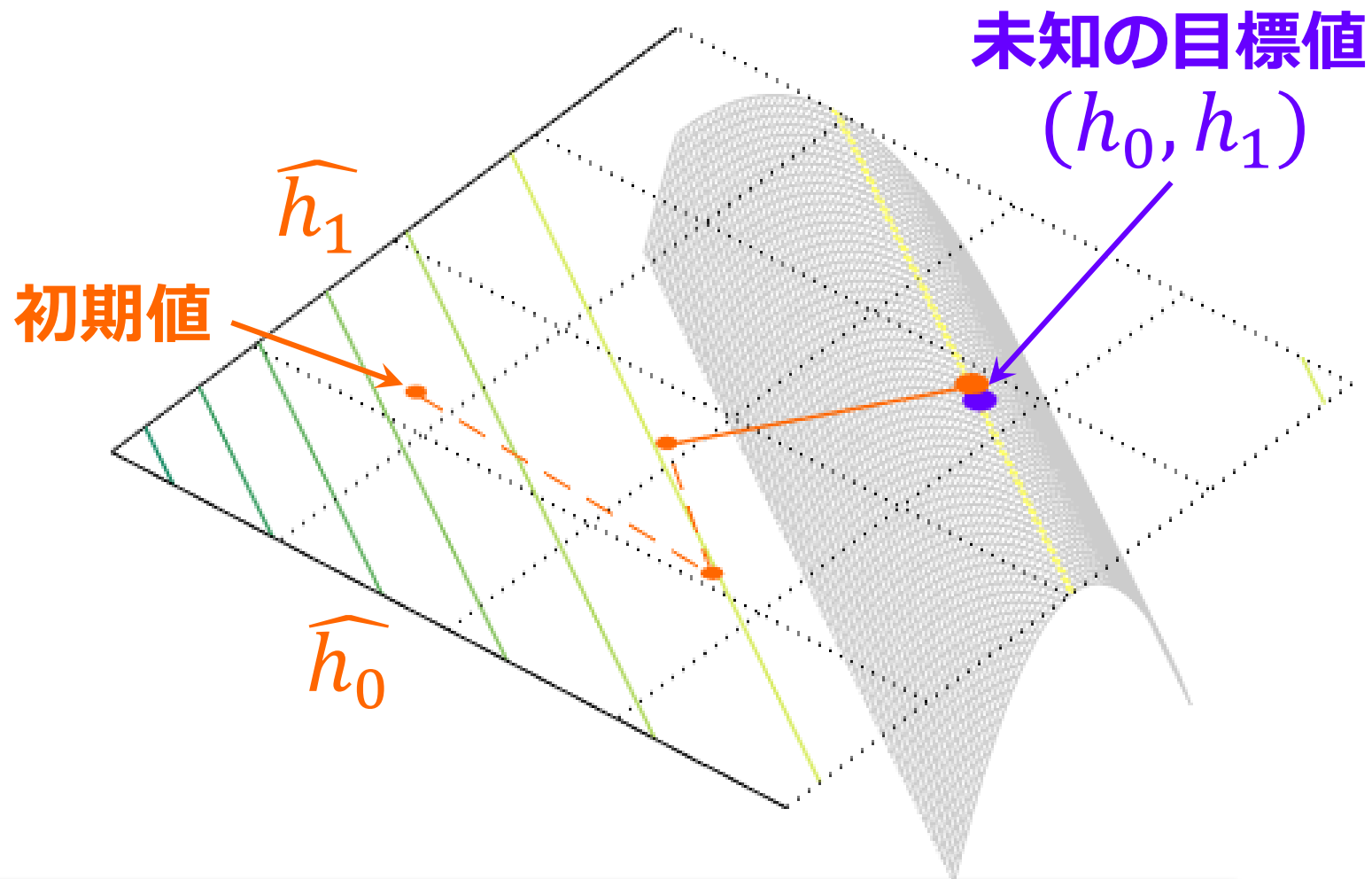
こちらではなく



こちら(底が1点ではない)



フィルタ係数の更新のイメージ



※上下反転表示

繰り返しフィルタ係数を更新することで
しだいに目標値へ収束していく

フィルタ係数の更新式

損失関数 $J = \frac{1}{2} e_n^2 = \frac{1}{2} \left([h_0 - \widehat{h}_0, h_1 - \widehat{h}_1] \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} \right)^2$

の勾配ベクトル $\begin{bmatrix} \frac{\partial J}{\partial \widehat{h}_0} \\ \frac{\partial J}{\partial \widehat{h}_1} \end{bmatrix} = -e_n \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}$ から更新の方向(逆向き)がまず分かる。

さらに、勾配ベクトルの大きさをスピーカ信号のパワで正規化すれば、最短距離でピンポイントに残差曲面の極小値へ

$\begin{bmatrix} \widehat{h}_0 \\ \widehat{h}_1 \\ \vdots \\ \widehat{h}_L \end{bmatrix} \xleftarrow{\text{更新}} \begin{bmatrix} \widehat{h}_0 \\ \widehat{h}_1 \\ \vdots \\ \widehat{h}_L \end{bmatrix} + \mu \frac{e_n}{\sum_{k=0}^L x_{n-k}^2 + \lambda} \begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-L} \end{bmatrix}$ NLMSアルゴリズム (学習同定法)

μ, λ : 学習制御パラメータ

3. エコークャンセリングの理想と現実

エコーキャンセリング寄りの立場

マイクに入るのはエコーだけ

- ・フィルタの学習も簡単
- ・そもそもエコーキャンセル不要？

Web会議などを利用する側の立場

マイクは発話するためにある

- ・エコーの他、発話音声(残したい音)
- 背景雑音(消したい音)が混合

音源分離やノイズキャンセリングとの融合

音の回り込みの特性

- ・線形(インパルス応答で表現可能)
- ・時不変(変動してもゆっくり)

会議の環境や使われ方は様々

- ・大音量で歪みが発生
- ・動き回りながら通話

伝達モデルの拡張

Web会議は1部屋で1組

- ・通話相手の部屋とは音響的に独立
- ・1組のスピーカ・マイク系を考慮

自席から個々人参加のパーソナル会議

- ・自分の声が隣のマイクへ
- ・隣のスピーカから自分のマイクへ

会議系全体での連携

4. まとめ

エコーキャンセリング技術

- **スピーカ・マイク間の音の伝達特性に基づき
マイク信号からエコーを除去**
- **時々刻々と変化する音の伝達特性を、
通話音声を用いてリアルタイムで絶えず学習**
- **エコーと上手に付き合いながら快適な通信会議を！**

ご清聴ありがとうございました