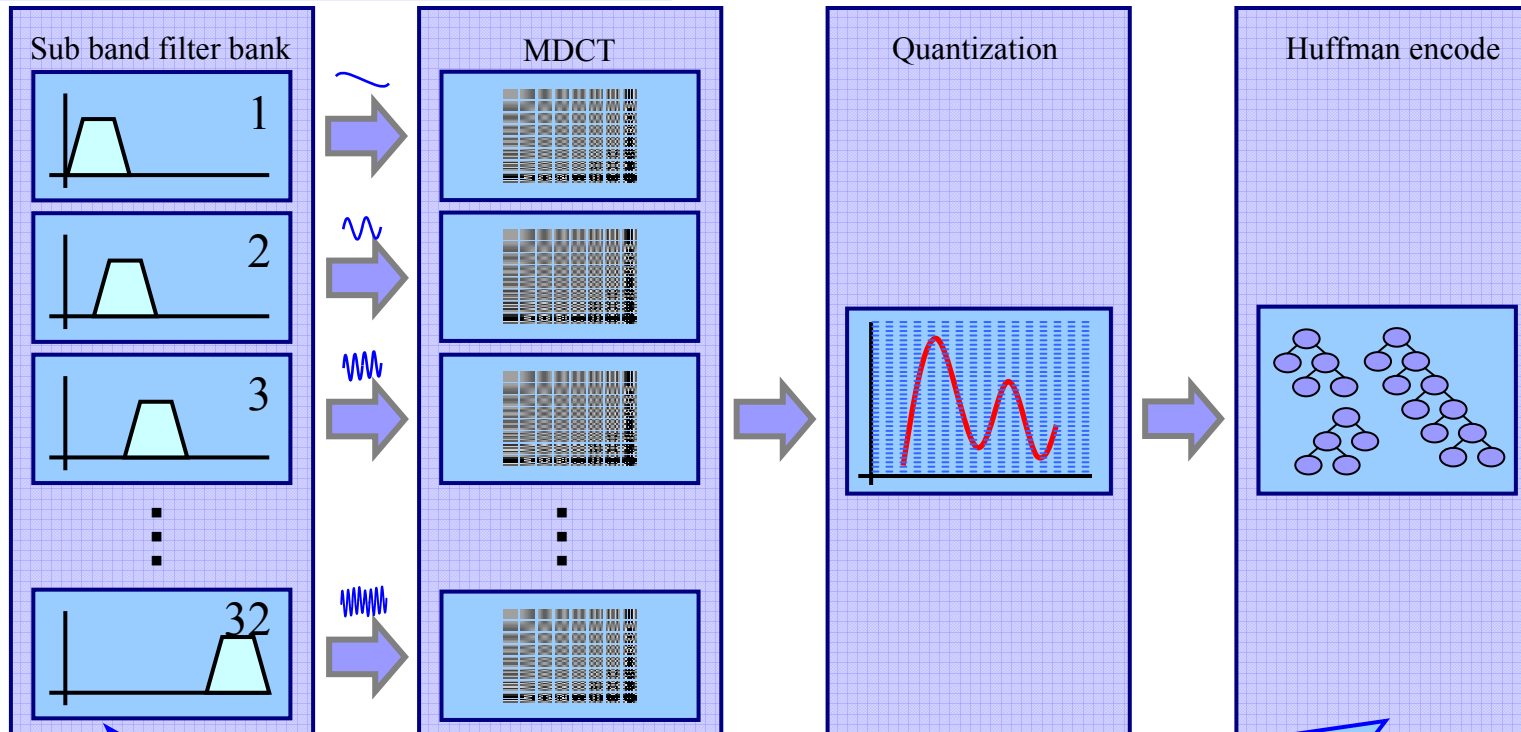


音声に対する 半開示スクランブル法

北海道大学大学院 情報科学研究科 メディアネットワーク専攻 メディア創生学研究室

--MP3符号化の概要--



- ▶ 一般にオーディオ信号は、量子化された値の出現頻度に偏りがある
- ▶ 量子化値が0に近い値の出現頻度が高く、絶対値の大きな量子化値の出現頻度は低い

- ▶ 1単位時間あたり32サブバンドの32サンプルを出力
- ▶ この処理を18回行い、各サブバンド毎に時系列に並べグラニューールを構成

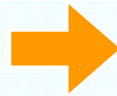
提案手法

音質の定量的な制御を可能とするオーディオ半開示スクランブル手法

- ▶ 一般に広く普及しているMP3 (MPEG-1/2 Audio Layer-3) が処理対象
- ▶ MP3符号化の過程でスクランブル処理を行う

目標のSNRを達成するための二乗誤差を下記の式で算出

1



$$\sum \{x(i) - y(i)\}^2 = \frac{\sum \{x(i)\}^2}{10^{\text{SNR} / 10}}$$

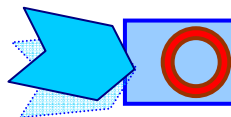
2

量子化後のMDCT係数を変更することで二乗誤差を発生させる

3

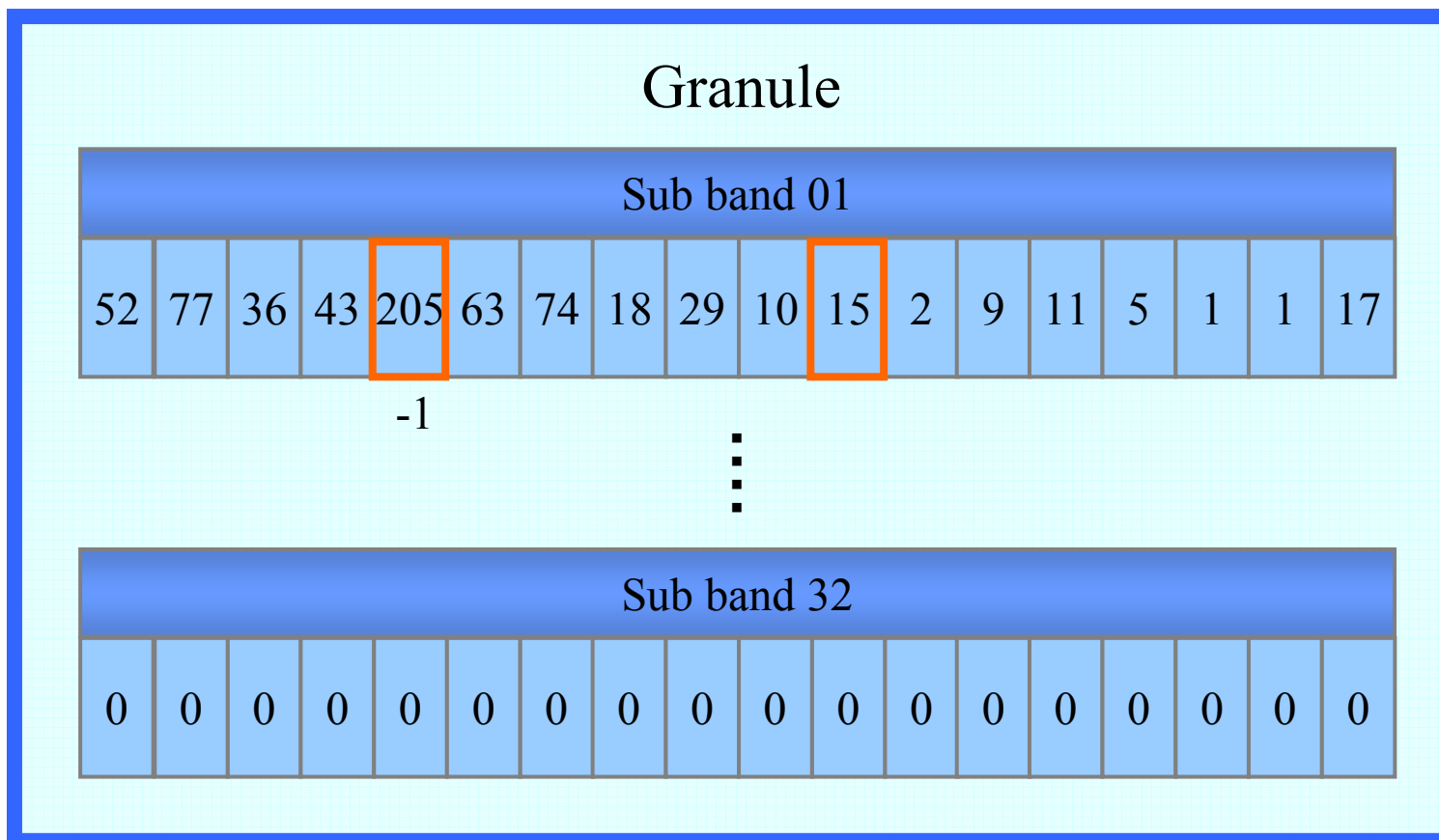
目標の二乗誤差を達成するまで繰り返す

目標のSNRを達成



定量的な音質制御が可能

提案手法 --MDCT係数の変更(増減)--



- ▶ ① 擬似乱数を用いて変更する係数を選択する(0は除外)
- ▶ ② 絶対値が小さくなる方向に係数を1だけ増減させる
- ▶ ③ 必要な二乗誤差を満足しなければ処理を繰り返す

提案手法 --MDCT係数の変更(入替)--

Granule

Sub band 01																	
52	77	36	43	205	63	74	18	29	10	15	2	9	11	5	1	1	17

⋮

Sub band 32																	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- ▶ ① **擬似乱数**を用いて変更する係数を選択する
- ▶ ② **一定の範囲内**で係数を交換する
- ▶ ③ **必要な二乗誤差**を満足しなければ処理を繰り返す

実験

実験内容

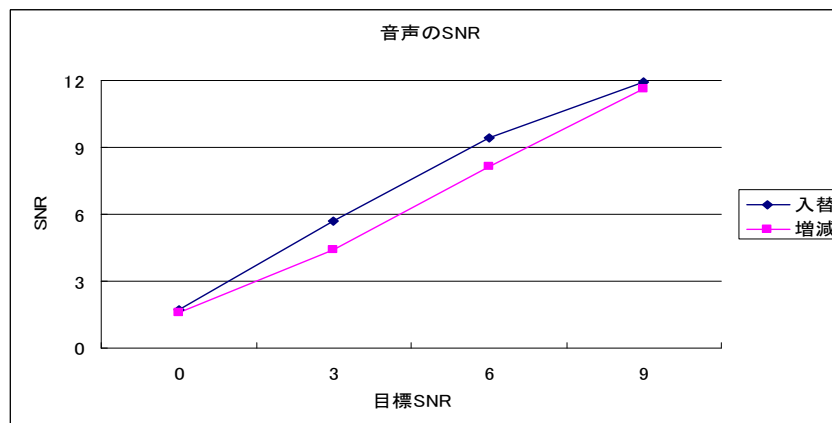
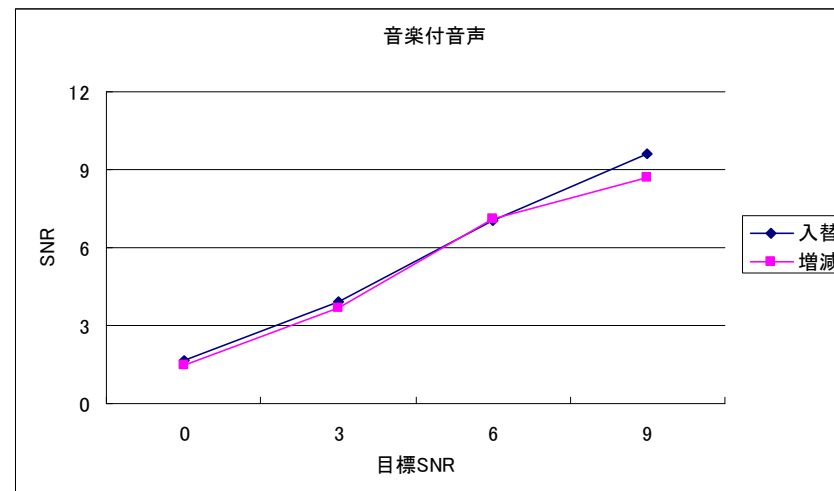
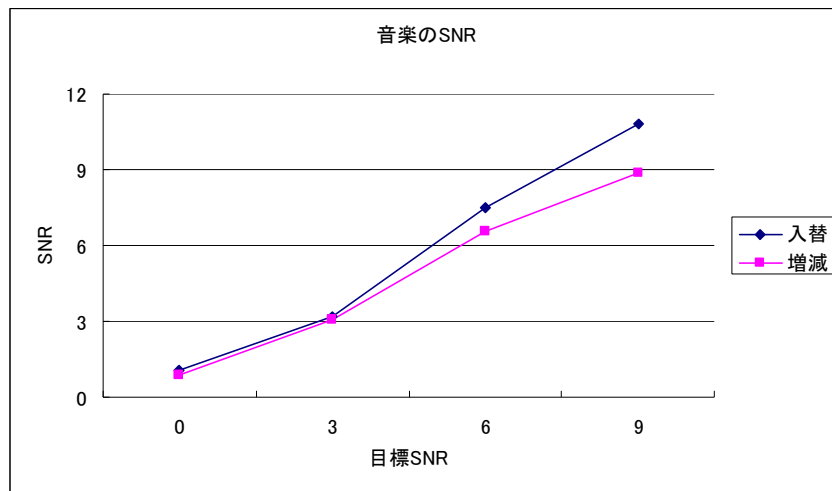
提案手法におけるSNRの制御性能の確認

実験条件

種類	音楽、音声、音楽付き音声
設定SNR	0、3、6、9

※SNRの算出は各オーディオ信号に対してMP3符号化のみを行った信号とMP3符号化に加え提案手法を適用した信号間で行った

実験結果



▶ MDCT係数に0が多い「音声」は
目標SNRよりやや高くなる傾向

▶ 入替の場合は目標SNRより高くなる傾向

実験

実験内容

提案手法を適用した各オーディオ信号の劣化度に関する主観評価実験

実験条件

種類 音楽、音声、音楽付き音声

設定SNR 0、3、6、9

評価者 非専門家15名

2重刺激尺度法[1]

無劣化オーディオ信号 無音 スクランブル信号

5秒

5秒

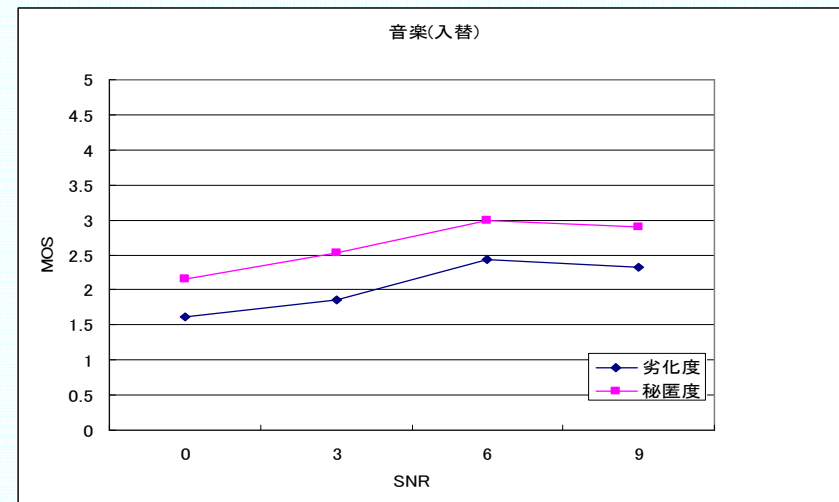
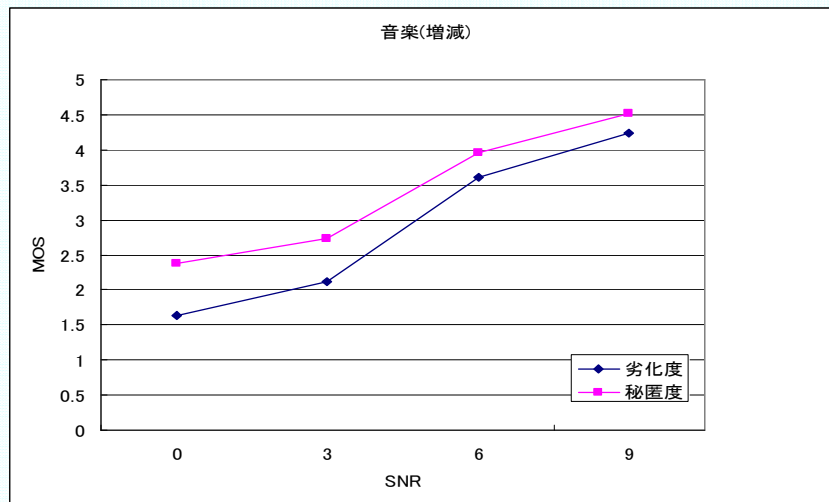
5	劣化がわからない	オリジナルと同等の価値がある
4	劣化がわかるが気にならない	オリジナルに近い価値がある
3	劣化が気になるが耳障りでない	オリジナルほどではないが価値がある
2	劣化が耳障りである	コンテンツの価値がほぼ失われている
1	劣化が非常に耳障りである	コンテンツの価値が失われている

※SNRの算出は各オーディオ信号に対してMP3符号化のみを行った信号とMP3符号化に加え提案手法を適用した信号間で行った

[1] 成田長人, 画像符号化の主観評価法に関する考察, "電子情報通信学会論文誌, vol.J77-B-I, no.2, pp.102--111, 1994.

実験結果

音楽に対して提案手法を適用した結果

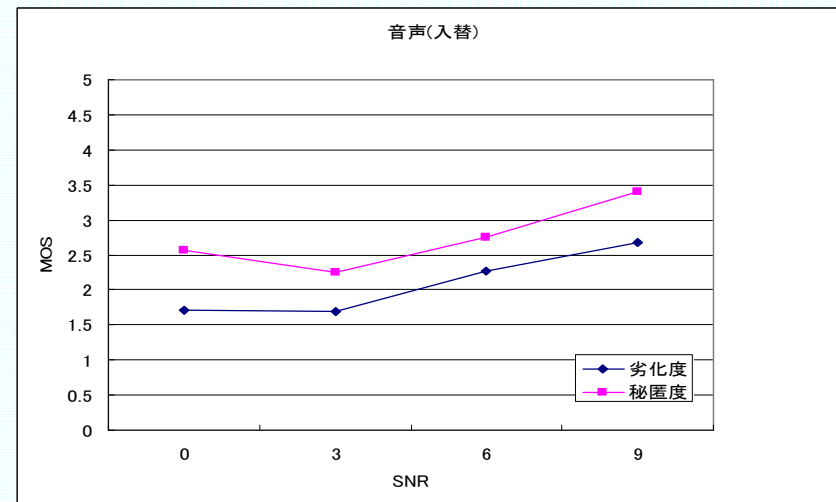
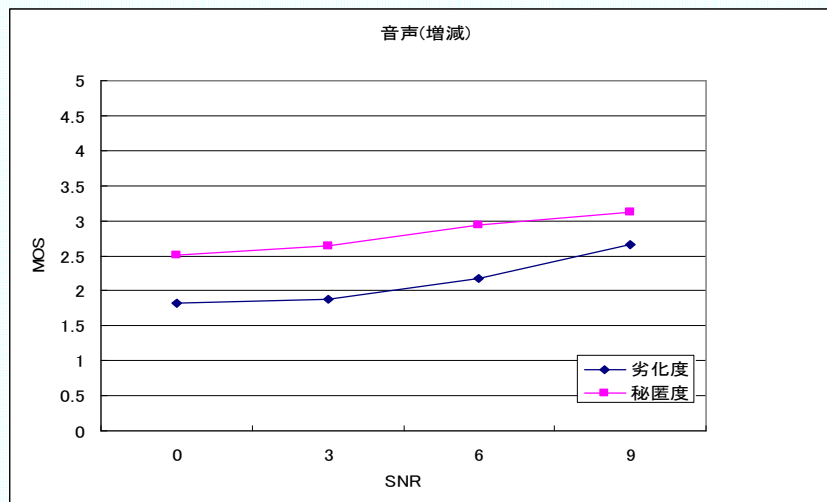


▶ MDCT係数の変更方法の違いでMOSの変化が大きく異なっている

➡ 音楽の場合では、入れ替えのほうが劣化を強く感じると考えられる

実験結果

音声に対して提案手法を適用した結果



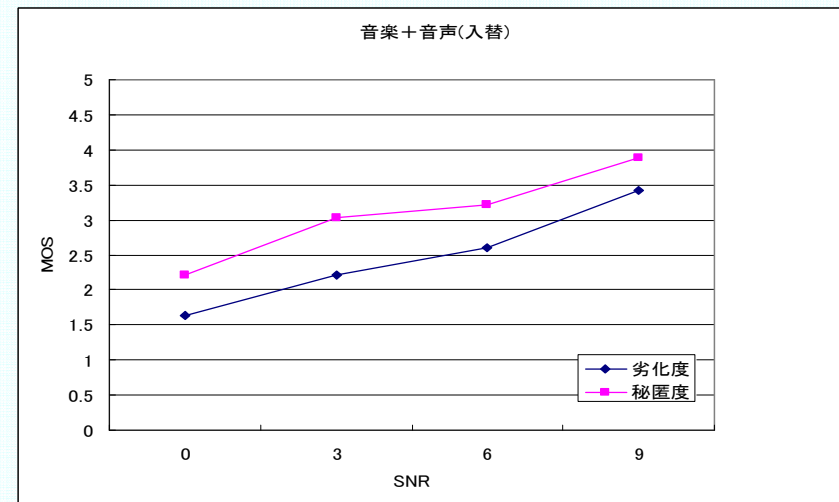
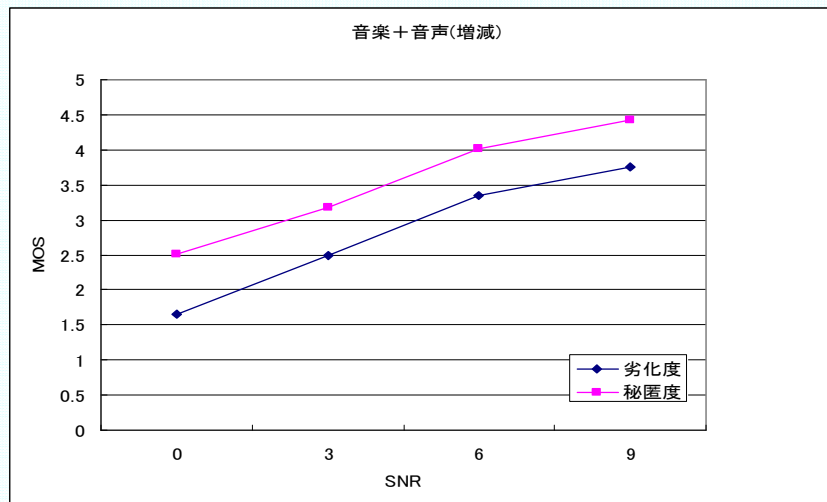
▶ 係数の変更方法の違いでMOSの変化が大きく異なっている様子は確認できない



音声の場合では、言葉を聞き取って意味を理解できるか否かが重視されると考えられる

実験結果

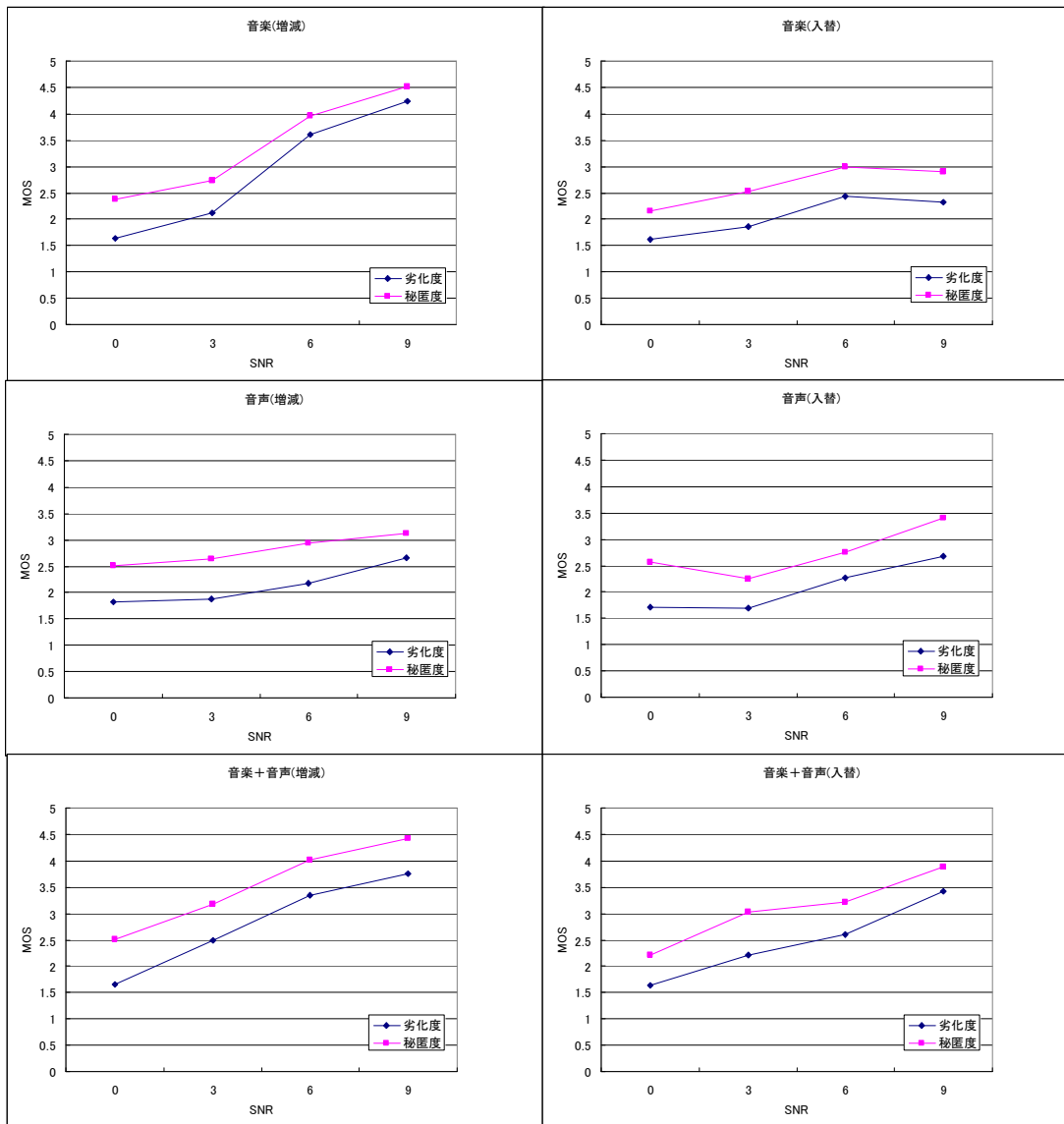
音楽付音声に対して提案手法を適用した結果



▶ MDCT係数の変更方法の違いでMOSの変化が大きく異なっている様子は確認できないが入れ替えの方がややMOSが低い傾向

➡ 音声そのものの劣化のみではなく、付与された音楽の劣化も間接的に音声の秘匿性を高めていると考えられる

実験結果



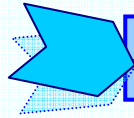
▶ SNRの上昇に伴い、劣化度および秘匿度のMOSも上昇する傾向

➡ 提案手法が定量的に音質を制御可能

▶ オーディオ信号の種類によって劣化の感じ方に違いがある

➡ オーディオ信号の特徴にあわせた適切な処理が必要

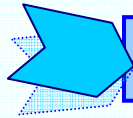
半開示スクランブル



コンテンツを保護しつつサンプルとして利用可能

提案手法

- ▶ SNRを音質の基準とし、定量的な音質の制御を行う



半開示スクランブルの強度が定量的に調節可能

考察

- ▶ オーディオ信号の種類の違いによる劣化の感じ方の違い
- ▶ オーディオ信号の種類に応じて適切な強度、スクランブル方法が異なる

今後の課題

- ▶ 一つのオーディオ信号内でスクランブル方法や目標SNRを変更可能とする