

# 逆転音声による音声・音楽の半開示スクランブルに関する検討

A Study on Partial-scrambling of Speech and Music by Reverse Speech

本多 麻衣<sup>\*1</sup> 河村 圭<sup>\*2</sup> 渡辺 裕<sup>\*2</sup> 富永 英義<sup>\*1\*2</sup>  
 Mai HONDA<sup>\*1</sup> Kei KAWAMURA<sup>\*2</sup> Hiroshi WATANABE<sup>\*2</sup> Hideyoshi TOMINAGA<sup>\*1\*2</sup>

<sup>\*1</sup> 早稲田大学理工学部 電子・情報通信学科 <sup>\*2</sup> 早稲田大学大学院 国際情報通信研究科  
<sup>\*1</sup> Dept. of Elec. Info. and Comm. Eng., WASEDA Univ. <sup>\*2</sup> Graduate School of GITS, WASEDA Univ.

## 1 はじめに

近年、音楽情報の取引が広範に普及する中、提供者が音楽情報にスクランブルをかけ、利用者が鍵を購入し解除するという仕組みは、不正コピーを防止し、確実な課金が可能となるという点で不可欠なものである。しかし、利用者は鍵を購入する前にその音楽情報を確認したいという要求がある [1]。

本稿では、音声スクランブルに逆転音声を適用するシステムを提案する。逆転音声について実データを用いた周波数解析と数式による理論解析を行う。

## 2 従来手法

既存の音楽半開示システムは特定の符号化に特化しており、必ずその方式で符号化しなければならない [1]。更に、多数存在する符号化の各々に特化したスクランブル方式を設計するのは、コストや手間がかかる。

逆転音声は、音声の時系列信号を周期的な区間に分け、分割した区間で信号の順序を逆転させる処理である。分割する区間の長さによって、音質変化の効果、ピッチ、抑揚が変化する効果、音韻の置換え効果が起こる [2]。

## 3 提案手法

本研究では、符号化方式によらない単一な半開示スクランブルを時間領域での簡単な処理で実現する逆転音声スクランブルシステムを提案する。

逆転音声スクランブルを解除するためには、逆転させる際に区切った区間長と開始点が分かっている必要があり、開始点が分からない音源を元に戻すことは困難である。また、逆転音声は区間長によっては音声の識別が可能であり、半開示システムを実現できる。本研究では再生器でスクランブルの有無の切り替えを行うシステムを想定する。

## 4 実験

### 4.1 実験条件

音声データには、男性アナウンサーによるニュース素材を使用し、その条件は、サンプルサイズ 16bit、チャンネル数 1 (モノラル)、サンプル周波数 22kHz、オーディオ形式は PCM である。本データを用いて逆転音声処理を行い、窓関数としてハミング窓を使用し、時間-周波数解析を行った。

ここで、逆転音声処理の際に分割する区間に含まれるサンプル数を区間内サンプル数  $N$  と定義する。実験条件を表 1 に示す。

### 4.2 理論解析

$N$  が小さいとき (音質変化)、原信号を  $x(t)$ 、逆転音声処理を施した信号を  $g(t)$  とし、それぞれのフーリエ変換を  $X(f)$ 、 $G(f)$  とすると、逆転音声信号  $g(t)$  は式 (1) のように表せる。

表 1 実験条件

	窓長	fft 長	オーバーラップ長
音質変化 ( $N \sim 66$ )	500	512	400
抑揚変化 ( $N=33 \sim 286$ )	2000	2048	1600
音韻置換 ( $N=5500 \sim$ )	500	512	400

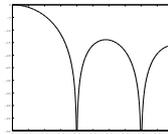


図 1 振幅 ( $N=5$ )

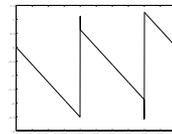


図 2 位相 ( $N=5$ )

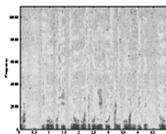


図 3 original

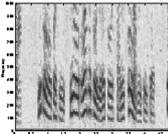


図 5 original

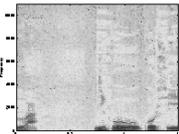


図 7 original

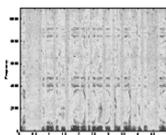


図 4 音質変化: $N=5$

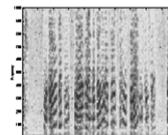


図 6 抑揚変化: $N=220$

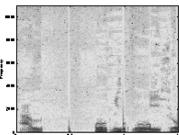


図 8 音韻置換: $N=5500$

$$g(t) = \sum_{i=0}^{N-1} x(t-2i) \sum_{n=-\infty}^{\infty} \delta(t-Nn-i) \quad (1)$$

これより、逆転音声処理を施した音声信号の周波数スペクトルはサンプリング周波数を  $f_s$  とおくと、

$$G(f) = \frac{1}{N} \sum_{m=0}^{N-1} e^{-jmf} \sum_{k=-\infty}^{\infty} X(f - k \frac{f_s}{N}) \quad (2)$$

と表せる。 $N=5$  のときの振幅と位相のグラフをそれぞれ図 1, 2 に示す。

### 4.3 結果と考察

原信号と  $N=5, 220, 5500$  のときの解析結果をそれぞれ図 3~8 に示す。ただし、時間範囲は  $N=5, 220$  のときは 5000[ms]、 $N=5500$  のときは 1500[ms] とし、周波数範囲は  $N=5, 5500$  のときは 11000[Hz]、 $N=220$  のときは 1000[Hz] とした。

音質変化が起こったとき、式 (2) より、逆転音声処理を施した信号は、サンプリング周波数を  $N$  で割った値毎に原信号を振幅と掛け合わせたものを足し合わせた信号となり、 $N$  の数だけパワーが強まる部分が発生する。 $N=5$  のときの図 4 では、特徴的な縞が  $N$  と同数発生している。

抑揚変化が起こった結果図 6 は周波数が周期的に反転しており、パワーの強い部分が反転したために、ピッチ、抑揚が変化した。音声の識別が可能であることから半開示スクランブルに適用できる。

$N$  が大きいときの結果図 8 ではパワーの強い部分が時間軸で移動しており、音韻が移動している。

## 5 まとめ

本稿では、逆転音声を用いて音楽・音声の半開示スクランブルを掛けるシステムを提案した。そして、実データを用いて逆転音声について周波数解析と数式による理論解析を行い、有効性を確認した。

### 参考文献

- [1] “音楽を自由にスクランブルする半開示技術を開発,” <http://www.ntt.co.jp/news/news97/970425b.html>
- [2] 片山泰男, “逆転音声について,” <http://home.catv.ne.jp/dd/pub/rev.html>