

PSOLA を用いたピッチ加工の評価実験

Evaluation Experiment of Pitch Modification Using PSOLA

A-4

田坂 直季[†]津田 浩利[†]小坂 直敏[†]Naoki TASAKA[†]Hiromichi TSUDA[†]Naotoshi OSAKA[†][†]東京電機大学[†]Tokyo Denki University

1. はじめに

TD-PSOLA 方式[1]は音声を対象にした時間伸縮およびピッチ加工の優れた方式である。本稿では、同方式を楽音に適用して音質評価実験をした結果について述べる。

2. PSOLA (ピッチ同期オーバーラップ加算)

ピッチ同期オーバーラップ加算(Pitch-Synchronous Overlap-Add)とは、音のピッチ周期と同期してOLA(Overlap-Add)を行う方式である。本稿では時間伸縮、ピッチ加工に関して優れた手法である TD-PSOLA を用いる。

2.3.1 ピッチ同期分析

以下、 n は離散時間、 s は分析短時間波形の番号、 u は合成短時間波形の番号として記す。

以下の式(1)を用いて、音のピッチに同期する窓を掛けることで、短時間波形に分解する。

$$x(s, n) = h_s(n)x(n-t(s)) \quad (1)$$

$x(n)$: 元のデジタル波形

$x(s, n)$: 分析短時間波形の系列

$h_s(n)$: ハニング窓の系列 (窓の長さは元の音のピッチ周期と比例して設定される)

$t(s)$: s 番目の分析時刻 (信号の有声上では元波形のピッチ周期レートで設定され、無声上では一定のレートで設定される)

2.3.2 ピッチ同期変更

以下の式(2)を用いて、合成したい音のピッチ周期と同期するように、それぞれの分析波形を合成波形へと変更し、時間伸縮やピッチ加工を行う。

$$y(u, n) = x(s, n), \tilde{t}(u) \rightarrow t(s) \quad (2)$$

$y(u, n)$: 合成音に対する短時間波形の系列

$\tilde{t}(u)$: u 番目の合成時刻 (それぞれ合成波形の時間伸縮と合うピッチ周期レートで設定される)

2.3.3 ピッチ同期オーバーラップ加算

以下の式(3)を用いて、合成時刻を元に、合成短時間波形をOLAすることで、合成波形を作成する。

$$y(n) = \sum_u y(u, n) \quad (3)$$

$y(n)$: 時刻 n の合成音

3. PSOLA によるピッチ加工の実験

PSOLA 方式を用いて、2つの楽音と2つの女性の音声に対して、3種類のピッチ加工を行い、音質の評価を行う。これらの音をそれぞれ3回ずつランダムに再生し、それぞれの音質を5段階で評価する。

被験者は成人男性 10 名とし、PSOLA によるピッチ加工の評価を行う。表 1 は、実験で用いる 16 個の音の一覧である。

図 1 に評価結果を示す。実験では楽音などのピッチが一定であるピッチ加工音に関して良好な結果が得られることが、分かった。逆に、ピッチが変動する女声に関しては、ピッチ加工の評価が著しく低い。

表 1 実験音源

	サンプル No	内容
使用する音	1-4	・ピアノ
	5-8	・チェロ
	9-12	・女声 (ピッチが変化)
	13-16	・女声 (ピッチが一定)
ピッチ加工方法	1	・原音
	2	・1/3oct 低
	3	・1/3oct 増
	4	・2/3oct 増

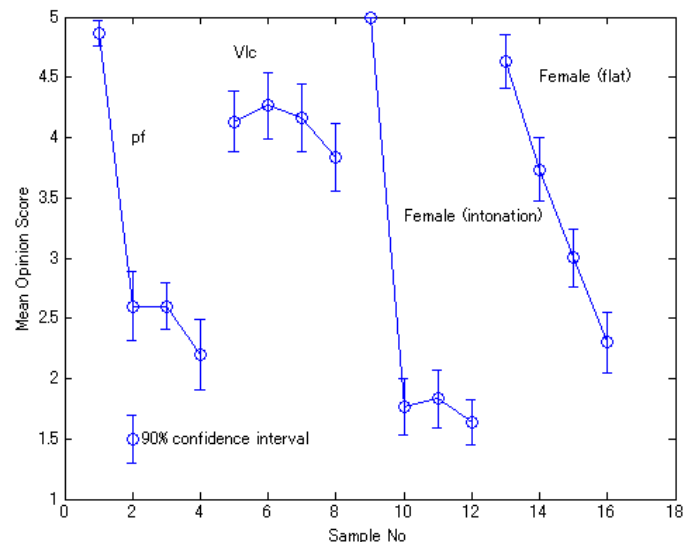


図 1 ピッチ加工の 5 段階評価

4. まとめ

今後はピッチ変動のある音声に対するピッチ抽出方式を改良していく。

参考文献

- [1] E. Moulines, F. Charpentier, Speech Communication 9, pp.453-467, 1990.
 [2] E. Moulines, J. Laroche, Speech Communication 16, pp.175-205, 1995.