

3K-8

# 時間間引きによる音声信号の時間軸圧縮に関する基礎検討

加藤 誠巳 藤原ひろみ 佐藤 恵臣  
(上智大学理工学部)

## 1. まえがき

音声信号の時間圧縮は速記録の作成、特定個所の検索、あるいはラジオ放送における時間調整など実用的に極めて重要である。テープレコーダに録音した音声を実に速いテープ速度で再生するとピッチが変化し了解性が極めて劣化する。そのため、従来から多数の研究が行われてきている<sup>(1)(2)</sup>がピッチの抽出を前提としているものや複雑な計算を伴うものも多い。単純な時間間引きの方法は誰でも考えつき極めて簡単であるにも拘らず余り検討されていないように見受けられる。本稿ではこの時間間引き法において良好な音質を得るための信号波形の相互接続法ならびに圧縮率と切り出し周期の関係を実験的に求めた結果について述べる。

## 2. 時間間引きの方法

ここでは処理の対象とする音声は一定速度でサンプリングされ、符号化されたものとする。この仮定はテープレコーダの再生にあつては再生速度に比例してサンプリング速度を上げることと同一である。図1(a)は原音声信号であり、圧縮率 $\rho$ の関数であるサンプル数 $N$ の周期で切り出しを行う。図1(b)は間引き窓であり原信号を圧縮率 $\rho$  ( $0 < \rho \leq 1$ )で圧縮する場合を示している。単純に間引き窓で原信号をゲートすると両端で信号の不連続を生じ、耳障りな雑音の原因となるので原信号の微係数が正である零交差点に一番近いサンプル点で切り出すものとする。この様な切り出しを行うと切り出されたデータ数は所望の $N\rho$ とはならないことが生じるので図1(c)に示すように間引き窓の前縁および後縁において夫々前後両方向に対し切り出し点を探索し、累積誤差ができるだけ少なくなる方の点を採用する。この様な時間間引きにより合成された時間圧縮波形を図1(d)に示す。

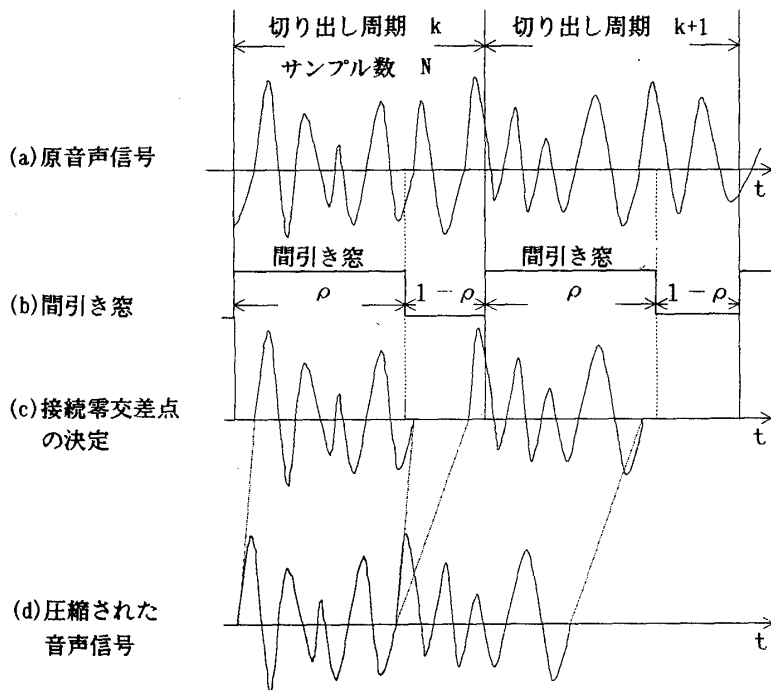


表1 切り出し周期

$\rho$	$N$	$N(1-\rho)$
1/1.1	3300	300
1/1.2	1800	300
1/1.5	900	300
1/1.8	675	300
1/1.9	633	300
1/2.0	600	300
1/2.5	500	300
1/3.0	450	300

図1 時間間引きの手法

### 3. 切り出し周期の決定

圧縮率 $\rho$ が大のときには切り出し周期のサンプル数 $N$ は大きく取り、 $\rho$ が小なるときには $N$ を小さくすべきことは容易に想像される。男声音、女声音さらには個人ごとに最適値は異なると考えられるが10kHzサンプルで $1/\rho$ が3以下の場合は間引くサンプル数 $N(1-\rho)$ を300程度の一定値にすると良好な結果が得られることが主観評価によって明らかとなった。その結果を表1に示す。

### 4. 時間圧縮された音声信号の特性評価

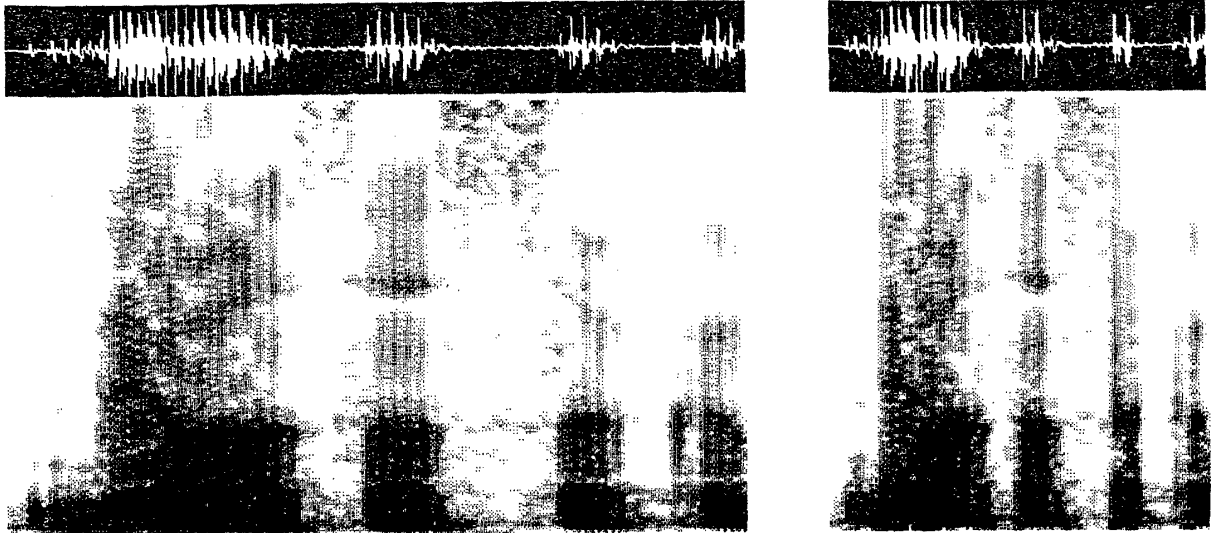
本方式を用いた時間圧縮は簡単であるにも拘らず極めて良好な音声が得られ、 $\rho=1/3$ でも十分了解性を有している。図2(a)に原音声の、図2(b)に $\rho=1/2$ とした時のサウンドスペクトログラムを示す。

原音声の仕様は次の通りである。

成人男性音声 : 「ではさっそく」

標本化周波数 : 10kHz

量子化ビット数 : 16ビット



(a)原波形のスペクトログラム

(b) $\rho=1/2$ のスペクトログラム

図2 スペクトログラムの比較

### 5. むすび

不連続性の発生を回避するべく正の微係数を有する零交差点近傍で相互接続を行う時間間引きによる音声信号の時間軸圧縮について検討を行い、圧縮率と切り出し周期の関係を実験的に求めた。本方式は極めて簡単であるにも拘らず良好な音質特性を有していることが明らかとなった。有益な御助言を戴いたNTT基礎研筭一彦氏に謝意を表す。

#### 参考文献

- (1) 井上, 中田: "音声波形時間軸伸縮方式", 音響学会講演論文集, 2-5-9(昭60-03).
- (2) 森田, 板倉: "自己相関法による音声の時間軸での伸縮方式とその評価", 信学技報, EA86-5(昭61).