

差分関数を利用した音声処理方式——SPAD——

准員 國枝 伸行[†] 正員 島村 徹也[†]
 正員 鈴木 誠史[†] 正員 八嶋 弘幸[†]

Speech Processing System by Use of Auto-Difference Function——SPAD——
 Nobuyuki KUNIEDA[†], Associate Member, Tetsuya SHIMAMURA[†],
 Jouji SUZUKI[†] and Hiroyuki YASHIMA[†], Members

[†] 埼玉大学工学部情報工学科, 浦和市
 Faculty of Engineering, Saitama University, Urawa-shi, 338 Japan

あらまし 雑音の重畳した音声の短時間差分関数の波形を逐次接続して出力することにより, 雑音レベルの低減を行うことができる. この方式は, 従来の自己相関関数を利用した音声処理方式 (SPAC) よりも単純な構成で, 同程度の雑音低減ができることを示す.

キーワード SPAC, 自己相関関数, 差分関数, 雑音低減

1. まえがき

自己相関関数を利用した音声処理方式 (SPAC) は, 音声信号の雑音レベルの低減をはじめ, 周波数帯域の圧縮・拡大などを行うことのできる音声処理方式である⁽¹⁾. しかし, SPACでは自己相関関数を計算する過程で周期波の振幅が2乗されるスペクトル2乗ひずみを生じる. また, 積分計算に時間を要するという欠点もある. そのため, 2乗ひずみを補正するイコライザの検討を行い⁽²⁾, 零交差波との相互相関関数を利用した音声処理方式 (SPOC) なども考案してきた⁽³⁾.

ここでは自己相関関数に代わって, 計算が簡単でスペクトル2乗ひずみの生じない差分関数を利用した音声処理方式 (SPAD: Speech Processing system by use of Auto-Difference function) を提案し, 従来のSPACやSPOCとの比較実験を行う.

2. 差分関数の利用

雑音の重畳した周期信号の自己相関関数では, 周期波の各周波数成分は保存されるが, 雑音成分は自己相関関数の原点付近に集中する. この部分を取り除いて波形を出力することによって, 雑音の低減を行うのがSPACの基本原理である. 音声のように時間的に変動する波形に対しては, 自己相関関数の代わりに短時間自己相関関数を用い, その1周期の波形を切り出して逐次接続して出力する. 図1にSPACの処理の流れを示す. SPACは, アルゴリズムが簡単で実用性の高い処理方式ではあるが, 自己相関関数の計算に時間がかかり, 2乗ひずみを補正するためのイコライザが必要

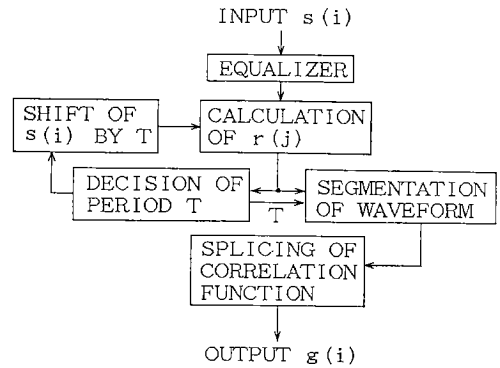


図1 SPACの処理過程
 Fig.1 Process of SPAC.

である. しかも, その理想的イコライザの構成は困難である. そこで自己相関関数の代わりに, 2乗ひずみがなく, 自己相関関数と同様の性質をもつ差分関数を利用することを考える.

差分関数は, 計算が簡単で自己相関関数と同様に周期性検出の能力が高いことから, これまでもピッチ抽出にしばしば利用されてきた^{(4),(5)}. しかし, これを波形として利用した例はない. また, 信号の振幅がガウス分布に従うとき, 正規化した自己相関関数 $R(\tau)$ と差分関数 $D(\tau)$ との間には, 比例定数 β を用いて

$$D(\tau) = \beta \sqrt{1 - R(\tau)} \quad (1)$$

という関係がある^{(4),(5)}. この式から, 差分関数についても原点付近に雑音成分が集中することがわかり, またその極値の位置も自己相関関数と一致していることもわかる. 従って, 差分関数を用いても, 自己相関関数と同様の雑音低減が期待できる.

そこで, シミュレーションによって雑音低減効果を求めることを考える. まず, 時系列信号 $s(i)$ の短時間自己相関関数 $r(j)$ と短時間差分関数 $d(j)$ を, 式(2), (3)で定義する.

$$r(j) = \frac{1}{N} \sum_{i=0}^{N-1} s(i) \cdot s(i+j) \quad j=0, 1, \dots, M \quad (2)$$

$$d(j) = \frac{1}{N} \sum_{i=0}^{N-1} |s(i) - s(i+j)| \quad j=0, 1, \dots, M \quad (3)$$

周期性のない雑音を入力信号とした場合, その自己相関関数の主要な振幅成分は原点付近に集中する. そのため, 相関関数の原点付近を除いた部分を出力信号とすれば, その出力雑音電力は入力雑音電力に比べて大幅に減少する. 入力雑音電力を P_n , 各処理後の出力雑音電力を P_{out} とし, 雑音低減比を式(4)で定義する.

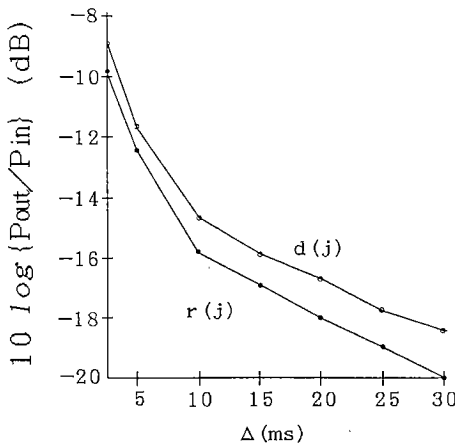


図2 雑音レベル低減効果
Fig.2 Effects of noise level reduction.

$$\gamma = 10 \log \frac{P_{out}}{P_{in}} \quad (4)$$

シミュレーションでは、0.2~3.4 kHz に帯域制限した白色ガウス雑音を入力信号（標本化周期 0.1 ms）として $r(j)$, $d(j)$ を求める。帯域幅と男声の平均的な基本周期を考慮して、 $j=10\sim 80$ の区間を出力信号とすることにす。その出力の分散から式(4)で定義した雑音低減比を計算する。

この雑音低減効果は、相関をとる計算時間によっても異なる。この時間を $\Delta = N \times$ (標本化周期) によって計算し、これと雑音低減効果との関係の様子を図2に示す⁽⁶⁾。なお、 $r(j)$ では理想的イコライザがあるものとしている。この図から、 $r(j)$ と $d(j)$ との差はわずか 1~2 dB であり、 $N=200$ のときでは -17 dB 程度の雑音低減が期待できることがわかる。

さて、正弦波の差分関数を計算すると全波整流波と同様の波形になる。振幅はすべて正の値であるため、これをそのまま切り出すと、音声入力の場合、直流成分が変動して不連続が生じる。そこで出力波形の平均値が 0 となるように振幅軸を移動して、1 周期の波形を切り出し、接続した信号を出力信号とする。しかし、正弦波入力で明らかなように、出力波に約 -12 dB の波形ひずみが生じる⁽⁶⁾。これは、全波整流波をフーリエ級数展開することによって計算できる。

このように入力信号が正弦波のとき、2 乗ひずみは生じないが高調波を生じる波形ひずみが生じる。しかし、SPAC のような 2 乗ひずみを解決するためのイコライザが不要であることは、大きな利点である。一方、

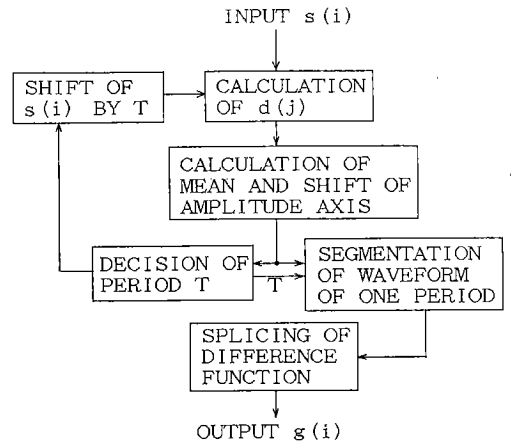


図3 SPADの処理過程
Fig.3 Process of SPAD.

音声のような調波成分からなる複合波の場合には、理論的な解析は困難であり、これらのひずみの影響がどの程度であるか、試聴実験による評価が必要となる。

3. 処理過程

SPAD の処理過程を図3で説明する。まず、入力信号 $s(i)$ の短時間差分関数 $d(j)$ を計算する。音声の基本周期数が存在する j の範囲内の $d(j)$ の最小値から周期 T を決定する。そして雑音成分の集中した $d(0)$ 付近を取り除いた 1 周期分を切り出し、その区間の平均値が 0 になるように振幅軸を移動させ、波形を出力する。次に $s(i)$ を T だけシフトさせて $d(j)$ を計算し、同様の処理を繰り返す。こうして出力された波形をつなぎ合わせていくことによって、入力信号の雑音レベルを低減した信号を得ることができる。

周期 T を決定する際、2 周期分を 1 周期と判断することもあるが、本方式では誤りにはならない。また無声音やポーズ区間では、周期性が見られないが、この場合にも同様に最小値によって T を定めて波形を切り出して接続する。

図1と図3の処理を比較すると、SPAD では波形の平均値を求め、振幅軸を移動する過程があるが、積和計算が不要である。また SPAD では、イコライザが不要であり、処理が単純化されている。

4. シミュレーション実験

標本化周波数 10 kHz で標本化した音声資料に白色雑音を重畳させ、式(2)、(3)での定数を $N=200$, $M=165$, 周期 T の検索の範囲は $j=30\sim 150$ として各処理を行った。また、SPOC の相互相関関数 $q(j)$ は式(5)で計算し、 M , N の定数も同様に設定した。

$$q(j) = \frac{1}{N} \sum_{i=0}^{N-1} s(i) \cdot \text{Sign}(s(i+j))$$

$$j=0, 1, \dots, M \quad (5)$$

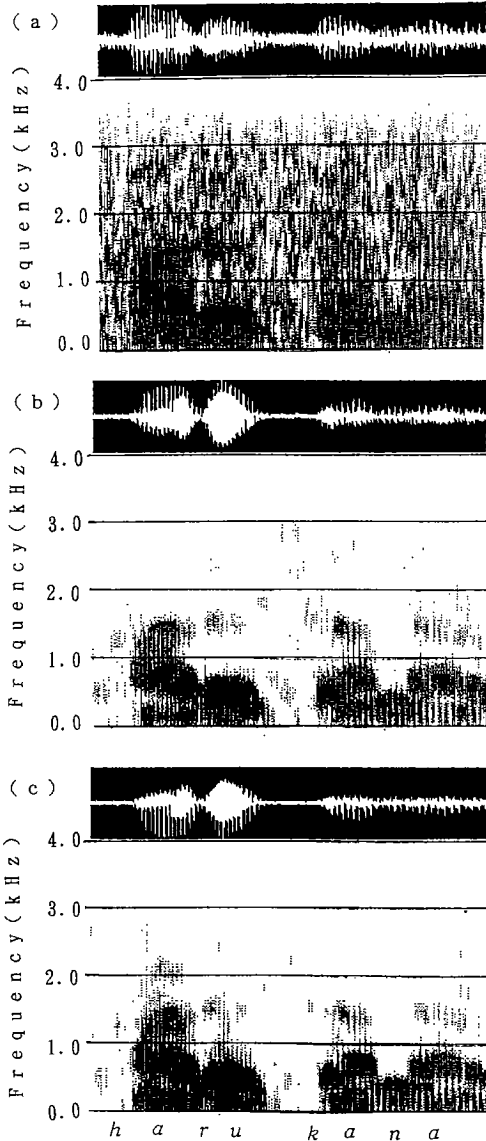
また、SPACの前処理としては、単に振幅の平方根をとるC1型イコライザと、周波数帯域を3チャンネル

に分割した後で平方根をとるC3型イコライザの2種類⁽²⁾で処理した。SN比0dBの場合についての処理前とSPAC(C1型)、SPADで処理した音声のスペクトログラムを図4に示す。SPACもSPADも同程度の雑音低減がなされている。ここでは示さないが、C3型イコライザを使ったSPACは、雑音低減がやや劣るものの、高い周波数成分の保存がSPADよりもよくなっている。

5. 試聴実験

SPADとSPAC、SPOCで処理した音声と対比較試験を行った。音声資料は、3名の男性が発声した短文に白色雑音を重畳したものであり、SN比は10~-10dBまでの5dBごとを対象とした。これらの音声に対しSPAC(C1型、C3型)、SPOC、SPADの各処理を行い、5人の被験者に各対6回ずつランダムに呈示した。評価基準は、「どちらが雑音が少なく聞きやすいか」である。

SN比0dBの結果を表1に示す。ここでoriginalとは処理を行う前の音声のことである。またこれらをサーストンの心理尺度(ケースV)に変換したものを図5に示す。この図によるといずれのSN比の場合にも、SPADの評価が良いことがわかる。特にSN比が0dB以下の音声については、雑音低減が最も大きいという評価が得られている。



(a)Original speech(SNR 0dB)
 (b)Processing speech by SPAC(C1)
 (c)Processing speech by SPAD

図4 処理前後のスペクトログラムと波形

Fig.4 Spectrograms and waveforms before and after processing.

表1 SN比0dBのときの対比較試験結果
 (xに対するyの選択率)

x \ y	original	C1	C3	SPOC	SPAD
original	.122	.178	.144	.133	
C1	.878	.500	.733	.400	
C3	.822	.500	.400	.278	
SPOC	.856	.267	.600	.311	
SPAD	.867	.600	.722	.689	

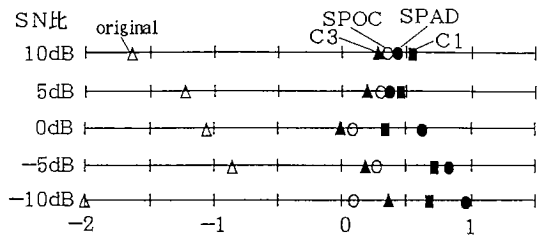


図5 心理尺度で表した対比較試験の結果

Fig.5 Quality of speech on psychological scale.

6. むすび

差分関数を利用した音声処理方式 (SPAD) は, SN 比が 0 dB 以下の音声はもちろん 10~0 dB の音声についても, 従来の SPAC や SPOC に劣らない雑音低減能力をもつことがわかった. SPAD の場合, 入力波形の短時間差分関数を計算し, その波形を切り出して出力するだけで, 極めて単純な処理である. しかも, 従来の SPAC で必要としたイコライザが不要であり, 積和計算がないことは大きな利点である. この点では, FFT を計算しなければならないスペクトル減算法よりも優れていると言える.

問題となる高調波ひずみに関しても, 入力音声の SN 比が高いときはその影響が目立つものと思われるが, SN 比の低い音声については, 今回の試聴実験の結果からさほど影響がないことがわかった. 従って, SN 比の低い音声の雑音レベルの低減方式としては, 有効な手段であると言える.

また, SPAC と同様の処理方式であるから, 雑音低減のみでなく, 周波数スペクトルの圧縮・拡大, 時間軸の伸長・短縮などを行える性質ももっている. すなわち狭帯域伝送, 通信品質の改善, 難聴者の通話補助などへの幅の広い応用も期待できる⁽¹⁾.

今後の課題として, 白色雑音だけではなく, 有色雑音の低減についての検討も行いたい. また今回は, 男性の音声のみを扱ったが, 女性の音声についても検討することが必要である. また, 音声の自然性についての評価を行うと共に, ハードウェアの作成も検討したい.

謝辞 本研究の一部は第 12 回国際通信研究奨励金によるものである.

文 献

- (1) 高杉敏男, 鈴木誠史, 田中良二: “自己相関関数を利用した音声処理方式 (SPAC) の機能と基本特性”, 信学論(A), **J62-A**, 3, pp. 175-182 (1979-03).
- (2) 鈴木誠史: “SPAC の前処理の検討”, 信学技報, **EA76-20** (1976-07).
- (3) 鈴木誠史: “相互相関関数を利用した音声処理方式”, 信学論(A), **J60-A**, pp. 338-339 (1977-03).
- (4) Ross M. J. Shaffer H. L., Cohen A., Freudberg R. and Manley H. J.: “Average Magnitude Difference Function Pitch Extractor”, IEEE Trans. Acoust., Speech & Signal Process., **ASSP-22**, 5, pp. 353-362 (1974).
- (5) 藤崎博也: “電子計算機による音声のピッチ抽出”, 情報と制御の研究, 2/3, pp. 68-77 (1963).
- (6) 鈴木誠史, 八嶋弘幸: “差分関数を使った音声処理の可能性”, 音響学会講演論文集, 2-6-5 (1991-03).

(平成 4 年 5 月 21 日受付, 6 月 25 日再受付)