

砂時計型ニューラルネットワークを用いた雑音除去フィルタの動特性の検討

1 P-6

並木寿枝 古本友紀 吉村宏紀 清水忠昭 井須尚紀 菅田一博

鳥取大学工学部

1. はじめに

補聴器への応用を目的として、我々は砂時計型ニューラルネットワーク (Sandglass type Neural Network: SNN)⁽¹⁾を用いた雑音除去フィルタを提案してきた⁽²⁾⁽³⁾。SNN とは、中間層のユニット数を入力・出力層のユニット数より少なくした階層型ニューラルネットワークであり、本研究では応答関数に線形関数を用いている。中間層ユニット数が 1 つの SNN を縦列に接続したネットワークを CSNN (Cascade SNN) と呼び、接続した SNN の個数を接続段数と呼ぶ。CSNN で構成した雑音除去フィルタ (CSNNRF) は、信号が入力される度に CSNN の学習を行い、Karhunen-Loève 変換を用いた雑音除去をリアルタイムで実現する。

CSNNRF は、入力信号に含まれる主要な主成分のみを抽出することによって雑音除去を行う。しかし、音声信号や雑音の大きさが経時的に変化する場合、最適な雑音除去を行うために CSNN の接続段数を適応的に変える必要がある。我々はこれまでに、雑音の大きさの変化に応じて適応的に接続段数を変化させる CSNNRF を提案し、音声の聞き取りやすさが向上することを示した⁽⁴⁾。本研究では、CSNNRF の適応的な変化に伴ってフィルタ特性がどのように変化するかを検討した。

2. 砂時計型ニューラルネットワーク雑音除去フィルタ (CSNNRF)

入力・出力層のユニット数を 20 個、中間層ユニット数 1 個の 3 層 SNN を 20 段に接続して CSNNRF を構成した (図 1)。各ユニットの応答関数は線形関数とした。CSNNRF の入力、音声信号系列を 10 点ずつシフトしながら 20 点ずつ第 1 段 SNN に入力した。2 段目以降の SNN には前段 SNN の誤差信号 (入力層ユニットへの入力-出力層ユニットの出力) を入力した。第 1 段から第 P 段 ($1 \leq P \leq 20$) までの出力信号を対応するユニットごとに加算したものを P 段接続 CSNNRF の出力とした。CSNNRF の出力を 10 点ずつシフトしながら対称な三角波窓をかけて加重平均し出力系列を得た。なお、SNN の学習には RLS 学習法を用いた。

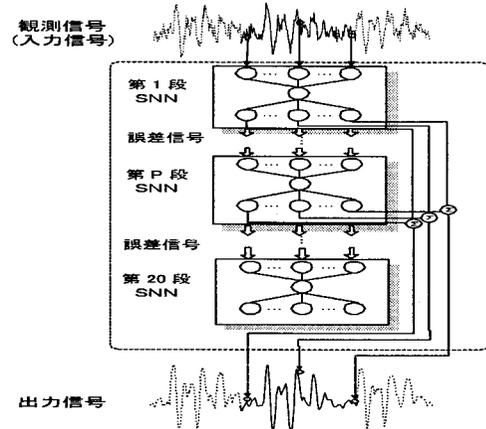


図1 CSNNRFの構成図

3. フィルタ特性

無雑音音声処理時のフィルタ特性

日本音響学会編集の「研究用連続音声データベース」(1991年, vol.2)より選択した文章音声標本(サンプリング周波数 16kHz, 量子化ビット数 16bits, データ長 2.8~9s)を入力信号とし、CSNN の学習を行った。連続音声を入力とし学習を行い、得られた結合係数を用いて P 段接続 ($P=1, 2, \dots, 20$) CSNNRF のフィルタ特性を求めた。フィルタ特性を求める際の入力には周期 1023 点の M 系列信号を用い、出力を 100 周期分加算平均して、1023 点離散フーリエ変換を施した。

時間長の長い摩擦子音と時間長の短い破裂子音および母音が入力された時の CSNNRF のフィルタ特性を調べた。連続音声の中の母音は定常部分で、破裂子音については破裂点付近で、摩擦子音については中央部分で CSNN の学習を打ち切った。また、子音については先行母音の影響異なる場合のフィルタ特性の比較を行い、先行母音の影響を調べた。

図 2 に母音 /i/ を入力した時点の 6, 8, 16 段接続 CSNNRF のフィルタ特性と /i/ のスペクトル包絡を示す。6 段では、/i/ の第 1 フォルマント周波数域を含む低周波域と第 2, 3 フォルマントの中間周波数域を強調するフィルタとなった。段数が多くなると低域の利得が平坦なフィルタ特性が得られた。図 3 に摩擦子音 /s/ (先行母音 /i/) を入力した時点の 6, 8, 16 段接続 CSNNRF のフィルタ特性と /s/ のスペクトルを示す。8 段 CSNNRF のフィルタ特性

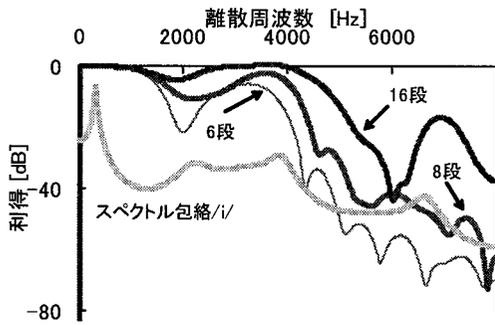


図2 音声/i/の無雑音処理時のCSNNRFのフィルタ特性

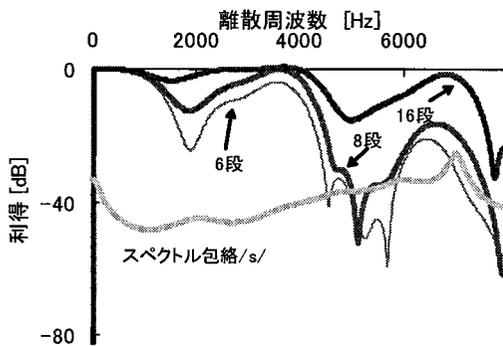


図3 音声/s/の無雑音処理時のCSNNRFのフィルタ特性(先行母音/i/)

では、 $/s/$ が最大のパワーを持つ高域周波数で利得が幾分高くなっているが、中間域では利得が低い。 $/s/$ が持つ周波数域を概ね通過させるためには16段階程度以上の接続段数が必要であった。また、先行母音によってフィルタ特性に相違が見られ、接続段数が少ないほどその差は顕著であった。

雑音付加音声処理時のフィルタ特性

雑音が付加された音声信号を入力した時のフィルタ特性を調べるため、平均SN比を-10dBから10dBまで5dBずつ変えた音声信号を作成した。音声標本は無雑音時と同一のものを使用し、雑音には定常な白色雑音を用いた。摩擦子音 $/s/$ (先行母音 $/i/$)を入力した時点の8段階接続CSNNRFのフィルタ特性を図4に示す。平均SN比が低い時にはフィルタ特性は平坦となり、 $/s/$ の持つ周波数域を多く通過させるフィルタとなった。一方、SN比が高いと、無雑音と同様に高域利得が低く、中間域の利得が更に低い。

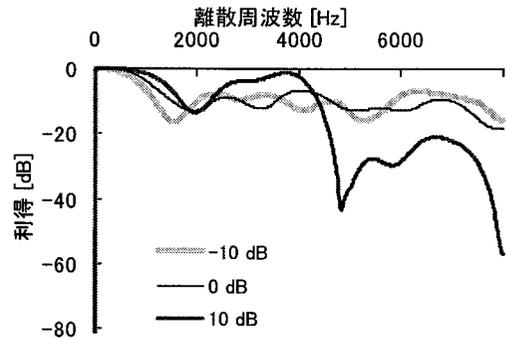


図4 音声/s/の雑音付加音声処理時の8段階CSNNRFのフィルタ特性

4. まとめ

無雑音の音声入力時、母音が入力されるとその母音の低次フォルマント周波数域周辺を通過させるフィルタとなり、子音が入力された時には、子音のみでなく先行母音の影響も受けたフィルタ特性となった。雑音が付加された音声入力時のフィルタ特性は、平均SN比が高い場合は高周波数域の利得が低く、子音の持つ高周波成分を遮断するフィルタとなるが、平均SN比が低い場合には、平坦な特性を得る。

CSNNRFのフィルタ特性は母音の影響を強く受けるため、子音を入力した際にも先行母音の影響で高域の利得が低くなるため、接続段数を母音入力時よりも増加させる必要があり、雑音除去効率が低下する。しかしながら、雑音が大きく、SN比が低い場合には、接続段数が少ないCSNNRFでも平坦なフィルタ特性が得られやすいので接続段数を減少させて雑音の増大を抑えることができる。

【参考文献】

- (1) 渡辺, 他: “KL変換用多段階接続パーセプトロン”, 信学論 (D-II), J75-D-II, 11, pp.1925-1932, 1992
- (2) 吉村, 他: “多段階接続砂時計型ニューラルネットワーク雑音除去フィルタを用いた適応的雑音除去”, 電学論C, Vol. 120, No.4, 2000
- (3) 佛崎, 他: “多段階接続砂時計型ニューラルネットワーク雑音除去フィルタで処理した音声の聞き取りやすさ”, 信学 98秋全大A-89, 1998
- (4) 古本, 他: “砂時計型ニューラルネットワークを用いた雑音強度依存型フィルタ”, 情報処理学会第61回全国大会講演論文集2分冊, 40-41, 2000