計算機アーキテクチャ　71－13
(1988. 7. 7)

ニューラルネットを用いた音声の特徴要素の抽出

Exploring Speech Features using Neural Networks

河原英紀
Hideki KAWAHARA
NTT 基礎研究所
NTT Basic Research Laboratories

あらまし　音声知覚過程の新しい研究手段としてのニューラルネットについて、従来の研究例と、これまでに我々が行なってきた方法を対比させて説明する。不特定話者の母音の識別の例では、自己相関関数から基底膜モデルにいたる様々な入力表現を用いて、ニューラルネットにより、どのような特徴が抽出されるかを紹介する。また、子音を含む音声の動的な特徴を表現する方法として、状態縮約の考え方を紹介し、具体的な音声の処理例を示す。

Abstract　An overview of neural network applications in speech perception and automatic speech recognition research is presented referring our current research activities and preceeding ones. Feature extraction characteristics in speaker-independent vowel recognition tasks using various input representations from autocorrelation coefficients to basilar membrane viblation patterns are duscussed in detail. The state compression technique which is intorduced to manipulate spatio-temporal patterns by neural networks is also expalined using actual speech data.

1　はじめに

特定の話者が離散的に発声した音声を認識することは、急速なマイクロエレクトロニクスの進歩もあり、今日では、実用的に安価に利用できる技術となっている。しかし、誰の話したことでも即座に聞き取ることのできる人間の能力が、今日の機械認識技術の遠く及ばないレベルにあることも事実である。このギャップを埋めることを狙って、様々の技法が開発され、試みられている。最近、急速に関心を集めているニューラルネットは、この音声認識の最終的な目標に挑戦するための、新しい手段を与えるものとして期待できる。ここでは、これまでに我々がニューラルネットを用いて行なってきた、音声の特徴抽出の研究相互の位置づけと今後の課題について、最近の新しい結果を交えながら、紹介する。

2　概要

我々は、これまで、人間の音声認識の機構を明らかにすることが、機械認識を狙う場合にも重要であるとの観点から、人間の言語音声知覚についての心理実験を行なってきた。様々な話者の音声に対する適応能力を調べた実験結果によれば、人間は非常に高度の能力を持っており、わずかに数音節を聞いただけで、いままで聞いたことの無い話者の音声であっても、その話者の音声に適した聞き方に変更できることが示されている [1]。この実験結果と物理的特性の比較により、5母音のホ

ルマント構造が、適応の鍵として用いられている可能性が高いことが、示唆されている。また、子音の知覚を詳細に調べた実験結果も、子音部から後続母音の定常部に到るまでの過渡部分が、子音の知覚に重要な情報を含んでいることを示している[2]。したがって、少なくとも日本語音声の研究に当たっては、母音および過渡部分からなる有声部分に含まれる特徴の抽出が重要な課題となる。

音声の特徴要素については、長い研究の歴史があり、多くの知見が積み重ねられて来ている。しかし、機械認識系の能力が、人間に比較してはるかに低いことは、「従来の音声の捉え方になんらかの見落としがある。」ことを示唆していると解釈することもできよう。Elman らの提案した、ニューラルネット（正確には、誤差逆向き伝播法 (BP)[3] を用いた、内部表現の探索）による音声研究の手法[4] は、先験的な知識を用いることなく、隠された規則性を抽出することを狙う方法であり、上で述べたような見落としの可能性を再検討しようとする問題意識を具体的な研究に結びつける有力な手段となる。このようにして抽出された規則性を、ネットワークの構造から読みとることは困難な場合が多いが、本報告では、いくつかの場合について、直接的な解釈の方法を与えている。

音声認識への適用は、ニューラルネット自身を研究する立場からも興味深い。音声は、時間的変化が本質的な情報を担っているが、時間的な情報の扱いは、現在のニューラルネット研究におけるホットな話題であり、音声への適用を目指した研究から新しいアイデアが生まれる可能性もある。時間パターンをニューラルネットを用いて処理することを目的に、時間原点の移動に対する不変性、時間軸の非線形な伸縮への耐力、時間間隔と影響度の関連に対する要請、を満たすように導入された状態縮約の概念[5] は、このようなものの一つである。

3 不特定話者の母音特徴要素の抽出

話者に依存しない母音の特徴の抽出と、音声知覚研究に対するニューラルネットワークの適用の可能性を探ることを目的として、一連の実験を行なった。これらの結果については、別に報告[6,7] したが、ここでは、全体像と、新しい構成法や解釈の方法を中心に説明する。ここで入力の表現法として採用された自己相関関数と LPC ケプストラムは、まったく同じ情報を異なった方法で表現しているだけであり、表現法の影響を評価するときに都合がよい。

3.1 自己相関関数入力

自己相関関数は、LPC 分析の基本となるパラメタであり、様々な美しい関係により、LPC 属のパラメタや、周波数領域のパラメタと結びつけられている。これらの関係を用いることにより、母音識別を行なうニューラルネットの直接的な構成方法と、学習により形成されたニューラルネットの機能の解釈の方法が、与えられる。

3.1.1 直接的構成法

自己相関関数を入力としたパーセプトロン型の回路網は、付録に示すように、枝の荷重で決まる仮想的な目標スペクトルと入力のスペクトルの類似度を計算する機構であると解釈することができる。したがって、母音認識を行なわせる場合の自明な解は、母音の各々に対応するユニットを用意し、入力からそれぞれのユニットにつながる枝の重みを、母音標準パターンの最尤スペクトルパラメタとすることである。各々の母音に対応するパターンが、パラメタ空間内で同一の分散の多次元正規分布に従う場合には、このような 2 層構造の回路網により形成される線形識別関数の組合せにより、最適な識別系を構成することができる。これは、パーセプトロンによる識別系に他ならない。

しかし、実際の母音の分布は多次元正規分布とは異なっているため、このような方法では、必要な複雑な識別曲面を構成することはできない。また、母音相互の類似性を考慮していないため、冗長な処理を行なっている可能性もある。異なった母音間の類似性を取り込むとともに、必要な複雑な識別曲面を形成

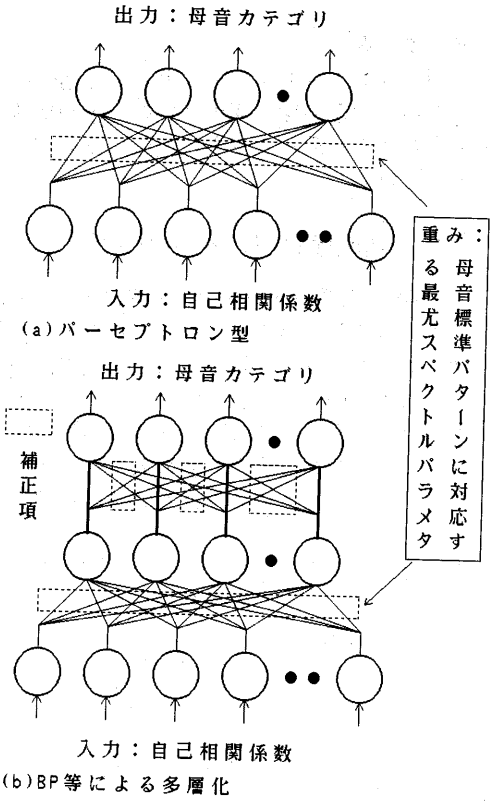

図 1: A trivial vowel discrimination scheme by a perceptron and a multi layer BP-network using autocorrelation inputs.

するためには、多層構造のネットワークを用いる必要がある。

最も直接的に複雑な識別曲面を形成するには、図1に示すように、母音識別用の2層構造のパーセプトロン型回路網の出力を、さらに補正のためのニューラルネットを通すことにより、最終出力を求めればよい。この場合、補正のためのネットワークの部分は、たかだか3層構造までを考えれば十分である。母音間の相互依存性については、このようなネットワーク全体を、BP等の方法を用いて訓練することにより、自動的に取り込まれることが期待できる。

3.1.2 重みからの知識の抽出

前節で説明したような知識をあらかじめ組み込まない場合に、学習によって、ニューラルネットの構造にどのような規則性が抽出されてくるかを調べることは、興味深い問題である。具体的には、隠れ層に含まれるユニットの数をパラメタとして、母音識別や、自己想起の課題について、BPによる学習を行なう。

多数の話者の発声した母音サンプルを用いてこのような実験を行なうと、3個の隠れユニットを用いるだけでも5母音を識別できることが示される。これは、5母音が3次元の空間の中に埋め込まれていることを示唆している。この場合、入力から隠れ層への枝の重みから計算される3個の仮想的な目標スペクトルは、5母音のいずれとも一致せず、自明ではない解が得られていることが分かる。

付録で説明する方法を用いて、入力から隠れ層への枝の重み全体を周波数領域で解釈することにより、各々の周波数成分の貢献度を表わす合成特性が求められる。図2に示した合成応答特性から、母音の識別に必要な情報が、主に4kHz以下の領域に存在していることが推測できる。このことは、教師信号として入力信号そのものを用いる自己想起学習の結果から求められた合成特性と比較することにより、より明らかとなる。このような、自己想起学習により形成されたネットワークについて計算される仮想的な目標スペクトルは、高域に高い貢献度を有しており、5母音のいずれとも、また、識別学習の場合に計算される目標スペクトルのいずれとも異なっている。

このことから、8kHzまでの帯域で見た場合には、スペクトルの大局的形状の主要な変動要因は、音韻性よりもむしろ話者性であることが推測できる。この状況を、図4に模式的に示す。

3.2 LPCケプストラム入力

LPCケプストラムは、音声の機械認識で広く用いられているパラメタであり、線形の伝達特性の影響を容易に分離できる性質を有している。入力としてLPCケプストラムを用いた実験の結果は、自己相関関数を用いた場合と比較すると、誤差の収束が早く、識別率もやや優れている。これは、周波数領域での各成分の強度を対数変換する処理が、音声を識別するための効果的な前処理であることを意

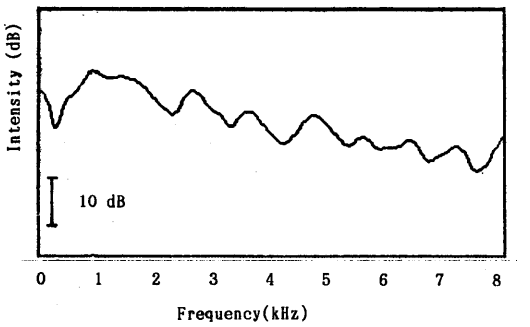

図 2: An example omposite response for vowel discrimination tasks with three hidden units.

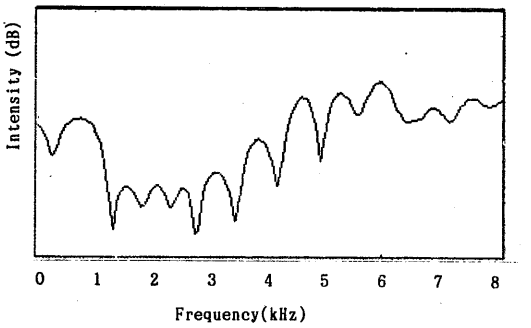

図 3: An example composite response for auto-association tasks with three hidden units.

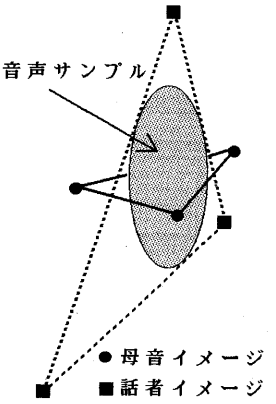

図 4: Spatial relations between virtual target responses for discrimination tasks and auto-association tasks.

味している。

LPC ケプストラムを入力とした場合の枝の重みと物理特性の間には、自己相関関数の場合のように分かりやすい関係は見いだされていない。

3.3 基底膜振動パターン入力

基底膜振動パターンは、実際に人間が利用している情報であることから、どのような特徴が抽出されるかを調べることは興味深い。この表現方法では、基底膜のスチフネスの位置依存性による周波数軸の非線形の伸縮と、高域における周波数分解能の低下により、自己相関関数入力で求められた周波数貢献度に対応した重みづけが自然に形成されている。また、有毛細胞と一次聴神経での変換を模擬するために、各点での振動のパワーを対数変換したものを入力としているが、これは、LPC ケプストラムを用いた実験で得られた知見を取り入れたことにもなっている。

基底膜を 70 個の区間に分割したシミュレーション結果を入力とした場合の識別率は、LPC ケプストラム入力の場合と同程度になることが示されており、基底膜振動パターンの空間周波数成分の低次の 20 個を用いた場合にも同程度の識別率が得られている[7]。これらの結果から、母音識別に必要な特徴は、振動パターンの大局的な形状に含まれていることがわかる。さらに、各々の空間周波数成分の貢献度を詳細に調べることにより、ネットワークには、なだらかな側抑制型の結合と類似の機能が形成されていることが、推測できる。

4 状態縮約による時間的特徴の表現

その明かな欠点にもかかわらず、多くの問題において、予想外の優れた結果を示してきた BP ではあるが、時間的パターンの処理に対する適用例は少ない。これは、BP 以外のモデルにも共通する傾向であり、現在のニューラルネットの弱点でもある。

ここで、大規模なニューラルネットの存在証明である人間における時間感覚の構造について調べることは参考になろう。表 1に、聴覚的な時間知覚の様式について、対応する時

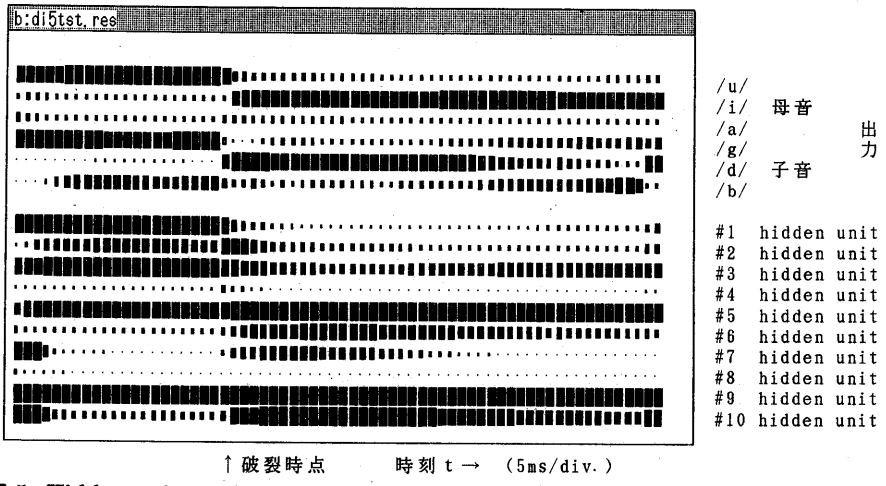

図 5: Hidden units and output units activation pattens for plosive monosyllabic inputs.

時間長	感覚・現象	対応する技術
$10\mu s$	・方向知覚	・信号処理
$1ms$	・音色による識別	
$10ms$	・母音の韻質	
	・逐次感	・パターン処理
	・生起順序の知覚	
$100ms$	・子音の識別	
	・調音結合の補償	
$1s$	・単語知覚	・記号処理
	・リズム感の上限	
	・文	
$10s$	・会話、文章	・知能処理

表 1: Hierarchycal stracture of temporal perception.

間スケールおよび技術領域を併せて示す。このように様々の時間感覚があることは、時間的特徴の表現に、階層的な構造を導入する必要性を示唆する。

時間的な依存関係が線形である場合には、フィードバックを有する線形システムにより、動特性を記述することが出来る。しかし、音声の場合にそのようなモデルが適合するのは、30ms 以下の、ほぼ定常と考えられる時間長の場合だけである。数 100ms 以上にわたる時間長の変化については、言語構造に起因する離れた要素間の依存関係や、調音器官の運動から声道形状を通じて音声波形に到る変換過程に含まれる非線形性のため、単純な線形モデルで記述することは困難である。

ニューラルネットは、非線形の予測機構と見なすことも出来るため[8]、非線形の変換機構に起因する動的特徴を表現する手段として、利用できる可能性がある。そのためには、ある時間範囲にわたる音声の情報を、ニューラルネットに入力するための表現方法が必要となる。このような観点から、音声処理に必要とされる、

- 時間原点の移動に対する不変性、
- 時間軸の非線形の伸縮への耐力、
- 時間間隔と影響度の関連に対する要請、

を満たす性質を持つ前処理として、状態縮約を用いる方法[5] が導入された。この方法の本質は、対数変換した時間軸の上の基底関数を用いて時間的な変化波形を展開するところにある。

他の表現方法については、文献[9] にまとめられている。

この状態縮約の考え方を用いて、有声破裂音 (/b//d//g/ と /a//i//u/ の組み合せからなる 9 種類の音節) の識別実験を行なった。識別系は、前回の報告で説明したように、聴覚の周波数分解能を模擬するフィルタバンクの出力時系列を、 5ms 毎に 15ms の期間の対数パワーに変換した後、各々の系列を、 4 種類 (10ms、20ms、 40ms、 80ms) の異なった時定数の臨界制動二次系の応答を有する IIR フィルタを用いて、 4 種類の時系列に分けた後に、 3 層

構造のニューラルネットに加える構成となっている。学習則としては、BPを用いた。出力としては、9種類の音節を直接9個の出力ユニットに対応させる局所表現と、6個の出力ユニットを、3個づつ母音と子音に対応させる分散表現を試みた。

予備実験の結果では、8個から10個の隠れユニットを用いることにより、十分な識別率が得られている。有声破裂音の識別を子音から母音への過渡部分を用いて学習させたネットワークに、単音節を入力した場合の、隠れユニットと出力ユニットの活性度の時間変化の例を図5に示す。過渡部分のみで学習させたにも関わらず、広い範囲にわたって安定な識別が行なわれている。

5 討論

最小限の先見的知識から出発して、徐々に知識を加えることにより、実際の人間の聴覚機構に類似した特徴抽出機構が組み立てられることを示した。この中で、繰り返し対数変換が見いだされることは、たいへん興味深い。

第一の対数変換は、周波数軸の変換である。この変換により、声道の長さの変化に基づくスペクトル概形の変形は、並行移動に置き換えられ、形状は保存される。第二の対数変換は、振幅軸の変換である。この変換により、口から耳までの伝播特性に基づく変換は、線形分離可能なものとなる。第三の対数変換は、時間軸の変換である。この変換により、発声速度の変化は、軸上での並行移動に置き換えられ、動きのパターンは保存される。むろん、これらの議論は近似的な記述であり、変形の大きさが数倍の範囲において成立するに過ぎないが、興味深い視点を提供することには変わりない。

時間軸の対数変換は、概念的な整理により導入されたものであるが、最近、聴覚の音の高さの2つの属性を説明するために対数螺旋上を聴覚信号が伝播するモデル[10]が、全く別の観点から導入されており、本質的な部分での関連が見つかる可能性もある。

2節で説明した、自己相関関数を入力とするニューラルネットと最尤スペクトルパラメタの関係は、ネットワークの構造によって決まるものであり、学習法則には依存しない。sigmoid型の連続関数を素子の非線形の飽和特性とする、多層パーセプトロン型の回路を用いれば、任意の連続写像を近似することができるが[11]、BPによってその近似が実現できる保証はない。これらの点を考え合わせれば、より直接的にネットワークのパラメタを決定する学習方法の必要性と可能性が見えてくる。

言語構造に起因する時間的・系列的な依存関係をニューラルネットで十分に表現できるかは、未解決の問題である。このような高次の過程の記述のレベルとして、ニューラルネットが適切であるかについては、知能の本質にも絡む問題でもあり、様々な議論が行なわれている。Elmanによる最近の試み[12]は、Jordanの導入したコンテクストユニットを用いる方法[13]により、この領域の構造を明らかにしようとしたものであり、ニューラルネットがここでも利用できる可能性を示している。

このような高次過程のモデルと、これまでに抽出してきた様々な知識を統合して、さらに詳細にわたる規則性を抽出することが、次のフェーズでの検討課題である。

6 まとめ

ニューラルネットを用いた音声の特徴抽出方法について、不特定話者による定常母音の分析の考え方と、子音から母音定常部にいたる過渡部分の処理方法を体系的に説明するとともに、ネットワークの機能を直接的に解釈できる方法について紹介し、従来の学習方法の改良への応用法を示した。

音声認識においては、空間的でかつ時間的に複雑な階層構造を有するパターンを処理することが必要であり、信号処理から知識処理までの多くのレベルの問題を含んでいる。現在のニューラルネットの弱点である時間的特徴、系列の処理と、記号処理についての様々な方法をテストするための課題としても、適切な問題であり、今後多くの試みがなされることを期待したい。

なお、本稿では、特に関連の深い文献のみを挙げたが、より詳しい情報は、著者による

解説記事等[9] を参照されたい。

謝辞： 本報告で引用した実験結果および必要なプログラムの多くは、当研究グループの入野研究主任との共同研究の結果であり、ここに感謝する。ここで紹介した考え方について、日頃討論していただいている、筧リーダおよびグループのメンバーに感謝する。自己相関関数と最尤スペクトルパラメタの関係については、板倉名大教授のコメントにヒントを得たものであり、感謝する。

参考文献

[1] 加藤、：音声知覚における話者への適応性の検討、音響学会誌、 vol.44, No.3, (1988.3).

[2] Furui, S.:"On the Role of Spectral Transition for Speech Perception," *J. Acoust. Soc. Amer.*, 80, 4, pp.1016-1025, 1986.

[3] Rumelhart, D.E. and J.L. McClelland (Eds.): PARALLEL DISTRIBUTED PROCESSING, M.I.T. press, 1986.

[4] Elman, J.L., and Zipser, D.:"Learning the Hidden Structure of Speech," *J. Acoust. Soc. Amer.*, 83, pp.1615-1626, April 1988.

[5] 河原、入野：状態縮約を用いた神経回路網による時空間パターン特徴抽出の検討、音響学会聴覚研究会資料、 H-88-15, (1988.3)

[6] 入野、河原：神経回路網による母音認識、音響学会聴覚研究会資料、 H-82-52, (1987.10).

[7] 入野、河原：基底膜振動を入力とした母音特徴抽出の検討、音講論、 3-P-15,(1988.3).

[8] 李、板倉：重回帰係数の適応化による中国語の声調認識、音講論、 1 2-10, (1988.3).

[9] 河原：神経回路網モデルと音声認識、人工知能学会誌、 vol.3, No.4, (1988.7).

[10] Patterson, R. D.:"Spiral Detection of Periodicity and the Spiral Form of Musical Scales," *Psychol.Music*, 14, pp.44-61 (1986).

[11] 船橋：ニューラル・ネットワークによる連続写像の近似的実現について、信学技報、 MBE88-9, (1988.4)

[12] Elman, J.L.:"Finding Structures in Time," *CRL Techical Report* 8801, April 1988.

[13] Jordan, M. I.:"Attractor Dynamics and Parallelism in a Connectionist Sequential Machine," *Proc. of the Eighth Annual Conf. of the Cognitive Society* , pp.531-546, August 1986.

[14] Itakura, F.:"Minimum Prediction Residual Principle Applied to Speech Recognition," *IEEE Trans.*, ASSP-23, pp.67-72, February 1975.

A 自己相関関数入力時の重みの解析

A.1 時間領域

自己相関係数を入力とする層状のニューラルネットを考えた場合、自己相関係数は時間領域のパラメタであることから、入力層と第2層を結ぶ枝の重み $\{w_{ij}\}_{j=0,\ldots,p}$ も、時間領域のパラメタとなる。ここで、第2層の特定のユニットへの入力 $I(i);(i = 1,\ldots,n)$は、音声波形のサンプル値 $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ から計算される自己相関係数 $\{v_j\}_{j=0,\ldots,p}$ を用いて、次のように表わされる。

$$I(i) = \sum_{j=0}^{p} w_{ij} v_j \tag{1}$$

ここで、

$$v_j = \sum_{t=1}^{N-|j|} x_t x_{t+|j|} / \sum_{t=1}^{N} x_t^2$$

である。ここで記号を $I_h \equiv I(i)$, $w_j \equiv w_{ij}$ のように置き換えることにより、 i 番目のユニットへの全入力は、次のように書ける。

$$I_h = \sum_{j=0}^{p} w_i v_j \tag{2}$$

一方、板倉により導入された最尤スペクトルパラメタ $\{A_i^p\}_{i=-p,\ldots,0,\ldots,p}$ を用いれば、正規化残差は、次のように表わされる。

$$I_d = \sum_{i=-p}^{p} A_i^p v_i \tag{3}$$

ここで、最尤スペクトルパラメタと線形予測係数 $\{\alpha_j^p\}_{i=0,\ldots,p}$ が次の式で結び付けられていることに留意する。

$$A_i^p = \sum_{j=0}^{p-|i|} \alpha_j^p \alpha_{j+|i|}^p, \; i = 0, \pm 1, \ldots, \pm p$$

後で、これらの関連を利用する。

ここで、(2) 式と (3) 式を比較することにより、枝の重みと最尤スペクトルパラメタを次のように対応づけることができる。

$$A_i^p = \begin{cases} w_{|i|}/2 & i \neq 0 \\ w_i & i = 0 \end{cases} \quad \text{for} \quad i = 0, \pm 1, \ldots, \pm p. \tag{4}$$

A.2 スペクトルマッチング

最尤スペクトルパラメタは、音声のスペクトル構造を自己回帰型確率過程とみなした場合の最尤推定から導出されたものであり、スペクトル領域では、次の量を最小化する。

$$d(T,R) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\{\log\frac{T(\omega)}{R(\omega)} + \frac{R(\omega)}{T(\omega)} - 1\}d\omega \tag{5}$$

ここで、T は、自己回帰型の目標スペクトルを、 Rは、入力音声のパワースペクトラムを表わす。

$$T(\omega) = \frac{1}{2\pi}\frac{\sigma_T^2}{\sum_{k=-p}^{p} A_k^p e^{jk\omega}}$$

$$R(\omega) = \sum_{i=-p}^{p} v_k \cos(k\omega).$$

パラメタ σ_T^2 は、予測残差を表わす。このスペクトルマッチング尺度と、前節の正規化残差の間には、次の関係があることから、正規化残差は、目標となるスペクトルと入力音声のパワースペクトラムとの距離を表わすものと考えることが出来る。

$$I_d = C_0 \log d(T,R) + C_1 \tag{6}$$

ここで、(4) 式に注目すると、学習により形成された枝の重みから、ネットワークが目標として設定したスペクトルを、次の式を用いて直接計算できることが分かる。

$$T(\omega) = \frac{C_2}{\sum_{k=-p}^{p} w_{|k|} e^{jk\omega} + w_0} \tag{7}$$

A.3 逆フィルタ

線形予測係数から構成される逆フィルタの出力 y_t は、次の式で与えられる。

$$y_t = \sum_{i=0}^{p} \alpha_i^p x_{t-i} \tag{8}$$

逆フィルタの出力の自己相関W_kは、最尤スペクトルパラメタを用いて、次のように求められる。

$$W_k = \frac{\sum_{t=1}^{N-|k|} y_t y_{t+|k|}}{\sum_{t=1}^{N} y_t^2}$$

$$= \frac{\sum_{t=1}^{N-|k|}\sum_{i=0}^{p}\sum_{j=0}^{p}\alpha_i^p\alpha_j^p x_{t+i}x_{t+j+|k|}}{\sum_{t=1}^{N}\sum_{i=0}^{p}\sum_{j=0}^{p}\alpha_i^p\alpha_j^p x_{t+i}x_{t+j}}$$

$$= \sum_{i=-p}^{p} A_i^p v_{|k|-i} \tag{9}$$

各々の信号のパワースペクトルを大文字で表わせば、これらには、次の関係がある。

$$Y(\omega) = \frac{X(\omega)}{T(\omega)} \tag{10}$$

ここで、(9) 式と (10) 式を比較することにより、 $1/T(\omega)$ を周波数領域での重みとみなすことができる。(7) 式を参照することにより、逆フィルタの応答関数$U(\omega) \equiv 1/T(\omega)$ を枝の重みから、次式により直接計算することが出来る。

$$U(\omega) = C_3(\sum_{k=-p}^{p} w_{|k|} e^{jk\omega} + w_0) \tag{11}$$

ニューラルネットによる学習は、フィルタの安定性を保証するものではないため、$T(\omega)$は、不安定になる場合もある。その場合でも、FIR フィルタである$U(\omega)$は、常に実現可能である。

A.4 合成応答

これまでは、単一のユニットへの入力について考慮してきたが、ネットワーク全体としての性質を表わす指標が必要である。学習により求められた重み$\{w_{ij}\}$ については、正定値である保証はないので、仮想的な逆フィルタ特性 $U(\omega)$ は負の値をとる可能性がある。そこで、次式で定義される合成応答を用いる。

$$G(\omega) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|U_i(\omega)|^2} \tag{12}$$

この合成応答は、各々の周波数成分の貢献度を表わすものとして、解釈できる。