

Title	時系列パターン認識機構の神経回路モデルに関する研究
Author(s)	山内, 康一郎
Citation	大阪大学, 1994, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3075131
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	山内康一郎
博士の専攻分野の名称	博士(工学)
学位記番号	第 11410 号
学位授与年月日	平成 6 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当 基礎工学研究科物理系専攻
学位論文名	時系列パターン認識機構の神経回路モデルに関する研究
論文審査委員	(主査) 教授 福島 邦彦 (副査) 教授 佐藤 俊輔 教授 北橋 忠宏 講師 倉田 耕治

論文内容の要旨

本論文は、人間の音声処理能力を模倣した、時系列パターン（音声）処理機構のモデルについて述べたものである。一般に、時系列パターン認識を行なうには、その時間伸縮に柔軟に対処するメカニズムが必要である。これまでも、この問題を克服するための時系列パターン認識システムが多く提案されてきた。その多くは、入力された時系列パターンから、どの特徴がどのような順序で現れたかという情報のみを検出することによってその時系列パターンを認識する。すなわち、時系列パターンの各特徴の継続時間長に関する情報を無視することによって、その時間伸縮に対処していた。その一方で、モールス信号のように、トーン信号の継続時間長の比の中に情報が含まれるような信号も存在する。このような信号をこれら従来のシステムによって識別することは難しい。しかし人間は、このモールス信号のように継続時間長の比の中に情報が含まれるようなパターンであっても、その時間伸縮に影響を受けずに正しく識別できる。

本論文ではこの人間の能力を模倣するため、時系列パターンを、その時間構造を保持しながら処理・認識する形式を採り、且つその時間伸縮に対処できる神経回路モデルを提案している。このモデルは、伝搬速度可変のタップ付き Delay-Line を入力層として持つネットワーク 3 個で構成されている。これらのネットワークはほぼ同じ構造を持つが、Delay-Line の速度比が異なり、それぞれ高速、中速、低速に設定されている。認識出力は主に中速の Delay-Line を入力層として持つネットワークから得られる。そして、全ての Delay-Line の伝搬速度をその速度比を保ったまま連動制御し、常に中速の Delay-Line 上のパターンが、標準パターンに近くなるようにしてそれを認識する。この速度の制御は、高速および低速の Delay-Line を持つネットワークの出力を比較することで行われる。

計算機シミュレーションでは、各特徴の継続時間の比が異なるような音声信号を模したパターンを用意し、それらが伸縮を受けても正しく区別できることを確認している。また、実際に様々なスピードで発音した単語音声を入力して、それらを正しく認識できることも確認した。

また、このモデルを構築するに当たって、神経回路による時系列パターンの学習手法を新たに提案している。この学習法では、時系列パターンをランダムな順序で提示されるだけで、その中のどの部分が重要でどの部分が一塊のカテゴリかを自動的に発見して記憶することが可能である。

これまでも、人間の音声知覚に関する作業仮説がいくつか提案されていたが、音声信号の時間伸縮に対処しながら、多くの心理現象を説明できるモデルは少なかった。ここで提案した時系列パターン認識モデルは、音声信号をそ

の時間構造を保持したまま処理し、且つその時間伸縮に対処できるという点で、人間の音声知覚のより良い作業仮説へと発展できる可能性がある。

論文審査の結果の要旨

音声信号をはじめとする時系列パターンを正しく認識するシステムを実現するためには、入力信号の時間伸縮に柔軟に対応するメカニズムが不可欠である。この問題の克服を目指してこれまでも種々のシステムが提案されてきたが、その多くは、時系列パターンに含まれる各特徴の継続時間の情報を無視することによって時間軸の伸縮に対処しようとしていた。したがって、モルス信号のように信号の継続時間が重要な情報を担っている信号を、正しく識別することは難しかった。しかし人間は、時間軸に伸縮があっても、信号の相対的な時間長は正しくつかまえることができる。つまり、入力された時系列パターンの時間構造を壊すことなくその時間伸縮に対応する能力を持っている。

そこで申請者は、生物の聴覚神経系に関する生理学的知見や、人間の音声処理に関する心理学的知見などを参考にして、従来型システムの問題点を克服する時系列パターン認識機構の神経回路モデルを提案し、その能力を計算機シミュレーションで確認した。このモデルは、ネオコグニロン（視覚パターン認識機構の神経回路モデル）型の神経回路の前端に、可変速度の遅延線を導入したような構造を持っており、入力された時空間パターンを遅延線によって空間パターンに展開して認識処理を行なっている。この遅延線を流れる信号の速度を、入力信号の速度に順応して変化させることによって、時間軸の伸縮の影響を吸収しようとするものである。

ところで、神経回路の学習法に対する研究は多いが、その大部分は空間パターンを対象にしたものであって時空間パターン覚えさせることのできるよい学習法はまだ見つかっていなかった。従来の学習法によって時系列パターンを学習させようとする時、時系列パターンのどの部分を学習すべきかを明示的に示す必要があったり、あるいは、膨大な数の細胞数を必要とするという問題が生ずることが多かった。そこで申請者は、教師なし学習の新しい手法を提唱し、上記の神経回路モデルの学習に用いた。この学習法を用いると神経回路は、呈示された時系列パターンの中から重要な特徴を自動的に発見し、記憶していくようになる。

申請者はさらに、この時系列パターン認識モデルを組み込んだ音声認識システムを構成した。このシステムは、生物の聴覚神経系の基底膜から大脳聴覚野にいたる経路で行なわれる特徴抽出機能を模擬した特徴抽出部と、上記の時系列パターン認識部とからなる。本モデルが音声パターンの伸縮やスペクトル形状の変形に対する汎化能力をもつことを、実際の音声信号を入力とする計算機シミュレーションによって実証した。例えば、男性が普通のスピードで発音した単語を学習させただけで、女性が様々なスピードで発音した単語も正しく認識できるようになる。まだ少数のサンプルを用いた実験ではあるが、発声速度や話者の違いに影響されることなく音声信号を正しく認識するシステムを実現できることを示唆している。

以上のように、本論文は音声信号を始めとする時空間パターンの認識システムの実現の糸口を示すものであり、学位論文として価値のあるものと認める。