

Multi-clustering Network for Speech Recognition on Parallel Processor

(並列プロセッサによる音声認識のための多層クラスタリングネットワークに関する研究)

学位論文内容の要旨

マルチメディア時代の到来と共に、その重要な要素のひとつとして、音声情報処理に関する研究がますます盛んに行われている。本論文では、全く新しい音声認識システムを提案する。音声認識の分野では認識率と処理時間が特に重要であり、本論文でもこの二点について特に綿密に議論を行う。

本論文の音声認識システムにおいては、音声認識のための中核となる技術として自己組織化クラスタリング手法が用いられている。このクラスタリング手法は、音声スペクトルの分布から、音声信号を適切なクラスタ集合に分類するものである。忘却係数を用いた適応化の技術を導入することにより、時間と共にクラスタ分布が変化するようなデータに対しても、クラスタ集合を適切に再構築でき、分布の変化に追従することが可能である。各クラスタは神経回路網におけるノードとして表現され、ノード融合の規則を導入することにより、ノード総数は最少に保たれる。本クラスタリング手法は Kohonen のそれより速く収束する特徴を持つ。

一方、時間遅れニューラルネットワーク (Time-Delay Neural Network: TDNN) は音声認識研究における強力な手法の一つである。TDNN の認識性能は優れているが、その三層構造に起因する処理速度の遅さには問題が残っている。本論文における最初の実験システムとして、従来の TDNN よりも高い処理速度を持つ、二層構造の TDNN を提案する。自己組織化クラスタリングにより生成された時間遅れ類似度ベクトルテーブルが、二層単純パーセプトロンに入力され、最終的な出力を得る。提案する二層構造 TDNN は、同一のテストデータを用いた従来の (三層)TDNN に比べて約 10 倍の処理速度を持つが、認識性能はごくわずかに低下する。

処理速度と認識率の双方を向上させるため、本論文では新しく多層クラスタリングネットワークを提案する。多層クラスタリングネットワークは二つの異なったクラスタリング層を持ち、それぞれのクラスタリングアルゴリズムは効率的に並列化することが可能である。第一層では並列化自己組織化クラスタリングが用いられ、第二層では新しく提案される制限付クラスタリングネットワークが用いられる。

並列マルチプロセッサシステムに第一層の自己組織化クラスタリングを導入する際に重要となるのは、各プロセッサ間の通信コストをなくするために、いかにして各ノードを独立にするかという問題である。提案する手法では、クラスタリングされるべき全ての入力ベクトルが、それぞれのノードに一度に、独立に入力される。各入力ベクトルに対する類似度を各ノードに内において比較し、類似度があらかじめ決められたしきい値以上となる入力ベクトルを、類似度値の高い順番に最大 N 個選択し、そのノードに属するものとする。このように全てのノードは他のノードとは独立に内部パラメータを更新する。

第二層である制限付クラスタリング層においては、一つのノードに一つのカテゴリラベルを割り当てる代わりに、いくつかのノード (近傍ノードグループ) に同一のラベルを割り

当てる。最適な並列処理を行うため、各近傍ノードグループにおける処理は、それぞれ別個のプロセッサに割り当てられる。「制限付クラスタリング」の名称は、第二層におけるクラスタ形成が、選択された近傍ノードグループの中のみで行われることに由来する。特定の入力に対してどの近傍ノードグループが選択されるべきかを決定するために、規範ベクトルが導入される。これにより、制限付クラスタリング層は、バックプロパゲーションのような学習をすることなしに、学習済みパーセプトロンのように動作する。加えて、この制限付クラスタリング層は従来のパーセプトロンよりも効率良く並列化できる。

複数の並列プロセッサを使用することにより、本論文で提案するシステムは従来のネットワークと比較して最小の計算時間で実現することができる。例えば、15個のプロセッサを持つ並列計算システムでは、単一プロセッサのシステムに比較して学習時間を12における認識率は、日本語の有声破裂子音で97.50が示され、マルチプロセッサを用いることにより処理速度の向上が確認された。

提案された多層クラスタリングネットワークは、将来的に連続音声に対しても適用できると考えられる。異なる音韻セットのための多層クラスタリングネットワークを用いた、大語彙音声単語認識が将来の課題である。本ネットワークは並列処理が有効であることが示されたことから、実時間にせまる認識が可能であると考えられる。

本論文は6章で構成されている。第1章では本研究の概要を述べ、音声認識における従来の研究を概説し、本研究との関連を述べる。

第2章では、本論文で議論される音声認識システムの中核となる自己組織化クラスタリング手法について述べる。

第3章では、音声認識に適した、二層構造の時間遅れニューラルネットワーク(TDNN)を新しく提案する。本論文で提案する二層構造のTDNNは、従来のTDNNよりも処理速度を向上するものである。いくつかの実験により、提案する二層TDNNの有効性が示されている。

本論文の音声認識システムでは自己組織化クラスタリング手法が用いられているため、高速な認識を行うには高速なクラスタリングが必要である。このことから第4章では、自己組織化クラスタリングの並列処理化について述べる。この並列化の効果は実験によって確認される。第一層に自己組織化クラスタリング層、第二層に単純パーセプトロンを持つ、ハイブリッド型二層ネットワークの音声認識システムが構築され、実験により高い認識率と高速な処理が実現されることを示される。これにより本実験での認識システム、とりわけ並列化された自己組織化クラスタリング手法の有効性が示される。

第5章では、音声認識のための全く新しい多層クラスタリングシステムを提案する。このシステムは並列化手法を巧みに導入した二つのクラスタリング手法を用いている。実験により、提案する新しい多層クラスタリングシステムは、より高い認識率を達成し、より高速な動作をすることが示される。これに加えて、本論文で提案する多層クラスタリングネットワークシステムの将来の方向性についても議論する。音素認識システムから連続音声認識システムへの拡張の可能性と、並列処理による実時間にせまる認識処理の可能性について論ずる。

最後に第6章では、本論文で提案されたいくつかの新しい音声認識システムについて、従来システムと比較しながら詳細に議論し、提案する多層クラスタリングシステムが最も優れたものの一つであることを結論する。

学位論文審査の要旨

主 査 教 授 梶 内 香 次
副 査 教 授 新 保 勝
副 査 教 授 青 木 由 直
副 査 助 教 授 宮 永 喜 一

学 位 論 文 題 名

Multi-clustering Network for Speech Recognition on Parallel Processor

(並列プロセッサによる音声認識のための多層クラスタリングネットワークに関する研究)

実用的な音声認識システムの実現を目指し、高速かつ高精度な音声認識手法が求められている。本論文は、このための有力な手法の一つとして知られている自己組織化クラスタリング手法を中核とし、クラスタリングネットワークの構成法、ならびに高速化のための並列計算機へのインプリメントなどの新しい手法を提案し、実験を行なって手法の有効性を確認したもので、その主要な成果は以下に要約される。

(1) 自己組織化クラスタリング層と単純パーセプトロンからなる2層構造の TDNN (TimeDelay Neural Network) による認識手法を提案し、従来の3層構造 TDNN と比較し、認識性能はわずかに低下するものの、約10倍の処理速度が得られることを確認した。

(2) 上の手法を発展させ、認識性能を向上させる手法として、第2層にもクラスタリングネットワークを用いる多層クラスタリングネットワークを提案した。さらに、この第2層のネットワーク構成法として、制限付クラスタリングネットワークを提案した。

(3) 上記、多層クラスタリングネットワークを並列計算機上に実現するための並列化法の検討を行ない、まず第1層の自己組織化クラスタリングネットワークにおいては、各ノードが他のノードとは独立に内部パラメータを更新することにより、並列各プロセッサ間の通信コストを最小にする方法を提案した。また、第2層の制限付クラスタリングネットワークにおいては、各近傍ノードグループにおける処理は、それぞれ別個のプロセッサに割り当てられるようにして最適な並列処理を可能にする方法を提案した。

(4) これを実際に並列計算機上にインプリメントし、実音声データを用いて認識実験を行なった。その結果、例えばプロセッサを15個とした場合、単一プロセッサのシステムと比較して学習時間を12%に削減でき、日本語有声破裂子音で97.50%、無声摩擦子音で98.75%の認識率が得られ、本手法の有効性が確認された。

これを要するに、著者は、並列処理に適した新しい多層クラスタリングネットワーク構成による音声認識手法を提案して実際に並列計算機上に実験システムを構築し、実音声データによる認識実験を行なってその有効性を確認したもので、音声情報処理工学ならびに信号処理工学の発展に寄与するところ大である。

よって著者は、北海道大学博士（工学）の学位を授与される資格あるものと認める。