

平成 21 年 5 月 18 日現在

研究種目：若手研究（B）
 研究期間：2007～2008
 課題番号：19700181
 研究課題名（和文） 相補的な識別器の生成とその統合に基づくパターン認識に関する研究
 研究課題名（英文） A study on a pattern recognition system based on the combination of complementary classifiers
 研究代表者
 小川 哲司（OGAWA, Tetsuji）
 早稲田大学・高等研究所・助教
 研究者番号：70386598

研究成果の概要：

本研究では、パターン認識の性能を向上させるために、複数の識別器を統合する方式について検討を行った。識別器の生成においては、提案法は識別器の誤り傾向が異なることを理論的に保証し（相補性を有する）、少量の識別器で高い性能を与える（効率性を有する）という特徴を有する。また、生成した複数の識別器の統合においては、提案法は識別器各々が潜在的に有する特性の影響を受けにくい（頑健性を有する）という特徴を有する。本提案方式を音声認識に適用し、有効性を確認した。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2007 年度	1,300,000	0	1,300,000
2008 年度	1,100,000	330,000	1,430,000
年度			
年度			
年度			
総計	2,400,000	330,000	2,730,000

研究分野：総合領域

科研費の分科・細目：情報学・知能情報処理・知能ロボティクス

キーワード：パターン認識，音声認識

1. 研究開始当初の背景

音声認識，顔・ジェスチャ認識，文字認識などでは、パターン認識技術が用いられている。高精度な認識を行うために、確率モデルや探索アルゴリズムの精密化などが試みられており、これまでに性能の向上が見られてきた。しかし近年、単一の識別器を高精度化しても大幅な性能の向上は得られなくなってきた。一方、単一の識別器のみを用いるのではなく、複数の識別器から得られる情報を統合することで、認識性能を向上させるという枠組みが検討され始めている。申請者も、

平成 17 年度から 18 年度の科研費若手研究（B）の助成を受け確率モデルの精密化に関して検討を行った際、提案したモデルと従来モデルの誤り傾向が異なり、両モデルの情報を単純に統合することで認識性能の大幅な向上を確認した。この経験からも、識別器の統合について理論・実装両面から更なる検討が必要であると確信するに至った。

複数の識別器を独立に構築してもそれらの誤りの傾向が類似していれば、情報を統合しても誤りは削減されない可能性が高い。識別器を統合するにあたっては、各々の識別性

能が高く、かつ誤り傾向が異なる（以降、この性質を「相補性」と呼ぶ）識別器を構築することが、識別器の統合を効率良くかつ高精度に行う上で、極めて重要である。また、相補的な識別器が構築できたとしても、情報の統合の方法によって誤りの削減の度合いが異なる。したがって、識別器の統合に基づいて性能を改善するアプローチにおいては、以下の2点が課題となる。

- (1) 相補性を有する識別器の生成方式
- (2) 生成された相補的な識別器から得られた情報の統合方式

2. 研究の目的

複数の識別器を統合することでパターン認識の性能を向上させるために、1で述べた識別器統合における2つの課題（相補性を有する識別器の生成法とその統合法）を解決することを旨とする。

具体的には、(1)従来のように、識別器の構成時にブースティングを利用するのではなく、識別器の前段に置かれる特徴変換器を構成する際にブースティングを導入することで、相補的な識別器を生成することを試みる。ここでは、生成された複数の識別器における相補性の有無を実験的に明らかにする。その上で、(2)生成された複数の識別器が各識別クラスに対して与える尤度のパターンを特徴量としたパターン認識（識別クラスに対して各々の識別器が与える尤度の値を結合して特徴ベクトルとする）により最終的なクラスを決定することで、識別器の統合を行う。（以降では、この特徴ベクトルが作る空間を「複数識別器により与えられる尤度パターン特徴空間」と表現する。）ここでは、識別器の線形結合に基づく統合法（従来法）と識別性能を比較することで、提案法の有効性を明らかにする。

以上により得られた相補的な識別器の統合法を、(3)音声認識システムに適用し評価を行うことで、実際のアプリケーションに対する提案法の有効性を明らかにする。

3. 研究の方法

- (1) 特徴変換にブースティングを利用した相補的な識別器の生成に関する検討

特徴変換器の生成にブースティングを適用することで、相補的な（誤り傾向が異なる）複数の識別器を生成する方式の定式化および実装を行った。ここでは、異分散線形判別分析（Heteroscedastic linear discriminant analysis; HLDA）に基づく特徴変換器を前段に有するパターン認識システムを考えた。誤りが生じやすいサンプルに重み付けを行った上で、変換後のサンプルの尤度が最大となるように HLDA の変換行列を推定することで、新たな特徴変換器を生成した。この際重

要となるのは「誤り易さ」の定量化であるが、本研究では誤り易さを表す汎用尺度であるベイズリスクを適用した。この枠組みは、相補性を理論的に保証するものであるが、生成された識別器の識別結果や誤り傾向を調査することで、実験的にもその検証を行った。

- (2) 識別器の統合手法に関する検討

全ての識別クラス（例えば単語）に対して統合対象となる全ての識別器が与える尤度のパターンを特徴量として階層的なパターン認識を行うことで識別器を統合する方式を実装した。このとき用いるパターン認識の枠組みとして、サポートベクターマシン（Support vector machine; SVM）や混合正規分布（Gaussian mixture model; GMM）などを適用して評価を行うことで、最適な尤度分布のモデル化手法について調査を行う。また、提案する統合法と、各々の識別器の尤度を線形結合した値を用いて最終的なクラスを決定する従来法との比較を行った。

- (3) 音声認識システムへの適用

上記(1),(2)を音声認識システムに適用し評価を行うことで、実際のアプリケーションにおける提案法の有効性を調査した。

4. 研究成果

<研究の主な成果>

- (1) 相補的な識別器の生成

相補的な（誤り傾向が異なる）複数の識別器を生成する枠組みとして、特徴変換器の生成にブースティングを適用した、重み付き異分散線形判別分析（Weighted HLDA; WHLDA）の定式化および実装を行った。WHLDA は識別器の相補性を理論的に保証するが、識別器の誤り傾向を調査することで実験的にもその検証を行った。

孤立単語音声認識実験により提案方式の性能を評価したところ、従来の識別器と WHLDA により生成された識別器はともに95%程度の高い認識率を与え、かつ誤り傾向が25%程度異なることがわかった。この結果は、提案法が個々の性能が高くかつ相補的な識別器を生成可能な、効率的な方式であることを示したものである。

- (2) 効果的な識別器統合方式

基本方式開発

全ての識別クラスに対して統合対象となる全ての識別器が与える尤度のパターンを特徴量として、静的な識別器（SVM や GMM）を用いたパターン認識を行うことで識別器を統合する方式の実装を行った。このとき、特徴ベクトルの次元数は、「識別クラス数×統合対象の識別器の数」となるので、音声認識のように識別クラス数が多いタスクでは、

GMM に基づいた枠組みではシステムの頑健性が低下し易く、良好な性能が得られなかった。そこで、本研究では識別器統合部におけるパターン認識器として SVM を採用した。ここでは、各々の識別器の尤度の線形結合値を用いて最終的なクラスを決定する従来法との比較を行った。

(1)と同様の孤立単語音声認識実験により評価を行ったところ、識別器の統合を行わない場合の認識誤りの 74% を削減し、また、従来の識別器の尤度と WHLDA により生成された識別器の尤度の線形結合値を用いて最終的なクラスを決定する方式 (late integration) の認識誤りを 71%、従来の特徴量と WHLDA で生成される特徴量を結合した特徴量を用いてパターン認識を行う方式 (early integration) の認識誤りを 72% 削減しており、提案法の有効性が示されたと言える。

尤度パターン特徴ベクトルの次元数削減

提案する識別器統合方式を音声認識に適用する場合、生成された複数の識別器 (hidden Markov model; HMM) が各識別クラス (単語) に与える尤度のパターンを特徴量として静的なパターン認識を行うことで、識別器を統合する。しかし、この枠組みでは特徴ベクトルの次元数が語彙数と一致するため、大語彙を扱うタスクでは特徴ベクトルの次元数が膨大になる。そこで、認識語彙のうち識別に寄与する単語クラスのみを選択し、得られたクラスに対する尤度のみを用いて特徴ベクトルを構成することで、尤度特徴ベクトルの次元数を削減する方式を開発した。

提案する次元圧縮法が認識性能に与える影響を調査するため、次元圧縮を行った尤度特徴空間上における SVM を用いたパターン認識を、単語音声認識システムのリスコアリングに適用した。その結果、4.6% に圧縮したベクトルを特徴量とした場合においても、次元圧縮を行わないシステムの認識性能を劣化させることなく、かつ従来の HMM を上回る性能を与えることがわかった。

合成音声の利用

識別器統合部に用いるパターン認識システムでは学習データとして単語音声を必要とするため、語彙数の増大に伴い学習データ不足の問題が生じる。この問題を解決するため、HMM 音声合成方式により生成した合成音声を学習データとして識別器統合部のパターン認識システムを構築することを試みた。識別器統合部では HMM が与える尤度パターンの空間上で識別を行うため、HMM 合成音声は本方式と整合が良い。この方法を用いることで、実際の単語発話を大量に収集することなく、提案方式を用いることが可能となった。

<研究成果の国内外における位置づけ>

ブースティングに基づく識別器の統合は近年盛んに行われており性能の改善が報告されているが、それら全ては識別器を構成するために (具体的には、確率モデルのパラメータを推定するために) ブースティングを用いている。その際、基の識別器で誤ったサンプルに対して重み付けを行っているため、識別器の実現上重視するサンプルと軽視するサンプルを作ることになる。そのため、新たに生成された識別器単体では、基の識別器と比較して誤り傾向は異なるものの、識別性能自体は低くなる傾向がある。このように、通常のブースティングでは生成される識別器単体の性能が低いため、統合するにあたり多くの識別器を生成する必要がある。それに対し、提案法では識別器の学習の前段である特徴量の変換器をブースティングに基づき生成した。このとき、提案法では特徴量の抽出法に差を作るだけであり、識別器の構成において軽視するサンプルはなく、新たに生成される相補的な識別器の性能も良い。このように、誤り傾向が異なり (相補性の理論的保証) かつ性能の良い識別器を統合するため、性能が向上するのは当然のこと、通常のブースティングと比較して少量の識別器の組み合わせで高い性能が得られることが期待できる (高い効率性)。

また、統合の枠組みとして、識別クラスに対して統合対象の識別器が与える尤度の分布を特徴量としたパターン認識を行う提案法は、識別器の特性 (全てのクラスに対する尤度の分布) そのものを識別クラスの情報として学習する枠組みであるので、識別器の特性 (例えば、過学習が生じやすい、クラスに対する誤り傾向に偏りがある、など) の影響に対して非常に頑健な性能を与えることが期待できる (高い頑健性)。一方、識別器の線形結合による統合 (従来法) では、統合対象となる識別器のうちの一つが、ある特定のクラスに不当に高い尤度を与えてしまうという特性を有していた場合、その特性が統合後の識別性能にも悪影響を及ぼす可能性がある。しかし、尤度の分布を学習する提案法においては、例えば統合対象の識別器が不適切な特性を有していたとしても、その不適切な特性も含めて識別クラスの情報として学習するため、性能の劣化につながりにくい。

<今後の展望>

提案方式は、音声認識のみならず、顔・ジェスチャ認識、文字認識などのようにタスクを限定せず、あらゆるパターン認識において適用可能である。また、上記で述べた、相補性、効率性、頑健性は、パターン認識に基づくアプリケーションを実用化するに当たり重要な利点となる。相補性は、識別器統合後

の性能向上の必要条件である。効率性は、計算量の削減に寄与する。頑健性は、適用するタスクやアプリケーション、環境の変動などに対してロバストに高い性能を与えるという性質に他ならない。このような、相補性、効率性、頑健性を兼ね備えたパターン認識の枠組みはこれまでに例を見ない極めて画期的なものであり、近年、性能の向上に陰りが見え始めたパターン認識技術の性能を抜本的に改善することを期待できるとともに、パターン認識技術の適用可能分野を格段に広げることが予想される。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

小川哲司, 小林哲則, “HMMの尤度パターンを利用したリスコアリングにおける次元圧縮法の検討,” 電子情報通信学会技術研究報告, vol.108, no.142, pp.73-78, July 2008. (査読無し)

加藤健一, 小川哲司, 小林哲則, “HMMにおける尤度パターンの非対称性を利用した音声認識,” 日本音響学会講演論文集, pp.209-211, March 2008. (査読無し)

加藤健一, 小川哲司, 小林哲則, “テンプレート群からの確率的距離を用いた階層的音声認識の検討,” 日本音響学会講演論文集, pp.147-150, Sept. 2007. (査読無し)

加藤健一, 小川哲司, 小林哲則, “マクロな時間構造を持つテンプレート群からの確率的距離を用いた階層型音声認識,” 電子情報通信学会技術研究報告, vol.107, no.116, pp.79-84, June 2007. (査読無し)

6. 研究組織

(1)研究代表者

小川 哲司 (OGAWA TETSUJI)

早稲田大学・高等研究所・助教

研究者番号: 70386598