氏　　　　名	樺(かば)　澤(さわ)　康(やす)　夫(お)
授 与 学 位	工　学　博　士
学位授与年月日	昭和50年3月25日
学位授与の根拠法規	学位規則第5条第1項
研究科，専攻の名称	東北大学大学院工学研究科 （博士課程）電気及通信工学専攻
学位論文題目	**認識系における学習理論に関する研究**
指 導 教 官	東北大学教授　大泉　充郎
論文審査委員	東北大学教授　大泉　充郎　　東北大学教授　本多　波雄 東北大学教授　木村　正行　　東北大学教授　野口　正一

論　文　内　容　要　旨

第1章　序　　論

パターン認識の問題は，工学的には電子計算機の発達に伴い，その入力装置として文字認識装置や音声認識装置の必要性が要求され多くの研究と実用化が行なわれている。しかしながら初期における予想に反して認識の能力は未だ十分実用に耐えないのが現状である。そこで，この立場に立って行われた研究が本研究であり，特に認識系の能力の改善に関する研究を次の二つの観点から行なったものである。その一つは，コンテキストを用いる認識系の導入であり，もう一つは認識系に学習能力を付与することである。

第2章　パターン認識系の構成法

与えられたパターンの認識に際し，そのパターンだけの情報に基づいて，そのパターンの帰属

カテゴリを決定する方式の認識系（コンテキストを用いない認識系）と，そのパターンの近辺のパターンの情報をも利用して帰属カテゴリを決定する方式の認識系（コンテキストを用いる認識系）とが考えられる。パターンの系列には，一般にそのパターンの間に依存関係があるので，コンテキストを用いることによって識別率の改善が期待される。

2.2では，コンテキストを用いない認識系の構成法について統計的決定論の立場で概述し，2.3においてはコンテキストを用いる認識系の構成法について論ずる。

パターン系列の依存性を考慮した識別規則として，認識対象のパターンX_tの前後にわたるパターン列$\{X_{t-\ell},\cdots\cdots, X_t,\cdots\cdots, X_{t+m}\}$の観測に基づく識別規則を$(\ell, m)$形識別規則として定式化し，規則の構成法と能力に関する基礎的性質について検討する。ℓおよびmに対して，(ℓ, m)形識別規則による最小誤識別率は非増加なること，またパターン系列が統計的に独立ならば，コンテキストを用いても識別率は改善されないことが与えられる。

また，時刻tにおいて既に観測された全パターンの情報を用いる識別規則を(t, m)形識別規則として定式化する。この規則は，許容される(ℓ, m)形識別規則の中で最適なものであるが，時間の経過と共に観測パターン数が増加するので，一般の場合には適用困難であるが，本研究では有限の記憶量で実現可能のための十分条件を与え，その物理的意味を明確に示している。また(t, m)形識別規則の繰返形アルゴリズムによる構成法を与え，ブロック図を示す。

第3章　パターン認識系の評価

第2章で与えたパターン認識系について，その能力の評価を平均誤識別率によって行なう。パターンとカテゴリの系列に対する条件の明確なモデルの下で，解析的手法及びシミュレーションによって，(ℓ, m)形及び(t, m)形識別規則の最小平均誤識別率を求め，グラフ及び表に示す。またカテゴリ間の依存関係の強い場合について，パターンの分布間のdivergenceを観測パターン数の関数として与え，パターンの相関の影響について論ずる。とくにパターン間の依存性が強くても，観測パターン数を増しても識別率の改善されない例を示し，評価の重要性について注意を喚起する。以下，主要な評価結果を示しておく。

1°　$\ell+m=N$一定ならば，$(0, N)$形あるいは$(N, 0)$形よりも，$(N/2, N/2)$形識別規則の方が識別率が高い。

2°　コンテキストを用いない識別規則による最小平均誤識別率が15％程度以上のときには，ℓおよびmを大きくしてもその割には識別率の改善は少ない。逆に15％程度以下のときには，ℓおよびmを大きくすることの効果が顕著である。これは，コンテキスト自体があいまいな時にはコンテキストを用いても識別率が改善されないことを意味する。

第４章　教師つき学習系の構成

第２章において与えた最適な識別規則を実現するためには，パターン・カテゴリ対系列の確率的構造に関する情報を必要とする。第４章および第５章においては，これらの確率的構造の学習法について論ずる。

学習系に入力される学習パターンの帰属カテゴリも同時に与えられる場合の学習を教師つき学習法，与えられない場合の学習を教師なし学習法と呼び，それぞれ第４章および第５章で議論する。

学習パターン列の統計的独立性を仮定した教師つき学習系の構成理論については，既に十分に論ぜられほぼ確立されているが，独立性を仮定しない場合の構成理論は確立されていない。

コンテキストを用いる識別系では，パターン系列の依存性を考慮しているのであるから，このような系の学習においては，学習パターン列の独立性を仮定することはできない。

従来，学習理論の構成に確率近似法が有力な道具であったが，学習パターン列の独立性の仮定を要するため，独立性を仮定しないときには，直ちに適用することは論理的には正しくない。たとえ形式的に適用しても，収束性の保証に問題が残る。このような例として，4.2において独立性の仮定の下で与えられた学習アルゴリズムを依存性のある学習パターン列に対して適用すると誤った結果が与えられることがあることを示す。

等相関係列における共分散行列の学習において，Basuらは，独立系列に対して共分散行列Σの最尤推定量となる推定量SnがΣではなくて$\Sigma - R$の最尤推定量であることを示している。ここでRは異なるパターンX_i，X_jの間の共分散を表わす。本論文においては，等相関系列の性質を詳しく調べ，Snが$\Sigma - R$の一致推定量であることを示す。即ち，$n \to \infty$のとき$S_n \to \Sigma - R$である。これらは，SnがΣの推定量として不適当であることを示し，従来の認識系の構成法では正しい結果が得られないことを示している。

4.3においては，学習パターン列の独立性を仮定しないパラメトリックな教師つき学習系の構成法とアルゴリズムの収束定理を与える。学習対象は，条件つき確率密度関数$f(X_t \mid X_{t-\ell}, \cdots\cdots, X_{t-Q}, W_{t+M}, \cdots\cdots, W_{t-Q-L})$で正規性を仮定する。ここで，$W_t$はパターン$X_t$の帰属カテゴリを表わし，Q，L，Mはある非負整数である。

アルゴリズムは，条件つき共分散行列および条件つき平均を得るためのベクトルおよび行列の学習を行ない，エルゴード性の仮定の下に真の解への収束が保証される。

また付録Ⅰに，計算機シミュレーションの結果の一例を示す。

4.4においては，学習パターン列の独立性を仮定しないノンパラメトリックな教師つき学習系の構成について述べる。4.4.1では事後確率$P(W_t \mid Y_t)$を学習対象とする。(ℓ, m)ーパターン列$Y_t = \{X_{t-\ell}, \cdots\cdots, X_{t+m}\}$の上の$\Phi$ー関数系を与え，これらの線形結合により事後確率を最小二乗の意味の最良近似解を学習するアルゴリズムを与え，Φー関数系の一次独立性を仮定

せずに，パターン・カテゴリ対系列のエルゴード性の下で最適解への収束を保証する。さらに逐次形アルゴリズムに変形し，確率近似法との関係について述べる。また付録Ⅱにおいて，Φ一関数系として線形関数及び3次関数系を与えたときの解の性質と平均誤識別率及び学習アルゴリズムの計算機シミュレーションについて示す。またた結果として，線形の（ℓ，m）形識別規則の有効性を示唆した。

4.4.2では，条件つき確率密度関数 $f(X_t \mid X_{t-\ell},\cdots\cdots, X_{t-Q}, W_{t-Q-L})$ を学習対象とする。パターン列 $\{X_t,\cdots\cdots, X_{t-Q}\}$ の上のΦ一関数系を与え，最小二乗に類似の評価関数の最適化に基づき学習アルゴリズムが構成され，エルゴード性の下に最適解への収束が保証される。またΦ一関数系の一次独立性を仮定していない。

以上の学習アルゴリズムはいずれもエルゴード性の仮定の下で収束が保証され，実行も容易である。また各学習系のブロック図を示した。

第5章　教師なし学習系の構成

学習系に入力される学習パターンの帰属カテゴリが与えられていない場合，パターン空間の確率的構造を学習する教師なし学習系の構成法について論ずる。このような学習問題では大別して2通りのアプローチがあり，一つはパターン空間の分割に関する損失汎関数を与えて，平均損失を最小とする分割を学習することによって直接的に識別系を構成するものであり，一つは各カテゴリに帰属するパターンの分布がある与えられた分布関数の族の元であるという知識の下に，パラメトリックに確率的構造を学習するものである。第一のアプローチでは損失汎関数の選択に大きな問題が残されている。また第二のアプローチでは，Yakowity やRobbins らのminimum distauce 法が代表的であるが，いずれも与えられた学習アルゴリズムの実行は非常に困難である。最適解の存在は示されているが実際的な探査のアルゴリズムが構成困難なことに起因している。この章では，第二のアプローチに従って教師なし学習系の構成法について述べる。

5.2では，このようなアプローチでは重要な概念である有限mixture分布およびその Iden－tifiability について述べる。教師なし学習の問題が有限mixture 分布の分解法に帰着され，分解の一意性を保証するものがIdentifiableの条件である。

5.3では，確率近似法による教師なし学習系の構成法を示す。まずパターン空間上の分布関数の間の距離として平均二乗距離を定義し，この距離に基づく2種のminimum distance 法を構成する。そして距離の最小化に対して確率近似法を適用して2種の学習アルゴリズムが与えられる。5.3.2では確率近似法とその収束定理を示し，さらに拘束条件に対する取扱法について述べ

る。これらは上のアルゴリズムの構成の有力手段である。

5.3.4では，アルゴリズムの有効性を示すため多次元正規分布関数の族に対する応用例を示す。さらに付録Ⅲに一次元正規分布関数の族に対する計算機シミュレーションの結果の一例を示す。

与えられた学習法は，正規分布の他，指数分布，コーシ分布，ガンマ分布，ポワソン分布及び負の二項分布などに対して適用可能であり，実行容易な学習アルゴリズムが与えられる。

なお，この章では学習パターン列の統計的独立性を仮定している。エルゴード性の仮定の下での確率近似法の収束の保証の問題が解決されれば，拡張可能である。

第6章　結　　論

本論文の結論を述べる。

以　上

審査結果の要旨

計算機の急速な進歩に伴い，情報処理方式の研究は著しい進歩をとげている。しかし，人間が行っているような知的活動，例えばパターン認識等の基本的な情報処理方式の研究は，従来から多くの人々の努力にもかゝわらず，十分な成果を挙げていない。このためこの方面の研究を強力に推進することは，今後の情報処理工学発展のためのきわめて緊急な課題である。

著者はこの立場に立って，パターン認識に関する基礎的研究を行って来た。本論文はその成果をまとたもので全編6章より成る。

第1章は序論である。第2章ではパターン認識系の構成理論を述べている。従来の研究では認識機械に与えられる入力は，簡単のためそれぞれ独立なパターン系列として研究されてきたが，実際には，それぞれのパターン間に依存性がある方が一般的である。この立場より著者はパターン系列の依存関係に関する情報を最大限に活用する新しい認識系の統一的な設計法を与え，認識系構成の上で有力な新しい知見を加えている。又この認識系構成の著しい特色は，学習機能を積極的に取り入れたことであり，これにより認識系に必要な記憶容量が著しく減少し，又処理方式も簡易化されている。

第3章は，第2章で与えたパターン認識系の評価の研究の成果であり，本認識系の評価を解析的手法とシミュレーションの両面から詳細に行っている。その結果，本認識系の有効性が確実に実証された。

第4章は，認識系構成の上で重要な学習方式の研究に関するものである。本章では，学習パターン列の統計的独立性を仮定せず，一般的な立場からパターン空間の確率構造を学習により逐次明らかにする学習アルゴリズムを，従来より広い条件のもとで与え，このアルゴリズムの収束性等の基本的問題についても厳密な検討を行っている。その結果，パターン系列が独立とした時と本研究との差が理論的に明確にされ，重要な結果を得ている。

第5章は教師なし学習系の構成法について述べたものである。従来教師なし学習系の研究は取り扱いが難しく，十分な研究が行われていないが，本研究ではこの系を構成するための必要条件について綿密に検討し，その上で確率近似法を用いた有効な認識系構成のための一般的方法を導いている。本研究により，従来の方法よりはるかに簡単な認識系の構成方法が得られたことは，重要な成果である。

第6章は結論である。

以上，要するに本論文は，情報工学上重要なパターン認識の研究に多くの重要な知見を与えたもので，情報工学の発展に寄与するところ少なくない。

よって，本論文は工学博士の学位論文として合格と認める。