

書換え規則を用いた時系列パターンの認識

国際電気研究所 樽松 明

1. はしがき 音声や手書き文字などの認識においては、入力パターンの正規化、連続パターンの区分分け、および前後のパターンによる影響の処理などが内題で、高い識別率を得るには、発声速度又は筆記速度のちがいによる継続時間の伸縮に適應でき、かつ前後パターンによる局所的変形を吸収できる認識方法をとる必要がある。

音声を含めて手書き文字や図形などの複雑なパターンを認識するアプローチの一つとして、パターンを部分パターンに分解し、部分パターンの性質と相互関係により、パターンを階層的に記述することによって、パターンを認識する方法が考えられている⁽¹⁾。この方法は構文を考慮したパターン認識とよばれるもので、この方法によると、パターンが複雑になれば部分パターンを必要に応じて階層を深く分析して利用するだけで、認識機械は大部分ははじめのものが利用できることと、

入カパターンを記号系列に変換し、高度のパターン処理が容易にできることになる。

構文を考慮したパターン認識の中には、①順序回路または有限オートマトンにより認識する方法、②辞書を利用する方法、③文法の書換え規則を用いる方法がある。①の方法では、状態の割当と修正が複雑でシステムの構成がやっかいであること、②の方法では、大容量の記憶装置を必要とすること、認識に要する時間がかかることなどの欠点があるのに対し、③の方法は、記述が簡明で、規則の作成、修正が簡単であること、連続パターンの認識がセグメンテーションを特に考えずに可能であることの利点がある。

書換え規則を用いたパターンの認識は、主に文字、図形を対象として行われ、図形を記述するのに適した文法と認識方法についての研究がすすめられている。^{(2)~(7)}

ここでは、時系列パターンを書換え規則を用いて認識する方法に関して、特に限定語彙の音声認識に適用した場合についてのべる。

2. 認識機構の構成⁽⁸⁾

書換え規則を利用した音声認識の機構は図1に示される構

成となる。観測部において入力音声は、たとえば周波数スペクトル分析などの方法により時間と周波数のパターンとして観測され、時間的に変化する周波数パターンを適当な時間間隔で抽出することにより、周波数パターンの時系列として表わされる。すなわち音声パターンは周波数パターンを部分パターンとして部分パターンの時間的なつながりとみなせる。周波数パターンは、あらかじめパターンの性質をもとにして決められたパターンの基本要素となる標準パターンのどれに対応するかを調べて分類し、標準パターンにつけた記号によって記号系列で表現される。この記号系列に書換え規則を適用し、部分系列の逐次的比較により系列を変換していく。このようにして認識される音声パターンはつぎの性質を有すると考えられる。

- (1) 部分パターンの一次元的なつながりとみなすことか

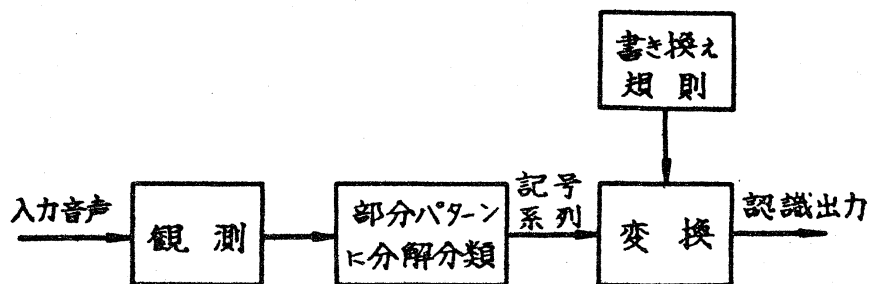


図1 書き換え規則を用いた音声認識の機構

きる。

(2) 部分パターンのいくつか連なったものが、音韻、音節、単語などの意味をもった特徴とみなされ、音声パターンはこのまとまりのつながりで構成される。

(3) 部分パターンの再帰的構造が音声パターンの特徴とはならない。すなわち、入力パターンの記号系列の集合が正規集合であるとみなせる。

(4) 記号の間に類似度が定義できる。

これらの性質を有するパターンは音声のほかの手書き文字を線分の継続とみなし、線分の指示方向を用いて記号系列に表現する場合がある。

パターンの基本要素につけられた記号を原始記号と呼ぶ。パターンの性質2に基づいて、一つのまとまりとみなせる原始記号の部分系列に名づけた記号を中間記号と呼ぶ。認識すべきパターンのカテゴリを表わす記号を類記号と呼ぶ。原始記号、中間記号、類記号の集合をそれぞれ S, Q, P で表わし、それぞれの要素は添字をつけて表わすことにする。

記号系列で表わされるパターンを parsing により認識する方法には、終端記号系列に順次書換え規則を適用し、書換えられた記号によってどのカテゴリに属するかを決定する方法と、始記号から出発し、入力パターンの記号系列に到達す

るまで書換え規則を適用し、成功したなら認識されたとする方法がある。前者は bottom up analysis とよばれる。後者は top down analysis とよばれる。ここでは、bottom up analysis により認識する方法をのべる。

書換え規則は、規則の数の減少と規則を参照するのに要する時間の短縮のために、階層構造とする。階層の深さは、パターンの複雑さによって異なる。ここでは、基本的と思われれる三段階の階層構造を取り扱う。図2はこのような書換え規則を用いて記号系列を認識する過程を示したものである。入力パターンの記号系列表現である原始記号系列は、系列の左から右に順次規則を適用して中間記号系列に変換される。この規則をオ1種規則と呼ぶ。つぎに、この中間記号系列はオ2種規則と呼ばれる規則により別の中間記号系列に変換される。最後に、オ3種規則と呼ばれる規則により類記号に変換

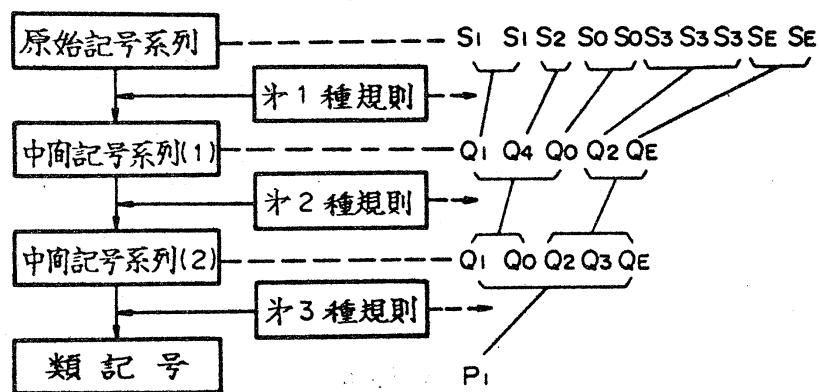


図2 書き換え規則を適用した記号系列の認識

され認識される。

3種類の規則の意味と形式をさらに詳しく述べる。なお、規則の表記法は句構造言語の生成文法の表記法に従うこととする。

〔オ1種規則〕 原始記号系列の部分系列の長さの正規化と、原始記号系列からまとまりとなる特徴を抽出する働きをする。この規則により、発声の速度の異なった音声パターンの正規化と音韻に相当する特徴の抽出を行なう。形式は、同じ中間記号に変換される原始記号の長さがいくらであっても同一とみなして一つの中間記号に変換する場合は、

$$Q_i \rightarrow S_i, \quad Q_i \rightarrow Q_i S_i \quad (Q_i \in Q, S_i \in S)$$

となり、同じ中間記号に変換される原始記号の長さが一定値(k)以上続くことが別の意味をもつ場合は、一定値以上続く場合に別の中間記号に変換してつぎの形式となる。

$$Q_i \rightarrow S_i, \quad Q_i \rightarrow S_i S_i, \dots, \quad Q_i \rightarrow \overbrace{S_i S_i \dots S_i}^{(k-1)},$$

$$Q_j \rightarrow \overbrace{S_i \dots S_i}^k, \quad Q_j \rightarrow Q_j S_i \quad (Q_i, Q_j \in Q, S_i \in S)$$

〔オ2種規則〕 パターンの生成の際に引き起こされた種々の変形の正規化と雑音の除去および調音結合のような隣接する前後の部分パターンによる影響の除去などの働きをする。形式は、中間記号の系列を標準的な中間記号系列に変換するもので、文脈に依存した形でつぎのように表わされる。

$$\alpha \gamma \beta \rightarrow \alpha \gamma' \beta \quad (\alpha, \beta, \gamma, \gamma' \in Q^*)$$

[オ3種規則] これはパターンの類を表わす標準的な記号系列(これを認識系列と呼ぶ)を検出する規則で、パターンの類を決定する働きをする。形式は、中間記号系列を類記号に変換する形でつぎのようになる。

$$P_i \rightarrow \theta \quad (P_i \in P, \theta \in Q^*)$$

書換え規則は、類ごとに規則を分離する構造と全体の類の規則を統合する構造の二つの構成法がある。前者は類ごとに並列的に規則を適用し、最後に判定して結果を出力するので規則の作成と修正は容易であるが、規則の数は多くなる。後者は複数の類で共通にある規則をまとめるため規則の数を少なくできる点と、連続パターンの認識をセグメンテーションを考えずにオ3種規則適用後の類記号を調べるだけのできる点が利点である。

ここでは、書換え規則の構成は、決定論的に適用する場合を考えた。SwainとFu⁽⁷⁾は、Stochastic Programmed Grammarにより記号系列を認識する方法を発表しており、興味深い。

3. 数字音声認識シミュレーション

書換え規則による認識方法を日本語数字音声に適用してシミュレーションを行なった。

入カ音声は18個のフィルタをもつ周波数分析装置で分析し、整流された出力をAD変換し、20msの間隔でサンプリングした。10回の発声による100個のデータのサンプルパターンとして、有声音部11個、無声音部2個、休止部1個、計14個の基本パターンを用いて周波数スペクトルパターンを分類して音声を記号系列に変換し、全部のサンプルパターンが正しく識別されるように規則を作成した。同一人の発声下、各数字単独に2回ずつ、2個の連続数字を数字の間を区切るように発声したものを2回ずつ、計220個の未知パターンを類ごとに構成した規則により認識した。結果は、「ヨン」、「ハチ」下各1個、「ロク」下3個の棄却が生じた他は正しく識別された。

つぎに、多様なパターンに対する適用として、異なる人の発声による認識実験を行った。10人の男性の発声した100個のデータをサンプルパターンとして、サンプリング長のスペクトルパターンを母音部24個、有声音部5個、無声音部4個、無声部1個の計34個の基本パターンに分類し、この分類に従って音声を記号系列に変換した。作成した規則により別の10人の男性の発声した100個のパターンを認識した結果のうち、類ごとに規則を構成した場合、正認識70個、誤認識2個、棄却28個であった。規則修正の手順を用

いて、合計 200 個のパターンを全部正しく識別するように規則を修正することになった。規則の数はオ1種 80 個、オ2種 84 個、オ3種 10 個で合計 174 個となった。

4. おすび

書換え規則を用いた時系列パターン認識の特徴は、書換え規則を階層構造としているので、規則の記述が簡明で直感的にわかりやすく、規則の作成、修正を機械的に行なうことができることである。この方法を数字音声の認識に適用した結果、同一人の発声の音声認識の場合、書換え規則を用いることにより発声の仕方の違いによるパターンの変動を吸収することができ、正しい認識が行なえることがわかった。

書換え規則を用いる方法についていくつかの考慮すべき点を以下に指摘する。

(1) パターンを特徴により分類して記号化する仕方と、書換え規則によるパターン認識の結果の良否とは密接な関係がある。パターンの特徴を十分よくとらえない記号化をすると、パターンの特徴をあらわす記号系列がとらえ難くなり、規則の適用度が低くなり、規則の数はよりふえることになる。

(2) ここで述べた認識方法は、オ3種規則であらわされる系列に一致したものが存在するときに認識結果を出す方法を

とっている。この方法で類の数が多くなると場合規則にあいまい性が少ないように規則を作成すると context の制約が厳しくなる。未知パターンの棄却率が高く認識能力が下る。規則のあいまい性があっても認識される類の出力が一方向性であるなら、類出力に優先順位をつけることにより正しく認識することができる。標準パターンの記号系列に一致した場合にのみ正認識とする規定をゆるめて、最も確率の高いものに判定を下すことを考える必要がある。

(3) 規則のデータ構造をリニアリストにしているので、規則を参照するのに時間がかかる。規則をグループに分割することにより認識に要する時間が短縮できる。パターンの記号系列が正規集合とみよせる場合は、順序回路で認識機構をハード化できるので、実時間の認識も可能である。

文献

- (1) T. G. Evans: "A grammar-controlled pattern analyzer", Proc. of IFIP 1968 p.1592
- (2) R. Kirsh: "Computer interpretation of English text and picture patterns", IEE Trans. EC-13, 4, p.363 (Aug. 1964). (3) P. Narasimham: "Syntax-directed interpretation of classes of pictures", Comm. ACM, 9, 3, p.166 (March 1966). (4) R. S. Ledly et al.: "Pattern recognition studies in the biomedical sciences", 1966, Proc. SJCC, p.441. (5) 森地: "最近のパターン認識研究", 情報処理, 11, 11, p.658 (AB45-11).
- (6) 阿部他: "文法によるパターン認識についての考察" 昭44信学全木, 82. (7) R. H. Swain et al.: "Stochastic Programmed Grammars for syntactic pattern recognition", Pattern Recognition, 4, p.83 (1972).
- (8) 樽井・武田・井上, "書換規則を用いた音声認識を行う場合の規則の構成法", 信学誌 Vol. 55-D, No. 2 (昭和47年2月).