

7E-5

m重マルコフモデルを用いた音節ラテイスからの候補絞り込み手法について

荒木 哲郎、村上 仁一、池原 悟

NTT情報通信処理研究所

1. はじめに

日本語音声認識の音響処理で解決出来ない音声認識の曖昧さを、日本語の音節、漢字等のマルコフ連鎖特性を用いて解決する方法について研究が行われ^{[3]-[5]}、大量の音節、漢字候補列の絞り込みを効率よく行う上で、Viterbiアルゴリズムが効果があることが示されている。そこではセグメンテーション誤りが無い(置換誤り)場合について、Viterbiアルゴリズムを2重マルコフモデルに拡張できることが指摘されている。今後は更に、セグメンテーション誤り(挿入、脱落誤り)を有する一般的な場合に対して、多重マルコフモデルを用いた効率的な候補絞り込み方法が期待される。そこで本報告では、Viterbiアルゴリズムを、一般のm重マルコフモデル及び、セグメンテーション誤りを有する場合に拡張する方法について述べる。

2. Viterbiの基本アルゴリズムと音節ラテイスの定義

Viterbiアルゴリズムは、離散時間の有限マルコフ過程と雑音を持つ無記憶の通信路において、通信路からの観測値Zから、マルコフ過程の出力Xを推定するものである。これはマルコフ過程の内部状態(出力Xと同一視する)と時間をそれぞれ縦軸、横軸に持つグラフ(trellisと呼ばれ、観測値Zを発生させる可能性のあるXのすべてを別ルートとしてあらわされるもの)に対して、次式の事後確率を最大にするようなルートをtrellis上から選ぶ方法を与えるものである。即ち、

$$P(\underline{X}|\underline{Z}) = P(\underline{X}, \underline{Z}) / P(\underline{Z}) = P(\underline{X}) \cdot P(\underline{Z}|\underline{X}) / P(\underline{Z}) \quad (1)$$

式(1)の分子を無記憶性、一重マルコフ性の条件及びを使いまた、その対数をとって、

$$s(x_i) = -\ln p(Z_i | X_i) - \ln p(X_{i+1} | X_i) \quad (2)$$

と表し、この値s(X)をtrellisの各枝に距離として与えることにより、trellis上で最短のルートを求める問題に帰着させる。この時アルゴリズム(以下ALと略す)はkを時刻変数(0が起点、Fが終点)、状態数をM、W(x_i^k)を時刻kの状態x_i^kにおける最短ルート、Q(x_i^k)をx_i^kにおける最短距離値とすると次のように表される。^[2]

[AL0] Vitebiの基本アルゴリズム: (初期値) k=0; W(x_i^0) = x_0; Q(x_i^0) = 0、(繰り返し) ① Q(x_i^{k+1}) = Min {Q(x_j^k) + (x_j^k)} ② W(x_i^{k+1}) = W(x_j^k)、(但し 0 < k < F, 1 < i < M)

このアルゴリズムによって、単重マルコフ過程の場合の最短

ルート(第一位候補)が求まる。本論文ではこれを多重のマルコフ過程で第n位までの候補を求める場合に拡張し、一般の連続音声認識における誤りを有する音節認識候補の絞り込みに適用する方法を示す。

そこで音声認識における音響処理の曖昧を、[1]のように音声の一つの基本単位である音韻/音節認識の曖昧さと、音節区間の切り出しの曖昧さ(セグメンテーション誤り)からなる音節ラテイス形式^[1]で定義する。すなわち横軸にセグメント番号を、また縦軸に音節候補位置をとって、該当位置に音節記号を付与したもので、入力X_iと出力X_iが1対1の時、音節マトリックス(図1-(a)) またn対mの時、音節ラテイス(図1-(b))と呼ぶ。ここでは文節境界では誤らないものとし、1文節単位の音節マトリックス/ラテイスを考える。次にこれらを正解音節の有無によりそれぞれ2通りに分け、以下の4つのクラスを考える。(A)正解音節を有する音節マトリックス、(B)正解音節を有する音節ラテイス、(C)正解音節が脱落した音節マトリックス、(D)正解音節が脱落した音節ラテイス。以下では、この内(A)-(C)のクラスについて、Viterbiを拡張した音節候補絞り込みアルゴリズムを示す。

3. m重マルコフ過程を用いた音節マトリックスからの候補絞り込み方法

先ず音節マトリックス(クラスA)に対して、2.で述べたViterbiアルゴリズムをm重マルコフ過程に拡張し、最尤な音節候補列を上位n位まで求める拡張アルゴリズム(これをEALと略す)を示す。

(EAL1) m重音節マルコフモデルへの拡張アルゴリズム: ①時刻kの各状態x_i^k, (1 < i < M)に、m個前の音節列の組み合わせx^{m-1} = (x_{i-m}, ..., x_{i-1})を割当てて。②状態x_i^kの先頭を除く後半のm-1個の音節列が、状態x_i^{k+1}、即ち(x_{i-m+1}, ..., x_i)の先頭のm-1個の音節列に一致するとき限りその間に遷移枝を設定し、その遷移枝に(2)式の一重マルコフ確率の代わり

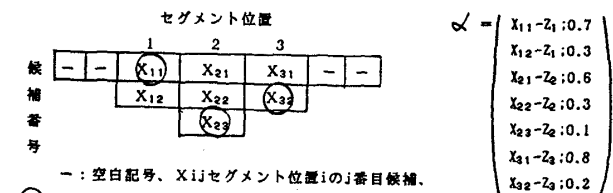


図1. (a) 音節マトリックスの例

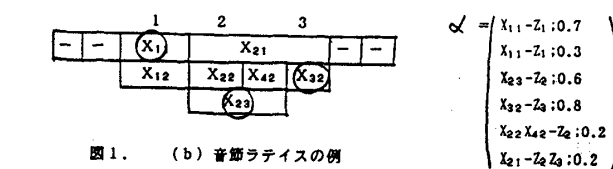


図1. (b) 音節ラテイスの例

A method for deciding a maximum likelihood candidate in syllable lattice using m-th order markov model
Tetsuo ARAKI, Jinichi MURAKAMI and Satoru IKEHARA
NTT Comm. and Info. processing Laboratories

に、 $-\ln p(x_i | x_{i-m}, \dots, x_{i-1})$ を割り当てたtrellisを作る。
 以上で本アルゴリズムにより、AL0の問題に帰着する。

ここで上位 $j(1 \leq j \leq n)$ 位の各ルートは、各時刻 k に於てある状態 x_i^k 、 $(1 \leq i \leq M)$ の j 位内に必ず存在することから、EAL1の③で、時刻 k の各状態 x_i^k 、 $(1 \leq i \leq M)$ において、上位 n 位までの距離値 $Q(x_i)$ と $W(x_i)$ をそれぞれ最終位置 F まで保存し最後に比較することにより、上位 n 位までのルート(音節候補列群)が求まる。2重マルコフの場合のtrellisの例を図2-(1)に示す。

4. 各種タイプのマルコフ連鎖確率による候補絞り込み

ここではマルコフ連鎖確率のタイプとして、通常のマルコフ連鎖確率(順方向)に加えて、逆方向、中間のものを次のように定義する。^[5]

- (1) 順方向 : $p(x_i | x_{i-m}, \dots, x_{i-1})$ 、(2) 逆方向 : $p(x_i | x_{i+1}, \dots, x_{i+m})$ 、(3) 中間 : $p(x_i | x_{i-r}, \dots, x_{i-1}, x_{i+1}, \dots, x_{r+1})$, $r = \lfloor m/2 \rfloor$

(EAL2) 逆、及び中間タイプの m 重マルコフ過程への拡張アルゴリズム : ①時刻 k の各状態 x_i^k 、 $(1 \leq i \leq M)$ に、 m 個後および前後の音節列 $(x_{i+1}, \dots, x_{i+m})$ および $(x_{i-r}, \dots, x_{i-1}, x_{i+1}, \dots, x_{r+1})$ を割り当てる。②EAL1の②で述べた同じ方法でtrellisを作成し、各遷移枝にそれぞれ上記の逆及び、中間タイプの m 重マルコフ連鎖確率値を割り当てる。以上で本アルゴリズムによって、AL0の問題に帰着する。

これは各状態 x_i^k 、 $(1 \leq i \leq M)$ におけるルート比較が、確率の対数値の総和形式での比較になる為、演算順序等によらず順方向と同じ方法で n 位までのルートが求まる。

5. その他の音節ラテイスクラスからの候補絞り込み

[1] 正解候補を有する音節ラテイス(クラスB)

2.及び3.の音節マトリックス(クラスA)では、 X_i と Z_i が1対1に対応するから、各時刻 k の状態 x_i^k 、 $(1 \leq i \leq M)$ で比較される各候補列は全て同一の長さであるが、音節ラテイスでは X_i と Z_i が n 対 m (脱落、挿入有り)に対応し、各候補列はそれぞれ長さが異なるため、各時点での最尤な音節列の選択基準を、正規化した $Q(x_i^k) = Q(x_i^k) / l$ 、(l は該当の音節長)で定めることにより次のように求まる。

(EAL3) 正解候補有りの音節ラテイスへの拡張アルゴリズム : ①最長の音節列に対応するセグメント番号を基準に、連結可能な m 個の音節列の組み合わせをEAL1の①のように各状態 x_i^k に割当てる。但し時刻 k のセグメントが事前のセグメントに継続する時は、新規に状態 x_i^{k+1} を設けずに x_i^k に縮退させる。②各状態間の遷移枝にEAL1と同様に(2)式を割当ててtrellisを作成する。以上本アルゴリズムにより、また正規化した距離値 $Q(x_i^k)$ を用いることにより、AL0の問題に帰着する。図2-(2)に、音節ラテイスのtrellisの例を示す。

[2] 正解候補が脱落した音節マトリックス(クラスC)

音響処理の結果から出力される各音節認識候補に対する認識確度情報(認識重み確率) U_i と各音節位置における

候補集合 P_i の補集合 P_i' に対してヒュリスティックに与えられるスコア値 V_i (一般に $U_i > V_i$)により次のように求まる。

(EAL4) 正解候補が欠落した音節マトリックスへの拡張アルゴリズム : ①各時刻 k の状態 x_i^k を、音節候補集合全体 $(P_i \cup P_i')$ に対してEAL1の①と同様に割り当てる。②各状態間の遷移枝に、式(2)の第2項を p_i に対しては U_i を、また P_i' に対しては V_i で置き換えたtrellisを作る。以上で本アルゴリズムにより、AL0の問題に帰着する。

6. あとがき

本報告では、Viterbiアルゴリズムを一般の m 重マルコフ、及びクラス(A)-(C)の音節ラテイスに適用して、最尤な音節候補列を求める方法について述べた。正解候補が欠落した音節ラテイスのクラスDの場合は、クラスCのように候補拡張出来ないで今後の研究課題である。

[文献]

- (1)中津他：“会話音声の機械認識における音響処理”，信学論，61-D,4pp261-268(1978)(2)Forney .G.D.：“The Viterbi Algorithm”，Proc. IEEE,61,3,pp268-278(1973)
- (3)荒木他：“2重マルコフモデルによる日本語の文節音節認識候補の曖昧さの解消効果”，情処論文誌，30,4,pp467-477(1989)(4)村上他：“2重マルコフ連鎖確率モデルを使用した単音節人力の改善”，信学技報，sp88-29,pp63-70(1988)(5)荒木他：“音節連鎖確率のタイプと音節認識候補絞り込み効果について”，信学全国大会，D-307,pp6-27(1989)

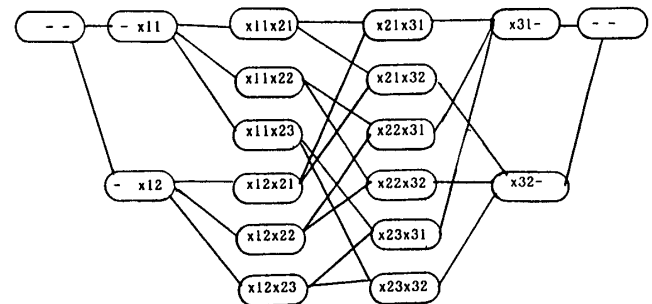


図2 (1) 音節マトリックス(a)のトレリス例

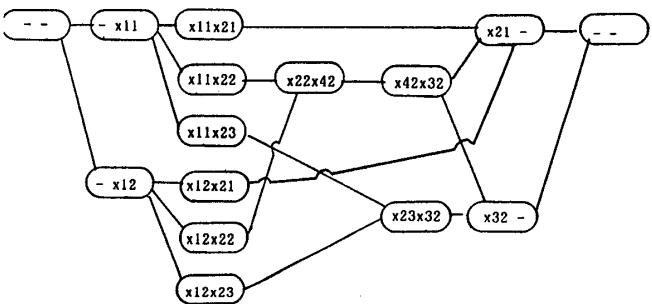


図2 (2) 音節ラテイス(b)のトレリス例