

# オンライン文字認識 HMM における混合ガウス分布出力機構に関する検討

渡辺 雄一郎\* 内田 誠一\* 迫江 博昭\*  
\*九州大学大学院 システム情報科学府

## 1 はじめに

オンライン文字データはペン運動情報の時系列として表現されるが、時間的にも空間的にも複雑な変動を伴う。したがって、音声認識での実績から HMM(隠れマルコフモデル)の適用が有効と考えられる。HMM を定義するパラメータとして、状態遷移確率と出力確率がある。前者は時間伸縮変動を規制し、後者は空間的特徴変動を規制する。後者の出力確率の方が、実際に観測される特徴に近いことから認識に与える影響が大きいとされている。複雑な空間的変動を一般的に記述する出力確率として混合ガウス分布が知られており、音声認識の分野では常識的に用いられている。

本研究では、オンライン文字認識 HMM において混合ガウス分布を用いることの有効性を確認し、さらに、新たなガウス分布出力機構の実装を検討する。

## 2 混合ガウス分布 HMM を用いたオンライン文字認識

線分情報を確率的に出力する状態(線分出力機構)を LR 型に配置した HMM を用いる(図 1)。HMM の状態は特徴量に関するなんらかの定常性を仮定しているため、文字を線分近似(例えばペンの移動方向)して考える。本稿では、線分出力機構における出力確率として混合ガウス分布を用いる。混合ガウス分布とは、複数のガウス分布(個々の分布をモードという)の 1 次結合として複雑な分布関数を近似するものである。このとき各状態  $j$  での出力確率  $b_j(x)$  は、

$$b_j(x) = \sum_{k=1}^K c_{jk} N(x, \mu_{jk}, U_{jk})$$

と表せる [1]。ただし、 $c_{jk}$  は各ガウス分布が生起する確率で、 $\sum_{k=1}^K c_{jk} = 1$  である。また、 $N(x, \mu, U)$  は平均ベクトル  $\mu$ 、共分散行列  $U$  の  $d$  次元単純ガウス分布である。混合ガウス分布出力を行う状態の実装は、各モードに対するガウス分布出力を行うサブ状態の並列配置で行われる。

認識では、ある入力文字パターン  $O = O_1, O_2, \dots, O_T$  が HMM  $\lambda$  から生成される確率(尤度)を  $P(O|\lambda)$  で表し、この尤度が最大となるカテゴリ  $\lambda$  を認識結果として出力する。本稿では、尤度の計算方法として前向きアルゴリズムを用いた [1]。

### 2.1 一貫型混合ガウス分布 HMM

オンライン文字認識の場合、自己回帰を伴う状態を線分に対応させると、各モードは線分方向のいくつかの代表値に対応すると考えられる。従来の混合ガウス分布 HMM では、モードの選択は回帰の度に独立に行われるため、モード選択の一貫性がない。一貫性を無視して異なったモードを選択すると折れ線や曲線を出力して、他の文字との尤度を高く評価し、結果として、クラス間分離能力の低下を来す可能性がある。この解決法として、各サブ状態に自己回帰ループをもうけ遷移確率を  $a_{jj}$  とするという構成により、モードの一貫性を付与することを考える(図 2)。

## 3 実験

認識実験には、研究室のデータ(筆記者 30 人分)と東京農工大学中川研究室が公開しているデータ(筆記者 120 人分)から抽出した 10 画の教育漢字全 83 文字種で筆順

の正しいものを対象とした。文字によってデータ数に大きくばらつきがあったため、データの少ない文字種は除外した。また、混合ガウス分布は混合数を 2 とした。

表 1 は、学習に研究室のデータを用いた時に、研究室データ・農工大データそれぞれを入力データとしたときの結果である。認識率からすれば、従来型、一貫型ともに認識率が高いことから、それぞれオンライン文字認識には有効な手法であるといえる。また、従来型で出見られた他の文字種への変形を規制するために一貫型混合ガウス分布を提案し、最もよい認識結果が得られたことがわかる。

## 4 まとめ

本稿では、オンライン文字認識 HMM において混合ガウス分布を用いることの有効性を確認した。また、新たなガウス分布出力機構として一貫型混合ガウス分布 HMM を提案し、0.06% ではあるが認識率の向上を確認した。参考文献

[1] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. IEEE, Vol. 77, No.2, pp. 257-284, Feb. 1989.

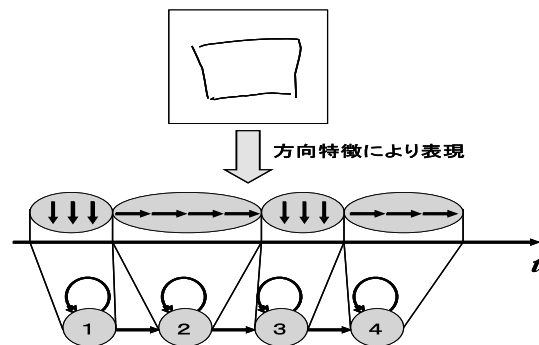


図 1: オンライン文字認識 HMM

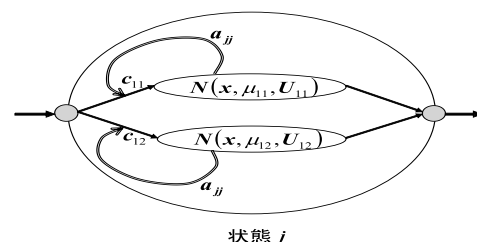


図 2: モードの一貫性を保存する混合ガウス分布 HMM

表 1: 認識結果

	研究室データ closed [1215]	農工大データ open [12398]
単純ガウス分布	99.9 % (1)	99.73 % (34)
従来型 混合ガウス分布	99.9 % (1)	99.75 % (31)
一貫型 混合ガウス分布	99.9 % (1)	99.81 % (24)

[ ]:総パターン数, ( ):誤認識数