

隠れ状態の継続時間長を考慮した確率モデルに関する調査

A Survey of Probabilistic Modeling Techniques of Hidden States and their Duration Distributions

黒川茂莉^{1*} 横山浩之¹ 吉井和佳² 麻生英樹²
Mori Kurokawa¹ Hiroyuki Yokoyama¹ Kazuyoshi Yoshii² Hideki Asoh²

¹ 株式会社 KDDI 研究所

¹ KDDI R&D Laboratories, inc.

² 産業技術総合研究所

² National Institute of Advanced Industrial Science and Technology

Abstract: Human activity data, so-called “life-log data”, has underlying contexts such as sleeping, driving a car, eating. In order to analyse and predict such data, the hidden Markov model is often used. However, the duration time of the hidden context state of HMM distributes exponentially and this is not suited for modeling the above contexts. In order for modeling these context flexibly, we introduce and compare probabilistic modeling techniques of hidden states with general duration distributions.

1 はじめに

ライフログなどの人間の行動データには、センサで計測した生体データや GPS による測位データなど、さまざまな時系列データが含まれる。こうした時系列データを元に、各ユーザに対して「今だけ、ここだけ、あなただけ」の適切な情報提供を行うには、休憩中、運動中、食事中といったユーザのコンテキストを隠れ状態とし、それを動的に推定することで、エージェントサービスを高度化するアプローチが考えられる。こうした時系列データを隠れマルコフモデルによってモデル化することがしばしば行われているが、上記のような隠れ状態の継続時間長はユーザやコンテキストによって千差万別であり、一般に、隠れ状態の継続時間長が指数分布に従う隠れマルコフモデルでは十分に精度良くモデル化できないことが多い。適切なモデル化のためには継続時間長の確率分布をより柔軟に取り扱う枠組みが必要とされている。そこで本稿では、継続時間が異なるコンテキストの確率モデルの検討のため、隠れ状態の継続時間長を考慮した確率モデルを調査し、比較する。

2 節では確率過程の基礎的概念を導入する。3 節では隠れマルコフモデル、4 節では継続時間長を考慮した隠れマルコフモデル、5 節で 4 節の各モデルの比較、6 節ではまとめを述べる。

2 確率過程の導入

継続時間長の分布を考慮した隠れマルコフモデルを説明する前に、必要となる確率過程の基礎的概念について説明する [1, 2]。

無限個の確率変数の集合 $\mathbf{X} = \{X_t, t \in \mathcal{T}\}$ を確率過程 (stochastic process) と呼ぶ。 \mathcal{T} は確率変数のインデックスの集合 (index set) であり、通常、時間と解釈される。 \mathcal{T} が可算集合の場合を離散時間 (discrete-time) の確率過程、連続集合の場合を連続時間 (continuous-time) の確率過程と呼ぶ。確率過程 \mathbf{X} のひとつの実現値 $\{x_t, t \in \mathcal{T}\}$ を標本路 (sample path, 見本路, 標本過程などとも訳される) と呼ぶ。

2.1 ポアソン過程と再生過程

定義 下記の条件を満たす確率過程 N_t を計数過程 (counting process) と呼ぶ。

1. $N_t \geq 0$.
2. $N_t \in \mathcal{Z}$.
3. $t_1 < t_2$ ならば $N_{t_1} \leq N_{t_2}$.

$t_1 < t_2$ に対して、 $N_{t_2} - N_{t_1}$ は、時間区間 $(t_1, t_2]$ の間に生じた事象の数と解釈でき、増分 (increments) と呼ばれる。ある計数過程が独立増分 (independent increment) を持つとは、任意の排他的な時間区間 $(t_1, t_2]$

*連絡先：株式会社 KDDI 研究所
〒356-8502 埼玉県ふじみ野市大原 2-1-15
E-mail: mo-kurokawa@kddilabs.jp

と $(t_3, t_4]$ に対して, 増分 $N_{t_2} - N_{t_1}$ と $N_{t_4} - N_{t_3}$ が独立であることを言う. ある計数過程が定常増分 (stationary increments) を持つとは, 任意の $t_1 < t_2$ と $s > 0$ に対して, 増分 $N_{t_2} - N_{t_1}$ と $N_{t_2+s} - N_{t_1+s}$ が同じ確率分布に従うことを言う.

二項過程 (binomial process) は, 代表的な離散時間の計数過程の一つであり, たとえば以下のように定義される.

定義 下記の条件を満たす離散時間の計数過程を二項過程と呼ぶ.

1. $N_0 = 0$
2. 独立・定常増分を持つ
3. 任意の $t, s > 0$ に対する増分が, ある事象生起確率 $p > 0$ の二項分布に従う. すなわち,

$$P\{N_{t+s} - N_t = n\} = \binom{s}{n} p^n (1-p)^{s-n}$$

ポアソン過程は, 二項過程を連続時間に自然に拡張した代表的な連続時間の計数過程の一つであり, 以下のように定義される.

定義 下記の条件を満たす連続時間の計数過程をポアソン過程と呼ぶ.

1. $N_0 = 0$
2. 独立・定常増分を持つ
3. 任意の $t, s > 0$ に対する増分が, ある正の数を $\lambda > 0$ として, 平均値 λs のポアソン分布に従う. すなわち,

$$P\{N_{t+s} - N_t = n\} = e^{-\lambda s} \frac{(\lambda s)^n}{n!}$$

計数過程 N_t に対して T_n を n 番目の事象が生起する時刻とし, D_n を $n-1$ 番目の事象と n 番目の事象の間の時間間隔とする. $\{D_n, n \geq 1\}$ は到着間隔 (あるいは事象生起間隔) (interarrival times) の系列と呼ばれる. 明らかに $D_1 = T_1$ である.

定理 パラメータ p の二項過程の事象生起間隔はパラメータ p の幾何分布に従う. すなわち, n によらずに, 任意の $k > 0$ について

$$P\{D_n = k\} = (1-p)^{k-1} p$$

が成り立つ.

定理 パラメータ λ のポアソン過程の事象生起間隔は, 平均が $1/\lambda$ の指数分布に従う. すなわち, n によらずに, 任意の $t > 0$ について

$$P\{D_n > t\} = e^{-\lambda t}$$

が成り立つ.

このように, 二項過程やポアソン過程の事象生起間隔の確率は間隔が長くなると指数的に減少するが, 現実の世界ではこの性質を持たない事象も多い. そうした事象をモデル化するために二項過程やポアソン過程を拡張したものが再生過程である.

定義 任意の $n > 1$ に対して事象生起間隔 D_n が独立に同一の確率分布に従うような計数過程を再生過程 (renewal process) と呼ぶ.

再生過程という名前の由来は, 事象の生起によってプロセスが一旦リセットされて, そこからまた新しいプロセスが始まる (再生) と考えられるためである. 再生過程によってモデル化される典型的な事象としては, 機械の故障, 車の通過, などがある.

定理 再生過程 N_t に対して, μ を事象生起間隔の平均とすると, 確率 1 で

$$N_t/t \rightarrow 1/\mu \quad (t \rightarrow \infty)$$

が成り立つ.

2.2 マルコフ過程

離散時間のマルコフ過程の定義は以下の通りである.

定義 任意の時刻 $t \geq 0$ における X_t の確率分布が以下の式を満たすとき, 確率過程 $\mathbf{X} = \{X_t, t \in \mathcal{T}\}$ はマルコフ過程と呼ばれる.

$$P(X_t | X_1, \dots, X_{t-1}) = P(X_t | X_{t-1})$$

マルコフ連鎖はマルコフ過程の一種であり, X_t のとる値が離散的な場合を言う.

また, 連続時間マルコフ過程の定義は以下の通りである.

定義 時刻 $0 < s_0 < s_1 < \dots < s_n < s$ について $X_t (t \geq 0)$ が以下の式を満たすとき, これを連続時間マルコフ過程と呼ぶ.

$$\begin{aligned} P(X_{s+t} = j | X_s = i, X_{s_n}, \dots, X_{s_0}) \\ = P(X_{s+t} = j | X_s = i) \end{aligned}$$

離散時間マルコフ過程, 連続時間マルコフ過程の両方について, 以下のチャップマン-コルモゴロフ方程式が成り立つ.

定理 マルコフ過程の推移確率 $P_t(i, j) \equiv P(X_t = j | X_0 = i)$ はチャップマン-コルモゴロフ方程式

$$P_{s+t}(i, j) = \sum_k P_s(i, k) P_t(k, j)$$

を満たす.

齊次的連続時間マルコフ過程では、推移確率 $P_t(i, j) \equiv P(X_t = j | X_0 = i)$ とその時間極限である推移率（強度と呼ぶことも多い） $q(i, j)$ を考える。推移率 $q(i, j)$ から推移確率 $P_t(i, j)$ の導出は、以下のチャップマン-コルモゴロフの方程式により与えられる。ここで、時刻 t において隠れ状態が状態 i から遷移するときの推移率の総和を $\lambda_i = \sum_{j \neq i} q(i, j)$ とする。

定義 以下のように Q 行列を定義する。

$$Q(i, j) \equiv \begin{cases} q(i, j) & (j \neq i) \\ -\lambda_i & (j = i) \end{cases}$$

定理 コルモゴロフの後向き方程式

$$P'_t = QP_t$$

証明

$$\begin{aligned} P'_t(i, j) &= \lim_{h \rightarrow 0} \frac{P_{t+h}(i, j) - P_t(i, j)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sum_k P_h(i, k) P_t(k, j) - P_t(i, j)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sum_{k \neq i} P_h(i, k) P_t(k, j)}{h} \\ &\quad + \lim_{h \rightarrow 0} \frac{(P_h(i, i) - 1) P_t(i, j)}{h} \\ &= \left(\sum_{k \neq i} \lim_{h \rightarrow 0} \frac{P_h(i, k)}{h} P_t(k, j) \right) \\ &\quad + \lim_{h \rightarrow 0} \frac{p_h(i, i) - 1}{h} P_t(i, j) \\ &= \left(\sum_{k \neq i} q(i, k) P_t(k, j) \right) - \lambda_i P_t(i, j) \end{aligned}$$

上式の $P_{t+h}(i, j) = \sum_k P_h(i, k) P_t(k, j)$ のところにチャップマン-コルモゴロフの方程式が使われている。 Q 行列の定義から上式は $P'_t = QP_t$ と記述できる。

以下のようにおけば、コルモゴロフの後向き方程式を満たす。

$$P_t = e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!} \quad (1)$$

2.3 セミマルコフ過程

定義 有限状態空間 $I = \{0, 1, \dots, m\}$ を持つ確率過程 $X_t(t \geq 0)$ を考え、これから導入される確率変数列 $Z_n(n \geq 0)$ と $T_n(n \geq 0)$ を考える。ここで、 $Z_n = i$ は X_t が n 回目の状態推移で状態 i にあること、 T_0 は観測開始時点、 T_n は n 回目の推移の時点を示す。これらが下記の条件を満たす場合

に、 $X_t(t \geq 0)$ をセミマルコフ過程 (semi-Markov process) と呼ぶ [3]。

$$\begin{aligned} P &(Z_{n+1} = j, T_{n+1} - T_n \leq t \\ &| Z_0, \dots, Z_n; T_0, \dots, T_n) \\ &= P(Z_{n+1} = j, T_{n+1} - T_n \leq t | Z_n) \end{aligned}$$

セミマルコフ過程において、時刻 t までの状態 i の訪問回数を $N_i(t)$ としたとき、 $\mathbf{N}(t) = (N_1(t), N_2(t), \dots)$ をマルコフ再生過程 (Markov renewal process) と呼ぶ。

3 隠れマルコフモデル

隠れマルコフモデル (hidden Markov models) は、マルコフ過程 (Markov process) に従って遷移する隠れ状態を有し、各状態に応じて記号が確率的に出力されるとするダイナミックベイジアンネットワークモデルである。

隠れマルコフモデルをダイナミックベイジアンネットワークのグラフとして表現すると、図1のようになる。ここで、 Z は隠れ変数であり、 X は出力変数である。出力変数は1つでも複数でも構わないが、図1では簡単のため1つとし X と表す。

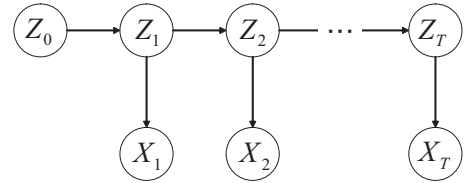


図1: 隠れマルコフモデルのダイナミックベイジアンネットワーク表現

隠れマルコフモデルの場合、図1に表現されているように任意の時刻の状態の確率分布が直前の状態のみに依存し、隠れ状態の系列 Z_1, Z_2, \dots, Z_T はマルコフ過程に従う。マルコフ過程で状態間の遷移確率分布が時刻 t に依存しない場合を、とくに齊次的マルコフ過程と呼ぶ。この場合、直前の状態を条件としたときの次の時刻の状態の確率分布は、時刻に依存せず $a_{ij} \equiv P(Z_t = j | Z_{t-1} = i)$ と置くことができる。

二項過程についてと同様に、隠れマルコフモデルでは、隠れ状態 i の継続時間を確率変数として d_i と表すと、継続時間長の分布は次の指数分布に従うことが分かる。

$$P(d_i) = (1 - a_{ii}) a_{ii}^{d_i - 1} \quad (2)$$

隠れマルコフモデルにおいては、通常、離散時間でのモデリングがおこなわれているが、Nodelman らが提案した連続時間ベイジアンネットワーク [4, 5] の特殊

形して連続時間隠れマルコフモデルを考えることもできる。連続時間ベイジアンネットワークは、確率変数間の依存関係に2節の斉次的連続時間マルコフ過程を仮定した確率モデルである。連続時間隠れマルコフモデルでは隠れ状態間の関係にのみ斉次的連続時間マルコフ過程を仮定することにより、連続時間ベイジアンネットワークの枠組みで隠れ状態の継続時間長を考慮することができる。連続時間ベイジアンネットワークでは、リンクのある変数間の関係にそれぞれ Q 行列を定義する。

連続時間隠れマルコフモデルでは、遷移確率の分布が斉次的連続時間マルコフ過程に従い、斉次的連続時間マルコフ過程では継続時間長の分布として指数分布を仮定しているため、継続時間長の分布の仮定は隠れマルコフモデルと変わらない。連続時間ベイジアンネットワークの拡張として、継続時間長の分布としてアーラン分布を仮定したものも提案されている [6]。

4 継続時間長の分布を考慮した隠れマルコフモデル

本節と次節では、Johnson[7]に従って、継続時間長の分布を考慮した隠れマルコフモデルについて述べる。継続時間長の分布を考慮した隠れマルコフモデルとしては、以下の3つのタイプがある。

1. 隠れセミマルコフモデル
2. Variable Transition 隠れマルコフモデル (VT 隠れマルコフモデル)
3. 隠れマルコフモデルの拡張

隠れセミマルコフモデル [8, 9, 10] は、隠れ状態間の遷移確率の分布 a_{ij} と各遷移に関連した直前状態の継続時間の分布 $p_{ij}(d)$ をそれぞれ定めるモデルである。Ferguson は離散時間の継続時間長の分布を導入し [8], Levinson はそれを連続時間の継続時間長の分布に対応できるように拡張した [9]。Ferguson, Levinson はそれぞれ、Variable duration (hidden Markov) model (VD 隠れマルコフモデル), Continuously variable duration hidden Markov model (CVD 隠れマルコフモデル) と呼んだが、後に隠れセミマルコフモデルという呼び名が一般的となった。また、セミ隠れマルコフモデルと呼ぶ場合もある。隠れセミマルコフモデルでは、2節のセミマルコフ過程の通り、任意の時刻の状態の確率分布が直前の状態のみならず直前の状態の継続時間長 d に依存すると仮定する。なお、以下の式では、継続時間長の確率過程 $D_n(n \geq 1)$ を考え、 $D_n = T_n - T_{n-1}(n \geq 1)$ とおいた。

$$\begin{aligned} & P(Z_n, D_n \leq d | Z_0, \dots, Z_{n-1}; D_1, \dots, D_{n-1}) \\ = & P(Z_n, D_n \leq d | Z_{n-1}) \end{aligned}$$

VT 隠れマルコフモデル [11] は、Duration-dependent transition hidden Markov models とも呼ばれることもあり、遷移確率が現在の状態の継続時間の長さの関数 $a_{ij}(d)$ として表現される。この方式は複数の研究者によって提案され、Ramesh らが提案した非斉次的隠れマルコフモデル [12] や Vaseghi らが提案した継続時間長に依存した状態遷移のモデル [13] などがある。

ここで、隠れセミマルコフモデルと VT 隠れマルコフモデルの異同点は、以下の通りである。

- 隠れセミマルコフモデルにおいては遷移確率と継続時間の分布を別々に考えるのに対して、VT 隠れマルコフモデルは一体として考える。
- 隠れセミマルコフモデルの遷移確率の分布は斉次的であるのに対し、VT 隠れマルコフモデルの遷移確率の分布は継続時間の関数として表現されるので非斉次的である。
- 隠れセミマルコフモデルは、継続時間を $[0, \infty)$ の範囲で考えることができるが、VT 隠れマルコフは継続時間を有限の範囲で考える必要がある。ただし、実際の適用では、隠れセミマルコフモデル、VT 隠れマルコフモデルのいずれについても、最大継続時間長をおくことがほとんどである。

隠れマルコフモデルの拡張 [14] では、隠れ状態の数を増やしたり、隠れ状態間の遷移関係のトポロジーを工夫したりすることにより、継続時間のばらつきを吸収することができる。隠れ状態間の遷移関係のトポロジーの例としては、図2のようなトポロジーがある。Johnson[7]によれば隠れマルコフモデルの拡張でも指数分布以外の継続時間長の分布を近似できることが実験的に示されている。

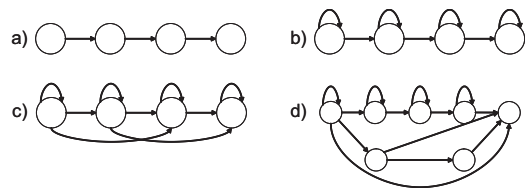


図 2: 隠れ状態の遷移関係のトポロジーの例 : a) left-to-right model, b) left-to-right model with self-loop, c) left-to-right model with self-loop and one-skip, d) an example for speech recognition[15]

5 継続時間長の分布を考慮した隠れマルコフモデルの比較

前節で紹介した3つのモデルについて、モデルの仮定、パラメータの推定方法、推定・推論の計算量の比較を

行う。

5.1 モデルの仮定

隠れセミマルコフモデルでは、遷移確率の分布と遷移に関連した直前状態の継続時間長の分布はそれぞれ独立に定められることを仮定している。継続時間長の分布は任意である。

VT 隠れマルコフモデルでは、遷移確率の分布が直前状態の継続時間長の関数として定められるので、隠れセミマルコフモデルとは形式的な違いがある。隠れセミマルコフモデルと同様に、継続時間長の分布は任意である。

隠れマルコフモデルの拡張の場合は隠れマルコフ過程と同じ仮定をおくが、問題に応じて遷移関係のトポロジーを工夫することにより図2のように様々な仮定を入れることができる。

5.2 パラメータの推定方式

前向き・後向きアルゴリズム（または、Baum-Welch アルゴリズム）は、潜在変数に関するパラメータを推定する EM アルゴリズムを隠れマルコフモデルに適用したアルゴリズムである。前向き・後向きアルゴリズムでは、前向き確率、後向き確率の計算を行い、その両方を使ってパラメータ推定を行う。時刻 t において隠れ状態が j のときの前向き確率 $\alpha_t(j)$ は、観測系列 o_1, o_2, \dots, o_t 、パラメータの集合を Θ とおいて、以下の式で定義される。

$$\alpha_t(j) \equiv P(o_1, o_2, \dots, o_t, X_t = j | \Theta) \quad (3)$$

隠れマルコフモデルの場合の前向き確率計算の式は以下の通りである。

$$\alpha_t(j) = \left(\sum_i \alpha_{t-1}(i) a_{ij} \right) b_j(o_t) \quad (4)$$

一方、隠れセミマルコフモデルの場合、前向き確率は以下の式のように計算する。

$$\alpha_t(j) = \sum_i \left(\sum_d \alpha_{t-d}(i) p_j(d) a_{ij} \prod_{s=t-d+1}^t b_j(o_s) \right) \quad (5)$$

また、VT 隠れマルコフモデルの前向き確率計算には複数の方式があるが、前向き確率計算は継続時間長に依存した前向き確率 $\alpha_t(i, d)$ を考える。継続時間長に依存した前向き確率 $\alpha_t(i, d)$ は、以下の式で定義される。

$$\alpha_t(j, d) \equiv P(o_1, o_2, \dots, o_t, X_t = j, d_t(j) = d | \Theta) \quad (6)$$

例えば、[16] では、前向き確率 $\alpha_t(i, d)$ は以下の式のように計算している。

$$\alpha_t(j, d) = \alpha_{t-1}(j, d+1) b_j(o_t) + \left(\sum_{i \neq j} \alpha_{t-1}(i, 1) a_{ij} \right) \cdot b_j(o_t) p_j(d) \quad (7)$$

5.3 推定・推論の計算量

前向き・後向きアルゴリズムの時間計算量のオーダーは全ての前向き確率を計算するための時間計算量のオーダーと一致する。また、フィルタリング、予測、等の計算も前向き確率計算を使って実行することができる。また、平滑化を行うアルゴリズムとして Viterbi アルゴリズムがあるが、これも同じ時間計算量となる。そこで、表1において各方式について前向き確率計算の時間計算量のオーダーを比較する。ここで、状態数を N 、事前状態数の平均を K 、観測数を T 、最大継続時間長を D 、観測データの尤度を計算するための演算回数を B とする。隠れマルコフモデルの拡張の場合は、隠れマルコフモデルに対して、トポロジーの工夫に応じて K の項の大きさに影響がある。

表 1: 推定・推論の計算量の比較

アルゴリズム	計算量
隠れマルコフモデル	$O((B+K)NT)$
隠れセミマルコフモデル	$O((B+K+D)NT)$
VT 隠れマルコフモデル ([12] の場合)	$O((B+KD)NT)$

6 むすび

本稿では、確率過程および隠れマルコフモデルについて概観するとともに、隠れ状態の継続時間長の柔軟なモデル化を可能にする3つのタイプのモデルを紹介した。どのモデルを採用するかは、モデル化の対象に想定される隠れ状態の種類や隠れ状態間の遷移関係、さらには遷移までの直前状態の継続時間について、どのような事前知識があるかによると考えられる。実際のモデル化に際しては、問題に合わせて状態遷移のトポロジーを工夫した上で、適切な状態継続時間長の分布を導入するような複合的なアプローチも有効ではないかと考えられる。

参考文献

- [1] R. Durrett. *Essentials of stochastic processes*. Springer New York, 1999 (今野他訳「確率過程の基礎」, シュプリンガー・ジャパン, 2005) .
- [2] S.M. Ross. *Stochastic Processes. 2nd*. John Wiley & Sons, 1996.
- [3] 広中平祐. 現代数理科学事典. 大阪書籍, 1991.
- [4] U.D. Nodelman. *Continuous time Bayesian networks*. PhD thesis, STANFORD UNIVERSITY, 2007.
- [5] U. Nodelman, C.R. Shelton, and D. Koller. Learning continuous time Bayesian networks. In *Proceedings of the 19th Annual Conference on Uncertainty in Artificial Intelligence (UAI-03)*, pp. 451–8, 2003.
- [6] K. Gopalratnam, H. Kautz, and D.S. Weld. Extending continuous time bayesian networks. In *Proceedings of the National Conference on Artificial Intelligence*, Vol. 20, p. 981. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005.
- [7] MT Johnson. Capacity and complexity of HMM duration modeling techniques. *IEEE Signal Processing Letters*, Vol. 12, No. 5, pp. 407–410, 2005.
- [8] J.D. Ferguson. Variable duration models for speech. In *Proceedings of the Symposium on the Application of hidden Markov models to Text and Speech*, Vol. 1, pp. 143–179, 1980.
- [9] S. Levinson. Continuously variable duration hidden Markov models for speech analysis. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'86.*, Vol. 11, 1986.
- [10] V. Barbu and N. Limnios. Semi-Markov Chains and Hidden Semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis. *Lecture Notes In Statistics*, p. 226, 2008.
- [11] YK Park, CK Un, and OW Kwon. Modeling acoustic transitions in speech by modified hidden Markovmodels with state duration and state duration-dependent observationprobabilities. *IEEE Transactions on Speech and Audio Processing*, Vol. 4, No. 5, pp. 389–392, 1996.
- [12] P. Ramesh and JG Wilpon. Modeling state durations in hidden Markov models for automatic-speech recognition. In *1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP-92.*, Vol. 1, 1992.
- [13] SV Vaseghi. State duration modelling in hidden Markov models. *Signal processing*, Vol. 41, No. 1, pp. 31–41, 1995.
- [14] M. Russell and R. Moore. Explicit modelling of state occupancy in hidden Markov models for automatic speech recognition. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'85.*, Vol. 10, 1985.
- [15] X. Wang, LFM ten Bosch, and LCW Pols. Integration of context-dependent durational knowledge intoHMM-based speech recognition. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, Vol. 2, 1996.
- [16] S.Z. Yu and H. Kobayashi. An efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE Signal Processing Letters*, Vol. 10, No. 1, pp. 11–14, 2003.