

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第5645777号
(P5645777)

(45) 発行日 平成26年12月24日(2014.12.24)

(24) 登録日 平成26年11月14日(2014.11.14)

(51) Int.Cl.

F I

H O 3 M 7/00 (2006.01)

H O 3 M 7/00

G O 6 T 7/00 (2006.01)

G O 6 T 7/00 3 5 O A

請求項の数 10 外国語出願 (全 22 頁)

(21) 出願番号	特願2011-183669 (P2011-183669)	(73) 特許権者	597067574
(22) 出願日	平成23年8月25日 (2011. 8. 25)		ミツビシ・エレクトリック・リサーチ・ラ
(65) 公開番号	特開2012-69103 (P2012-69103A)		ボラトリーズ・インコーポレイテッド
(43) 公開日	平成24年4月5日 (2012. 4. 5)		アメリカ合衆国、マサチューセッツ州、ケ
審査請求日	平成26年6月18日 (2014. 6. 18)		ンブリッジ、ブロードウェイ 201
(31) 優先権主張番号	12/872, 054		201 BROADWAY, CAMBR
(32) 優先日	平成22年8月31日 (2010. 8. 31)		IDGE, MASSACHUSETTS
(33) 優先権主張国	米国 (US)		02139, U. S. A.
早期審査対象出願		(74) 代理人	100110423
			弁理士 曾我 道治
		(74) 代理人	100094695
			弁理士 鈴木 憲七
		(74) 代理人	100111648
			弁理士 梶並 順
		最終頁に続く	

(54) 【発明の名称】 圧縮された状態シーケンスを求めるための方法

(57) 【特許請求の範囲】

【請求項 1】

圧縮復号部を有するプロセッサにより、入力データシーケンス “ x ” から出力圧縮状態シーケンス

【数 1】

\hat{s}

を求めるための方法であって、圧縮状態シーケンス “ s₀ ” の確率は、

【数 2】

$$p(s = s_0 | x) = \sum_{\forall y: \mathcal{C}(y) = s_0} p(y | x),$$

10

に従って前記入力データシーケンスの全てのセグメンテーション “ y ” を周辺化することによって与えられ、ここで、シンボル

【数 3】

\mathcal{C}

は圧縮関数を表し、該方法は、

20

前記プロセッサへ前記入力データシーケンスを入力するステップと、

$p(c = c_0 | x)$ に従う前記入力データシーケンス “ x ” = $\{x_1, x_2, \dots, x_T\}$ に対応する可能な圧縮状態シーケンスの長さ c の確率を求めるステップと、

前記出力圧縮状態シーケンス

【数 4】

$$\hat{s}$$

の長さ

【数 5】

$$\hat{c}$$

10

を推定するステップであって、前記長さ

【数 6】

$$\hat{c}$$

は

【数 7】

$$\hat{c} = \arg \max p(c = c_0 | x)$$

に従う最大確率を有するものと、

20

【数 8】

$$p(s_i = j | x, \hat{c}),$$

に従う長さ

【数 9】

$$\hat{c}$$

を有する出力圧縮状態シーケンスにおいて各状態 s_i の確率を求めるステップであって、
ここで、

30

【数 10】

$$i = 1, \dots, \hat{c}$$

であるものと、

【数 11】

$$\hat{s}_i = \arg \max p(s_i = j | x, \hat{c})$$

に従う最大確率を有する各圧縮状態

【数 12】

$$\hat{s}_i$$

40

を推定して、前記出力圧縮状態シーケンス

【数 13】

$$\hat{s}$$

を推定するステップと、

推定された前記出力圧縮状態シーケンスを前記プロセッサから出力するステップと、
を備えた、方法。

【請求項 2】

50

前記入力データシーケンスがロボットの移動環境の情報であり、前記出力圧縮状態シーケンスはロボットを追跡する、請求項 1 に記載の方法。

【請求項 3】

前記入力データシーケンスが連続する入力点間の量子化された角度であり、前記出力圧縮状態シーケンスは単一ストロークの文字を認識する、請求項 1 に記載の方法。

【請求項 4】

前記入力データシーケンスが形状コンテキスト特徴に対応する観測値であり、前記出力圧縮状態シーケンスは手書きの単語を認識する、請求項 1 に記載の方法。

【請求項 5】

前記入力データシーケンスが車両及びトレーラーの画像であり、前記出力圧縮状態シーケンスは前記車両のタイヤ及びトレーラーの数をカウントして料金を求める、請求項 1 に記載の方法。

10

【請求項 6】

前記入力データシーケンスが発話された単語であり、前記出力圧縮状態シーケンスは発話を認識する、請求項 1 に記載の方法。

【請求項 7】

前記出力圧縮状態シーケンスにおける各前記状態は確率を示すラベルを含む、請求項 1 に記載の方法。

【請求項 8】

前記出力圧縮状態シーケンスは隠れマルコフモデルを用いて求められる、請求項 1 に記載の方法。

20

【請求項 9】

前記出力圧縮状態シーケンスは条件付き確率場を用いて求められる、請求項 1 に記載の方法。

【請求項 10】

前記入力データシーケンスは時系列である、請求項 1 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、包括的にはデータの入力シーケンスから状態シーケンスを求めるための方法に関し、より詳細には圧縮状態シーケンスを求めることに関する。

30

【背景技術】

【0002】

シーケンシャルデータにラベルを割り当てることは、発話認識、計算言語学、コンピュータービジョン、及びロボット工学等の用途における一般的な問題である。たとえば、品詞を分析するとき、課題は、言語の文法構造を考えることによって単語のシーケンスにタグ付けすることであり、たとえば動詞 - 動詞 - 名詞 - 名詞 - 動詞 - 形容詞は英語において可能性が極めて低い文法シーケンスである一方、名詞 - 動詞 - 副詞はそうではない。同様に、発話認識において、単語又は音素は、特にそれらの順序に関して、基礎となる言語の或る特定の発音規則に従う。

40

【0003】

同様に、基礎となる言語の文法によって強制される構造を利用することによって手書きの文字 (character) のシーケンスに表音文字 (letter) 及び数字を割り当てることができる。これらの例において、シーケンシャルパターンが重要であり、シーケンシャルパターンを利用して大きなデータセットから情報を抽出することができる。

【0004】

そのような問題を解決するための 2 つの一般的なモデルは、隠れマルコフモデル (HMM) 及び条件付き確率場 (CRF) である。これらのモデルは非常に強力であるが、データのタイプが異なると、用途に特有の変更が必要となるので、モデルが様々な拡張されることになる。

50

【 0 0 0 5 】

たとえば、半マルコフ型 C R F は、セグメンテーション問題へのより一般的な解法であり、データのセグメントにおける非マルコフ型状態遷移を可能にし、個々のサンプルではなくシーケンスに直接ラベルを割り当てる。

【 0 0 0 6 】

別の方法は、現実的でない状態遷移を防ぐために状態持続性を有するシステムの非パラメトリック事前確率を説明している。この方法は状態持続性をもたらし、無限状態空間内で遷移確率をトレーニングすることも可能にする。上記の例において、最終状態シーケンス推測の基本課題は、正確な状態シーケンスを推定することである。

【 0 0 0 7 】

10

しかしながら、多くの用途において、これは必須の目標ではない。そうではなく、目標は状態シーケンスの何らかの確定関数を推定することである。特に、目標は、各状態における滞留時間を考慮することなく状態遷移を追跡することである。

【 0 0 0 8 】

例示的な用途では、人の動きが追跡されるが、このとき、「座る (s)」、「飛び跳ねる (j)」、「歩く (w)」、及び「走る (r)」等の状態間の正確な遷移は曖昧でありかつ重要でないが、特定の順序で発生した状態の固有のシーケンスの検出が重要である。

【 0 0 0 9 】

たとえば、人間の動きの例示的なグラウンドトゥース状態シーケンスは “ y ” = { s , s , j , j , j , w , w , r , r } であり、データの入力シーケンスは “ x ” = { x₁ , x₂ , . . . , x₉ } である。目標は、確定関数 c o m p r e s s の出力を正確に予測することであり、ここで、c o m p r e s s (“ y ”) = { s , j , w , r } である。すなわち、圧縮シーケンスにおいて重複した状態が取り除かれている。

20

【 0 0 1 0 】

さらに、j から w へちょうど遷移する時に、状態「 w 」の最初の出現を「 j 」に変換することによって、予測状態シーケンス “ y ’ ” = { s , s , j , j , j , j , w , r , r } が取得される場合、従来のアプリケーションではエラーとなるが、圧縮状態シーケンス推測を用いるアプリケーションでは、c o m p r e s s (“ y ”) = c o m p r e s s (“ y ’ ”) であるので、エラーとならない。

【 0 0 1 1 】

30

逆に、予測シーケンスが “ y ’ ’ ” = { s , s , j , j , w , j , w , r , r } である場合、1つの状態のみが “ y ” と異なる場合であっても、このアプリケーションには致命的なエラーとなる。ここで、状態遷移の曖昧性がこの問題の唯一の特徴ではなく、圧縮された出力の長さが知られておらず、かつ任意であり、たとえば人間の動作中にいくつかの固有の動作が出現順に発生したかは不明確である。

【 0 0 1 2 】

滞留時間を考えることなくシーケンス内の固有な状態をカウントすることに関心が持たれている状態カウントプロセスを含めて、状態シーケンス推測には、そのような特殊な処置を必要とするいくつかの他の問題が存在する。この発明者らの知る限り、これは、機械学習用途において大部分対処されていない問題である。

40

【 0 0 1 3 】

ビデオ解釈アプリケーションのための圧縮シーケンスが説明されている。正確な状態遷移は曖昧であり、ビデオ解釈には区別できる状態のみが重要である。この方法は、状態の数が非常に多数であることに起因して、非常に限定された領域にのみ適用可能であり、確率モデルを用いることができない。

【 0 0 1 4 】

別のビデオ解釈アプリケーションでは、単純な遷移 - コストモデルが用いられ、同じ状態への状態遷移はコストを有しないと仮定されるのに対し、全ての他の可能な遷移は同じコスト K を有すると仮定される。これは同じ状態への全ての遷移について重みゼロを有し、他の全ての遷移について重みとして同じコスト K を有する確率シーケンシャルモデルを

50

トレーニングすることと非常に類似している。これは、状態遷移のシーケンシャルモデリングが破壊されるので多くの用途において完全に非現実的である。

【0015】

図1は従来の圧縮状態シーケンス推測方法を示している。データの入力シーケンス“ x ” = { x_1, x_2, \dots, x_T } 101及び事前にトレーニングされたHMM/CRFパラメーター{ θ_j, μ_k } 106を所与とすると、HMM/CRF復号方法102は、データ101に対応する完全な状態シーケンス“ y ” = { y_1, y_2, \dots, y_T } 103を予測する。ここで、インデックス T のインクリメントは時間ステップである。次に、確定圧縮関数(deterministic compress function) f 104が完全状態シーケンス103に適用され、固有な状態の圧縮シーケンス“ s ” = f (“ y ”) = { s_1, s_2, \dots, s_c } 105が求められる。圧縮状態シーケンスにおいて、完全状態シーケンス103内の全ての重複する状態は除去される。

10

【発明の概要】

【発明が解決しようとする課題】

【0016】

隠れマルコフモデル(HMM)及び条件付き確率場(CRF)はシーケンシャルデータ“ x ”をモデリングするための非常に一般的な技法である。CRF及びHMMを用いる状態シーケンス推測手順によって、データの入力シーケンス“ x ”を所与として状態シーケンス“ y ”の推定が可能になる。

【0017】

20

しかしながら、多くの用途において、正確な(完全な)状態シーケンスの推定は必ずしも目標ではない。そうではなく、目標は状態シーケンス“ x ”の或る確定関数 f を推定して、圧縮状態シーケンス“ y ”を得ることである。そのような用途において、従来の状態シーケンス推測技法によって状態シーケンスを推定し、その後、該推定から関数マッピングを求めることは必ずしも最適ではない。

【課題を解決するための手段】

【0018】

したがって、この発明はデータの入力シーケンス“ x ”から圧縮状態シーケンス“ s ”を直接求める。

【0019】

30

特に、確定関数 f (“ x ”)が固有の状態遷移のみを追跡し、各状態の滞留時間を追跡しない特定の問題を考える。

【0020】

従来の圧縮状態シーケンス推測技法よりも性能が優れた多項式時間圧縮状態シーケンス推測方法を提供する。この発明が解決する問題は、ロボット追跡、文字認識、手書き認識、及び発話認識等の多くの異種の用途において一般に生じる。

【発明の効果】

【0021】

CRF及びHMMのための従来の状態シーケンス推測技法は、正確な予測に焦点を当てるが、これは目標が各状態の滞留時間を追跡することではなく、状態遷移を追跡することであるいくつかの用途では必要とされない。この発明は、この問題を解決するように明示的に設計された多項式時間状態シーケンス推測手順を提供する。

40

【図面の簡単な説明】

【0022】

【図1】従来のHMM/CRF復号及びそれに続く確定関数 f の適用のブロック図である。これによって圧縮状態シーケンスが生成される。

【図2A】この発明の実施の形態による、圧縮状態シーケンスを直接出力する圧縮HMM/CRF復号のブロック図である。

【図2B】この発明の実施の形態による、圧縮状態シーケンスを直接出力する圧縮HMM/CRF復号のブロック図である。

50

【図 3 A】この発明の実施の形態による、圧縮状態シーケンス推測方法の概略図である。

【図 3 B】この発明の実施の形態による、圧縮状態シーケンス推測方法の概略図である。

【図 3 C】この発明の実施の形態による、圧縮状態シーケンス推測方法の概略図である。

【図 4 A】この発明の実施の形態による、ロボットロケーション追跡アプリケーションの概略図である。

【図 4 B】この発明の実施の形態による、ロボットロケーション追跡アプリケーションの概略図である。

【図 4 C】時間の関数としての動きのグラフである。

【発明を実施するための形態】

【 0 0 2 3 】

10

図 2 A 及び図 2 B は、この発明の実施の形態による圧縮状態シーケンス推測方法を示している。データの入力シーケンス “ x ” = { x₁, x₂, . . . , x_T } 201 を所与とする。ここでデータは T によってインデックス付けされている。たとえば、シーケンスは時系列である。圧縮復号方法 202 は、トレーニングされた HMM モデル又は CRF モデルのパラメーター { λ_j, μ_k } 204 を用いて、推定圧縮状態シーケンス “ s ” = { s₁, s₂, . . . , s_c } 203 を直接出力する。本方法のステップは、当該技術分野において既知のようにメモリ及び入力／出力インターフェースに接続されたプロセッサ 200 において実行することができる。

【 0 0 2 4 】

条件付き確率場及び状態シーケンス推測技法

20

直鎖条件付き確率場 (CRF) 及び隠れマルコフモデル (HMM) は、シーケンシャルデータをラベル付けするのに用いることができる 2 つの確率モデルである。データがラベル付けされた後、状態シーケンスを求めることができる。選択された特徴関数に依拠して直鎖 CRF 及び HMM を交換可能に用いることができる。この発明は CRF の手順に焦点を当てる。

【 0 0 2 5 】

しかしながら、手順は HMM にも等しく適用可能である。この発明の方法にとって、いずれの特定のモデルを用いるかは重要ではない。すなわち、結果は一方から他方に一般化することができる。本明細書では直鎖 CRF を用いるが、本方法は HMM にも適用可能である。

30

【 0 0 2 6 】

従来の条件付き確率場

条件付き分布 p (“ x ” | “ y ”) をモデル化するための直鎖 CRF を以下のように定義することができる。

【 0 0 2 7 】

【数 1】

$$p(y|x) = \frac{1}{Z(x)} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t), \quad (1)$$

40

$$\Psi(y_t, y_{t-1}, x_t) = \exp \left\{ \sum_j \lambda_j g_j(y_{t-1}, y_t, x) + \sum_k \mu_k u_k(y_t, x) \right\}, \quad (2)$$

【 0 0 2 8 】

ここで、 Ψ はポテンシャル関数であり、 g は或る状態から次の状態への遷移特徴関数 (transition feature function) であり、 u は特徴関数であり、 λ 及び μ はトレーニングプロセスによって推定されるパラメーターであり、 Z は入力データシーケンスの関数である

50

正規化係数である。

【0029】

従来の状態シーケンス推測技法

データシーケンスの従来のラベル付けは、同時密度 " y^* " = $\arg \max_y p("y" | "x")$ を用い、ここで関数 " $\arg \max$ " は最大値を返す。解は、再帰 $t(j) = \max_i (j, i, x_t)_{t-1}$ を用いてビタビ手順により求めることができる。これは、最大積法則に基づいて最可能経路を伝播する。しかしながら、多くの用途において、完全なシーケンスを正確に予測することは非常に困難であるので、個々の予測が用いられる。これは、Forward-Backward (前向き後ろ向き) と呼ばれる動的プログラミング手順を適用することによって、周辺分布 $p(y_{i,t} | "x")$ から $y_{i,t}$ を予測することにより達成することができる。前方再帰は $t(j) = i(j, i, x_t)_{t-1}(i)$ であり、ここで $t(j)$ は前方変数であり、後方再帰は $t(i) = i_{t+1}(j, i, x_{t+1})_{t+1}(j)$ であり、ここで $t(i)$ は後方変数である。次に、これらの変数を用いることによって周辺確率を求めることができる。

10

【0030】

制約付き前方手順を用いて、前方再帰を制約付き状態シーケンスに制約することによって特定の状態シーケンスにおける信頼度を求めることができる。制約集合 $Y' = \{y_q, \dots, y_r\}$ を所与とすると、変更された前方手順を用いて任意のデータシーケンスが Y' を満たす確率を求めることができる。

20

【0031】

変更された前方再帰は、全ての $y_{t+1} \in Y'$ について、

【0032】

【数2】

$$\hat{\alpha}_t(j) = \begin{cases} \sum_i \Psi(j, i, x_t) \hat{\alpha}_{t-1}(i) & j \simeq y_{t+1} \text{ の場合} \\ 0 & \text{その他の場合} \end{cases} \quad (3)$$

【0033】

であり、ここで、演算子

30

【0034】

【数3】

$$j \simeq y_{t+1}$$

【0035】

は " j は制約 y_{t+1} に従う" として定義される。インデックス T において、特定の制約の信頼度は Z' / Z である。ここで、制約付き格子因子

【0036】

【数4】

$$Z' = \sum_i \hat{\alpha}_T(i)$$

40

【0037】

は制約付き前方変数を用いて求められ、制約なし格子因子 $Z = \sum_i \alpha_T(i)$ は制約なし前方変数を用いて求められる。同様に、周辺確率 $p(y_{i,t} | "x")$ を推定している間、集合 Y' 内に単一の制約 $y_{i,t}$ を有することによって同じ方法が辿られる。

【0038】

圧縮状態シーケンス推測

この発明の実施の形態は、取得データから状態の圧縮シーケンスを推測するための方法を提供する。人間の動きを予測する1つの用途例では、状態の圧縮シーケンスを、" s " = $\text{compress}("y")$ として定義する。ここで、 s 、 j 、 w 、 r は任意の状態で

50

ある。“y” = { s , s , j , j , j , w , w , r , r } である場合、“s” = c o m p r e s s (“ y ”) = { s , j , w , r } である。これ以降、シンボル

【 0 0 3 9 】

【 数 5 】



【 0 0 4 0 】

を用いて関数 c o m p r e s s を表す。この発明の目標は、圧縮シーケンス “ s ” を可能な限り正確に予測することである。確率 $p(“s” | “x”)$ を求めるための枠組みを構築する。

10

【 0 0 4 1 】

圧縮状態シーケンス “ s₀ ” の確率が全てのセグメンテーション “ y ” にわたる周辺化によって与えられる。ここで圧縮表現 “ s₀ ” の確率は、

【 0 0 4 2 】

【 数 6 】

$$p(s = s_0 | x) = \sum_{\forall y: \mathcal{C}(y) = s_0} p(y | x) \quad , \quad (4)$$

20

【 0 0 4 3 】

であり、シンボル

【 0 0 4 4 】

【 数 7 】



【 0 0 4 5 】

は圧縮関数を表す。

【 0 0 4 6 】

確率 $p(“y” | “x”)$ を求めることは効率的であるが、式 (4) の総和は指数関数的な演算数 M^T を含む。ここで M は状態数であり、 T は完全状態シーケンス “ y ” の長さであり、これは扱い難い。したがって、動的プログラミング前向き後ろ向き手順を用いて $p(“s” = “s_0” | “x”)$ を求めるための多項式時間手順を提供する。

30

【 0 0 4 7 】

ここでこの発明の圧縮状態シーケンス推測方法の基本ステップを記述する。

【 0 0 4 8 】

確率 $p(“s” = “s_0” | “x”)$ を求める

式 (4) において、圧縮シーケンス “ s₀ ” に対応する指数関数的な個数の可能な完全状態シーケンス “ y ” が存在し得る。この問題を克服するために、前方再帰を適用しながらシーケンス “ s ” の遷移を追跡することを説明する。

40

【 0 0 4 9 】

図 3 A に示すように、シーケンス “ s₀ ” の長さは $c = | “s_0” |$ である。表記を容易にするために、“ s₀ ” の個々の項を s_i と呼ぶ。ここで、 $i = 1, \dots, c$ である。シーケンス “ s₀ ” 内の正確に $i - 1$ 個の遷移を t まで追跡する前方変数

【 0 0 5 0 】

【 数 8 】

$$\bar{\alpha}_t(i)$$

【 0 0 5 1 】

を定義する。インデックス $t - 1$ からインデックス t まで、前方変数は以下に基づいて更

50

新される。

入力データ x_t について同じ状態に留まり（矢印 3 1 1 で示される）、

状態 s_i から $s_i + 1$ への遷移を行う（矢印 3 1 2 で示される）。

【 0 0 5 2 】

動的プログラミング法は以下の通りである。

【 0 0 5 3 】

シーケンス “ s_0 ” の確率 $p(“s” = “s_0” | “x”)$ は

【 0 0 5 4 】

【数 9】

$$p(s = s_0 | x) = \sum_{\forall y: \mathcal{C}(y) = s_0} p(y | x) \propto \sum_{\forall y: \mathcal{C}(y) = s_0} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t),$$

10

【 0 0 5 5 】

である。これは再帰

【 0 0 5 6 】

【数 1 0】

$$\bar{\alpha}_t(i) = \Psi(s_i, s_{i-1}, x_t) \bar{\alpha}_{t-1}(i-1) + \Psi(s_i, s_i, x_t) \bar{\alpha}_{t-1}(i),$$

20

【 0 0 5 7 】

によって求められる。ここで、 $i = 1, \dots, c$ である。

【 0 0 5 8 】

インデックス（時刻） T において、

【 0 0 5 9 】

【数 1 1】

$$\bar{\alpha}_T(c) = \sum_{\forall y: \mathcal{C}(y) = s_0} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t).$$

30

【 0 0 6 0 】

を得る。

【 0 0 6 1 】

この再帰によって、格子正規化係数（lattice normalization factor）

【 0 0 6 2 】

【数 1 2】

$$Z(s_0) = \bar{\alpha}_T(c)$$

【 0 0 6 3 】

を求める。正規化係数 Z は、 $p(“s” = “s_0” | “x”) = Z(“s_0”) / Z$ によって $Z(“s_0”)$ を確率に変換する。正規化係数 Z の計算は以下に説明される。この方法の証明が付録において与えられている。

【 0 0 6 4 】

$p(c = c_0 | “x”)$ を求める

データ “ x ” の入力シーケンスを所与とすると、圧縮シーケンス “ s ” を求める最初のステップは、シーケンス “ s ” の長さ c を求めることである。ここで、 c は、状態遷移が存在しないことを意味する 1 から、最大で、全ての 1 つ 1 つのインクリメントにおいて遷移が存在することを意味するシーケンス長 T までの値を取ることができる。全ての $c > T$ について、 $p(c = c_0 | “x”) = 0$ である。長さ i の圧縮シーケンスの集合は “ S_{\perp} ”

50

”であり、すなわち $i = 1, \dots, T$ について “ S_i ” = { “ s ” : | “ s ” | = i }
である。 $i \neq j$ について

【 0 0 6 5 】

【 数 1 3 】

$$S_i \cap S_j = \emptyset,$$

【 0 0 6 6 】

である場合、確率 $p(c = c_0 | \text{“ } x \text{”})$ は

【 0 0 6 7 】

【 数 1 4 】

10

$$p(c = c_0 | x) = p(s \in S_{c_0} | x) = \sum_{\forall s': |s'| = c_0} p(s = s' | x) \propto \sum_{\forall s': |s'| = c_0} \sum_{\forall y: \mathcal{C}(y) = s'} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t) \quad (5)$$

【 0 0 6 8 】

である。

【 0 0 6 9 】

長さ c の 1 つの可能なシーケンス “ s_0 ” の確率は $p(\text{“ } s \text{”} = \text{“ } s_0 \text{”} | \text{“ } x \text{”})$ である。図 3 B に示すように、2 つの可能な圧縮シーケンス “ s_1 ” 及び “ s_2 ” を有する
場合、

20

【 0 0 7 0 】

【 数 1 5 】

$$p(s = s_1 | x) + p(s = s_2 | x) \propto \bar{\alpha}_T(c)_{s1} + \bar{\alpha}_T(c)_{s2}$$

【 0 0 7 1 】

である。ここで、

【 0 0 7 2 】

【 数 1 6 】

$$\bar{\alpha}_T(c)_{si}$$

30

【 0 0 7 3 】

は、シーケンス s_i に前方再帰が適用されたことを意味する。しかしながら、これらの 2
つのシーケンスは、圧縮領域の 1 つの点においてのみ異なる。同じ格子によってシーケン
スを表し、複数の計算を回避することを可能にするために、テーブルに

【 0 0 7 4 】

【 数 1 7 】

$$\bar{\alpha}_t$$

【 0 0 7 5 】

40

のベクトル表現を表す。前方変数 $\bar{\alpha}_t$ は M 次元であり、以前の

【 0 0 7 6 】

【 数 1 8 】

$$\bar{\alpha}_t$$

【 0 0 7 7 】

は c 次元である。この発明の

【 0 0 7 8 】

【数 1 9】

$$\hat{\alpha}_t$$

【 0 0 7 9】

は $c \times M$ 次元である。

【 0 0 8 0】

格子に対する全ての制約の集合は、 $Q = \{ \dots, q_{i-1}, q_i, q_{i+1}, \dots \}$ であり、ここで各制約 q_i はテーブルの非ゼロエントリの 1 組の座標である。たとえば、特定の圧縮シーケンス “ s_0 ” の場合、集合は $\{ (1, s_1), (2, s_2), \dots, (c, s_c) \}$ に対応する。特定の集合 “ S_i ” の場合、この集合は高さ I を有するテーブルの全ての座標に対応し、

10

【 0 0 8 1】

【数 2 0】

$$Q_{S_i}$$

【 0 0 8 2】

で表される。所与の制約集合 Q のための再帰は以下の通りである。

【 0 0 8 3】

【数 2 1】

$$\hat{\alpha}_t(i, j)_Q = \begin{cases} \Psi(j, j, x_t) \hat{\alpha}_{t-1}(i, j)_Q + \sum_{k: k \neq j} \Psi(j, k, x_t) \hat{\alpha}_{t-1}(i-1, k)_Q & (i, j) \in Q \text{ の場合} \\ 0 & \text{その他の場合} \end{cases}$$

20

. (6)

【 0 0 8 4】

図 3 B に示すように、この再帰はテーブル内の全ての非ゼロエントリを通して伝播し、全てのゼロエントリを無視する。なぜなら、ゼロエントリは格子 Q 内に含まれないためである。

30

【 0 0 8 5】

図 3 C は、制約がインデックス $t-1$ から t 及び t から $t+1$ の格子の全てのロケーションに含まれる再帰を示している。

【 0 0 8 6】

式 (6) の再帰は、集合 Q によって以下のように定義される全ての圧縮シーケンスの確率を求める。

【 0 0 8 7】

【数 2 2】

$$p(Q|\mathbf{x}) \propto Z(Q) = \sum_j \hat{\alpha}_T(c_0, j)_Q$$

40

. (7)

【 0 0 8 8】

制約集合表記

【 0 0 8 9】

【数 2 3】

$$Q_{S_{c_0}}$$

【 0 0 9 0】

を用いると、確率 $p(c = c_0 | \mathbf{x})$ は、

50

【 0 0 9 1 】

【 数 2 4 】

$$p(c=c_0|\mathbf{x}) \propto \sum_{\forall s: |s|=c_0} \sum_{\forall y: \mathcal{C}(y)=s} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t) = Z(Q_{s_{c_0}}) = \sum_j \hat{\alpha}_T(c_0, j) Q_{s_{c_0}}. \quad (8)$$

【 0 0 9 2 】

である。

10

【 0 0 9 3 】

これは、制約集合

【 0 0 9 4 】

【 数 2 5 】

$$Q_{s_{c_0}}$$

【 0 0 9 5 】

を用いて式 (6) において再帰を適用し、行 c_0 におけるエントリの和を求めることに
 対応する。上述したように、 $c_0 > T$ 又は $c_0 < 1$ のとき、 $p(c=c_0 | "x") = 0$ で
 ある。この手順を制約集合

20

【 0 0 9 6 】

【 数 2 6 】

$$Q_{s_T}$$

【 0 0 9 7 】

に対して実行する場合、テーブル

【 0 0 9 8 】

【 数 2 7 】

$$\hat{\alpha}_T(i, j) Q_{s_T}$$

30

【 0 0 9 9 】

の行の和によって全ての格子因子

【 0 1 0 0 】

【 数 2 8 】

$$Z(Q_{s_i})$$

【 0 1 0 1 】

が同時に得られる。このテーブルの総和は正規化係数 Z に等しく、これは $p("s" = "s_0" | "x") = Z("s_0") = Z$ 及び

【 0 1 0 2 】

【 数 2 9 】

40

$$p(c=c_0|\mathbf{x}) = Z(Q_{s_{c_0}})/Z$$

【 0 1 0 3 】

を求めるのに必要である。この識別情報は、“ s ”の全ての可能な長さ及び組み合わせに
 ついて、 Z が格子因子の総和に等しいということから得られる。

【 0 1 0 4 】

$p(s_i = j | "x", c)$ を求める

周辺分布 $p(s_i = j | "x", c)$ を求めるために、行 i におけるエントリを除いて

50

、高さ c を有するテーブルの全てのエントリを含めることによって、制約集合 $Q_{i,j}$ を構築する。次に、この集合に (i, j) を加える。この特定の制約集合構成は、長さ c 及び $s_i = j$ の全ての可能な圧縮シーケンス構成を含む。このとき、周辺確率は、 $p(s_i = j | "x", c) = Z(Q_{i,j}) / \sum_j Z(Q_{i,j})$ である。

【0105】

圧縮シーケンス

次に、上記のステップを適用し、以下を最大にすることによって、圧縮状態シーケンスが求められる。

【0106】

【数30】

$$\hat{c} = \arg \max_c p(c = c_0 | x)$$

【0107】

によって c を推定し、

【0108】

【数31】

$$\hat{s}_j = \arg \max_j p(s_j = j | x, \hat{c})$$

【0109】

によって s_j を推定する。

【0110】

データから状態シーケンスを生成するための方法

上記に続いて、図2Bは方法202をより詳細に示している。

最初に、 $p = (c | "x")$ に従って入力データシーケンス $"x" = \{x_1, x_2, \dots, x_T\}$ 201 に対応する、可能な圧縮状態シーケンスの長さ c 261 の確率を求める(251)。ここで、 $i = 1, \dots, T$ である。

次に、出力圧縮状態シーケンス

【0111】

【数32】

$$\hat{s}$$

【0112】

203の長さ

【0113】

【数33】

$$\hat{c}$$

【0114】

262を推定する(252)。ここで、長さ

【0115】

【数34】

$$\hat{c}$$

【0116】

は $\arg \max p(c | "x")$ に従う最大確率を有し、関数 $\arg \max$ は最大確率を返す。

次に、

【0117】

10

20

30

40

【数 3 5】

$$p(s_j | \mathbf{x}, \hat{c})$$

【 0 1 1 8】

に従う長さ

【 0 1 1 9】

【数 3 6】

$$\hat{c}$$

【 0 1 2 0】

を有する出力圧縮状態シーケンス

10

【 0 1 2 1】

【数 3 7】

$$\hat{s}$$

【 0 1 2 2】

内の各状態 s_j 2 6 3 の確率を求める (2 5 3)。ここで、

【 0 1 2 3】

【数 3 8】

$$j = 1, \dots, \hat{c}$$

20

【 0 1 2 4】

である。

そして最後に、

【 0 1 2 5】

【数 3 9】

$$\arg \max p(s_j | \mathbf{x}, \hat{c})$$

【 0 1 2 6】

に従う最大確率を有する各圧縮状態

【 0 1 2 7】

30

【数 4 0】

$$\hat{s}_j$$

【 0 1 2 8】

を推定して (2 5 4)、出力圧縮状態シーケンス

【 0 1 2 9】

【数 4 1】

$$\hat{\mathbf{s}},$$

【 0 1 3 0】

40

を推定する。これらのステップはプロセッサにおいて実行される。

【 0 1 3 1】

応用形態

この発明の圧縮状態シーケンス推測方法の以下の応用形態を説明する。

【 0 1 3 2】

移動ロボットのロケーション追跡

移動ロボットへの応用形態において、ロボットはグリッドを用いて環境内で追跡される。特に環境を発掘する目的で、シーケンシャルモデルが頻繁に用いられる。この応用形態では、正確なグリッド反復は重要でなく、ロボットが順番に横切ったロケーションを示す圧縮バージョンが重要である。

50

【 0 1 3 3 】

例示的なロボット応用形態について図 4 A に示されているように、ロボットは該ロボットが部屋の色（青（B）、緑（G）、黄（Y）、赤（r））、及びそのロケーションへの動きを妨げる障害物（ブロック）を検知することができる環境内を移動する。

【 0 1 3 4 】

全てのステップにおいて、ロボットは移動{上、下、左、右}を試みる。ここでロボットは方向をランダムに選択する。意図される方向に障害物が存在する場合、ロボットは再度試行する。

【 0 1 3 5 】

図 4 B は、ロボットの一例示の経路 4 0 1 を示している。この問題において、状態は（ x, y ）座標におけるロボットのロケーションを指し、観測値は観測される色を指し、これはセンサー、たとえばカメラの出力である。

10

【 0 1 3 6 】

図 4 C は、図 4 B における経路 4 0 1 について、観測される色の観点で、時間の関数として動きを示している。不都合なことに、色検出的中率は不完全である可能性があり、たとえば平均 8 0 % である。この場合、ロボットは異なるロケーションにいるにも関わらず、図 4 C の雑音 4 0 2 によって示すようにセンサー出力は誤ったものとなる。しかしながら、圧縮状態シーケンス推測方法は、カメラ観測値のシーケンスに関して、ロボットが順番に横切る正しいロケーションを予測する。

【 0 1 3 7 】

20

単一ストローク文字認識

この応用形態は、スマートフォン、携帯情報端末（PDA）、及びタブレットPC等の、タッチスクリーンを用いて文字及びグリフの入力を可能にする電子デバイスと共に用いることができる。一般に、問題は、通常 1 つの連続ストロークでタッチスクリーン上に描かれた形状を認識することである。

【 0 1 3 8 】

この応用形態では、状態は方向{上、右、下、左}を指し、観測値は連続する入力点間の量子化された角度を指す。一般に、現在のシステムは、この目的のために確率的有限状態マシン又はHMMを用いる。通例、HMM又は有限状態マシンは単一の文字に関してトレーニングされる。次に、モデルを用いたデータ尤度試験によって 1 つのクラスが選択される。この方法の 1 つの欠点は、任意の形状を扱う能力が限られていることである。全ての文字について単一のモデルをトレーニングして、そのモデルを用いることによって状態を復号することができる。それにもかかわらず、単一の状態誤りによって予測全体が誤ったものとなる可能性がある。このため、雑音に耐性のある状態シーケンス推測方法を提供する。さらに、1 つの状態から別の状態への移行は通常曖昧であるので、状態遷移における曖昧性も問題である。

30

【 0 1 3 9 】

しかしながら、この発明の方法は正確な遷移を推定しない。その代わり、この発明の方法は、このタスクには十分である遷移シーケンスを生成する。観測値、たとえばユーザーインターフェースから取得される連続点間の量子化された角度を所与とすると、この発明の圧縮状態シーケンス推測方法は、全ての方向を順番に生成し、これは特定の文字を示す。

40

【 0 1 4 0 】

手書きの単語認識

手書きの認識は通常、最初に手書きをセグメント化し、次にそれらのセグメントをサポートベクターマシン（SVM）等のマルチクラス分類器を用いて認識することによって実行される。多くの場合に、言語の文法も用いられる。しかしながら、セグメント化は前処理として行われることが通常想定される。

【 0 1 4 1 】

この発明の方法によれば、文字はセグメント化されず、シーケンスの長さは知られるこ

50

とも必要とされることもない。なぜなら、手書きの単語は異なる長さを有する可能性があるためである。入力の手書きは画像の形態をとる。この応用形態では、各インクリメントは手書きの画像内の1つの垂直列に対応する。状態は対応するセグメントを指し、観測値は形状コンテキスト特徴に対応する。

【0142】

特徴を抽出する間、まずスライディングウィンドウ技法を用いることによって画像から重複するパッチを取る。画像内の列毎に1つのパッチが取られる。次に、形状コンテキスト記述子を適用する。次に、K平均クラスタリングを適用して、トレーニングデータを用いて辞書をトレーニングする。また、トレーニングされた辞書内のパッチ毎にヒストグラムを生成する。トレーニングされた辞書は、パッチから抽出された各点を該辞書内のクラスター中心にマッピングする。CRFモデルはインクリメント毎の観測値としての形状コンテキスト特徴のヒストグラムを用いてトレーニングされ、状態はその観測値を生成する文字に対応する。

10

【0143】

分類している間、同じマッピング手順を実行し、トレーニングセットとの共有空間、すなわちヒストグラム空間を有するようにする。試験を通じたスライディングウィンドウを通じて求められた観測値、たとえば形状コンテキスト特徴のヒストグラムを所与とすると、圧縮状態シーケンス推測は、シーケンス内の所与の単語における文字を予測する。

【0144】

料金収受

20

電子料金収受(ETC)の応用形態において、目標は、カメラ又は3Dセンサーを用いて車両のタイヤ及びトレーラーの数をカウントすることである。圧縮状態シーケンス推測方法は、経時的にこれらの観測値を用いて、全ての状態、たとえば車体、タイヤ、牽引棒を推定し、これによって、料金を求めるための、車両のタイヤ及びトレーラーの数に関する情報がもたらされる。

【0145】

発話認識

発話認識において、HMM、並びにビタビ及びForward-Backward等の従来の状態シーケンス推測方法は、多くの場合に、発話された単語をテキストに変換するために用いられる。オーディオ信号の基礎にある状態は音素であり、一般的に、2つの状態間の正確な遷移点は曖昧である。しかしながら、発話された単語の速度に基づいて、オーディオ信号の長さは同じ単語について変動する可能性がある。圧縮状態シーケンス推測は、正確な遷移点を見つけることなく発話された単語を音素に復号する。これは発話認識には十分な情報である。

30

【0146】

付録

証明。一般性を損なうことなく、 $s_{0,1} = 1$ 、 $s_{0,2} = 2$ 、 $s_{0,3} = 3$ 、 \dots 、 $s_{0,c} = c$ とし、 t_1 、 t_2 から t_{c-1} までを状態遷移時間とする。すなわち t_1 は $s_{0,1} = 1$ から $s_{0,2} = 2$ への遷移である。

【0147】

40

【数 4 2】

$$p(s=s_0|x) = \sum_{\forall y: \mathcal{C}(y)=s_0} p(y|x) \quad (11)$$

$$\propto \sum_{\forall y: \mathcal{C}(y)=s_0} \prod_{t=1}^T \Psi(y_t, y_{t-1}, x_t) \quad (12)$$

$$= \sum_{0 < t_1 < t_2 < \dots t_{c-1} \leq T} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \quad (13)$$

$$\left(\prod_{t=t_2+1}^{t_3-1} \Psi(3, 3, x_t) \right) \Psi(4, 3, x_{t_3}) \dots \Psi(c, c-1, x_{t_{c-1}}) \left(\prod_{t=t_{c-1}+1}^T \Psi(c, c, x_t) \right) \\ = \Psi(c, c-1, x_T) \left\{ \sum_{0 < t_1 < t_2 < \dots t_{c-2} \leq T-1} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \right. \quad (14)$$

$$\left. \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \dots \Psi(c-1, c-2, x_{t_{c-2}}) \left(\prod_{t=t_{c-2}+1}^{T-1} \Psi(c-1, c-1, x_t) \right) \right\}$$

$$+ \Psi(c, c, x_T) \left\{ \sum_{0 < t_1 < t_2 < \dots t_{c-1} \leq T-1} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \right. \quad (15)$$

$$\left. \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \dots \Psi(c, c-1, x_{t_{c-1}}) \left(\prod_{t=t_{c-1}+1}^{T-1} \Psi(c, c, x_t) \right) \right\}$$

【0 1 4 8】

式(13)において、分配法則を用いることによって式(12)を書き換える。式(14)において、単に時間Tを計算から除外することによって総和を2つのケースに分割する。(1)第1の部分は、時刻Tにおいてc-1からcへの遷移が存在するケースを考え、(2)第2の部分は時刻Tにおいて遷移が存在せず、このためc-1からcへの遷移がTより前であり、時刻Tにおいて以前の状態cが反復され则认为する。次に、(オーバーバー) $\tau(c)$ としてのs領域の前方変数を以下のように定義する。

【0 1 4 9】

10

20

30

40

【数 4 3】

$$\begin{aligned} \bar{\alpha}_T(c) = & \sum_{0 < t_1 < t_2 < \dots < t_{c-1} \leq T} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \\ & \left(\prod_{t=t_2+1}^{t_3-1} \Psi(3, 3, x_t) \right) \Psi(4, 3, x_{t_3}) \dots \Psi(c, c-1, x_{t_{c-1}}) \left(\prod_{t=t_{c-1}+1}^T \Psi(c, c, x_t) \right) \end{aligned} \quad (15)$$

【0 1 5 0】

このとき、式(14)における2つの総和項をこれらの前方変数に関して以下のように書き換えることができることが明らかである。 10

【0 1 5 1】

【数 4 4】

$$\begin{aligned} \bar{\alpha}_{T-1}(c-1) = & \left\{ \sum_{0 < t_1 < t_2 < \dots < t_{c-2} \leq T-1} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \right. \\ & \left. \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \dots \Psi(c-1, c-2, x_{t_{c-2}}) \left(\prod_{t=t_{c-2}+1}^{T-1} \Psi(c-1, c-1, x_t) \right) \right\} \end{aligned} \quad (16)$$

$$\begin{aligned} \bar{\alpha}_{T-1}(c) = & \left\{ \sum_{0 < t_1 < t_2 < \dots < t_{c-1} \leq T-1} \left(\prod_{t=1}^{t_1-1} \Psi(1, 1, x_t) \right) \Psi(2, 1, x_{t_1}) \right. \\ & \left. \left(\prod_{t=t_1+1}^{t_2-1} \Psi(2, 2, x_t) \right) \Psi(3, 2, x_{t_2}) \dots \Psi(c, c-1, x_{t_{c-1}}) \left(\prod_{t=t_{c-1}+1}^{T-1} \Psi(c, c, x_t) \right) \right\} \end{aligned} \quad (17) \quad 20$$

【0 1 5 2】

最後に、再帰式

【0 1 5 3】

【数 4 5】

$$\bar{\alpha}_T(c) = \Psi(c, c-1, x_T) \bar{\alpha}_{T-1}(c-1) + \Psi(c, c, x_T) \bar{\alpha}_{T-1}(c) \quad 30$$

【0 1 5 4】

を得る。

【0 1 5 5】

この証明は全ての長さcについて有効である。換言すれば、1からc-1までの信号をこの発明の信号と見なすことができる。このため、任意の長さにおいて、時刻Tにおける再帰は以下のように書くことができる。

【0 1 5 6】

【数 4 6】

$$\bar{\alpha}_T(i) = \Psi(i, i-1, x_T) \bar{\alpha}_{T-1}(i-1) + \Psi(i, i, x_T) \bar{\alpha}_{T-1}(i) \quad 40$$

【0 1 5 7】

さらに、これは、Tに戻って再帰することによって任意のt = 1, . . . , Tにも同様に一般化することができる。このため、最終形式は、

【0 1 5 8】

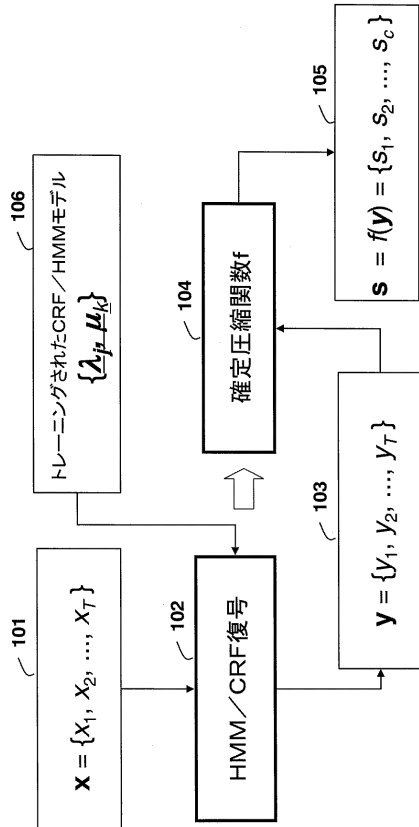
【数 4 7】

$$\bar{\alpha}_t(i) = \Psi(i, i-1, x_t) \bar{\alpha}_{t-1}(i-1) + \Psi(i, i, x_t) \bar{\alpha}_{t-1}(i)$$

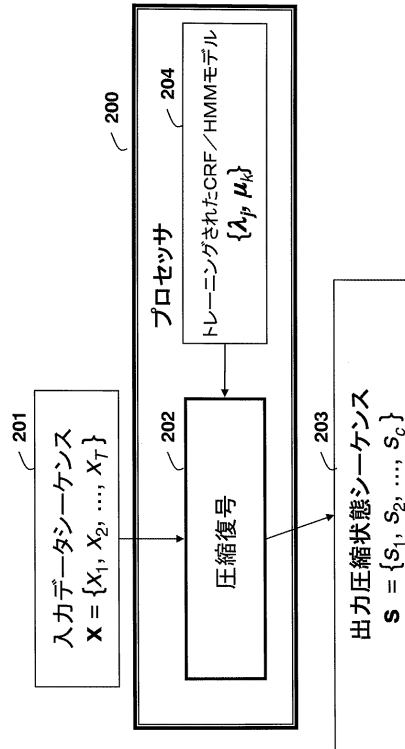
【0 1 5 9】

となる。

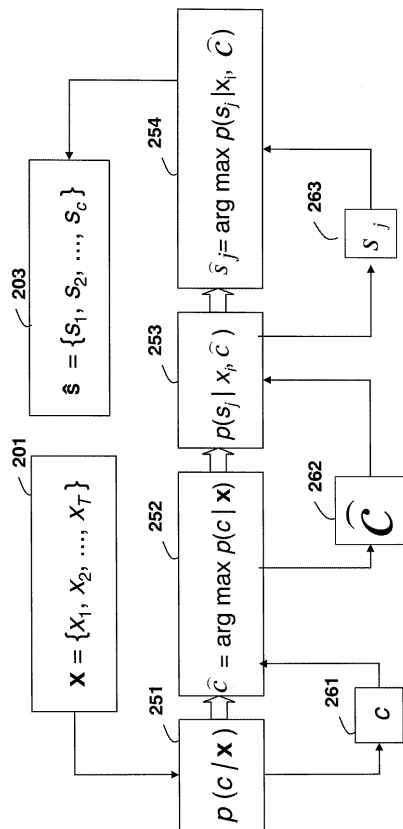
【図 1】



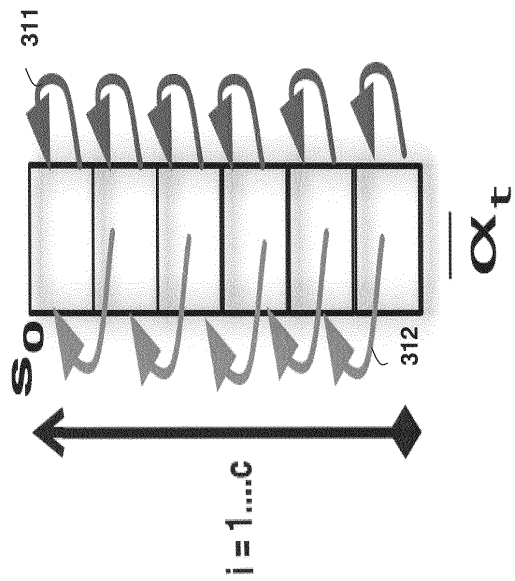
【図 2 A】



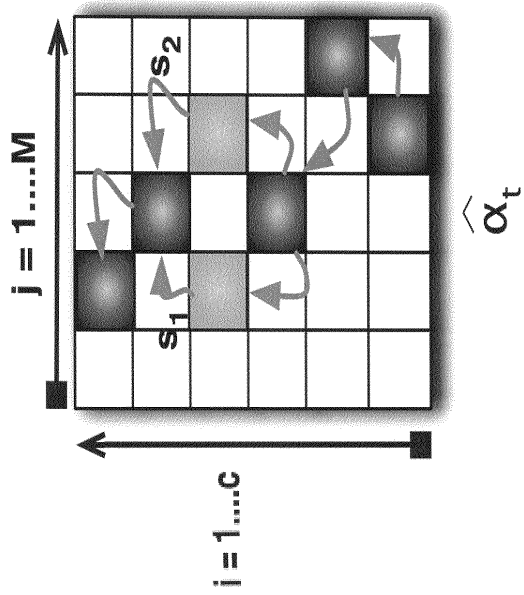
【図 2 B】



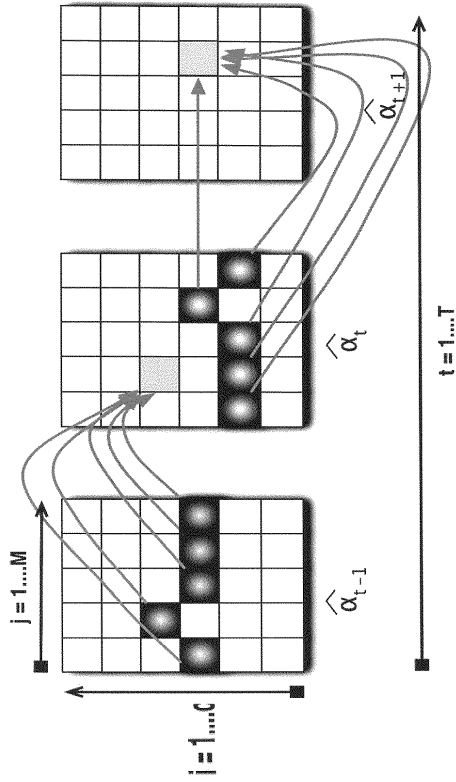
【図 3 A】



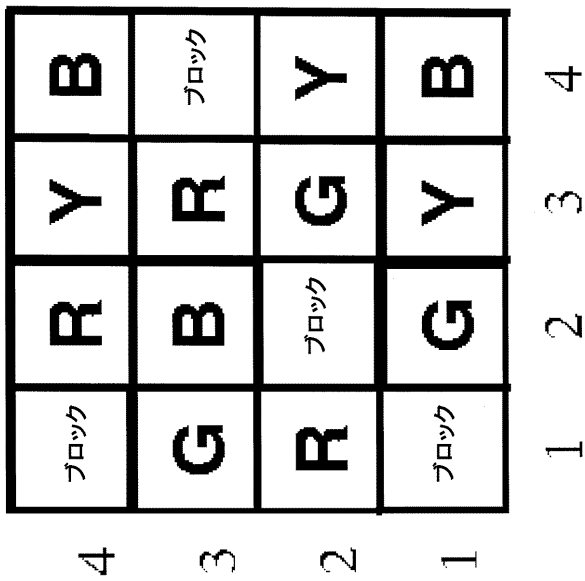
【図 3 B】



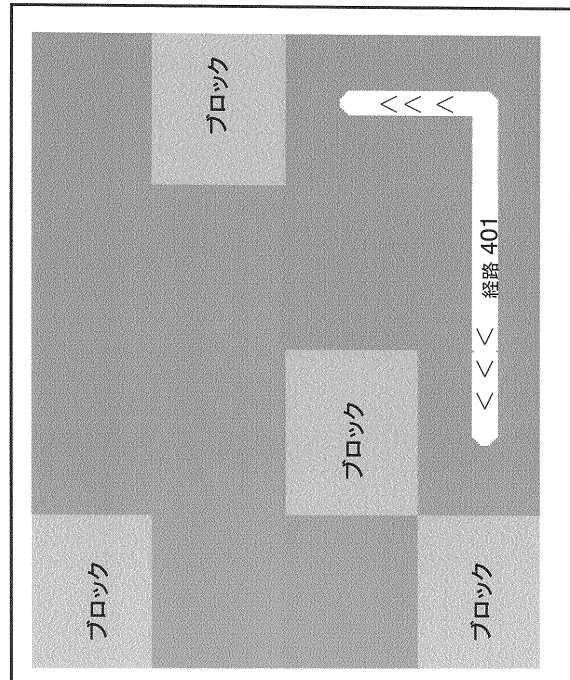
【図 3 C】



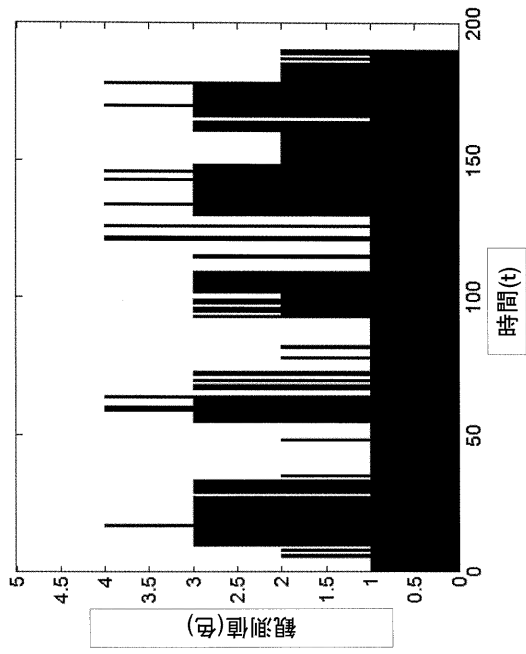
【図 4 A】



【図 4 B】



【図 4 C】



フロントページの続き

- (74)代理人 100122437
弁理士 大宅 一宏
- (74)代理人 100147566
弁理士 上田 俊一
- (74)代理人 100161171
弁理士 吉田 潤一郎
- (74)代理人 100161115
弁理士 飯野 智史
- (72)発明者 ジュネイト・オンジェル・トゥゼル
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ヒューロン・アヴェニュー 700、アパートメント 17エム
- (72)発明者 グンゴール・ポラトゥカン
アメリカ合衆国、ニュージャージー州、ピスカタウェイ、マーヴィン・レーン 29

審査官 佐田 宏史

- (56)参考文献 特開平8 - 167853 (JP, A)
特開平7 - 230294 (JP, A)
特開2009 - 276886 (JP, A)
高良 富夫、外2名, “ 遺伝的アルゴリズムを用いる多数状態マルコフモデルの状態数の決定 ”
 , 電子情報通信学会論文誌, 日本, 社団法人電子情報通信学会, 1997年 5月25日, Vol.
J80-D-II, No.5, p.1308-1310
- (58)調査した分野(Int.Cl., DB名)
H03M 7/00
G06T 7/00 - 7/60