

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-26596

(P2010-26596A)

(43) 公開日 平成22年2月4日(2010.2.4)

(51) Int.Cl.

G06N 5/04 (2006.01)

F I

G06N 5/04 550 J

テーマコード (参考)

審査請求 未請求 請求項の数 11 O L (全 20 頁)

(21) 出願番号	特願2008-184075 (P2008-184075)	(71) 出願人	000003078
(22) 出願日	平成20年7月15日 (2008.7.15)		株式会社東芝
			東京都港区芝浦一丁目1番1号
		(74) 代理人	100075812
			弁理士 吉武 賢次
		(74) 代理人	100082991
			弁理士 佐藤 泰和
		(74) 代理人	100096921
			弁理士 吉元 弘
		(74) 代理人	100103263
			弁理士 川崎 康
		(74) 代理人	100118876
			弁理士 鈴木 順生
		(72) 発明者	佐 藤 誠
			東京都港区芝浦一丁目1番1号 株式会社東芝内

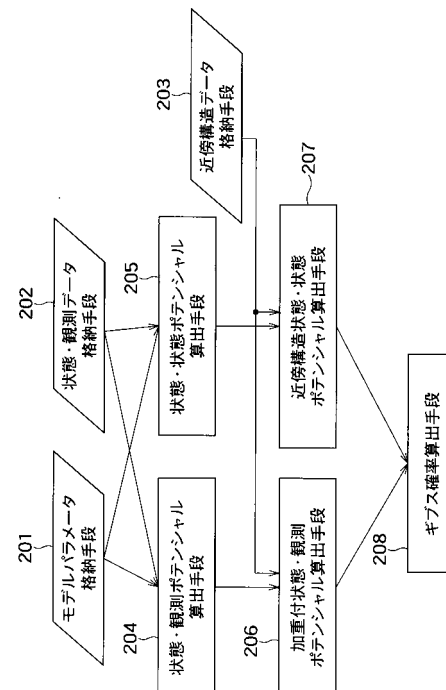
(54) 【発明の名称】 地物確率モデル装置、地物確率モデル処理方法およびコンピュータプログラム

(57) 【要約】

【課題】 地物の重みや地物間の近傍構造に基づく、地物間の依存関係を反映した混合モデルを実現する。

【解決手段】 本発明の地物確率モデル装置は、複数の地物について観測値と状態値とを格納する手段と、状態値毎に観測値の確率分布のパラメータを格納する手段と、対象地物の観測値と前記状態値の乖離度を表す状態・観測ポテンシャルを状態値毎に計算する手段と、前記状態値と前記対象地物の隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを状態値毎に計算する手段と、対象地物の重み係数を状態・観測ポテンシャルに掛け合わせ、対象地物およびその隣接地物間の隣接状況データおよび各々の重み係数を指数とする近傍構造関数に状態・状態ポテンシャルを掛け合わせ加算することにより前記状態値のエネルギーを状態値毎に計算する手段と、前記状態値毎に計算された前記エネルギーに基づきギブス分布計算を行い前記状態値毎の発生確率を計算する手段とを備える。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第 1 ～ 第 n の状態値の中から選択された状態値とを格納する状態・観測データ格納手段と、

前記第 1 ～ 第 n の状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段と、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段と、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段と、

前記複数の地物から対象地物を選択する対象地物選択手段と、

前記対象地物について前記第 1 ～ 第 n の状態値を順次、対象状態値として選択する状態選択手段と、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算する状態・観測ポテンシャル計算手段と、

前記対象状態値と、前記対象地物に隣接する対象隣接地物の状態値との関係に基づき、前記対象状態値と、前記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算する状態・状態ポテンシャル計算手段と、

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせるにより加重付状態・観測ポテンシャルを計算する加重付状態・観測ポテンシャル計算手段と、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせるにより近傍構造状態・状態ポテンシャルを計算する近傍構造状態・状態ポテンシャル計算手段と、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算するエネルギー計算手段と、

前記対象地物について、前記第 1 ～ 第 n の状態値毎に計算された前記エネルギーに基づきギブス分布計算を行うことにより前記第 1 ～ 第 n の状態値のうち少なくとも 1 つについてその発生確率を計算する確率計算手段と、

を備えた地物確率モデル装置。

【請求項 2】

前記近傍構造関数は、前記対象地物の前記重み係数と、前記対象隣接地物の重み係数とを加えたものを 2 で割った値を算出する

ことを特徴とする請求項 1 に記載の地物確率モデル装置

【請求項 3】

前記近傍構造関数は、前記対象地物の前記重み係数の平方根と、前記対象隣接地物の前記重み係数の平方根とを掛け合わせた値を算出する

ことを特徴とする請求項 1 に記載の地物確率モデル装置。

【請求項 4】

前記近傍構造データは、前記複数の地物の各々毎に、前記地物と前記隣接地物との隣接長に基づく隣接度係数を含み、

前記近傍構造関数は、前記対象地物の前記重み係数の平方根、前記対象隣接地物の前記重み係数の平方根、前記対象地物および前記対象隣接地物間の前記隣接度係数を掛け合わせた値を算出する、

ことを特徴とする請求項 1 に記載の地物確率モデル装置。

【請求項 5】

前記近傍構造データは、前記複数の地物の各々毎に、前記地物と前記隣接地物との隣接長に基づく隣接度係数を含み、

10

20

30

40

50

前記近傍構造関数は、

前記対象地物の前記重み係数と、前記対象地物および前記対象隣接地物間の前記隣接度係数とを掛け合わせたものと、

前記対象隣接地物の前記重み係数と、前記対象隣接地物と前記対象地物間の前記隣接度係数とを掛け合わせたもの

を加えた値を算出する、

ことを特徴とする請求項 1 に記載の地物確率モデル装置。

【請求項 6】

前記近傍構造データは、前記複数の地物の各々毎に、前記地物と前記隣接地物との距離を表す距離係数を含み、

前記近傍構造関数は、前記対象地物の前記重み係数と前記対象隣接地物の前記重み係数とを掛け合わせたものを、前記対象地物と前記対象隣接地物間の前記距離係数の平方で割った値を算出する、

ことを特徴とする請求項 1 に記載の地物確率モデル装置。

【請求項 7】

前記近傍構造データは、前記複数の地物の各々毎に、前記地物と前記隣接地物との隣接長に基づく隣接度係数と、前記地物と前記隣接地物との距離を表す距離係数とを含み、

前記近傍構造関数は、前記対象地物の前記重み係数、前記対象隣接地物の前記重み係数、および、前記対象地物と前記対象隣接地物間の前記隣接度係数を掛け合わせたものを、前記対象地物と前記対象隣接地物間の前記距離係数の平方で割った値を算出する、

ことを特徴とする請求項 1 に記載の地物確率モデル装置。

【請求項 8】

前記確率計算手段は、前記第 1 ～ 第 n の状態値のそれぞれについて前記発生確率を計算し、最も値の大きい発生確率を、前記対象地物の状態値として採択する

ことを特徴とする請求項 1 ないし 8 のいずれか一項に記載の地物確率モデル装置。

【請求項 9】

前記確率計算手段は、前記第 1 ～ 第 n の状態値毎に計算された前記エネルギーをそれぞれ $U_1 \sim U_n$ としたとき、前記第 n の状態値の発生確率を、

【数 1】

$$\frac{\exp(U_n)}{\sum_{x=1}^n \exp(U_x)}$$

により計算する

ことを特徴とする請求項 1 ないし 8 のいずれか一項に記載の地物確率モデル装置。

【請求項 10】

地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第 1 ～ 第 n の状態値の中から選択された状態値とを格納する状態・観測データ格納手段と、

前記第 1 ～ 第 n の状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段と、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段と、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段と、

を用意し、

前記複数の地物から対象地物を選択し、

前記対象地物について前記第 1 ～ 第 n の状態値を順次、対象状態値として選択し、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算し、

10

20

30

40

50

前記対象状態値と前記対象地物に隣接する対象隣接地物の状態値との関係に基づき、前記対象状態値と、前記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算し、

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせるにより加重付状態・観測ポテンシャルを計算し、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせるにより近傍構造状態・状態ポテンシャルを計算し、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算し、

前記対象地物について、前記第1～第nの状態値毎に計算された前記エネルギーに基づきギブス分布計算を行うことにより前記第1～第nの状態値のうち少なくとも1つについてその発生確率を計算する、

地物確率モデル処理方法。

【請求項11】

地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第1～第nの状態値の中から選択された状態値とを格納する状態・観測データ格納手段からデータを読み出すステップと、

前記第1～第nの状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段からデータを読み出すステップと、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段からデータを読み出すステップと、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段からデータを読み出すステップと、

前記複数の地物から対象地物を選択するステップと、

前記対象地物について前記第1～第nの状態値を順次、対象状態値として選択するステップと、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算するステップと、

前記対象状態値と前記対象地物に隣接する対象隣接地物の状態値との関係に基づき、前記対象状態値と、前記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算するステップと、

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせるにより加重付状態・観測ポテンシャルを計算するステップと、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせるにより近傍構造状態・状態ポテンシャルを計算するステップと、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算するステップと、

前記対象地物について、前記第1～第nの状態値毎に計算された前記エネルギーに基づきギブス分布計算を行うことにより前記第1～第nの状態値のうち少なくとも1つについてその発生確率を計算するステップと、

をコンピュータに実行させるためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

20

30

40

50

本発明は、パラメータを用いて地物に関するなんらかの変数の振る舞いをモデル化する地物確率モデル装置、地物確率モデル処理方法およびコンピュータプログラムに関する。

【背景技術】

【0002】

分析対象物の何らかの性質が数値として表すことができ、その性質が確率的な振る舞いをするとき、数値化された性質は確率変数と呼ばれる。いくつかのモデルパラメータ θ を用いて1つ以上の確率変数の振る舞いを簡潔に表すためにさまざまな確率モデルが提案されている。

【0003】

例えば、式(1)の正規分布モデルは平均(μ)と標準偏差(σ)という2つのモデルパラメータ($\theta=\{\mu, \sigma\}$)を用いて確率変数 X が値 x をとりうる確率を記述することが可能な確率モデルがある。

【数1】

$$\Pr(X=x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) \quad (1)$$

【0004】

確率モデルのパラメータは、モデル化したい性質について数値化した情報を複数の対象物について収集した学習データから決定(学習)することができる。

【0005】

正規分布やポアソン分布などの単純な確率分布を用いて複雑な確率的現象をモデル化するために、単純な確率分布(コンポーネント)を2つ以上組み合わせた混合モデルと呼ばれる確率モデルが用いられる。混合モデルにおいて、各事例の振る舞いは複数存在する単純な確率分布のいずれかによって説明できればよい。混合モデルでは変数の値に応じて適切なコンポーネントを決定してからそのコンポーネントのパラメータを用いて確率計算をすることができる。

【0006】

店舗の売り上げや地価、人口動態などの地表上で生じる現象に関する確率変数について確率モデルによって振る舞いを正確に表現できれば、対象地域における店舗の売り上げや地価の予測、人口動態の構造解析など様々な応用を行うことができる。そのためには、地理空間的な現象のための地理空間的な確率モデルが有効となる。そして、複雑な地理空間的な現象のためには地理空間的な混合モデルが有効となる。

【0007】

複雑な地理空間的な現象をモデル化するためには混合モデルを用いることが有効であるが、地理空間的な現象には空間依存性が存在する場合がある。例えば、対象エリアの各地点における機器の故障発生について数値化した故障指数という確率変数が(単純な正規分布では表現しきれないといった)複雑な振る舞いを示すため、混合モデルを用いてモデル化する場合を考える。故障指数の分布が単純な正規分布にならないのは、塩害という隠れた空間的要因が存在するためであり、海に近いエリアと海から遠いエリアとでは故障指数の確率分布が異なるからだとする。そのような場合、ある地点が塩害エリアに含まれるとき隣の地点も塩害エリアに含まれる可能性は高いという点を考慮しつつ、確率計算に用いるコンポーネントを決定しなければならない。空間依存性が存在する複雑な地理空間的な現象をモデル化する際に、データに含まれるノイズの悪影響を避けるためには、このような対象間の依存関係を考慮した混合モデルである依存混合モデルが必要になる。

【0008】

マルコフ確率場(Markov Random Field)という依存関係を考慮することが可能な確率モデルを利用した依存混合モデル装置は例えば非特許文献1に示されている。マルコフ確率場では対象間の依存関係をポテンシャル関数と呼ばれる関数で表現する。一般に、マルコフ確率場におけるポテンシャル関数はIsingモデルやPottsモデル(共に非特許文献2参照)

10

20

30

40

50

と呼ばれるものが用いられている。

【非特許文献1】 Simultaneous Parameter Estimation and Segmentation of Gibbs Random Fields Using Simulated Annealing, Sridhar Lakshmanan他, IEEE Transaction on Pattern Analysis and Machine Intelligence. vol. 11, no. 8, 1989.

【非特許文献2】 Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction, Gerhard Winkler, Springer, 2003, ISBN-10: 3540442138

【発明の開示】

【発明が解決しようとする課題】

【0009】

10

上で述べたようにマルコフ確率場を用いた依存混合モデルは画像モデル化のために既に提案されている。しかし、このモデルでは、均一な対象物が格子状に配置された場合のために提案されたポテンシャル関数であるIsingモデルやPottsモデルを用いているため、対象の重みや対象間の隣接関係(近傍構造と呼ぶ)がまちまちである場合にはそれらの情報を依存関係に反映させることができないという問題点があった。

【0010】

本発明は以上のような問題を解決するためになされたものであり、その目的は、地物の重みや地物間の近傍構造に基づく、地物間の依存関係を反映した混合モデルを実現するための、地理空間データ向け地理確率モデル装置、地物確率モデル処理方法およびコンピュータプログラムを提供することにある。

20

【課題を解決するための手段】

【0011】

本発明の一態様としての地物確率モデル装置は、

地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第1～第nの状態値の中から選択された状態値とを格納する状態・観測データ格納手段と、

前記第1～第nの状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段と、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段と、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段と、

30

前記複数の地物から対象地物を選択する対象地物選択手段と、

前記対象地物について前記第1～第nの状態値を順次、対象状態値として選択する状態選択手段と、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算する状態・観測ポテンシャル計算手段と、

前記対象状態値と前記対象地物に隣接する対象隣接地物の状態値との関係に基づき、前記対象状態値と、前記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算する状態・状態ポテンシャル計算手段と、

40

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせるにより加重付状態・観測ポテンシャルを計算する加重付状態・観測ポテンシャル計算手段と、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせるにより近傍構造状態・状態ポテンシャルを計算する近傍構造状態・状態ポテンシャル計算手段と、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算するエネルギー計算手段と、

前記対象地物について、前記第1～第nの状態値毎に計算された前記エネルギーに基づ

50

きギブス分布計算を行うことにより前記第1～第nの状態値のうち少なくとも1つについてその発生確率を計算する確率計算手段と、
を備える。

【0012】

本発明の一態様としての地物確率モデル処理方法は、
地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第1～第nの状態値の中から選択された状態値とを格納する状態・観測データ格納手段と、

前記第1～第nの状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段と、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段と、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段と、

を用意し、

前記複数の地物から対象地物を選択し、

前記対象地物について前記第1～第nの状態値を順次、対象状態値として選択し、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算し、

前記対象状態値と前記対象地物に隣接する対象隣接地物の状態値との関係に基づき、前記対象状態値と、前記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算し、

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせるにより加重付状態・観測ポテンシャルを計算し、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせるにより近傍構造状態・状態ポテンシャルを計算し、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算し、

前記対象地物について、前記第1～第nの状態値毎に計算された前記エネルギーに基づきギブス分布計算を行うことにより前記第1～第nの状態値のうち少なくとも1つについてその発生確率を計算する。

【0013】

本発明の一態様としてのコンピュータプログラムは、

地理空間上の複数の地物のそれぞれについて、前記地物から観測された観測値と、前記地物の状態を表す、第1～第nの状態値の中から選択された状態値とを格納する状態・観測データ格納手段からデータを読み出すステップと、

前記第1～第nの状態値毎に、前記観測値に関する確率分布または確率モデル、のパラメータを格納するパラメータ格納手段からデータを読み出すステップと、

前記複数の地物の重み係数をそれぞれ格納する重み係数格納手段からデータを読み出すステップと、

前記複数の地物の各々毎に、前記地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段からデータを読み出すステップと、

前記複数の地物から対象地物を選択するステップと、

前記対象地物について前記第1～第nの状態値を順次、対象状態値として選択するステップと、

前記対象状態値に対応する前記パラメータと前記対象地物の観測値とを用いて、前記対象地物の観測値と前記対象状態値の乖離度を表す状態・観測ポテンシャルを計算するステップと、

前記対象状態値と前記対象隣接地物の状態値との関係に基づき、前記対象状態値と、前

10

20

30

40

50

記対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算するステップと、

前記対象地物の重み係数を前記状態・観測ポテンシャルに掛け合わせることににより加重付状態・観測ポテンシャルを計算するステップと、

前記対象地物の重み係数と、前記対象隣接地物の重み係数と、前記対象地物および前記対象隣接地物間に関する前記隣接状況データとのうち、少なくとも前記対象地物の重み係数と、前記対象隣接地物の重み係数とを引数とした近傍構造関数に、前記状態・状態ポテンシャルを掛け合わせることににより近傍構造状態・状態ポテンシャルを計算するステップと、

前記加重付状態・観測ポテンシャルと、前記近傍構造状態・状態ポテンシャルを加算することにより前記対象状態値のエネルギーを計算するステップと、

前記対象地物について、前記第1～第nの状態値毎に計算された前記エネルギーに基づきギブス分布計算を行うことににより前記第1～第nの状態値のうち少なくとも1つについてその発生確率を計算するステップと、

をコンピュータに実行させる。

【発明の効果】

【0014】

本発明により、地物の重みや地物間の近傍構造に基づく、地物間の依存関係を反映した混合モデルを実現できる。

【発明を実施するための最良の形態】

【0015】

以下、図面に基づいて、本発明の実施の形態について説明する。

【0016】

本実施の形態は、地物の地理空間的な依存関係を考慮した依存混合モデル装置、特に、PC (Personal Computer) やエレベータの部品などの故障分析やエリアマーケティングにおける地域類型化分析などにおいて利用することが可能な、地物の重みや、地物間の隣接度合いを反映した依存混合モデル装置に関するものである。

【0017】

図2は、本発明に関わる地物確率モデル装置の一実施の形態を示した構成図である。図2に示されるように、この本発明に関わる地物確率モデル装置は、モデルパラメータ格納手段201、状態・観測データ格納手段202、近傍構造データ格納手段203、状態・観測ポテンシャル算出手段204、状態・状態ポテンシャル算出手段205、加重付状態・観測ポテンシャル算出手段206、近傍構造状態・状態ポテンシャル算出手段207、および、ギブス確率算出手段208を備えている。

【0018】

図2の各手段はたとえばプログラムモジュールとして実現することができ、この場合、各プログラムモジュールを含むプログラムを図1に示すコンピュータシステムにおいて実行することで各手段による機能を実現することができる。このコンピュータシステムには、プログラム命令を実行するCPU102、メモリ等の主記憶装置103、ハードディスク、磁気ディスク装置または光磁気ディスク装置等の外部記憶装置104、ユーザによるデータ入力を行う入力装置105、ユーザにデータ表示を行う表示装置106およびこれらを互いに接続するバス101が備わっている。プログラムは外部記憶装置104に保存されており、CPU102がこのプログラムを主記憶装置103に展開し、展開したプログラムを逐次読み出し実行することにより手段204～208が実現される。手段201～203は、主記憶装置103または外部記憶装置104によって実現される。入力装置105は、ユーザにより各種指示またはデータを入力する。表示手段106は、手段204～208により得られた演算データまたは手段201～203内のデータを表示する。

【0019】

モデルパラメータ格納手段201には、モデルパラメータとして、(1)モデル化対象の各地物について状態値と観測値との乖離度を表す状態・観測ポテンシャルを算出するため

の観測モデルパラメータ、および、(2) 各地物における状態値と、各地物の近傍地物における状態値との乖離度を表す状態・状態ポテンシャルを算出するための状態依存性パラメータとが、格納される。

【0020】

観測モデルパラメータは、状態値と観測値の違い（乖離度）をポテンシャルとして数値化するためのパラメータであり、具体的には、正規分布やポアソン分布などの確率分布のパラメータ、生存曲線など確率モデルのパラメータなどである。観測モデルパラメータは状態値毎に用意されている。各状態値のそれぞれに対応する観測モデルパラメータを用いて、観測値の確率を求め、求めた確率に応じた値をポテンシャルとして数値化することができる。

10

【0021】

また、状態依存性パラメータは状態値と状態値の違い（乖離度）をポテンシャルとして数値化するためのパラメータである。例えば、PottsモデルやIsingモデルでは1つのパラメータ β が用いられ、たとえばPottsモデルでは、2つの状態値が異なるときはポテンシャルは $-\beta$ 、同じときは0とできる。ただしこれは一例であり、パラメータは複数用いても良い。たとえば、複数のパラメータ β_{ab} , β_{bc} , β_{ca} を用意し、状態aと状態bのポテンシャルは $-\beta_{ab}$ 、状態bと状態cのポテンシャルは $-\beta_{bc}$ 、状態cと状態aのポテンシャルは $-\beta_{ca}$ 、2つの状態が同じときは0として、ポテンシャルを数値化してもよい。

【0022】

図3はモデルパラメータ（観測モデルパラメータおよび状態依存性パラメータ）の例と、観測モデルパラメータに基づく確率分布（ここでは正規分布）のグラフ表示例とを示したものである。

20

【0023】

図3の301は観測モデルパラメータの例であり、2つのコンポーネントからなる混合正規モデルのパラメータが格納されている。301において、コンポーネントMaは平均(μ)=1.0、標準偏差(σ)=1.0の正規分布であり、コンポーネントMbは $\mu=3.0$ 、 $\sigma=2.0$ の正規分布である。図3の303は、301の混合正規モデル（コンポーネントMa、Mb）の正規分布をグラフ形式で表したものである。図3の302は状態依存性パラメータの例であり、Pottsモデルの連続値係数 $\beta=0.2$ が格納されている。

【0024】

30

状態・観測データ格納手段202には、地理空間上の各地物であるサイトについて、サイトID、観測された観測変数値、状態値が格納される。状態値および観測値は連続値でも離散値でもよく、また、多次元でも良い。状態値としてはたとえば購買傾向を示す値もしくはラベルがあり、観測値としてはたとえば商品の売り上げ比率などがあり得る。

【0025】

状態値には、値が固定され変動の余地がないものと、変動の余地があるもの（不定のもの）の2種類が存在してよい。マルコフ確率場の確率計算の際に一般に用いられるICMアルゴリズムやMCMCアルゴリズム（非特許文献2）では、変動の余地のある変数の発生確率を変動の余地のない変数に対する条件付確率として反復的に算出する処理が必要になる。

【0026】

40

図5は状態・観測データの例であり、3つのサイト $\{s1, s2, s3\}$ における観測値(y)と状態値(z)が格納されている。ここで、状態値 $x^{\wedge}$ とは値がxで変動の余地がないことを表し、xとは、条件付確率を知りたい値がxで変動の余地があることを表している。変動の余地があるかどうかは、たとえば処理を行う時点で処理アルゴリズムまたは分析者が決定する。図5の例では、条件付確率、

$$P(z3=a \mid z1=a, z2=b, y=\{1.2, 4.5, 2.0\}) \quad (2)$$

という確率を計算したいことを意味している。ここで、 $P(X=x)$ は確率変数Xが値xを取る確率を表す関数である。また、 $P(X=x|Y=y)$ は確率変数Yをyに固定した条件の下で確率変数X

50

が値xを取る確率を表す関数である。

【0027】

図5においてID付けされたサイトは地理空間上の地物に対応しており、何らかの位置関係を持って配置されている。図4はサイト{s1,s2,s3}の空間配置の例であり、隣接する3つのポリゴンを表している。これらの地物の空間配置状況について幾つかの観点から特徴抽出して得られた数値を集めたものが近傍構造データである。近傍構造データ格納手段203にはこのように地物の空間配置状況を数値化したデータが格納される。

【0028】

地物の種類としては点（たとえば県庁所在地の位置など）、線（道路など）、ポリゴン等、どのようなタイプでも良い。点の場合は、たとえば所定の半径内に属する点、または最も近い上位k個の点を近傍点（隣接点）としてもよい。線またはポリゴンの場合は、実際に隣接している線またはポリゴンを隣接線または隣接ポリゴンとしてよい。

【0029】

ここで、近傍構造データには少なくとも地物の重み係数が含まれている必要がある。重み係数としては、例えば、地物の面積または人口などを用いることができる。また、近傍構造データには、地物間の距離、ポリゴンとポリゴンの隣接線分長、または当該隣接線分長に基づく隣接度などの、隣接状況データが含まれていてよい。近傍構造データ格納手段203は、各地物の重み係数をそれぞれ格納する重み係数格納手段と、各地物の各々毎に、地物に隣接する隣接地物との隣接状況を表す隣接状況データを格納する隣接状況データ格納手段を含む。

【0030】

図6は図5の3つのポリゴンの空間配置状況について数値化した近傍構造データの例である。

【0031】

図6の601は、各サイトについて、重み係数(w)と、隣接周長(l)とを表している。重み計数はサイトの大きさを表し、隣接周長は、隣接サイトとの隣接長を足し合わせた値を表す。例えばサイトs1の重み係数はw1=3.0であり、サイトs1が他のサイトs2,s3と隣接している部分の長さの合計はl1=5.0である。

【0032】

図6の602は、各サイトの重心間の距離をサイト間距離(r)として表している。例えばサイトs1とサイトs2のサイト間距離はr12=1.1である。

【0033】

図6の603は、隣接する2つサイトの隣接線分長(D)を表している。例えばサイトs1とサイトs2の隣接線分長はD12=1.0である。

【0034】

図6の604は、隣接する2つのサイトの隣接度を表すサイト隣接度(d)を表している。サイト隣接度には様々なものを用いることが可能であり、例えば603に示した隣接線分長そのものを用いてもよいし、または隣接線分長を定数倍したものなどを用いることが可能である。そのほかにも、例えば、

$$d_{12} = D_{12} / Dave_{12} \quad (3)$$

などを用いることも可能である。ここで、Dave12はサイトs1およびサイトs2について、隣接線分長の平均をとった値である。図6の604では式(3)を用いてサイト隣接度を算出している。例えばサイトs1とサイトs2との隣接線分長は1.00、サイトs1とサイトs3との隣接線分長は4.00、サイトs2とサイトs3との隣接線分長は1.00であるため、 $Dave_{12} = (1+1+4)/4 = 1.75$ 、 $D_{12}=1$ となり、したがって $d_{12} = 1/1.75 = 0.57$ となる。この値がサイトs1とサイトs2の隣接度（あるいは隣接度係数）を表している。

【0035】

状態・観測ポテンシャル算出手段204では、モデルパラメータ格納手段201に格納された

観測モデルパラメータ301を用いて、状態変動の余地のあるサイト（対象地物）がとりうるすべての状態値について、状態値と観測値との乖離度を表す状態・観測ポテンシャルを算出する。

【0036】

図7は図3の観測モデルパラメータ301と図5の状態・観測データ501を用い算出された状態・観測ポテンシャルの例701を示す。

【0037】

図5において、状態変動の余地があるサイトはs3のみであり、サイトs3が取りうる状態値はa,bの2つ（ $z3 \in \{a,b\}$ ）であるので、2種類の状態・観測ポテンシャルを算出すればよい。

【0038】

$z3=a$ の状態・観測ポテンシャルは、状態値 $z3=a$ と観測値 $y3=2.0$ との乖離度を表し、乖離度が大きいほど値が小さくならなければならない。混合モデルの場合は、例えば、 $\log(P(y|\theta(z)))$ によって状態・観測ポテンシャルを算出することが可能である。ここで、 $\theta(z)$ は z の値に対応した混合モデルのコンポーネントのパラメータである。図3の正規分布の場合は、式(4)によってポテンシャルを求めることが可能である。

【数2】

$$\log\left(\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(y-\mu)^2}{\sigma^2}\right)\right) = -\frac{(y-\mu)^2}{\sigma^2} - \log(\sqrt{2\pi}\sigma) \quad (4)$$

よって、例えば、 $y=2$ と $z=a$ との乖離を表すポテンシャルは下記式によって-1.92と算出することが可能である（なお、コンポーネントMaのパラメータは図3より $\mu=1.0$ 、 $\sigma=1.0$ ）。

【数3】

$$-\frac{(2.0-1.0)^2}{1.0^2} - \log(\sqrt{2\pi}) \doteq -1.92$$

【0039】

上記状態・観測ポテンシャル算出手段204は、対象地物を選択する対象地物選択手段と、対象地物が取りうるすべての状態値（第1～第nの状態値）を順次、対象状態値として選択する状態選択手段とを含み、対象状態値に対応する観測モデルパラメータと対象地物の観測値とを元に、対象状態値と観測値の乖離度を表す状態・観測ポテンシャルを計算する。対象地物選択手段は、たとえば図5の例の場合、「^」が付いていない地物を対象地物として選択する。

【0040】

なお、他のサイト（ここではサイトs1,s2）の状態値が無視できる場合には、図7の状態・観測ポテンシャル701のみからギブス分布（ $\exp(U)$ ）を用いて $P(z3=a)$ を算出できる。ギブス確率算出処理に従って、エネルギーUを指数関数に代入した $\exp(U)$ を用いて、

$$P(z3=a) = \exp(U(z3=a))/Z \quad (5)$$

によって確率値を算出することが可能である。ここで、状態値 z のエネルギー $U(z)$ は z に関するすべてのポテンシャル $H(z)$ を足し合わせたものである。また、 Z はすべての状態値におけるエネルギー U' に関して $\exp(U')$ を足し合わせたもの（ $\sum \exp(U')$ ）である。

【0041】

他のサイトの状態値が無視できるということは、状態・観測ポテンシャルがエネルギーに等しい（すなわち $U(z)=H1(z)$ ）ということであるので、図7の状態・観測ポテンシャル701を用いると、 $z3=a$ となる条件付確率は、

10

20

30

40

50

$$\begin{aligned}
 P(z_3=a) &= \exp(U(z_3=a)) / (\exp(U(z_3=a)) + \exp(U(z_3=b))) \\
 &= \exp(H1(z_3=a)) / (\exp(H1(z_3=a)) + \exp(H1(z_3=b))) \\
 &= \exp(-1.92) / (\exp(-1.92) + \exp(-1.86)) \quad 0.49
 \end{aligned}$$

と算出できる（図7の702参照）。なお、本発明では、他のサイトの状態値が無視できない場合を取り扱っているため、上記のようにして条件付確率を求めるのではないことに注意する。

【0042】

状態・状態ポテンシャル算出手段205では、モデルパラメータ格納手段201に格納された状態依存性パラメータ302を用いて、状態変動の余地のあるサイト（対象地物）がとりうるすべての状態値毎に、当該状態値と隣接サイトの状態値との乖離度を表す状態・状態ポテンシャルを算出する。状態・状態ポテンシャルは、2つの状態値の関係に基づいて値が決まる。状態・状態ポテンシャル算出手段205では、例えば、離散値状態変数の場合にはPottsモデル、真偽値状態変数の場合にはIsingモデルなどを、状態・状態ポテンシャル関数として用いることができる。

【0043】

図8は図3の状態依存性パラメータ302と図5の状態・観測データを用いて算出された状態・状態ポテンシャルの例801を示す。

【0044】

図5において、変動の余地があるサイトはs3のみであり、s3の状態値は $z_3 \in \{a, b\}$ であるので、2種類の状態値について隣接する2つのサイトとの状態・状態ポテンシャルを算出すればよい（ $2 \times 2 = 4$ 通り）。状態・状態ポテンシャル関数としては、ここではPottsモデルを用いる。Pottsモデルでは、隣接するサイトの状態値が同じ場合はポテンシャルが0、異なる場合はポテンシャルが $-\beta$ となる。図8の例（上から2行目）では、サイト1の状態値がa（ $z_1=a$ ）、サイト3の状態値がb（ $z_3=b$ ）の場合、これらの状態値は互いに異なるため、状態・状態ポテンシャルは $-\beta = -0.2$ となる。

【0045】

上記状態・状態ポテンシャル算出手段205は、対象地物を選択する対象地物選択手段と、対象地物がとりうるすべての状態値（第1～第nの状態値）を順次、対象状態値として選択する状態選択手段とを備え、対象状態値と前記対象地物に隣接する対象隣接地物の状態値との乖離度を表す状態・状態ポテンシャルを計算する。対象地物選択手段は、たとえば図5の例の場合、「^」が付いていない地物を対象地物として選択する。

【0046】

なお、通常のマルコフ確率場では、上で述べた状態・観測ポテンシャルH1と状態・状態ポテンシャルH2を用いてエネルギーを計算し、確率値を算出する。図9に通常のマルコフ確率場におけるエネルギー901と、算出された確率値902とを示す。図9において、例えば、エネルギーは、

$$\begin{aligned}
 U(z_3=a) &= H1(z_3=a) + H2(z_3=a \mid z_1=a) + H2(z_3=a \mid z_2=b) \\
 &= -1.92 + 0 - 0.2 = -2.12 \\
 U(z_3=b) &= H1(z_3=b) + H2(z_3=b \mid z_1=a) + H2(z_3=b \mid z_2=b) \\
 &= -1.86 - 0.2 + 0 = -2.06
 \end{aligned}$$

と算出できる。そして、式(5)を用いて、

$$\begin{aligned}
 P(z_3=a) &= \exp(U(z_3=a)) / (\exp(U(z_3=a)) + \exp(U(z_3=b))) \\
 &= \exp(-2.12) / (\exp(-2.12) + \exp(-2.06)) \quad 0.49
 \end{aligned}$$

10

20

30

40

50

と確率値が算出できる。

【0047】

この確率値は、前述した、他のサイトの影響を無視した場合と同じ確率になっている。これは、サイトs3の値に関してサイトs1の影響とサイトs2の影響が打ち消しあっているためである。しかし、図4のポリゴンの隣接関係を見ると、サイトs3はサイトs2より面積も大きく、また隣接度合いが強いサイトs1からの影響をサイトs2からよりも強く受けるとみなすことは自然である。

【0048】

そこで、本発明ではそのようなサイトの隣接状況等を考慮したポテンシャル関数を導入することにより、近傍構造を反映したマルコフ確率場に基づく確率モデル装置を実現する。これは特に以下に説明する加重付状態・観測ポテンシャル算出手段206、近傍構造状態・状態ポテンシャル算出手段207および近傍構造データ格納手段203によって実現される。

【0049】

加重付状態・観測ポテンシャル算出手段206では、状態・観測ポテンシャル算出手段204によって算出された状態・観測ポテンシャルに、近傍構造データ格納手段203に格納された対応するサイトの重み係数を掛け合わせるにより加重付状態・観測ポテンシャルを算出する。すなわち、

$$H1'(z(i)) = w(i) * H1(z(i)) \quad (6)$$

によってサイトの状態 $z(i)$ に対する加重付状態・観測ポテンシャルを算出する。

【0050】

図10は図7の状態・観測ポテンシャル701と図6の近傍構造データ（ここでは601に示す重み係数）を用いて算出された加重付状態・観測ポテンシャルの例1001を示す。サイトs3に対応する重み係数は3.0であるので、図10では図7のポテンシャルがそれぞれ3.0倍された値が示される。

【0051】

近傍構造状態・状態ポテンシャル算出手段207では、近傍構造データ格納手段203に格納された近傍構造データを用いて、近傍構造パラメータを引数とする（少なくとも重み係数を引数とする）近傍構造関数を計算することによって近傍構造係数を算出し、算出した近傍構造係数を、状態・状態ポテンシャル算出手段205によって算出された状態・状態ポテンシャルに掛け合わせ、これにより近傍構造状態・状態ポテンシャルを算出する。すなわち、

$$H2'(z(i), z(j)) = G(i, j) * H2(z(i), z(j)) \quad (7)$$

によって、始点サイト s_i と終点サイト s_j のときのポテンシャルを算出する。ここで、 $G(i, j)$ は近傍構造関数である。近傍構造関数 $G(i, j)$ は以下の制約（A）～（C）を満たすものとする。

（A）近傍構造関数 $G(i, j)$ は必ず引数として始点サイトの重み係数 $w(i)$ と終点サイトの重み係数 $w(j)$ を取らなければならない。

（B）また、近傍構造関数 $G(i, j)$ は、 $G(i, j) = G(j, i)$ をみたすという意味の対称関数でなければならない。

（C）さらに、近傍構造関数 $G(i, j)$ は、状態値 $z(i)$ または $z(j)$ によって異なる値を出力してはならない。すなわち、 $G(i, j)$ は $z(i)$ または $z(j)$ を引数としてはならない。

【0052】

以下、近傍構造関数の例をいくつか列挙して示す。

【0053】

（1）重み係数 $\{w(i)\}$ のみを用いた近傍構造関数としては、例えば、

10

20

30

40

【数 4】

$$G(i, j) = \frac{w(i) + w(j)}{2} \quad (8)$$

$$G(i, j) = \sqrt{w(i) * w(j)} \quad (9)$$

などが挙げられる。このような関数を用いることにより、サイトの重みを考慮したポテンシャルの算出が可能になる。

【0054】

(2) 次に、重み係数 $\{w(i)\}$ とサイト隣接度 $\{d(i, j)\}$ を用いた近傍構造関数としては、例えば、 10

【数 5】

$$G(i, j) = \sqrt{w(i) * w(j)} * d(i, j) \quad (10)$$

$$G(i, j) = w(i) * d(i, j) + w(j) * d(j, i) \quad (11)$$

などが挙げられる。このような関数を用いることにより、サイトの重みとサイト間の隣接度とを考慮したポテンシャルの算出が可能になる。

【0055】

(3) また、重み係数とサイト間距離 $\{r(i, j)\}$ を用いた近傍構造関数としては、例えば、 20

【数 6】

$$G(i, j) = \frac{w(i) * w(j)}{r(i, j)^2} \quad (12)$$

などが挙げられる。このような関数を用いることにより、サイトの重みとサイト間の距離とを考慮したポテンシャルの算出が可能になる。

【0056】

(4) さらに、重み係数、サイト隣接度、サイト間距離を用いた近傍構造関数としては、例えば、 30

【数 7】

$$G(i, j) = \frac{w(i) * w(j) * d(i, j)}{r(i, j)^2} \quad (13)$$

などが挙げられる。このような関数を用いることにより、サイトの重み、サイト間の隣接度、およびサイト間の距離を考慮したポテンシャルの算出が可能になる。

【0057】

図11は図8の状態・状態ポテンシャル801と図6の近傍構造データとを用いて算出された近傍構造状態・状態ポテンシャルの例1101を示す。ここで、近傍構造関数は式(13)を用いた。図11の2行目の計算例を示すと、近傍構造係数は 40

【数 8】

$$G(1,3) = \frac{w(1) * w(3) * d(1,3)}{r(1,3)^2} = \frac{3.0 * 3.0 * 1.6}{0.9^2} \doteq 17.8$$

と計算できるので、 $H2'(z3=b|z1=a) = G(1,3) * H2(z3=b|z1=a) = 17.8 * (-0.2) = -3.56$ として、近傍構造状態・状態ポテンシャルが計算できる。

【0058】

ギブス確率算出手段208では、加重付状態・観測ポテンシャル算出手段206において算出 50

された加重付状態・観測ポテンシャルと、近傍構造状態・状態ポテンシャル算出手段207において算出された近傍構造状態・状態ポテンシャルとを用いて、対象地物（サイト）について、状態値毎のエネルギーを算出し、その後、式(5)のギブス関数値計算を用いることにより状態値毎の確率値を算出する。ギブス確率算出手段208は、対象地物について、状態値毎のエネルギーを計算するエネルギー計算手段と、状態値毎に計算されたエネルギーに基づいて各状態値の少なくとも1つについてその発生確率を計算する確率計算手段とを含んでいる。

【0059】

図12は図10の加重付状態・観測ポテンシャルと図11の近傍構造状態・状態ポテンシャルとを用いて算出された、エネルギー1201と確率値1202とを示している。

10

【0060】

図12のエネルギー1201を用いて $z_3=a$ の条件付確率（すなわちサイト3が状態値aを取る条件付き確率）を算出すると、

$$\begin{aligned} P(z_3=a) &= \exp(U'(z_3=a)) / (\exp(U'(z_3=a)) + \exp(U'(z_3=b))) \\ &= \exp(-6.10) / (\exp(-6.10) + \exp(-9.14)) \quad 0.95 \end{aligned}$$

となる。ただし、

$$\begin{aligned} U'(z_3=a) &= H1'(z_3=a) + H2'(z_3=a|z_1=a) + H2'(z_3=a|z_2=b) = -5.76 + 0 - 0.34 = -6.10 \\ U'(z_3=b) &= H1'(z_3=b) + H2'(z_3=b|z_1=a) + H2'(z_3=b|z_2=b) = -5.59 - 3.56 + 0 = -9.14 \end{aligned}$$

20

である。

【0061】

この確率値0.95は、図9の902に示す確率値0.49（通常のマルコフ確率場の場合）よりも大きい。これは、本発明の確率モデル装置では、近傍ポリゴンとの隣接状況や、対象地物および隣接地物の重み等を考慮することにより、サイトs3へのサイトs1からの影響度が相対的に強くなったためである。このように、本発明の確率モデル装置では、近傍構造を考慮した依存関係をもつ確率分布を表現することができる。

【0062】

30

以下、既存手法と比較して本発明の優位性を示す。

【0063】

仮にサイトs3に対して確率モデルMaを割り当てるのが正しい場合を考える。すると、サイトs3の観測値 y_3 は、図3の301から、平均値=1.0、標準偏差1.0の正規分布に従って値が決まることになる。そこで、30点分の y_3 の値と、それぞれの y_3 における既存のポテンシャル関数を用いた条件付確率値(図9の902参照)、および、それぞれの y_3 における本発明のポテンシャル関数を用いた条件付確率値(図12の1202参照)を図14に示す。

【0064】

図14では既存のポテンシャル関数を用いた場合には25/30の割当て正解率なのに対して、本発明のポテンシャル関数を用いた場合には29/30の割当て正解率となり、既存手法よりも、正確なモデル割当てが可能であることが確認できる。ただし閾値は0.5とし、0.5以上は正解、0.5未満は不正解としている。なお、確率モデルMaから生成された観測値であっても、モデルMbから発生すると考えたほうが自然な値が、ある確率で生じてしまうが（閾値未満の値をもつID）、そのような場合にも、近隣サイトの情報を用いてMbではなくMaを割り当てることにより、既存手法に比して、空間的な依存関係を考慮した正しいモデル割当てを実現できる。このように、本発明によって地理空間的な近傍構造を考慮することにより地理空間的な現象を正確に表現することが可能な地物確率モデル装置を実現できる。

40

【0065】

図13は、図2に示した地物確率モデル装置による処理の流れを示すフローチャートであ

50

る。

【0066】

CPU102が、外部記憶装置104から、モデルパラメータ（観測モデルパラメータおよび状態依存性パラメータ）（図3参照）、状態観測データ（図5参照）、近傍構造データ（図6参照）を読み出し、主記憶装置103へ格納する（ステップ1301）。

【0067】

CPU102（状態・観測ポテンシャル算出手段204）は、状態観測データにおいて状態変動の余地のある地物を対象地物として特定し、対象地物を取りうる各状態値のそれぞれについて、各状態値のそれぞれに対応する観測モデルパラメータ（確率分布または確率モデルのパラメータ）を用いて、当該状態値と観測値の乖離度を表す状態・観測ポテンシャルを算出する（ステップ1302）。 10

【0068】

CPU102（状態・状態ポテンシャル算出手段205）は、上記状態依存性パラメータを用いて、対象地物を取りうる各状態値のそれぞれについて、対象地物の状態値と、対象地物の隣接地物（対象隣接地物）の状態値との乖離度を表す状態・状態ポテンシャルを算出する（ステップ1303）。

【0069】

CPU102（加重付状態・観測ポテンシャル算出手段206）は、上記各状態値のそれぞれについて算出された状態・観測ポテンシャルに、上記近傍構造データに含まれる対応する地物の重み係数を掛け合わせるにより加重付状態・観測ポテンシャルを算出する（ステップ1304）。 20

【0070】

CPU102（近傍構造状態・状態ポテンシャル算出手段207）は、上記近傍構造データを用いて、近傍構造パラメータを引数とする（少なくとも重み係数を引数とする）近傍構造関数を計算することによって近傍構造係数を算出し、算出した近傍構造係数を、状態・状態ポテンシャルに掛け合わせるにより近傍構造状態・状態ポテンシャルを算出する（ステップ1305）。

【0071】

CPU102（ギブス確率算出手段208）は、上記加重付状態・観測ポテンシャルと、上記近傍構造状態・状態ポテンシャルとを加算することにより、対象地物について状態値毎のエネルギーを算出し、その後、式(5)のギブス関数値計算を用いることにより状態値毎の確率値を算出する（ステップ1306）。 30

【0072】

これまで説明した本発明の確率モデル装置を用いて、複数の状態値候補の中から最も確率の高い状態値を探索する状態値推論装置を実現することが可能である。そのような装置では、例えば、対象地物について、状態値候補 $\{z_1, z_2, z_3, \dots, z_n\}$ のすべての状態値について本発明の確率モデル装置によって確率値を算出し、最も高い確率値を示す状態値を決定すれば良い。そして、この処理を、CPU102において、対象地物を変更しながら（状態変動の余地のあるサイトを変更しながら）、繰り返し行うことにより、地理空間的な現象を正確に表現した地物確率モデルを生成できる。例えば、パーソナルコンピュータやエレベータの部品、自動車の部品など機器の故障や劣化において、環境要因や地域による居住者の使用習慣が影響する場合には、本発明の地物確率モデル装置を用いることにより、データに含まれるノイズの悪影響に頑健な予測モデル（地物確率モデル）を実現することができる。 40

【0073】

また、1つ以上の観測値に基づいて似た傾向を示す地域をグループ化し（すなわち似た傾向を示す地域に同じ状態値を割り当て）、対象領域を幾つかのセグメントに分割する（同一状態値をもちかつ隣接する領域郡を1つのセグメントとする）という地域類型化分析にも本発明の地物確率モデル装置を用いることができる。この場合も、例えば、対象地物について、状態値候補 $\{z_1, z_2, z_3, \dots, z_n\}$ の状態値について本発明の確率モデル装置によっ 50

て確率値を算出し、最も高い確率値を示す状態値を決定すれば良い。そして、この処理を、CPU102において、対象地物を変更しながら（状態変動の余地のあるサイトを変更しながら）、繰り返し行うことにより、各地物についての状態値を決定して、地理空間的な現象を正確に表現したセグメント化が可能となる。上記地域類型化分析は、例えば複数の商品の売り上げ比率をエリアごとに集計したデータに基づいて購買傾向の似た地域セグメントを発見し、販売促進計画に利用するなどというエリアマーケティングなどにおいても必要になる。地域類型化分析では、依存関係を考慮しないモデルを用いてしまうと、データに含まれるノイズの悪影響のため、ある地域がどのクラス（状態値）に含まれるのかという地域判定の結果がモザイク上になってしまう。すると、セグメントの境界線がはっきりせず、どのような要因で地域がセグメントに分かれているのか分析者が把握しづらい。本発明の近傍構造を考慮した方法を用いることで、境界線がよりはっきりとなるような各地物に対する状態値の決定がなされることが期待され、よってデータに含まれるノイズの悪影響を緩和した地域類型化分析が実現できる。

10

【0074】

さらに、本発明の地物確率モデル装置を用いて地域類型化分析を行って、機器の劣化傾向の似た地域セグメントを発見し、機器の制御ロジックを地域セグメントごとに変更することにより、例えばある部品が壊れやすい地域ではその部品に負荷がかかりにくい制御ロジックを導入するなどという、地域特性に応じた最適制御を行う地域別制御システムを実現できる。

20

【0075】

本発明により、地物の重みや対象間の近傍構造を反映した依存関係を表現することが可能な地物確率モデル装置が実現できる。

【0076】

なお、以上に説明した地物確率モデル装置は、例えば、汎用のコンピュータ装置を基本ハードウェアとして用いることでも実現することが可能である。すなわち、地物確率モデル装置における状態・観測ポテンシャル算出手段、状態・状態ポテンシャル算出手段、加重付状態・観測ポテンシャル算出手段、近傍構造状態・状態ポテンシャル算出手段、ギブス確率算出手段は、上記のコンピュータ装置に搭載されたプロセッサにプログラムを実行させることにより実現することができる。このとき、管理装置は、上記のプログラムをコンピュータ装置にあらかじめインストールすることで実現してもよいし、CD-ROMなどの記憶媒体に記憶して、あるいはネットワークを介して上記のプログラムを配布して、このプログラムをコンピュータ装置に適宜インストールすることで実現してもよい。また、モデルパラメータ格納手段、状態・観測データ格納手段、近傍構造データ格納手段は、上記のコンピュータ装置に内蔵あるいは外付けされたメモリ、ハードディスクもしくはCD-R、CD-RW、DVD-RAM、DVD-Rなどの記憶媒体などを適宜利用して実現することができる。

30

【図面の簡単な説明】

【0077】

【図1】本発明の実行環境を示したコンピュータシステムを示す図。

【図2】本発明に関わる地物確率モデル装置の一実施の形態を示した構成図。

40

【図3】モデルパラメータ（観測モデルパラメータおよび状態依存性パラメータ）の例と、観測モデルパラメータのグラフ表示例とを示す図。

【図4】サイト $\{s_1, s_2, s_3\}$ の空間配置の例を示す図。

【図5】状態・観測データの例を示す図。

【図6】図5の3つのポリゴンの空間配置状況について数値化した近傍構造データの例を示す図。

【図7】状態・観測ポテンシャルの例を示す図。

【図8】状態・状態ポテンシャルの例を示す図。

【図9】通常のマルコフ確率場におけるエネルギーと、算出された確率値との例を示す図。

50

【図 10】加重付状態・観測ポテンシャルの例を示す図。

【図 11】近傍構造状態・状態ポテンシャルの例を示す図。

【図 12】図10の加重付状態・観測ポテンシャルと図11の近傍構造状態・状態ポテンシャルとを用いた場合の、エネルギーと確率値とを示している。

【図 13】図2の地物確率モデル装置により行なわれる処理の流れを示すフローチャート

【図 14】既存手法と比較して本発明の優位性を示す図。

【符号の説明】

【0078】

101：バス

102：CPU

103：主記憶装置

104：外部記憶装置

105：入力装置

106：表示装置

201：モデルパラメータ格納手段

202：状態・観測データ格納手段

203：近傍構造データ格納手段

204：状態・観測ポテンシャル算出手段

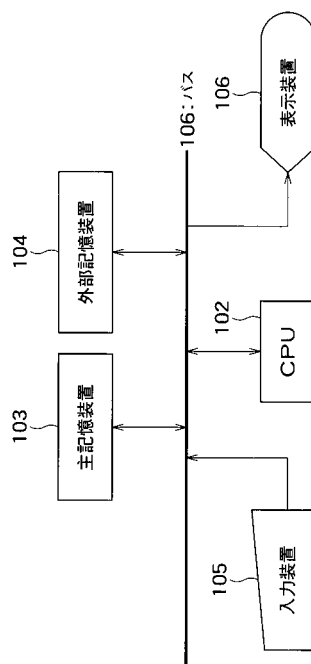
205：状態・状態ポテンシャル算出手段

206：加重付状態・観測ポテンシャル算出手段

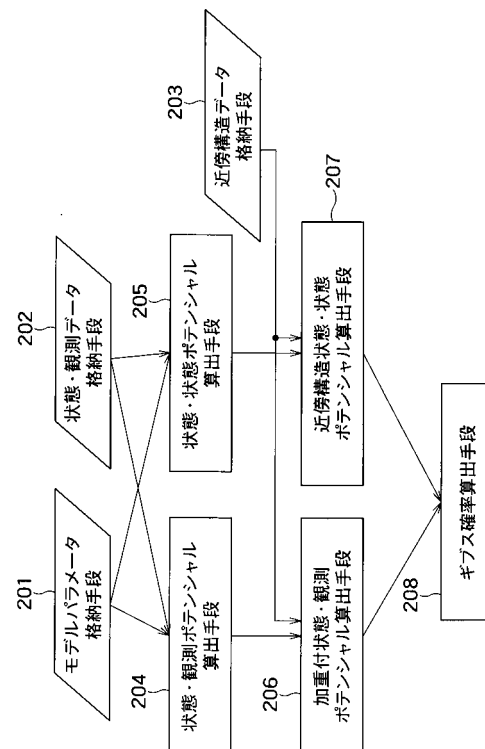
207：近傍構造状態・状態ポテンシャル算出手段

208：ギブス確率算出手段

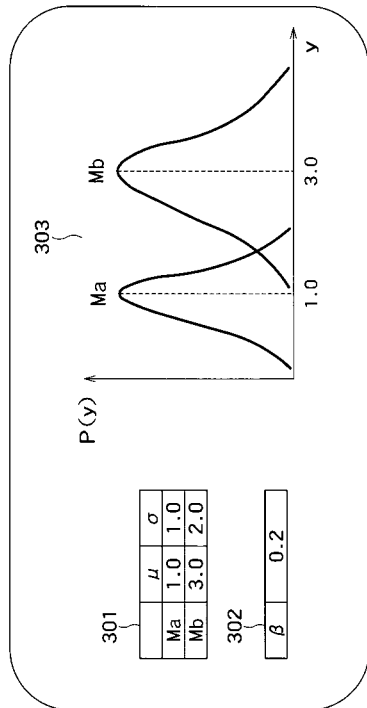
【図 1】



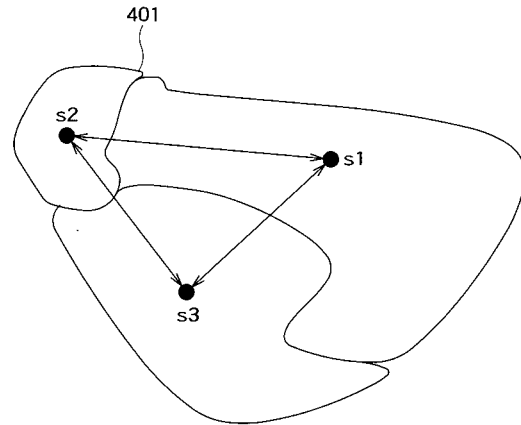
【図 2】



【図 3】



【図 4】



【図 5】

サイト	観測値(y)	状態(z)
s1	1.2	$\hat{a}$
s2	4.5	$\hat{b}$
s3	2.0	a

【図 6】

601			602		
サイト	重み(w)	隣接周長(l)	r*1	r*2	r*3
s1	3.0	5.0	r1*	-	1.1
s2	1.0	2.0	r2*	-	-
s3	3.0	5.0	r3*	-	-

603			604		
D*1	D*2	D*3	d*1	d*2	d*3
D1*	-	1.00	d1*	-	0.57
D2*	-	-	d2*	-	-
D3*	-	-	d3*	-	-

【図 9】

901		エネルギー(1)
$U(z3=a z1=a, z2=b)$		-2.12
$U(z3=b z1=a, z2=b)$		-2.06

902		条件付確率2
$P(z3=a z1=a, z2=b)$		0.49

【図 7】

701		状態・観測ポテンシャル
$H1(z3=a)$		-1.92
$H1(z3=b)$		-1.86

702		条件付確率1
$P(z3=a)$		0.49

【図 10】

1001		加重付き状態・観測ポテンシャル
$H1'(z3=a)$		-5.76
$H1'(z3=b)$		-5.59

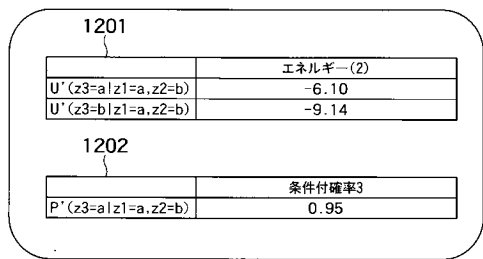
【図 11】

1101		近傍構造状態・状態ポテンシャル
$H2'(z3=a z1=a)$		0.00
$H2'(z3=b z1=a)$		-3.56
$H2'(z3=a z2=b)$		-0.34
$H2'(z3=b z2=b)$		0.00

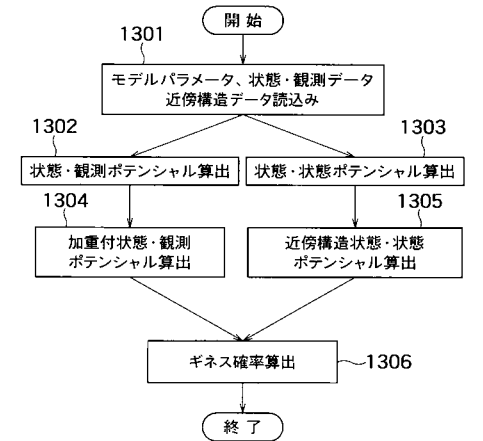
【図 8】

801		状態・状態ポテンシャル
$H2(z3=a z1=a)$		0.00
$H2(z3=b z1=a)$		-0.20
$H2(z3=a z2=b)$		-0.20
$H2(z3=b z2=b)$		0.00

【図 1 2】



【図 1 3】



【図 1 4】

ID	s3の観測値	既存手法	本発明
		$P(s3=a z1=a,z2=b)$	$P'(s3=a z1=a,z2=b)$
1	0.10	0.8795	0.9999
2	1.16	0.8190	0.9996
3	1.45	0.7490	0.9985
4	2.29	0.2990	0.6584
5	0.20	0.8822	0.9999
6	2.03	0.4666	0.9433
7	1.64	0.6793	0.9958
8	0.63	0.8767	0.9999
9	-0.10	0.8681	0.9999
10	2.35	0.2671	0.5461
11	1.08	0.8327	0.9997
12	-0.73	0.7654	0.9988
13	-0.25	0.8545	0.9998
14	0.51	0.8813	0.9999
15	1.75	0.6268	0.9916
16	0.22	0.8826	0.9999
17	1.06	0.8365	0.9997
18	0.43	0.8828	0.9999
19	-0.43	0.8300	0.9997
20	0.11	0.8797	0.9999
21	1.66	0.6689	0.9951
22	1.25	0.8015	0.9994
23	0.97	0.8484	0.9998
24	-0.61	0.7958	0.9993
25	1.40	0.7643	0.9988
26	1.63	0.6819	0.9959
27	2.00	0.4847	0.9539
28	2.74	0.0903	0.0237
29	0.90	0.8568	0.9998
30	0.58	0.8786	0.9999
正解率		25/30	29/30