

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年6月16日(16.06.2022)



(10) 国際公開番号

WO 2022/123743 A1

(51) 国際特許分類:

G06N 20/00 (2019.01)

(21) 国際出願番号: PCT/JP2020/046123

(22) 国際出願日: 2020年12月10日(10.12.2020)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 田中 佑典 (TANAKA, Yusuke); 〒1808585 東京都武蔵野市緑町3丁目9-1

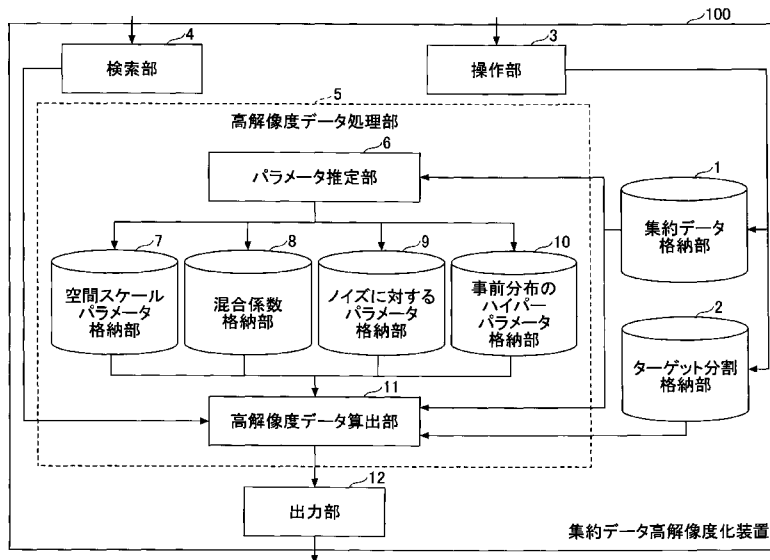
1 NTT 知的財産センタ内 Tokyo (JP). 岩田 具治(IWATA, Tomoharu); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 NTT 知的財産センタ内 Tokyo (JP). 倉島 健(KURASHIMA, Takeshi); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 NTT 知的財産センタ内 Tokyo (JP). 戸田 浩之(TODA, Hiroyuki); 〒1808585 東京都武蔵野市緑町3丁目9-1 1 NTT 知的財産センタ内 Tokyo (JP).

(74) 代理人: 伊東 忠重, 外(ITO, Tadashige et al.); 〒1000005 東京都千代田区丸の内二丁目1番1号 丸の内 M Y P L A Z A (明治安田生命ビル) 16階 Tokyo (JP).

(54) Title: PARAMETER ESTIMATION DEVICE, AGGREGATED DATA RESOLUTION INCREASING DEVICE, PARAMETER ESTIMATION METHOD, AGGREGATED DATA RESOLUTION INCREASING METHOD, AND PROGRAM

(54) 発明の名称: パラメータ推定装置、集約データ高解像度化装置、パラメータ推定方法、集約データ高解像度化方法、及びプログラム

[図1]



- 1 Aggregated data storage unit
- 2 Target division storage unit
- 3 Operation unit
- 4 Search unit
- 5 High-resolution data processing unit
- 6 Parameter estimation unit
- 7 Spatial scale parameter storage unit
- 8 Mixing coefficient storage unit
- 9 Noise parameter storage unit
- 10 Prior distribution hyperparameter storage unit
- 11 High-resolution data calculation unit
- 12 Output unit
- 100 Aggregated data resolution increasing device

(57) Abstract: Provided is a parameter estimation device for estimating a plurality of parameters to be used for calculating high-resolution data from data aggregated at coarse granularity, the parameter estimation device comprising: a parameter estimation unit that estimates a plurality of parameters, which are unknown variables in a model based on a multivariate Gaussian process, so as to maximize a marginal likelihood on the assumption that actually measured aggregated data is generated from the model, the multivariate Gaussian process representing, by linear mixing, a plurality of latent Gaussian processes for a plurality of types of pieces of aggregated data in a plurality of domains; and a storage unit that stores the plurality of parameters including hyperparameters of a prior distribution for mixing coefficients to be used for the

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

linear mixing.

(57) 要約: 粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定装置であって、実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、前記複数のパラメータを格納する格納部と、を備え、前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含むパラメータ推定装置。

明 細 書

発明の名称：

パラメータ推定装置、集約データ高解像度化装置、パラメータ推定方法、
集約データ高解像度化方法、及びプログラム

技術分野

[0001] 本発明は、粗い粒度に集約された集約データから高解像度データを推定する技術に関連するものである。

背景技術

[0002] 近年、政府や企業などが都市環境や事業の改善を目的として、様々な種類の空間データ（貧困度、大気汚染度、犯罪数、人口、交通量など）を収集し、公開している。空間データとは、位置情報（緯度経度など）とそれに紐づく何らかの値とのペアで与えられるデータを指す。

[0003] しかし、このような空間データは、収集コストが高く十分なサンプル数を確保することが難しい。そのため、ある程度粗い粒度の領域（住所や地域など）において集約されて提供されることが多い。このようなデータを以下では「集約データ」と呼ぶ。より効果的な都市環境の改善のためには、できる限り高解像度のデータを入手することが望ましい。例えば、貧困度が高い地域や大気汚染度が高い地域を詳細に絞りこむことで、より適切な介入が可能となる。

[0004] したがって、低解像度の集約データから高解像度なデータを予測するという問題は重要である。従来技術では、ターゲットとする低解像度の集約データとは別に各々の解像度を持つ様々な種類の集約データを用意し、多変量ガウス過程に基づいて同時にモデル化することで低解像度のデータに対して高精度な予測が実現されている（非特許文献3）。更に、異なるドメイン（都市など）におけるデータを活用する試みもなされている（非特許文献2）。

先行技術文献

非特許文献

[0005] 非特許文献1: D. Kingma and M. Welling. Auto-encoding variational Bayes. In ICLR, 2014.

非特許文献2: Y. Tanaka, T. Tanaka, T. Iwata, T. Kurashima, M. Okawa, Y. Akagi, and H. Toda. Spatially aggregated Gaussian processes with multivariate areal outputs. In NeurIPS, pages 3000–3010, 2019.

非特許文献3: F. Yousefi, M. T. Smith, and M. A. Alvarez. Multi-task learning for aggregated data using Gaussian processes. In NeurIPS, pages 15050–15060, 2019. 8

発明の概要

発明が解決しようとする課題

[0006] 非特許文献2、3に開示されている従来技術では、ガウス過程を基礎として、複数種類の集約データを多変量ガウス過程に基づいて同時にモデル化することでターゲットとする高解像度データを予測する。このとき、非特許文献2に開示された技術では、ドメイン（都市など）が異なる場合でも、ガウス過程のパラメータをドメイン間で共有することにより、複数ドメインにおけるデータを活用する。しかし、当該技術にはドメイン間の類似度が考慮できないという問題点がある。

[0007] すなわち、従来技術では、異なるドメインにおいて同じ種類のデータの組が存在したとしても、それらのデータ間の依存関係の強さはドメイン毎に独立であることを仮定している。しかし、ドメインによっては、低解像度のデータしか入手できない場合も少なくない。そのような場合に、データ間の依存関係の強さを適切に推定することが難しいという問題が存在した。

[0008] 本発明は、上記の点に鑑みてなされたものであり、複数ドメインにおける多種多様なデータを活用して、高解像度データの高精度な予測を実現することを目的とする。

課題を解決するための手段

[0009] 開示の技術によれば、粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定

装置であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、

前記複数のパラメータを格納する格納部と、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

パラメータ推定装置が提供される。

発明の効果

[0010] 開示の技術によれば、複数ドメインにおける多種多様なデータを活用して、高解像度データの高精度な予測を実現することが可能となる。

図面の簡単な説明

[0011] [図1]本発明の実施の形態における集約データ高解像度化装置の構成図である。

[図2]装置のハードウェア構成例を示す図である。

[図3]全体の処理の流れを説明するための図である。

[図4]本発明の実施の形態における検索部への検索例と、出力部からの出力例を示す図である。

発明を実施するための形態

[0012] 以下、図面を参照して本発明の実施の形態（本実施の形態）を説明する。

以下で説明する実施の形態は一例に過ぎず、本発明が適用される実施の形態は、以下の実施の形態に限られるわけではない。

[0013] （実施の形態の概要）

本実施の形態では、ドメイン間の類似度が考慮できないという従来技術の問題を解決するために、データ間の依存関係を表すパラメータ（混合係数）に対する事前分布を導入し、それを加味したパラメータ推定手段を導入して

いる。これにより、ドメイン間の類似度を同時に推定しつつ、複数ドメインにおける多種多様なデータを活用して、高解像度データの高精度な予測を実現することが可能となる。より具体的には、下記の手段１と手段２が導入される。

[0014] (１) 手段１：混合係数に対する事前分布を導入した、集約データのための多変量ガウス過程モデル

手段１においては、「空間スケールパラメータ」と、「混合係数」と、「ノイズに対するパラメータ」と、「事前分布のハイパーパラメータ」とを未知変数として、複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程に基づいて、集約データの値を該当領域におけるガウス過程の積分値によりモデル化する。実際に観測された集約データが上記のモデルから生成されたものと仮定して、周辺尤度を最大化するように未知変数を推定する。

[0015] 手段１に対する効果は下記のとおりである。

[0016] 複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程モデルに基づいて、各ドメインにおける複数のデータを同時にモデル化し、更に、混合係数に対する事前分布を導入する。これにより、ドメイン間の類似度を推定しつつ、「空間スケールパラメータ」、及び、「混合係数」を複数ドメインにおけるデータを活用して学習することができる。

[0017] (１) 手段２：変分ベイズ法による効率的なパラメータ推定手段

手段２においては、手段１における多変量ガウス過程モデルに加えて、変分ベイズ法に基づいて、未知変数を推定する。

[0018] 手段２に対する効果は下記のとおりである。

[0019] 未知変数を推定する際には、周辺尤度を計算する必要があるが、実際には解析的に計算することは難しい。そこで任意の近似的な推定手段を用いることが考えられるが、選択肢の一つとして変分ベイズ法に基づく学習方法を導出する。これにより、より効率的に未知変数を学習することができる。

[0020] 本実施の形態では、上述した手段を組み込んだ集約データ高解像度化装置１００が提供される。集約データ高解像度化装置１００は、任意の領域で集

約されたデータ（以下では簡単のため集約データと呼ぶ）全般を対象とすることができ、使用するデータの種類（貧困度、大気汚染度や交通量など）やデータの入力空間の次元数 $d \in \{1, \dots\}$ に依存せず、それらに対して柔軟に対応し、高解像度データの推定を行うことができる。

[0021] $v = 1, \dots, V$ をドメインの引数とし、 v 番目のドメインにおけるデータの種類集合を S_v とする。ここで、全ドメインにおけるデータの種類の集合の和集合を S とすると、 $S_v \subset S$ である。以下では、簡略化のため全てのドメインにおいて S 種類のデータが存在するとし、 $s = 1, \dots, S$ をデータ種類の引数として定式化を行う。ただし、上記の定義のようにドメインによって得られるデータの種類の異なっても同様にして定式化が可能である。

[0022] 本実施の形態における集約データ高解像度化装置 100 は、複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程モデルに基づいて、集約データの値を該当領域におけるガウス過程の積分値によりモデル化し、混合係数に対する事前分布を設定した上で、変分ベイズ法に基づいて未知変数を推定し、高解像度データを予測する。

[0023] 以下、集約データ高解像度化装置 100 の構成と動作について詳細に説明する。なお、以下では、主に集約データとして二次元空間上で集約されたデータを実施例として説明をするが、本発明は、任意の次元数における空間上において集約されたデータに対して適用することができる。例えば、一次元空間を考えた場合、センサデータ等の時系列データが任意の時間間隔毎に集約されたものに対応する。

[0024] （装置構成例）

図 1 に、本実施の形態における集約データ高解像度化装置 100 の構成図を示す。同図に示す集約データ高解像度化装置 100 は、集約データ格納部 1、ターゲット分割格納部 2、操作部 3、検索部 4、高解像度データ処理部 5、パラメータ推定部 6、空間スケールパラメータ格納部 7、混合係数格納部 8、ノイズに対するパラメータ格納部 9、事前分布のハイパーパラメータ格納部 10、高解像度データ算出部 11、出力部 12 を有する。各部の動作

詳細については後述する。

[0025] 集約データ高解像度化装置１００は、複数の装置（コンピュータ）から構成されてもよいし、１つの装置で構成されてもよい。また、集約データ高解像度化装置１００を集約データ高解像度化システムと呼んでもよい。また、集約データ高解像度化装置１００をパラメータ推定装置と呼んでもよい。また、図１において、集約データ格納部１、ターゲット分割格納部２以外の機能部からなる装置を集約データ高解像度化装置１００と呼んでもよい。

[0026] また、高解像度化を行う機能を含まずにパラメータ推定を行う機能（つまり、パラメータ推定部６）を有する装置をパラメータ推定装置と呼んでもよい。また、パラメータ推定を行う機能を含まずに、高解像度化を行う機能（つまり、高解像度データ算出部１１）を有する装置を集約データ高解像度化装置と呼んでもよい。

[0027] （ハードウェア構成例）

上述した集約データ高解像度化装置、パラメータ推定装置（これらを総称して装置と呼ぶ）はいずれも、例えば、コンピュータに、本実施の形態で説明する処理内容を記述したプログラムを実行させることにより実現可能である。なお、この「コンピュータ」は、物理マシンであってもよいし、クラウド上の仮想マシンであってもよい。仮想マシンを使用する場合、ここで説明する「ハードウェア」は仮想的なハードウェアである。

[0028] 上記プログラムは、コンピュータが読み取り可能な記録媒体（可搬メモリ等）に記録して、保存したり、配布したりすることが可能である。また、上記プログラムをインターネットや電子メール等、ネットワークを通して提供することも可能である。

[0029] 図２は、上記コンピュータのハードウェア構成例を示す図である。図２のコンピュータは、それぞれバスＢで相互に接続されているドライブ装置１０００、補助記憶装置１００２、メモリ装置１００３、ＣＰＵ１００４、インタフェース装置１００５、表示装置１００６、入力装置１００７、出力装置１００８等を有する。

[0030] 当該コンピュータでの処理を実現するプログラムは、例えば、CD-ROM又はメモリカード等の記録媒体1001によって提供される。プログラムを記憶した記録媒体1001がドライブ装置1000にセットされると、プログラムが記録媒体1001からドライブ装置1000を介して補助記憶装置1002にインストールされる。但し、プログラムのインストールは必ずしも記録媒体1001より行う必要はなく、ネットワークを介して他のコンピュータよりダウンロードするようにしてもよい。補助記憶装置1002は、インストールされたプログラムを格納すると共に、必要なファイルやデータ等を格納する。

[0031] メモリ装置1003は、プログラムの起動指示があった場合に、補助記憶装置1002からプログラムを読み出して格納する。CPU1004は、メモリ装置1003に格納されたプログラムに従って、当該装置に係る機能を実現する。インタフェース装置1005は、ネットワークに接続するためのインタフェースとして用いられる。表示装置1006はプログラムによるGUI(Graphical User Interface)等を表示する。入力装置1007はキーボード及びマウス、ボタン、又はタッチパネル等で構成され、様々な操作指示を入力させるために用いられる。出力装置1008は演算結果を出力する。

[0032] (集約データ高解像度化装置100の処理動作)

まず、図3を参照して、全体の処理の流れを説明する。

[0033] S101において、パラメータ推定部6が、複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程モデルに基づいて、集約データの値を該当領域におけるガウス過程の積分値によりモデル化し、混合係数に対する事前分布を設定した上で、変分ベイズ法に基づいて未知変数である複数のパラメータを推定する。

[0034] S102において、高解像度データ算出部11が、S101で算出された複数のパラメータを用いて、集約データから高解像度データを算出し、出力部12から出力する。

[0035] 以下、集約データ高解像度化装置 100 を構成する各部の機能及び処理動作について説明する。

[0036] <集約データ格納部 1>

集約データ格納部 1 は、集約データ高解像度化装置 100 によって解析され得るデータを格納しており、高解像度データ処理部 5 からの要求にしたがって、データを読み出し、該当のデータを高解像度データ処理部 5 に送信する。

[0037] v 番目のドメインにおける入力空間を $X_v \subset \mathbb{R}^d$ とし、 $x \in X_v$ を入力変数とする。例えば、 $d = 2$ の場合は、 X_v は v 番目の都市全体に、 x は緯度経度等に対応する。 v 番目のドメインの s 番目のデータに対し、 P_{vs} を該当データが紐づく分割を表すものとする。ここで分割とは、例えば住所や地域による都市の分割等に対応する。また、 N_{vs} は分割 P_{vs} に含まれる領域数を表すものとする。領域の引数 $n = 1, \dots, N_{vs}$ に対して、 n 番目の領域を $R_{vsn} \in P_{vs}$ とする。 s 番目のデータに含まれる n 番目の観測を、領域 R_{vsn} と値 $y_{vsn} \in \mathbb{R}$ の組として (R_{vsn}, y_{vsn}) と表す。集約データ格納部 1 に格納されるデータは、

[0038] [数1]

$$\{(R_{vsn}, y_{vsn}) \mid v = 1, \dots, V; s = 1, \dots, S; n = 1, \dots, N_{vs}\} \quad (1)$$

である。つまり、集約データ格納部 1 には、ドメイン毎、データの種類毎、領域毎の領域と値が格納されている。集約データ格納部 1 は、具体的には、例えば、Web サーバや、データベースを具備するデータベースサーバ等によって実現することができる。集約データ格納部 1 が、1 つのコンピュータ内の記憶装置であってもよい。

[0039] <ターゲット分割格納部 2>

ターゲット分割格納部 2 は、集約データ高解像度化装置 100 によって出

力され得る分割を格納しており、高解像度データ処理部5からの要求にしたがって、データを読み出し、該当のデータを高解像度データ処理部5に送信する。 P_{target} をターゲットとする分割を表すものとする。 P_{target} に含まれる領域の1つを R_{target} と表す。ここで、ターゲット分割は任意のものを使用することが可能であり、住所や地域に基づく分割や使用者が設定した任意のサイズのメッシュなどが考えられる。ターゲット分割格納部2は、例えば、Webサーバや、データベースを具備するデータベースサーバ等により実現される。ターゲット分割格納部2が、1つのコンピュータ内の記憶装置であってもよい。

[0040] <操作部3>

操作部3は、集約データ格納部1、及び、ターゲット分割格納部2のデータに対するユーザからの各種操作を受け付ける。各種操作とは、格納された情報を登録、修正、削除する操作等である。操作部3の入力手段は、キーボードやマウス、メニュー画面、タッチパネルによるもの等、どのようなものでもよい。操作部3は、例えば、マウス等の入力手段のデバイスドライバや、メニュー画面の制御ソフトウェアで実現される。

[0041] <検索部4>

検索部4は、高解像度化を行う対象とするデータ種別、及び、ターゲット分割を受け付ける。検索部4で指定されたデータ、及び、ターゲット分割に対して、集約データ高解像度化装置100によって予測された高解像度データを出力する。なお、検索部4の入力手段は、キーボードやマウス、メニュー画面、タッチパネルによるもの等、どのようなものでもよい。検索部4は、マウス等の入力手段のデバイスドライバや、メニュー画面の制御ソフトウェアで実現され得る。

[0042] <高解像度データ処理部5>

図1に示すとおり、高解像度データ処理部5は、パラメータ推定部6、空間スケールパラメータ格納部7、混合係数格納部8、ノイズに対するパラメータ格納部9、事前分布のハイパーパラメータ格納部10、高解像度データ

算出部 11 を有する。

[0043] 高解像度データ処理部 5 では、パラメータ推定部 6 により、集約データ格納部 1 に格納されたデータを学習データとして、空間スケールパラメータ、混合係数、ノイズに対するパラメータ、事前分布のハイパーパラメータを推定する。そして、これらのパラメータを用いて予測された高解像度データを出力する。

[0044] 高解像度データ処理部 5 では、上記の未知変数を組み込んだ上で、複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程モデルに基づいて、複数ドメインにおける集約データをモデル化し、観測データが与えられたとしたときの周辺尤度に基づいて未知変数を推定する。その後、ガウス過程の予測分布に基づいて高解像度データを算出する。

[0045] <パラメータ推定部 6>

以下、パラメータ推定部 6 において処理されるガウス過程モデルとパラメータ推定法について詳細に説明する。まず、複数の潜在ガウス過程の線形混合で表現された多変量ガウス過程の定式化を行う。 $V \times L$ 個の独立なガウス過程を

[0046] [数2]

$$g_{vl}(\mathbf{x}) \sim \mathcal{GP}(0, \gamma_l(\mathbf{x}, \mathbf{x}')), \quad v = 1, \dots, V; l = 1, \dots, L \quad (2)$$

とする。ここで、 $\mathbf{X} = \mathbf{X}_1 \cup \mathbf{X}_2 \dots \dots \cup \mathbf{X}_V$ としたとき、 $\gamma_l(\mathbf{x}, \mathbf{x}')$: $\mathbf{X} \times \mathbf{X} \rightarrow \mathbb{R}$ は l (エル) 個目の相関関数であり、任意のものを使用できる。ここでは、相関関数は

[0047]

[数3]

$$\gamma_l(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2\beta_l^2} \|\mathbf{x} - \mathbf{x}'\|^2\right) \quad (3)$$

を用いるものとする。ここで、 β_l は l （エル）個目の相関関数の空間スケールパラメータである。 $v = 1, \dots, V$ に対して、 $f_{vs}(\mathbf{x})$ を各ドメインの s 番目の集約データに対するガウス過程とし、 S 変量ガウス過程 $\mathbf{f}_v(\mathbf{x}) = (f_{v1}(\mathbf{x}), \dots, f_{vS}(\mathbf{x}))^\top$ を L 個の独立なガウス過程の線型混合として

[0048] [数4]

$$\mathbf{f}_v(\mathbf{x}) = \mathbf{W}_v \mathbf{g}_v(\mathbf{x}) + \mathbf{n}_v(\mathbf{x}), \quad v = 1, \dots, V \quad (4)$$

と表す。ここで、 $\mathbf{g}_v(\mathbf{x}) = (g_{v1}(\mathbf{x}), \dots, g_{vL}(\mathbf{x}))^\top$ を表し、 $\mathbf{W}_v$ は $S \times L$ の混合行列であり、 (s, l) 要素である $w_{vsl} \in \mathbb{R}$ は混合係数を表す。また、 $\mathbf{n}_v(\mathbf{x})$ は各データに対するノイズ過程を表し、 S 変量の平均0のガウス過程であり、

[0049] [数5]

$$\mathbf{n}_v(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \mathbf{\Lambda}_v(\mathbf{x}, \mathbf{x}')), \quad v = 1, \dots, V \quad (5)$$

と表される。ここで、ベクトル $\mathbf{0}$ は全要素に0（ゼロ）を持つベクトルであり、 $\mathbf{\Lambda}_v(\mathbf{x}, \mathbf{x}')$ は

[0050]

[数6]

$$\mathbf{\Lambda}_v(\mathbf{x}, \mathbf{x}') = \text{diag}(\lambda_{v1}(\mathbf{x}, \mathbf{x}'), \dots, \lambda_{vS}(\mathbf{x}, \mathbf{x}')), \quad v = 1, \dots, V \quad (6)$$

とする。 $\lambda_{vs}(\mathbf{x}, \mathbf{x}') : X_v \times X_v \rightarrow \mathbb{R}$ は v 番目のドメインの s 番目のデータに対するノイズ過程の相関関数であり、任意のものを使用できる。ここでは、

[0051] [数7]

$$\lambda_{vs}(\mathbf{x}, \mathbf{x}') = \alpha_{vs}^2 \exp\left(-\frac{1}{2\kappa_{vs}^2} \|\mathbf{x} - \mathbf{x}'\|^2\right) \quad (7)$$

を用いるものとする。式(4)における $g_v(\mathbf{x})$ と $n_v(\mathbf{x})$ は積分消去を行うことができ、 S 変量ガウス過程は

[0052] [数8]

$$\mathbf{f}_v(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, \mathbf{K}_v(\mathbf{x}, \mathbf{x}')), \quad v = 1, \dots, V \quad (8)$$

と書ける。ここで、 $\mathbf{K}_v(\mathbf{x}, \mathbf{x}') : X \times X \rightarrow \mathbb{R}^{S \times S}$ は相関行列を表し、

[0053] [数9]

$$\mathbf{K}_v(\mathbf{x}, \mathbf{x}') = \mathbf{W}_v \mathbf{\Gamma}(\mathbf{x}, \mathbf{x}') \mathbf{W}_v^\top + \mathbf{\Lambda}_v(\mathbf{x}, \mathbf{x}'), \quad v = 1, \dots, V \quad (9)$$

となる。ここで、 $\mathbf{\Gamma}(\mathbf{x}, \mathbf{x}') = \text{diag}(\gamma_1(\mathbf{x}, \mathbf{x}'), \dots, \gamma_S(\mathbf{x}, \mathbf{x}'))$ 、

$\gamma_L(x, x')$ とした。

[0054] 次に、混合係数に対する事前分布の導入を行う。混合係数 w_{vsl} に対する事前分布を

$$w_{vsl} \sim p(w_{sl})$$

とする。 $p(w_{sl})$ は任意の分布を用いることができるが、ここではガウス分布を用いて

[0055] [数10]

$$w_{vsl} \sim \mathcal{N}(w_{vsl} | \bar{w}_{sl}, \eta^2) \quad (10)$$

とする。ここで、 $\bar{w}_{sl}$ 及び η^2 は混合係数に対する事前分布のハイパーパラメータである。なお、明細書のテキストの記載の便宜上、文字の頭に置かれる記号は、例えば「 $\bar{w}_{sl}$ 」のように、文字の前に記載している。

[0056] 次に、集約データの観測モデルについて説明する。集約データの値を該当領域におけるガウス過程の積分値を平均に持つガウス分布の実現値で表現したものが観測モデルである。 $v = 1, \dots, V$ に対して同様のモデルを用いるため、 v 番目のドメインに対する定式化について説明する。 s 番目のガウス過程から生成される N_{vs} 次元の観測ベクトルを $y_{vs} = (y_{vs1}, \dots, y_{vsN_{vs}})^T$ と表し、 S 個のガウス過程から生成される観測ベクトルをまとめて

[0057] [数11]

$$y_v = \begin{pmatrix} y_{v1} \\ y_{v2} \\ \vdots \\ y_{vS} \end{pmatrix} \quad (11)$$

と表す。 y_v は多次元ガウス分布

[0058] [数12]

$$y_v | f_v(x) \sim \mathcal{N} \left(y_v | \int_{X_v} A_v(x) f_v(x) dx, \Sigma_v \right) \quad (12)$$

に従うと仮定する。ここで、 $N_v = \sum_{s=1}^S N_{vs}$ として、 $A(x) : X_v \rightarrow \mathbb{R}^{N_v \times S}$ であり、

[0059] [数13]

$$A_v(x) = \begin{pmatrix} a_{v1}(x) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & a_{v2}(x) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & a_{vS}(x) \end{pmatrix} \quad (13)$$

とする。ここで、 $a_{vs}(x) = (a_{vs1}(x), \dots, a_{vsN_{vs}}(x))^T$ とした。なお、明細書での記載の便宜上、 $a_{vsN_{vs}}(x)$ と記載しているが、 $a_{vsN_{vs}}$ に関して、これは、「 vsN_{vs} 」を a に下付きで付けたものを意図している。 $a_{vsn}(x)$ は任意のものを使用することができ、 $a_{vsn}(x)$ の設定の仕方によって各領域における集約の仕方を変えることができる。ここでは各領域 R_{vsn} において領域平均をした結果、観測が得られる場合を考える。このとき $a_{vsn}(x)$ は

[0060] [数14]

$$a_{vsn}(x) = \frac{\mathbb{1}(x \in \mathcal{R}_{vsn})}{\int_{X_v} \mathbb{1}(x' \in \mathcal{R}_{vsn}) dx'} \quad (14)$$

と書ける。ここで、 $1(\cdot)$ は指示関数であり、 C が真のとき $1(C) = 1$ を出力し、そうでないとき $1(C) = 0$ を出力する。また、

[0061] [数15]

$$\Sigma_v = \begin{pmatrix} \sigma_{v1}^2 \mathbf{I} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \sigma_{v2}^2 \mathbf{I} & \dots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \dots & \sigma_{vS}^2 \mathbf{I} \end{pmatrix} \quad (15)$$

とし、 σ_{vs}^2 は s 番目のデータに対するノイズ分散パラメータである。ここで、 $\mathbf{I}$ は単位行列であり、 $\mathbf{O}$ は全要素を 0 とする行列である。

[0062] 次に、各種パラメータを学習する方法について説明する。上記の式 (12) のガウス過程モデル (観測モデル) から観測データが生成されたと仮定したとき、対数周辺尤度は

[0063] [数16]

$$\ln p(\{\mathbf{y}_v\}) = \ln \prod_v \iint p(\mathbf{y}_v, \mathbf{W}_v, \mathbf{f}_v(\mathbf{x})) d\mathbf{f}_v(\mathbf{x}) d\mathbf{W}_v \quad (16)$$

[0064] [数17]

$$= \sum_v \ln \int \left[\int p(\mathbf{y}_v | \mathbf{W}_v, \mathbf{f}_v(\mathbf{x})) p(\mathbf{f}_v(\mathbf{x})) d\mathbf{f}_v(\mathbf{x}) \right] p(\mathbf{W}_v) d\mathbf{W}_v \quad (17)$$

[0065]

[数18]

$$= \sum_v \ln \int p(\mathbf{y}_v | \mathbf{W}_v) p(\mathbf{W}_v) d\mathbf{W}_v \quad (18)$$

と記述できる。ここで、

[0066] [数19]

$$p(\mathbf{y}_v | \mathbf{W}_v) = \mathcal{N}(\mathbf{y}_v | \mathbf{0}, \mathbf{C}_v) \quad (19)$$

と解析的に計算可能である。また、 $\mathbf{C}_v$ は $N_v \times N_v$ の相関行列であり、

[0067] [数20]

$$\mathbf{C}_v = \iint_{\mathbf{X}_v \times \mathbf{X}_v} \mathbf{A}_v(\mathbf{x}) \mathbf{K}_v(\mathbf{x}, \mathbf{x}') \mathbf{A}_v(\mathbf{x}')^\top d\mathbf{x} d\mathbf{x}' + \Sigma_v \quad (20)$$

である。式(20)における二重積分の計算について説明する。入力空間 $\mathbf{X}_v$ の次元数が $d = 1$ の場合には、非特許文献3と同様の方法によって解析的に計算が可能である。 $d > 1$ の場合には、積分を解析的に計算することは困難なので、非特許文献2と同様の方法によって離散近似を行い数値的に計算する。式(18)における $p(\mathbf{W}_v)$ は

[0068] [数21]

$$p(\mathbf{W}_v) = \prod_{s,l} \mathcal{N}(w_{vsl} | \bar{w}_{sl}, \eta^2) \quad (21)$$

である。式（１８）を最大化することで各種パラメータの解を得ることができる。ここで、 W_v についての積分は解析的には行えないため、任意の近似方法を用いる。例えば、数値近似、モンテカルロ近似、変分近似等である。近似後の目的関数を最大化するための方法としては、任意の連続最適化手法が使用可能である。例えばBFGS法を用いて解くことができる。

[0069] 以下では、式（１８）における W_v についての積分を、変分近似を用いて近似する場合について説明する。本手法は変分ベイズ法と呼ばれる。変分ベイズ法では式（１８）の変分下限（Evidence lower bound）を算出し、それを目的関数としてパラメータ推定を行う。イェンセンの不等式を用いて、変分下限は

[0070] [数22]

$$\ln p(\{y_v\}) \geq \sum_v \int Q(W_v) \ln \frac{p(y_v | W_v) p(W_v)}{Q(W_v)} dW_v \quad (22)$$

[0071] [数23]

$$= \sum_v \mathbb{E} [\ln p(y_v | W_v)] - KL [Q(W_v) || p(W_v)] \quad (23)$$

と記述できる。ここで、 $Q(W_v)$ は提案分布と呼ばれ、 W_v の事後分布を近似するために導入される。 $Q(W_v)$ は任意の確率分布を設定することができるが、ここでは簡単のため W_v の各要素 $w_{v,i}$ が独立なガウス分布に従うものとする。このような近似は平均場近似と呼ばれ、

[0072]

[数24]

$$Q(\mathbf{W}_v) = \prod_{s,l} \mathcal{N}(w_{vsl} | \bar{w}'_{vsl}, \eta'_{vsl}{}^2) \quad (24)$$

と表される。ここで、 $\bar{w}'_{vsl}$ 及び $\eta'_{vsl}{}^2$ は変分パラメータである。このとき、式(23)の第二項におけるKLダイバージェンスは解析的に計算ができるが、第一項の期待値は解析的に計算が困難である。そこで以下のようにサンプリングを用いた近似を行う。

[0073] [数25]

$$\mathbb{E}[\ln p(\mathbf{y}_v | \mathbf{W}_v)] = \int Q(\mathbf{W}_v) \ln p(\mathbf{y}_v | \mathbf{W}_v) d\mathbf{W}_v \quad (25)$$

[0074] [数26]

$$\approx \ln p(\mathbf{y}_v | \hat{\mathbf{W}}_v). \quad (26)$$

ここで、 $\hat{\mathbf{W}}_v \sim Q(\mathbf{W}_v)$ である。また、変分パラメータの推定を可能にするために非特許文献1と同様にReparameterization trickを用いて $\hat{\mathbf{W}}_v$ の各成分を

[0075] [数27]

$$\hat{w}_{vsl} = \bar{w}'_{vsl} + \epsilon \cdot \eta'_{vsl}{}^2 \quad (27)$$

とする。ここで、 $\varepsilon \sim N(0, 1)$ とする。

[0076] 以上より、パラメータ推定部6は、式(23)を最大化するように各種パラメータを推定する演算を行って、各種パラメータを求める。最大化のためには、任意の連続最適化手法を使用することができる。

[0077] パラメータ推定部6が推定すべきパラメータをまとめると下記のとおりである。

- [0078] ・空間スケールパラメータ $\{\beta_l \mid l = 1, \dots, L\}$
- ・混合係数 $\{W_v \mid v = 1, \dots, V\}$
- ・ノイズに対するパラメータ $\{\alpha_{vs} \mid v = 1, \dots, V; s = 1, \dots, S\}$ と $\{\kappa_{vs} \mid v = 1, \dots, V; s = 1, \dots, S\}$ と $\{\Sigma_v \mid v = 1, \dots, V\}$
- ・事前分布のハイパーパラメータ $\{-w_{s,l} \mid s = 1, \dots, S; l = 1, \dots, L\}$ と η^2

なお、変分パラメータは上記パラメータを推定するための補助変数として機能し、後述する高解像度データ算出部11においては使用されない。推定されたパラメータは、下記の各格納部に格納される。なお、複数のパラメータを格納する格納部が1つだけ備えられることとしてもよい。

[0079] <空間スケールパラメータ格納部7>

空間スケールパラメータ格納部7は、パラメータ推定部6で求めた $\{\beta_l \mid l = 1, \dots, L\}$ を格納する。空間スケールパラメータ格納部6は、この情報が保存され、復元可能なものであればどのようなものでもよい。例えば、空間スケールパラメータ格納部6は、データベースや、あらかじめ備えられた汎用的な記憶装置（メモリやハードディスク装置）の特定領域であってもよい。

[0080] <混合係数格納部8>

混合係数格納部8は、パラメータ推定部6で求めた $\{W_v \mid v = 1, \dots, V\}$ を格納する。混合係数格納部8は、この情報が保存され、復元可能なものであればどのようなものでもよい。例えば、混合係数格納部8は、データ

ベースや、あらかじめ備えられた汎用的な記憶装置（メモリやハードディスク装置）の特定領域であってもよい。

[0081] <ノイズに対するパラメータ格納部 9>

イズに対するパラメータ格納部 9 は、パラメータ推定部 6 で求めた $\{\alpha_{vs} \mid v=1, \dots, V; s=1, \dots, S\}$ 、 $\{\kappa_{vs} \mid v=1, \dots, V; s=1, \dots, S\}$ 、 $\{\Sigma_v \mid v=1, \dots, V\}$ を格納する。ノイズに対するパラメータ格納部 9 は、この情報が保存され、復元可能なものであればどのようなものでもよい。例えば、ノイズに対するパラメータ格納部 9 は、データベースや、あらかじめ備えられた汎用的な記憶装置（メモリやハードディスク装置）の特定領域であってもよい。

[0082] <事前分布のハイパーパラメータ格納部 10>

事前分布のハイパーパラメータ格納部 10 は、パラメータ推定部 6 で求めた $\{-w_{s,l} \mid s=1, \dots, S; l=1, \dots, L\}$ 、 η^2 を格納する。事前分布のハイパーパラメータ格納部 10 は、この情報が保存され、復元可能なものであればどのようなものでもよい。事前分布のハイパーパラメータ格納部 10 は、例えば、データベースや、あらかじめ備えられた汎用的な記憶装置（メモリやハードディスク装置）の特定領域であってもよい。

[0083] 上記の各格納部に格納されているパラメータを使用して、高解像度データ算出部 11 が、高解像度データの算出を行う。以下、高解像度データ算出部 11 の処理内容を詳細に説明する。

[0084] <高解像度データ算出部 11>

高解像度データ算出部 11 は、検索部 4 で指定された集約データ、及び、ターゲット分割に対して、学習済みの各種パラメータを用いて高解像度データを算出し、出力部 12 へ渡す。以下に、高解像度データの算出法について説明する。ドメインによらず事後分布の導出は同様に行うことができるので、以下では v 番目のドメインについて説明をする。まず、 v 番目のドメインの S 変量ガウス過程 $f_v(x)$ の事後プロセスを導出する。事後プロセス $f^*_v(x)$ は、

[0085] [数28]

$$\mathbf{f}_v^*(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}_v^*(\mathbf{x}), \mathbf{K}_v^*(\mathbf{x}, \mathbf{x}')) \quad (28)$$

と書ける。ここで、 $\mathbf{m}_v^*(\mathbf{x}) : X_v \rightarrow \mathbb{R}^S$ は平均ベクトルを表し、 $\mathbf{K}_v^*(\mathbf{x}, \mathbf{x}') : X_v \times X_v \rightarrow \mathbb{R}^{S \times S}$ は相関行列を表す。次に、 $\mathbf{H}_v(\mathbf{x}) : X_v \rightarrow \mathbb{R}^{N_v \times S}$ を

[0086] [数29]

$$\mathbf{H}_v(\mathbf{x}) = \int_{X_v} \mathbf{A}_v(\mathbf{x}') \mathbf{K}_v(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \quad (29)$$

と置く。 $\mathbf{H}_v(\mathbf{x})$ を用いて、

[0087] [数30]

$$\mathbf{m}_v^*(\mathbf{x}) = \mathbf{H}_v(\mathbf{x})^\top \mathbf{C}_v^{-1} \mathbf{y}_v, \quad (30)$$

[0088] [数31]

$$\mathbf{K}_v^*(\mathbf{x}, \mathbf{x}') = \mathbf{K}_v(\mathbf{x}, \mathbf{x}') - \mathbf{H}_v(\mathbf{x})^\top \mathbf{C}_v^{-1} \mathbf{H}_v(\mathbf{x}') \quad (31)$$

と記述できる。高解像度データ算出部11は、事後平均(30)をターゲット分割 P^{target} に含まれる各領域 R^{target} において積分することによって

、目的の高解像度データを算出する。

[0089] <出力部 1 2>

出力部 1 2 は、高解像度データ算出部 1 1 からの情報に基づいて、高解像度データを出力する。ここで、出力とは、ディスプレイへの表示、プリンタへの印字、音出力、外部装置への送信等を含む概念である。出力部 1 2 は、ディスプレイやスピーカ等の出力デバイスを含むと考えても含まないと考えてもよい。出力部 1 2 は、出力デバイスのドライバソフト又は、出力デバイスのドライバソフトと出力デバイス等で実現され得る。

[0090] 図 4 に出力例を示す。検索部 3 から対象とするドメインの引数、集約データの引数、及び、ターゲット分割を受け取り、それに応じて、出力部 1 2 において高解像度データの可視化結果が表示される。

[0091] 可視化結果における濃淡は、例えば、データ値に比例して決まる。このような出力例を用いて、例えば、貧困度が高い領域や大気汚染度が高い領域を詳細に絞りこみ、より適切な介入を行うための検討に活用できる。

[0092] (実施の形態の効果)

以上、説明したとおり、本実施の形態では、データ間の依存関係を表すパラメータ（混合係数）に対する事前分布を導入し、それを加味したパラメータ推定を行うこととしているので、ドメイン間の類似度を同時に推定しつつ、複数ドメインにおける多種多様なデータを活用して、高解像度データの高精度な予測を実現することが可能となる。

[0093] (付記)

本明細書には、少なくとも下記各項のパラメータ推定装置、集約データ高解像度化装置、パラメータ推定方法、集約データ高解像度化方法、及びプログラムが開示されている。

(第 1 項)

粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定装置であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約

データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、

前記複数のパラメータを格納する格納部と、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

パラメータ推定装置。

(第2項)

前記モデルは、ドメインの空間を分割することより得られる複数の領域における各領域の集約データの値を、該当領域における多変量ガウス過程の積分値でモデル化したモデルである

第1項に記載のパラメータ推定装置。

(第3項)

前記パラメータ推測部は、変分ベイズ法を用いて前記複数のパラメータを推定する

第1項又は第2項に記載のパラメータ推定装置。

(第4項)

第1項ないし第3項のうちいずれか1項に記載のパラメータ推定装置により推定された前記複数のパラメータを用いて、集約データから高解像度データを算出する高解像度データ算出部

を備える集約データ高解像度化装置。

(第5項)

粗い粒度に集約された集約データから高解像度データを算出する集約データ高解像度化装置であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するよ

うに、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、

前記複数のパラメータを格納する格納部と、

前記複数のパラメータを用いて、集約データから高解像度データを算出する高解像度データ算出部と、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

集約データ高解像度化装置。

(第6項)

粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定装置が実行するパラメータ推定方法であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するステップと、

前記複数のパラメータを格納部に格納するステップと、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

パラメータ推定方法。

(第7項)

粗い粒度に集約された集約データから高解像度データを算出する集約データ高解像度化装置が実行する集約データ高解像度化方法であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するステ

ップと、

前記複数のパラメータを格納部に格納するステップと、

前記複数のパラメータを用いて、集約データから高解像度データを算出するステップと、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

集約データ高解像度化方法。

(第8項)

コンピュータを、第1項ないし第3項のうちいずれか1項に記載のパラメータ推定装置における各部として機能させるためのプログラム、又は、第5項に記載の集約データ高解像度化装置における各部として機能させるためのプログラム。

[0094] 以上、本実施の形態について説明したが、本発明はかかる特定の実施形態に限定されるものではなく、特許請求の範囲に記載された本発明の要旨の範囲内において、種々の変形・変更が可能である。

符号の説明

[0095] 100 集約データ高解像度化装置

- 1 集約データ格納部
- 2 ターゲット分割格納部
- 3 操作部
- 4 検索部
- 5 高解像度データ処理部
- 6 パラメータ推定部
- 7 空間スケールパラメータ格納部
- 8 混合係数格納部
- 9 ノイズに対するパラメータ格納部
- 10 事前分布のハイパーパラメータ格納部
- 11 高解像度データ算出部

1 2 出力部

1 0 0 0 ドライブ装置

1 0 0 1 記録媒体

1 0 0 2 補助記憶装置

1 0 0 3 メモリ装置

1 0 0 4 C P U

1 0 0 5 インタフェース装置

1 0 0 6 表示装置

1 0 0 7 入力装置

請求の範囲

- [請求項1] 粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定装置であって、
- 実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、
- 前記複数のパラメータを格納する格納部と、を備え、
- 前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む
- パラメータ推定装置。
- [請求項2] 前記モデルは、ドメインの空間を分割することより得られる複数の領域における各領域の集約データの値を、該当領域における多変量ガウス過程の積分値でモデル化したモデルである
- 請求項1に記載のパラメータ推定装置。
- [請求項3] 前記パラメータ推測部は、変分ベイズ法を用いて前記複数のパラメータを推定する
- 請求項1又は2に記載のパラメータ推定装置。
- [請求項4] 請求項1ないし3のうちいずれか1項に記載のパラメータ推定装置により推定された前記複数のパラメータを用いて、集約データから高解像度データを算出する高解像度データ算出部
- を備える集約データ高解像度化装置。
- [請求項5] 粗い粒度に集約された集約データから高解像度データを算出する集約データ高解像度化装置であって、
- 実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多

変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するパラメータ推測部と、

前記複数のパラメータを格納する格納部と、

前記複数のパラメータを用いて、集約データから高解像度データを算出する高解像度データ算出部と、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

集約データ高解像度化装置。

[請求項6] 粗い粒度に集約された集約データから高解像度データを算出するために使用される複数のパラメータを推定するパラメータ推定装置が実行するパラメータ推定方法であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数のパラメータを推定するステップと、

前記複数のパラメータを格納部に格納するステップと、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

パラメータ推定方法。

[請求項7] 粗い粒度に集約された集約データから高解像度データを算出する集約データ高解像度化装置が実行する集約データ高解像度化方法であって、

実際に観測された集約データが、複数ドメインにおける複数の種類の集約データに対する複数の潜在ガウス過程を線形混合で表現した多変量ガウス過程に基づくモデルから生成されたものと仮定して、周辺尤度を最大化するように、前記モデルにおける未知変数である複数の

パラメータを推定するステップと、

前記複数のパラメータを格納部に格納するステップと、

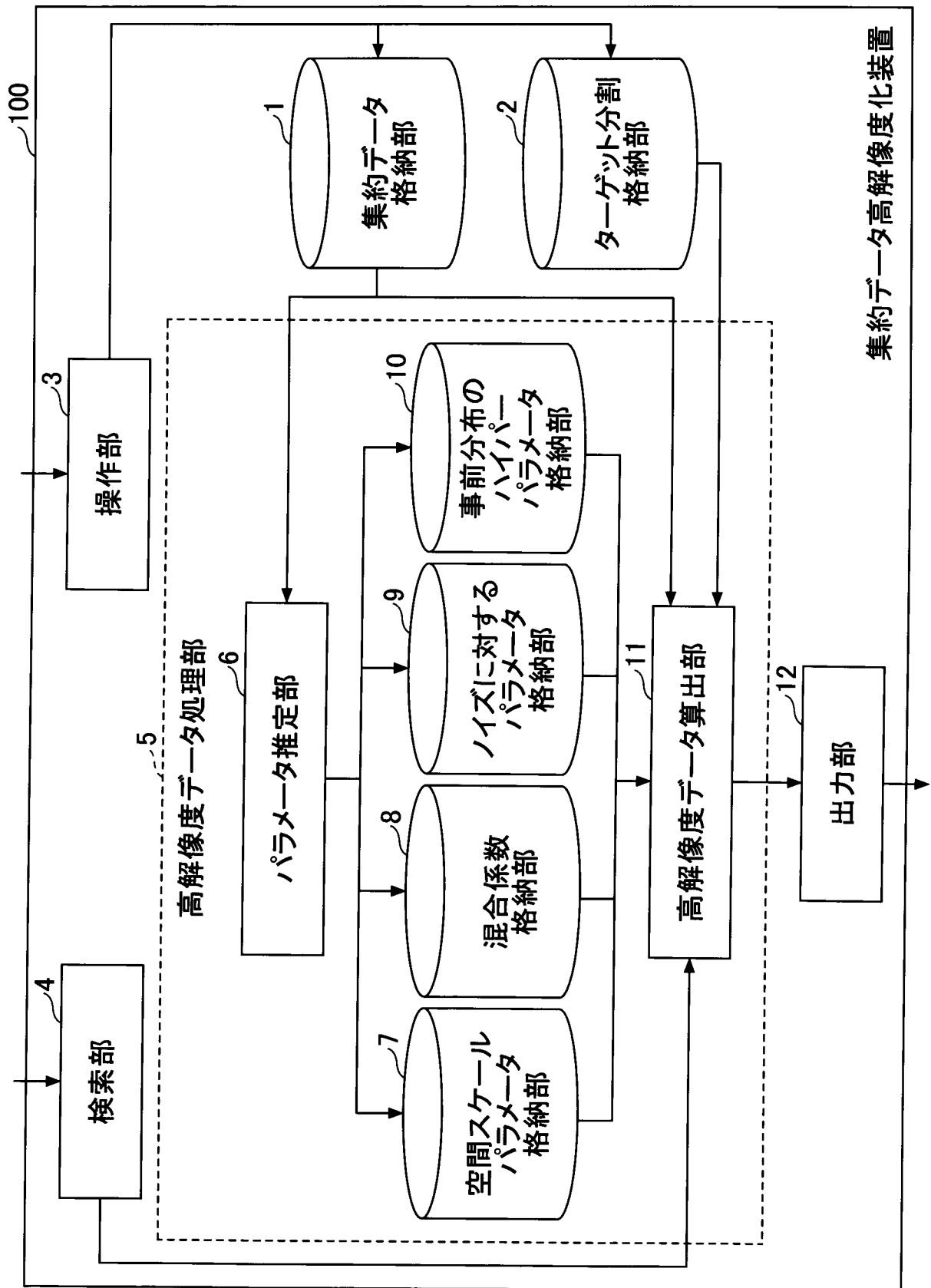
前記複数のパラメータを用いて、集約データから高解像度データを算出するステップと、を備え、

前記複数のパラメータは、前記線形混合で用いられる混合係数に対する事前分布のハイパーパラメータを含む

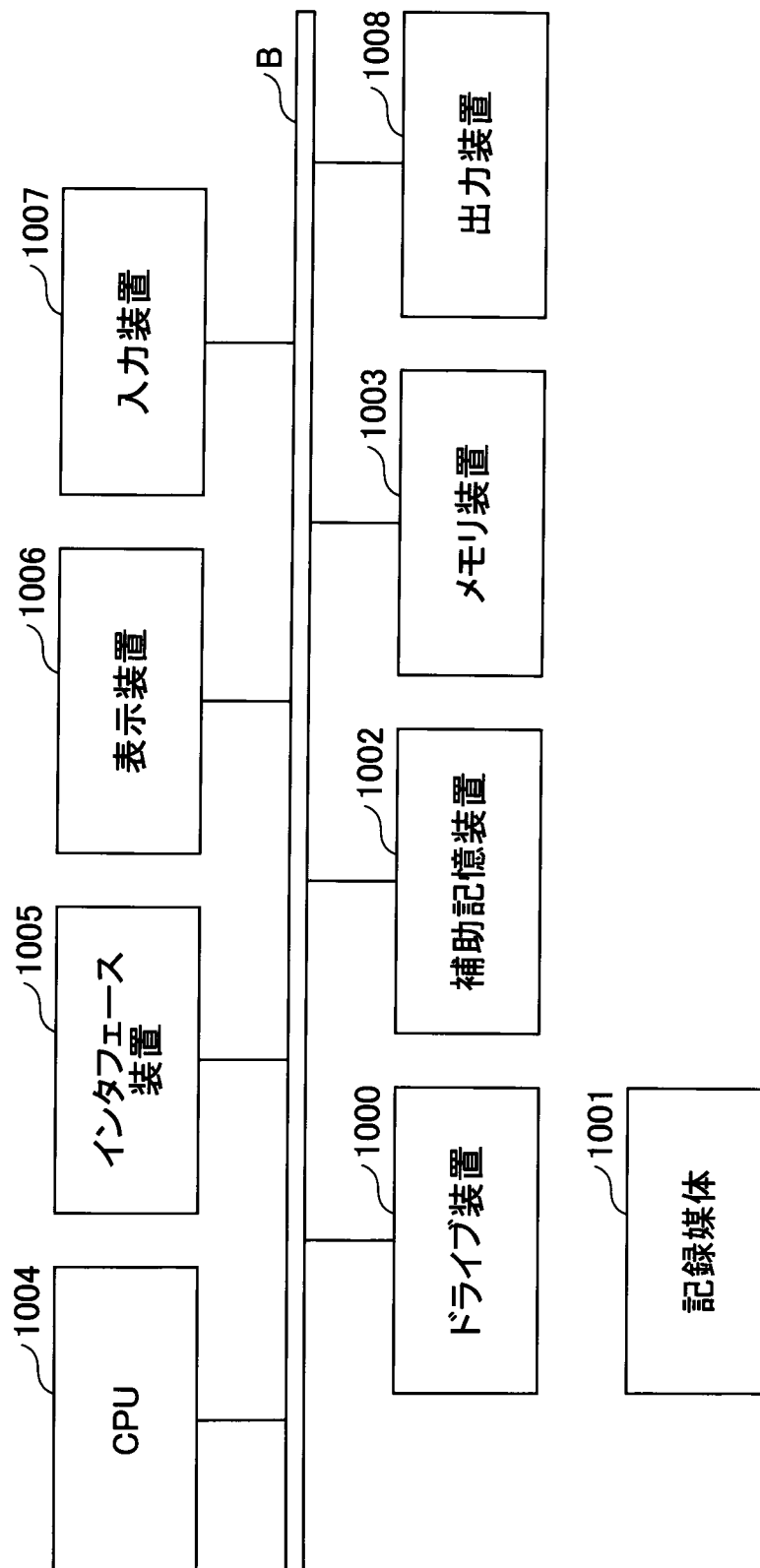
集約データ高解像度化方法。

[請求項8] コンピュータを、請求項1ないし3のうちいずれか1項に記載のパラメータ推定装置における各部として機能させるためのプログラム、又は、請求項5に記載の集約データ高解像度化装置における各部として機能させるためのプログラム。

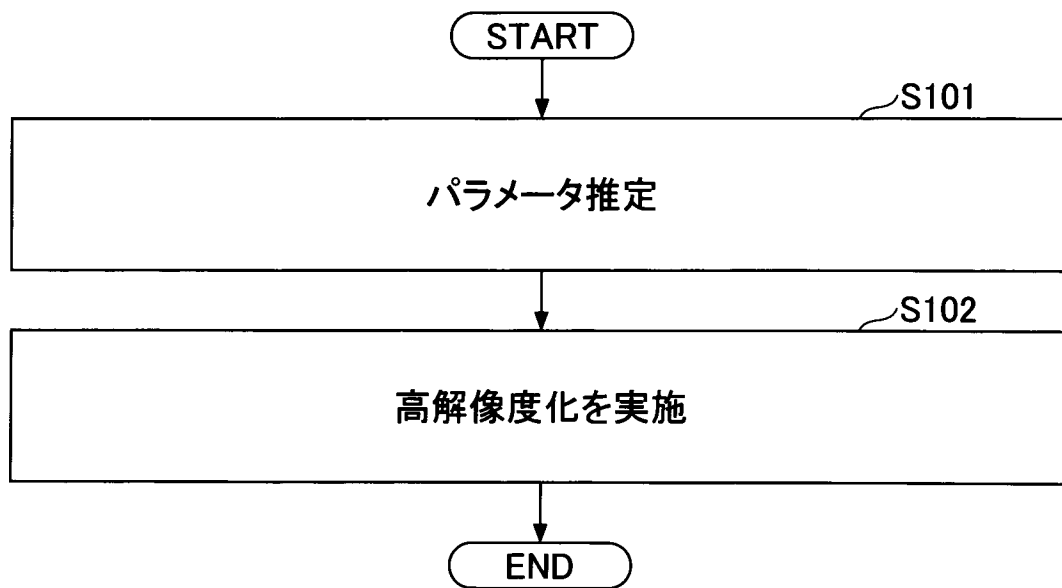
[図1]



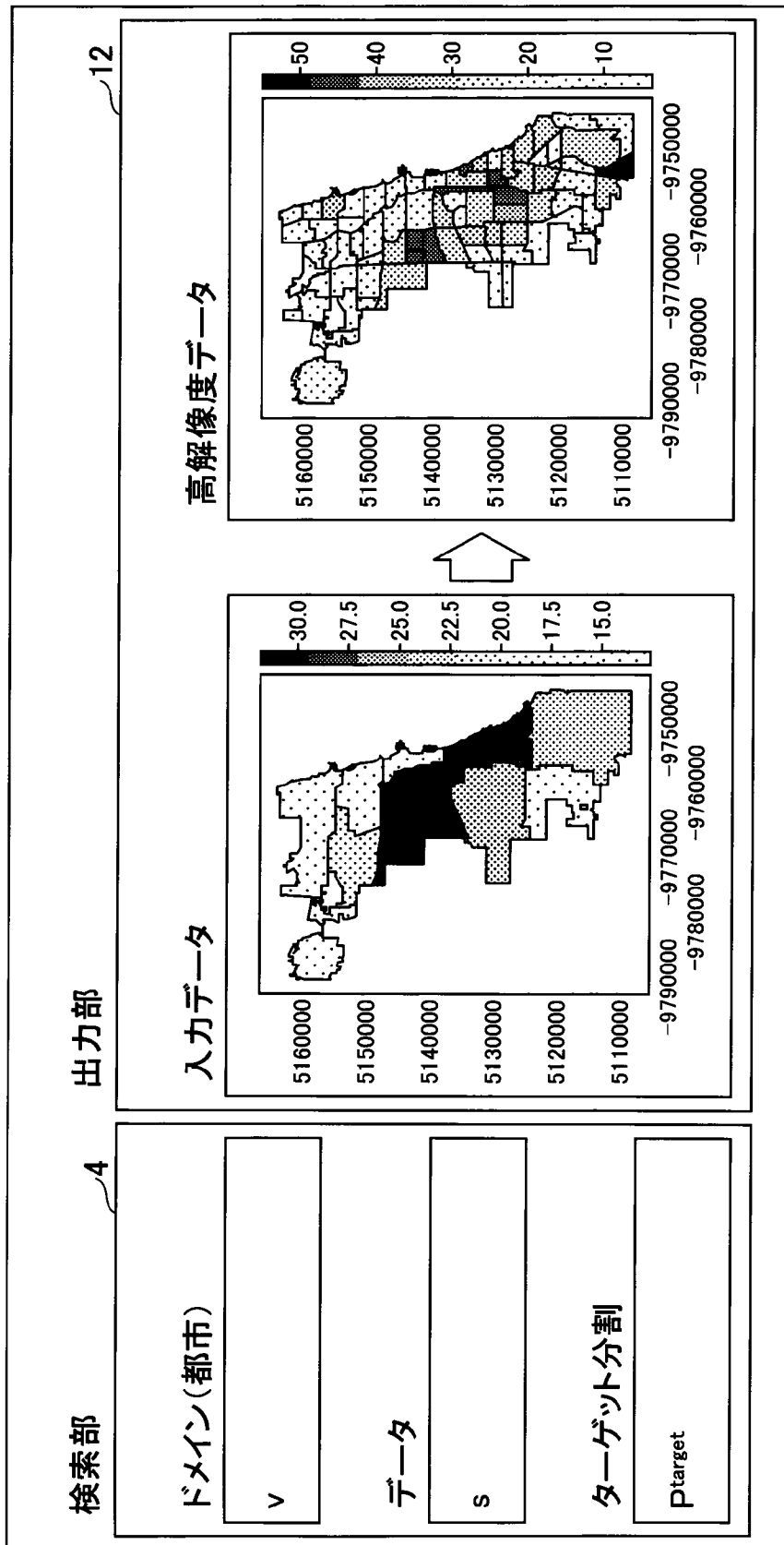
[図2]



[図3]



[図4]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/046123

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. G06N20/00 (2019.01) i

FI: G06N20/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl. G06N20/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2017-033198 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 09 February 2017 (2017-02-09), paragraphs [0012]-[0015], [0022]-[0024], [0043]-[0050]	1-8
A	EP 2530626 A1 (BAE SYSTEMS PLC.) 05 December 2012 (2012-12-05), paragraphs [0002], [0003], [0084]-[0111]	1-8

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
22.01.2021

Date of mailing of the international search report
02.02.2021

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2020/046123

Patent Documents referred to in the Report	Publication Date	Patent Family	Publication Date
JP 2017-033198 A	09.02.2017	(Family: none)	
EP 2530626 A1	05.12.2012	US 2014/0101083 A1	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06N 20/00(2019.01)i FI: G06N20/00														
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） G06N20/00 最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2021年</td> </tr> </table> 国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2021年	日本国実用新案登録公報	1996 - 2021年	日本国登録実用新案公報	1994 - 2021年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2021年													
日本国実用新案登録公報	1996 - 2021年													
日本国登録実用新案公報	1994 - 2021年													
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	JP 2017-033198 A（日本電信電話株式会社）09.02.2017（2017 - 02 - 09） [0012]-[0015], [0022]-[0024], [0043]-[0050]	1-8												
A	EP 2530626 A1（BAE SYSTEMS PLC.）05.12.2012（2012 - 12 - 05） [0002]-[0003], [0084]-[0111]	1-8												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 22.01.2021	国際調査報告の発送日 02.02.2021													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 渡部 博樹 5B 3868 電話番号 03-3581-1101 内線 3545													

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2020/046123

引用文献			公表日	パテントファミリー文献	公表日
JP	2017-033198	A	09.02.2017	(ファミリーなし)	
EP	2530626	A1	05.12.2012	US	2014/0101083 A1