

(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
2010年4月15日(15.04.2010)

PCT



(10) 国際公開番号

WO 2010/041355 A1

(51) 国際特許分類:
G05B 23/02 (2006.01) G06Q 50/00 (2006.01)

市戸塚区吉田町292番地株式会社日立製作所生産技術研究所内 Kanagawa (JP).

(21) 国際出願番号: PCT/JP2009/002391

(22) 国際出願日: 2009年5月29日(29.05.2009)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2008-263030 2008年10月9日(09.10.2008) JP

(71) 出願人(米国を除く全ての指定国について): 株式会社日立製作所(HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 前田俊二(MAEDA, Shunji) [JP/JP]; 〒2440817 神奈川県横浜市戸塚区吉田町292番地株式会社日立製作所生産技術研究所内 Kanagawa (JP). 渋谷久恵(SHIBUYA, Hisae) [JP/JP]; 〒2440817 神奈川県横浜

(74) 代理人: ポレール特許業務法人(Polaire I.P.C.); 〒1040032 東京都中央区八丁堀二丁目7番1号 Tokyo (JP).

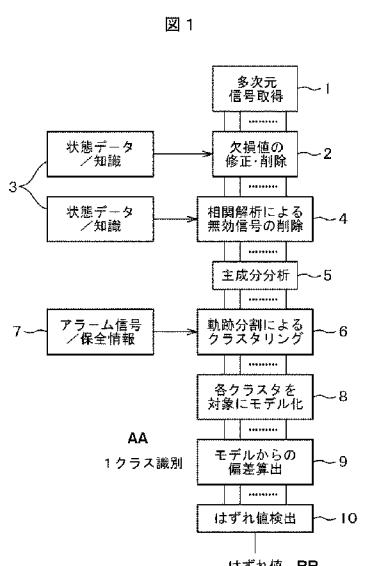
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN,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

[続葉有]

(54) Title: ERROR DETECTION METHOD AND SYSTEM

(54) 発明の名称: 異常検知方法及びシステム



- 1 Acquire multi-dimensional signal
- 2 Correct / eliminate lost values
- 3 State data / knowledge
- 4 Eliminate redundant signals by correlation analysis
- 5 Analyze main components
- 6 Divide traces into clusters
- 7 Alarm signal / maintenance information
- 8 Create model of each cluster
- 9 Calculate the deviation from the model
- 10 Determine the offset value
- AA One-class classification
- BB Offset value

(57) Abstract: Provided are a method which permits complete training data and data with added errors, and enables the early and accurate discovery of errors in facilities such as a plant, and a system thereof. To achieve the objectives, (1) the behavior of temporal data is observed over time, and the trace is divided into clusters; (2) the divided cluster groups are modeled in partial spaces, and the offset values are calculated as error candidates; (3) the training data are used (compare, reference, etc.) for reference to determine the state transitions caused by the changes over time, the environmental changes, the maintenance (parts replacement), and the operation states; and (4) the modeling is a partial space method such as regression analysis or projection distance method of every N data removing N data items ($N = 0, 1, 2, \dots$) (for example, when $N = 1$, one error data item is considered to have been added, this data is removed, then the modeling is performed), or a local partial space method. Linear fitting in regression analysis is equivalent to the lowest order regression analysis.

(57) 要約: 学習データの完全性、異常の混入を許容でき、プラントなどの設備において異常の早期・高精度な発見を可能とする方法及びそのシステムを提供する。上記目的を達成するために、本発明は、(1) 時間的なデータの振舞いに着目し、時間を追って軌跡をクラスタに分割する。(2) 分割したクラスタ群に対して、部分空間でモデルし、はずれ値を異常候補として算出する。(3) 学習データをリファレンスとして活用(比較・参照など)し、経時変化、環境変動、保守(部品交換)、稼動状態による状態遷移を把握する。(4) モデル化は、データの N 個抜き ($N=0, 1, 2, \dots$) の回帰分析法や投影距離法等の部分空間法(例えば $N=1$ の場合は、異常データが 1 個混入していると考え、これを除いてモデル化する)、或いは局部部分空間法によるものとする。なお、回帰分析法における直線の当てはめは、最も低次の回帰分析に相当する。



(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, SE, SI, SK, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG). 添付公開書類:
— 國際調查報告（條約第 21 条(3)）

明 細 書

発明の名称：異常検知方法及びシステム

技術分野

[0001] 本発明は、プラントや設備などの異常を早期に検知する方法に関する。

背景技術

[0002] 電力会社では、ガスタービンの廃熱などを利用して地域暖房用温水を供給したり、工場向けに高圧蒸気や低圧蒸気を供給したりしている。石油化学会社では、ガスタービンなどを電源設備として運転している。このようにガスタービンなどを用いた各種プラントや設備において、その異常を早期に発見することは、社会へのダメージを最小限に抑えることができ、極めて重要である。

[0003] ガスタービンや蒸気タービンのみならず、水力発電所での水車、原子力発電所の原子炉、風力発電所の風車、航空機や重機のエンジン、鉄道車両、エスカレータ、エレベータ、機器・部品レベルでも、搭載電池の劣化・寿命など、早期に異常を発見しなければならない設備は枚挙に暇がない。最近では、健康管理のため、脳波測定・診断に見られるように、人体に対する異常（各種症状）検知も重要になりつつある。

[0004] このため、例えば米国のSmartSignal社では、US6,952,662やUS6,975,962に記載のように、おもにエンジンを対象に、異常検知の業務をサービスしている。そこでは、過去のデータをデータベース（DB）としてもっておき、観測データと過去の学習データとの類似度を独自の方法で計算し、類似度の高いデータの線形結合により推定値を算出して、推定値と観測データのはずれ度合いを出力する。General Electric社のように、US6,216,066出願内容を見ると、異常検知をk-meansクラスタリングにより検出している例もある。

先行技術文献

特許文献

[0005] 特許文献1：US6,952,662

特許文献2：US6,975,962

特許文献3：US6,216,066

非特許文献

- [0006] 非特許文献1：Stephan W. Wegerich ; Nonparametric modeling of vibration signal features for equipment health monitoring、Aerospace Conference, 2003. Proceedings. 2003 IEEE, Volume 7, Issue, 2003 Page(s) :3113 – 3121

発明の概要

発明が解決しようとする課題

- [0007] SmartSignal社が用いている手法では、データベースに格納する過去の学習データは様々な状態を網羅的に含む必要がある。もし、学習データにない観測データが観察されると、これらはすべて学習データには包含されないものとして扱われ、はずれ値と判断され、正常信号においても異常と判定されてしまい、検査の信頼性が著しく低下してしまう。このため、ユーザは、過去のあらゆる状態のデータすべてをDBとして蓄えることが必須になる。
- [0008] 一方で、学習データに異常が混入されている場合、異常を表している観測データとの乖離度が低くなり、これを見逃してしまう。このため、学習データに異常が含まれないよう、十分なるチェックが必要である。
- [0009] このように、SmartSignal社が提案する学習データに基づく方法では、ユーザは、網羅的なデータ収集と異常の排除という負荷を背負うことになる。特に、経時変化や、周囲の環境変動、部品交換などの保守作業の有無などに、きめこまかく対応する必要がある。このような対応を行うことは、実質難しく、不可能な場合が多い。
- [0010] General Electric社の方法では、k-meansクラスタリングのため、信号の挙動を見ておらず、その点で本質的な異常検知にはなっていない。
- [0011] そこで、本発明の目的は、上記課題を解決し、学習データの不完全性や異常の混入を許容し、ユーザ負荷を軽減することを図り、さらに早期に異常を高感度に検知することが可能な異常検知方法およびシステムを提供すること

にある。

課題を解決するための手段

[0012] 上記目的を達成するために、本発明は、（1）時間的なデータの振舞いに着目し、時間を追って軌跡をクラスタに分割する、（2）分割したクラスタ群に対して、部分空間でモデルし、はずれ値を異常候補として算出する。（3）学習データをリファレンスとして活用（比較・参照など）し、経時変化、環境変動、保守（部品交換）、稼動状態による状態遷移を把握する。（4）モデル化は、データのN個抜き（N=0, 1, 2, ···）の回帰分析法や投影距離法等の部分空間法（例えばN=1の場合は、異常データが1個混入していると考え、これを除いてモデル化する）、或いは局所部分空間法によるものとする。

[0013] さらに、（5）部分空間法など、複数の識別器の出力を統合して。異常判断を行う。なお、回帰分析法における直線の当てはめは、最も低次の回帰分析に相当する。

発明の効果

[0014] 本発明によれば、学習データが完全でなくとも、異常の混入を許容でき、プラントなどの設備において異常の早期・高精度な発見を可能とする。即ち、SmartSignal社のように正常部の学習、複数の状態に応じた網羅的データ収集の必要性がない。

図面の簡単な説明

[0015] [図1]本発明の第1の実施例を示すチャートである。

[図2]多次元時系列信号の例である。

[図3]相関行列の例である。

[図4]軌跡分割クラスタリングの適用例である。

[図5]軌跡分割クラスタリングの適用例である。

[図6]軌跡分割クラスタリングの適用例である。

[図7]軌跡分割クラスタリング実施時のモデルとの偏差の例である。

[図8]軌跡分割クラスタリング未実施時のモデルとの偏差の例である。

- [図9]特徴空間にてデータ表示をした一例である。
- [図10]特徴空間にてデータ表示をした他の例である。
- [図11]特徴空間にてデータ表示をしたさらに他の例である。
- [図12]特徴空間にてデータ表示をあいたさらに他の例である。
- [図13]本発明の第2の実施例を示すチャートである。
- [図14]本発明の第3の実施例を示すチャートである。
- [図15]アラーム信号のヒストグラム例である。
- [図16]本発明の第4の実施例を示すチャートである。
- [図17]Wavelet（変換）解析の例である。
- [図18]本発明の第5の実施例を示すチャートである。
- [図19]散布図解析および相互相関解析の例である。
- [図20]本発明の第6の実施例を示すチャートである。
- [図21]時間・周波数解析の例である。
- [図22]本発明の第7の実施例を示すチャートである。
- [図23]本発明の第7の実施例の詳細を支援するチャートである。
- [図24]部分空間法の一例である。
- [図25]Wavelet変換の説明図である。
- [図26]異常診断との関係を説明する図である。
- [図27]ハードウェア構成図
- [図28]局所部分空間法の適用例である。
- [図29]投影距離法、局所部分空間法の適用例である。
- [図30]複数識別器の統合による識別例である。
- [図31]複数識別器の統合による異常検知例である。

発明を実施するための形態

[0016] 以下実施例に従って、本発明の内容を詳細に説明する。

実施例 1

[0017] 説明を簡略化するために、本願においては、同一の構成要素には同一の符号を付加している。本発明の実施の一形態を図1に、対象とする信号の例を図

2に示す。対象信号は、図2に示すような複数の時系列信号である。ここでは、系列1、2、3、4という4種類の信号を表している。実際には、信号は4種類に限るものではなく、数百から数千と言った数になる場合もある。

- [0018] 各信号が、対象プラントや設備に設けられた複数のセンサからの出力に相当する。例えば、シリンダ、オイル、冷却水などの温度、オイルや冷却水の圧力、軸の回転速度、室温、運転時間などが、いろいろなセンサから、1日に数回ないしリアルタイム等、定めた間隔で観測される。出力や状態を表すのみならず、何かを制御するための制御信号（入力）の場合もある。ON/OFF制御のことであれば、一定値になるように制御する場合もある。これらのデータは、お互いに相関が高いものや低いものがある。これら全ての信号が、対象になり得る。これらのデータを見て、異常の有無を判断する。ここでは、多次元時系列信号として扱う。
- [0019] 図1に示した異常検知方法について説明する。まず、多次元時系列信号を多次元信号取得部1にて取得する。次に、取得した多次元時系列信号が、欠損している場合があるため、欠損値の修正・削除を欠損値の修正・削除部2にて行う。欠損の修正は、例えば、前後のデータの置換えや移動平均の置換えが一般的である。削除は、多くのデータが同時に0にリセットされた時など、データとしての異常を排除するものである。欠損値の修正・削除は、設備の状態やエンジニアの知識を状態データ／知識3というDBに蓄えておき、これをもとに行う場合もある。次に、修正・削除された多次元時系列信号に関して、相関解析による無効信号の削除を、相関解析による無効信号の削除部4にて行う。これは、図3にて相関行列の例を示すように、多次元時系列信号に対して相関解析を行い、相関値が1に近い複数の信号がある場合など、極めて類似性が高い場合に、これらは冗長だとして、この複数の信号から重複する信号を削除し、重複しないものを残すものである。この場合も、状態データ／知識3に格納した情報に基づいて、削除する。
- [0020] 次に、データの次元削減を主成分分析部5にて行う。ここでは、主成分分析により、M次元の多次元時系列信号を、次元数rのr次元多次元時系列信

号に線形変換する。主成分分析は、ばらつき最大となる軸を生成するものである。KL変換でも構わない。次元数 r は、主成分分析により求めた固有値を降順に並べ、大きい方から加算した固有値を全固有値の和で割り算した累積寄与率なる値に基づいて決める。

[0021] 次に、 r 次元の多次元時系列信号に対して、軌跡分割によるクラスタリングを軌跡分割によるクラスタリング部 6 にて行う。図 4 にその模様を示す。図 4 の左上の 3 次元表示（特徴空間と呼ぶ）は、主成分分析後の r 次元の多次元時系列信号を寄与率の高い 3 次元で表示したものである。この状態では、まだ対象設備の状態が複雑なものとして観測されていることが分かる。図 4 の残りの八つの 3 次元表示は、軌跡を時間に沿って追跡し、クラスタリングを実施したものであり、各クラスタを表している。クラスタリングは、時間に沿ってデータ間の距離が定めたしきい値を超えると、別のクラスタとし、しきい値を超えない場合は、同じクラスタとして扱うものである。これにより、クラスタ 1、3、9、10、17 は、運転 ON の状態のクラスタであり、クラスタ 6、14、20 は、運転 OFF の状態のクラスタに別れていることが分かる。クラスタ 2 など、図示していないクラスタは、過渡期のものである。これらのクラスタを分析すると、運転 ON の状態では、線状に軌跡移動していることが伺え、運転 OFF の状態では、不安定な軌跡移動していることが伺える。このように、軌跡分割によるクラスタリングは、いくつかのメリットがあることが分かる。

運転 ON の状態、運転 OFF の状態など、複数の状態に分類できる。

運転 ON の状態に見られるように、これらのクラスタは、例えば線形など、低次元のモデルで表現できる。

これらのクラスタリングは、設備のアラーム信号や保全情報を加味して、これらのひも付きとして、実施してもよい。具体的には、各クラスタに、アラーム信号などの情報が属性として付加される。

[0022] 図 5 に、特徴空間にて、クラスタリングによりラベルを付けた状態のほかの例を表す。図 6 は、一つの時系列信号の上に、クラスタリングの結果を表

示したものである。この場合、クラスタは 16 個生成でき、時系列信号が、16 個のクラスタに分割されている様子が分かる。運転時間（累積時間）も重ねて表示した。水平になっている部分は、運転 OFF である。運転 ON と運転 OFF が、精度良く分離できていることが分かる。

[0023] 次に、クラスタリングされた各クラスタを対象に、低次元部分空間でのモデル化を、各クラスタを対象にモデル化部 8 にて行う。正常部限定である必要はなく、異常が混入することは問題ない。ここでは、例えば、回帰分析によりモデル化を行う。回帰分析の一般式は、下記のとおりである。「y」が、クラスタごとの r 次元の多次元時系列信号に相当する。「X」が、y を説明する変数である。「 \hat{y} 」がモデルになる。「e」は偏差である。

y : 目的変数 (r 列)

b : 回帰係数 (1+p列)

X : 説明変数行列 (r 行、1+p列)

$\|y - Xb\| \Rightarrow \min$

$b = (X' X)^{-1} X' y$ (' は転置を表す)

$\hat{y} = Xb = X(X' X)^{-1} X' y$ (説明変数の影響を表す部分)

$e = y - \hat{y}$ (\hat{y} で近似できない部分。説明変数の影響を取り除いた部分)

ただし、rank $X = p+1$

ここでは、各クラスタの r 次元の多次元時系列信号に対し、データの N 個抜き (N=0, 1, 2, ···) の回帰分析を行う。例えば N=1 の場合は、異常信号が 1 種混入していると考え、これを除いた信号を「X」としてモデル化する。N = 0 の場合は、全 r 次元多次元時系列信号を扱うことになる。

[0024] 回帰分析以外に、CLAFIC 法や投影距離法等の部分空間法を適用してもよい。そして、モデルからの偏差をモデルからの偏差算出部 9 にて求める。図 2-4 に一般的な CLAFIC 法を図解して示す。2 クラス、2 次元パターンの場合を示す。各クラスの部分空間、即ち、ここでは 1 次元の直線として表される部分空間を求める。

[0025] 一般的には、各クラスのデータの自己相関行列を固有値分解して、固有ベ

クトルを基底として求める。値が大きい、上位何個かの固有値に対応する固有ベクトルを用いる。未知パターン q （最新の観測パターン）が入力されると、部分空間への正射影の長さ、或いは部分空間への投影距離を求める。そして、正射影の長さが最大、或いは投影距離が短いクラスに、未知パターン（最新の観測パターン） q を分類する。

[0026] 図24では未知パターン q （最新の観測パターン）は、クラスAに分類される。図2に示した多次元時系列信号では、基本的に正常部を対象とするため、1クラス識別の問題になるので（図1にて図示）、クラスAを正常部とし、未知パターン q （最新の観測パターン）からクラスAまでの距離を求めて、これを偏差とする。そして、偏差が大きいと、はずれ値と判断する。このような部分空間法では、異常値が若干混ざっていても、次元削減し、部分空間にした時点で、その影響が緩和される。部分空間法適用のメリットである。

[0027] なお、投影距離法では、各クラスの重心を原点とする。各クラスの共分散行列にKL展開を適用して得られた固有ベクトルを基底として用いる。いろいろな部分空間法が立案されているが、距離尺度を有するものならば、はずれ度合いが算出可能である。なお、密度の場合も、その大小により、はずれ度合いを判断可能である。CLAF10法は、正射影の長さを求ることから、類似度尺度である。

[0028] このように、部分空間にて距離や類似度を計算し、はずれ度合いを評価することになる。投影距離法などの部分空間法は、距離に基づく識別器のため、異常データが利用できる場合の学習法として、辞書パターンを更新するベクトル量子化や距離関数を学習するメトリック学習を使うことができる。

[0029] また、局所部分空間法と呼ぶ、未知パターン q （最新の観測パターン）に近い k 個の多次元時系列信号を求め、各クラスの最近傍パターンが原点となるような線形多様体を生成し、その線形多様体への投影距離が最小となるクラスに未知パターンを分類する方法も適用できる（図31の局所部分空間法の枠内参照）。局所部分空間法も部分空間法の一種である。

- [0030] 既に説明したクラスタリング後の各クラスタに対して、局所部分空間法を適用することになる。 k は、パラメータである。異常検知では、先ほどと同様に、1クラス識別の問題になるので、大多数のデータが属するクラスAを正常部とし、未知パターン q （最新の観測パターン）からクラスAまでの距離を求めて、これを偏差とする。
- [0031] この手法では、例えば、 k 個の多次元時系列信号を用いて形成される部分空間への、未知パターン q （最新の観測パターン）からの正射影した点を推定値として算出することもできる（図31の局所部分空間法の枠内に説明した推定値なるデータ）。また、 k 個の多次元時系列信号を、未知パターン q （最新の観測パターン）に近い順に並べ替え、その距離に反比例した重み付けを行って、各信号の推定値を算出することもできる。投影距離法などでも、同様に推定値を算出できる。
- [0032] パラメータ k は通常は1種類に定めるが、パラメータ k をいくつか変えて実行すると、類似度に応じて対象データを選択することになり、それらの結果から総合的に判断できるため、一層効果的である。局所部分空間法では、クラスタ内に選択したデータを対象とするため、異常値が若干混ざっていても、局所部分空間にした時点で、その影響が大きく緩和される。
- [0033] 局所部分空間法の「局所」という概念は、回帰分析にも適用可能である。即ち、「 y 」として、観測未知パターン q に近い k 個の多次元時系列信号を求め、この y のモデルとして、「 \hat{y} 」を求め、偏差「 e 」を算出する。
- [0034] なお、単に、1クラス識別の問題と考えれば、1クラスサポートベクターマシンなどの識別器も適用可能である。この場合、高次空間に写像する、radial basis functionなどのカーネル化が使えることになる。1クラスサポートベクターマシンでは、原点に近い側が、はずれ値、即ち異常になる。ただし、サポートベクターマシンは、特徴量の次元は大きくても対応できるが、学習データ数が増えると計算量が膨大となるという欠点もある。
- [0035] このため、MIRU2007（画像の認識・理解シンポジウム、Meeting on Image Recognition and Understanding 2007）にて発表されている、「IS-2-10

加藤丈和、野口真身、和田俊和（和歌山大）、酒井薰、前田俊二（日立）；「パターンの近接性に基づく1クラス識別器」などの手法も適用可能であり、この場合、学習データ数が増えても、計算量は膨大なものとならないというメリットがある。

[0036] 次に、回帰分析を例にとり、実験例を説明する。図7に、 $N=0$ とし、線形回帰分析により、 r 次元多次元時系列信号をモデル化し、そのモデルと実測値の偏差を図示した例を示す。図8は、参考に、軌跡分割によるクラスタリングを実施しない場合である。図7の場合、偏差が大きいのは、運転OFFの区間、および、運転ONの区間で時系列信号が振動的な振舞いをしている場合である。最後に、はずれ値検出部10にてはずれ値を求める。ここでは、しきい値との大小をチェックする。検出された異常信号は、主成分分析後であるため、これを逆変換し、元の信号がどのような割合で合成されて、異常と判断されたかを確認することもできる。

[0037] このように、軌跡分割によるクラスタリングを中心に、低次元モデルで多次元時系列信号を表現することにより、複雑な状態を分解でき、簡単なモデルで表現できるため、現象を理解しやすいという利点がある。また、モデルを設定するため、SmartSignal社の方法のように完全に、データを完備する必要はない。データの欠落があっても良いというメリットがある。

[0038] 次に、局所部分空間法の適用例を図28に示す。信号を前半と後半に分け（交差確認と呼ぶ検証方法に従った）、それぞれを学習データとして、残りのデータまでの距離を求めた例である。パラメータ k は10とした。 k をいくつか変えて、それらの多数決をとれば、安定な結果が得られる（後述するバギングという手法と同様の考えに基づく）。この局所部分空間法では、データのN個抜きが、自動で行われるというメリットがある。同図適用例では、運転OFF中の不規則な挙動が、検知されている。

[0039] 上記例では、クラスタリングの必要性も緩和されるが、観測データが属しているクラスタ以外を学習データとし、このデータと観測データに局所部分空間法を適用してもよい。この方法によれば、別クラスタからの乖離度を評

価できる。投影距離法も同様である。図29にそれらの例を示す。観測データが属しているクラスタ以外を学習データとした。この考えは、時系列データのように類似データが連續する場合に、最も類似したデータを「局所」領域から排除できるため、効果的である。なお、データのN個抜きは、特徴量（センサ信号）として説明したが、時間軸方向のデータであってもよい。

- [0040] 次に、データの表現形態を、いくつか図を用いて説明する。図9に幾つかの例を示す。図9の左側の図は、主成分分析後の r 次元時系列信号を2次元表示したものである。データの振舞いを可視化した例となる。図9の右側の図は、軌跡分割によるクラスタリングを実施し、クラスタを図示したものである。クラスタ毎に、簡単な低次のモデル（ここでは直線）で表現した例である。図10の左側の図は、データの動く速さが分かるように図示した例である。後述のWavelet解析を適用すれば、速さ、即ち周波数も分析可能、そして多変量として扱うことが可能となる。図10の右側の図は、図9の右側の図で示したモデルからの偏差が分かるように表示した例である。
- [0041] 図11の左側の図は、別の例である。距離基準などに基づき類似と判定したクラスタをマージし（同図では、隣接したクラスタのマージを示した）、マージ後のモデルを示し、かつモデルからの偏差を図示した例である。図11の右側の図は、状態を表現したものである。状態A、B、Cという3種類の状態を分けて表示した。状態を分けて考えると、図12の左側の図に示すように、状態Aの変化などを図示可能となる。
- [0042] 図6の例で考えると、同じ運転ONの状態でも、運転OFFの前後で、異なる挙動を示しており、これらを特徴空間の中で表現できることとなる。図12の右側の図は、過去の学習データから得られたモデル（低次の部分空間）からの変化を示したものであり、状態変化を観察できる。このように、データを加工し、加工したデータをユーザに示し、現在の状況を可視化することにより、より良い理解を促すことができる。
- [0043] 次に、別の実施例を説明する。既に説明したブロックは省略する。図13に異常検知方法を示す。ここでは、各クラスタに対し、 r 次元多次元時系列

信をランダムに定めた個数分選択する。ランダム選択により、全ての信号を使う場合に見えなかった特性が現れる有効ではない信号を除く全ての組み合わせより、短時間で計算できるという利点がある。

さらに、時間軸方向に r 次元多次元時系列信号をランダムに定めた個数分選択するという選択も考えられる。ここでは、クラスタを単位とすることもあるが、クラスタ内を区分し、これらを定めた個数分、ランダムに選択する。

- [0044] 図 14 に別の実施例を示す。アラーム信号を処理する部分 12 を付加した。図 15 (a) に示すようにアラーム信号の発生履歴を取得する。そして、そのヒストグラムを表示する。頻度が高い区間は、異常の度合いが高いことは容易に想像ができる。そこで、図 15 (b) に示すように、ヒストグラムの頻度も考慮して、図 1 に示す異常特定部 13 にて、アラーム信号とはずれ値を組み合わせて異常の度合いや信頼度を付加したり、異常判定を行う。
- [0045] 図 16 に別の実施例を示す。Wavelet (変換) 解析を付加した例である。Wavelet 解析信号付与部 14 において、M 次元の多次元時系列信号を対象に、図 17 に示す Wavelet 解析を行い、これらの信号を M 次元の多次元時系列信号に加える。M 次元の多次元時系列信号と置換することも可能である。このような新規に加えられたり、置換されたりした多次元時系列信号を対象に、局所部分空間法などの識別器により、異常を検知する。なお、図 17 (a) は後で説明する図 25 におけるスケール 1 の信号に該当し、図 17 (b) は後で説明する図 25 におけるスケール 8 の変動に該当し、図 17 (c) は図 25 におけるスケール 4 の変動に該当し、図 17 (d) は図 25 におけるスケール 2 の変動に該当する。
- [0046] Wavelet 解析は、多重解像度表現を与えるものである。図 25 に Wavelet 解析を図解して示す。スケール 1 の信号がもとの信号である。これを順次隣と加算してスケール 2 の信号を作り、元の信号との差を算出して、スケール 2 の変動信号を生成する。順次、これを繰り返すと、最後に、スケール 8 の一

定値の信号とその変動信号が得られ、結局、元の信号は、スケール2、4、8の各変動信号とスケール8の直流信号に分解できる。従って、このようなスケール2、4、8の各変動信号を、新たな特徴的信号とみなして、多次元時系列信号に追加して扱う。

- [0047] パルスやインパルスなどの非定常信号では、フーリエ変換を行って得られる周波数スペクトルは全域に広がり、個々の信号について特徴を抽出することが困難である。時間に局在したスペクトルを得ることができるWavelet変換は、化学プロセスのような、パルスやインパルスなどを含む、非定常な信号を多く含むデータを対象にした場合に都合がよい。
- [0048] また、一次遅れをもつ系では、時系列の状態だけでは、そのパターンを観測困難であるが、時間・周波数領域上では、識別可能な特徴が現れることがあり、Wavelet変換が有効な場合が多い。
- [0049] なお、Wavelet解析の応用は、電気学会編集で、新誠一著の「ウェーブレット解析の産業応用」2005年、朝倉出版にくわしい。化学プラントの制御系診断、空調プラント制御における異常検知、セメントの焼成プロセスの異常監視、ガラス溶解炉制御など、多くの対象に適用されている。
- [0050] 本実施例における、従来技術との相違点は、Wavelet解析を多重解像度表現として扱い、もとの多次元時系列信号の情報を、Wavelet変換により顕在化した点にある。その上で、これらを多変量として扱うことにより、異常が微弱な段階から早期に検知できる。即ち、予兆として早期に検知可能となる。
- [0051] Wavelet解析を分類に適用したほかの例を説明する。異常データが少なからずあり、教示できる場合の例である。使用する記号を以下に示す。
- [0052]

[数1]

N : 1 系列のデータ長 (サンプル数)

m : AR モデルの次数

K : クラスの数

N_p : トレーニングデータの系列数

y_t : 時系列信号 ($t = 1 \sim N$)

$$z_{t-i} = [y_{t-1}, y_{t-2}, \dots, y_{t-m}]^T$$

$$Z_N = \{y_1, y_2, \dots, y_N\}$$

Z_N^j : j 番目の観測値系列

C_i : クラス i ($i = 1 \sim K$)

f_j : AR 係数 ($j = 1 \sim m$)

$$\Phi = [f_1, f_2, \dots, f_m]^T$$

$$q = [\Phi^T, r]^T$$

$P_r(C_i)$: クラス i の事前確率

$P(C_i | Z_N)$: クラス i の事後確率

$p(Z_N | C_i)$: クラス C_i に関する時系列データ Z_N の条件付き確率密度関数

$p(Z_N | q)$: パラメータ q に関する時系列データ Z_N の条件付き確率密度関数

[0053] まず、以下のモデルを仮定する。

[0054]

[数2]

$$y_t = \sum_{j=1}^m \phi_j y_{t-j} + v_t = \Phi^T z_{t-1} + v_t, \quad m+1 \leq t \leq N$$

ただし、 v_t は平均値 0、分散 ρ のガウス性白色系列である。

ここで、 $p(Z_N|\theta)$ を考える。

$$p(Z_N|\theta) = p(y_N, y_{N-1}, \dots, y_1|\theta)$$

$$= p(y_1, y_2, \dots, y_m, \theta) p(\theta) \times \prod_{t=m+1}^N p(y_t|y_{t-1}, y_{t-2}, \dots, y_1, \theta)$$

y_1, y_2, \dots, y_m は、観測値として用い、 $p(y_1, y_2, \dots, y_m, \theta) = p(y_1, y_2, \dots, y_m)p(\theta)$ と仮定すると、

$$p(Z_N|\theta) = p(y_1, y_2, \dots, y_m) \prod_{t=m+1}^N p(y_t|y_{t-1}, y_{t-2}, \dots, y_1, \theta)$$

ここで、 v_t はガウス分布であると仮定しているので、 $p(y_t|y_{t-1}, \dots, y_1, \theta)$ もガウス分布となり、

$$p(y_t|y_{t-1}, \dots, y_1, \theta) = \frac{1}{\sqrt{2\pi\rho}} \exp\left(-\frac{1}{2\rho}(y_t - \Phi^T z_{t-1})^2\right)$$

従って、 $p(Z_N|\theta)$ は次式で表される。

$$p(Z_N|\theta) = p(y_1, y_2, \dots, y_m) \left(\frac{1}{2\pi\rho} \right)^{(N-m)/2} \exp\left(-\frac{1}{2\rho} \sum_{t=m+1}^N (y_t - \Phi^T z_{t-1})^2\right)$$

パラメータの最尤推定値は、対数尤度 $\ln p(Z_N|\theta)$ を、 Φ, ρ により偏微分し 0 とおいて、次のように求められる。

$$\hat{\Phi} = \left(\sum_{t=m+1}^N z_{t-1} z_{t-1}^T \right)^{-1} \left(\sum_{t=m+1}^N z_{t-1} y_t \right)$$

$$\hat{\rho} = \frac{1}{N-m} \sum_{t=m+1}^N (y_t - \hat{\Phi}^T z_{t-1})^2$$

[0055] 次に、判別式としてつぎの式を考える。

[0056]

[数3]

$$I^* = \operatorname{Arg} \max_i \ln p(C_i | Z_N)$$

上式は観測値系列 Z_N が得られたときのクラス C_i の事後確率が最大になるクラスを判別クラスとして採用する。

$$I^* = \operatorname{Arg} \max_i \ln p(Z_N | C_i) P_r(C_i)$$

ここで、システムを決定するパラメータ θ は各クラスにおいて一定であり、
 $p(\theta | C_i) = \delta(\theta - \theta_i)$ と仮定すると、 $p(Z_N | C_i)$ は、次のようになる。

$$\begin{aligned} p(Z_N | C_i) &= \int p(Z_N, \theta | C_i) d\theta = \int p(Z_N | \theta, C_i) p(\theta | C_i) d\theta \\ &= \int p(Z_N | \theta, C_i) \delta(\theta - \theta_i) d\theta = p(Z_N | \theta_i, C_i) = p(Z_N | \theta_i) \end{aligned}$$

従って、判別式は次のようになる。

$$I^* = \operatorname{Arg} \max_i \ln p(Z_N | \theta_i) P_r(C_i)$$

[0057] また、パラメータ推定に関しては、次式で定義される尤度を最大にする最尤推定法を用いることとする。

[0058] [数4]

$$L = \sum_{j=1}^{N_p} \ln p(Z_N^j | \theta_i)$$

ただし、 Z_N^j は j 番目の観測値系列であり、 N_p はトレーニングデータの個数である。

以上の結果より、判別式、パラメータ推定式は、次のようになる。

[0059] [判別式]

[0060] [数5]

$$I^* = \operatorname{Arg} \max_i \left(-\frac{N-m}{2} \ln \hat{\rho}_i + \ln p_r(C_i) - \frac{1}{2\hat{\rho}_i} \sum_{t=m+1}^N (y_t - \hat{\Phi}_i^T z_{t-1})^2 \right)$$

[パラメータ推定式]

$$\hat{\Phi}_i = \left(\sum_{j=1}^{N_p} \sum_{t=m+1}^N z_{t-1}^j z_{t-1}^{jT} \right)^{-1} \left(\sum_{j=1}^{N_p} \sum_{t=m+1}^N z_{t-1}^j y_t^j \right)$$

$$\hat{\rho}_i = \frac{1}{N_p(N-m)} \sum_{j=1}^{N_p} \sum_{t=m+1}^N (y_t^j - \hat{\Phi}_i^T z_{t-1}^j)^2$$

- [0061] 上式に基づいて、分類が可能になる。特に、上記はマルチクラスへの分類の例であるが、クラス数Kを2とすれば、2分類問題、すなわち異常検知の問題となる。そして、上記観測値にWavelet解析の結果を適用すれば、時間に局在した異常をも対象とした異常検知を行うことができる。このように、異常データが少なからずあり、教示できる場合には、統計的な意味で、より精度の高い異常検知を行うことができる。
- [0062] 図18に別の実施例を示す。散布図・相関解析部15を付加した例である。r次元の多次元時系列信号を対象に、散布図・相関解析を行った例を図19に示す。図19の相互相関解析では、遅れのラグ(lag)を考慮している。通常、相互相関関数の最大値の位置をラグと呼ぶ。この定義に従えば、二つの現象に関する時間のずれは、相互相関関数のラグに等しい。
- [0063] ラグの正負は、二つの現象のどちらが早く起こるかにより決まる。このような散布図解析や相互相関解析の結果は、時系列信号間の相関を表すものであるが、各クラスタの特徴付けにも活用でき、クラスタ間の類似性の判断指標となり得るものである。例えば、ラグの量の一致度により、クラスタ間の類似性を判断する。これにより、図11に示した類似したクラスタのマージ等が可能となる。マージ後のデータを用いてモデル化する。なお、マージの方法は、他の手法であってもよい。
- [0064] 図20に別の実施例を示す。時間・周波数解析部16を付加した例である。r次元の多次元時系列信号を対象に、時間・周波数解析を行った例を図21に示す。時間・周波数解析、或いは散布図・相関解析を行い、これらの信号をM次元の多次元時系列信号に加える、またはM次元の多次元時系列信号と置換することも可能である。
- [0065] 図22に別の実施例を示す。学習データのDB17とモデル化(1)18を付加した例である。図23に、その詳細を示す。モデル化(1)により、学習データを対象にそれを複数のモデルとしてモデル化し、観測データとの類似性を判断して該当モデルを当てはめ、観測データとの偏差を算出する。モデル化(2)は、図1と同様の部分であり、これから、観測データから得

られるモデルとの偏差を算出する。

[0066] そして、モデル化（1）（2）のそれぞれの偏差から、状態変化を算出し、総合偏差を算出する。ここで、モデル化（1）（2）を均等に扱うことも可能であるが、重み付けをしてもよい。即ち、学習データを基本と考えれば、モデル（1）の重みを大きくし、観測データを基本と考えれば、モデル（2）の重みを大きくする。

[0067] 図12に示した表現に従えば、モデル（1）で構成した部分空間モデルをクラスタ間で比較すれば、そしてそれらが本来同一状態のクラスタならば、その状態変化を知ることができる。そして、観測データの部分空間モデルがそれから移動していれば、状態変化を読み取ることができる。状態変化が、部品交換など意図するものであれば、即ち、設計側が知っており、それによる変化を許容すべきであれば、モデル（1）の重みを小さくし、モデル（2）の重みを大きくする。状態変化が、意図しないものであれば、モデル（1）の重みを大きくする。例えば、パラメータ α を、モデル（1）の重みとして使えば、

$$\alpha \times \text{モデル（1）} + (1 - \alpha) \times \text{モデル（2）}$$

として定式化できる。

[0068] モデル（1）の重みを、古いものほど小さくするという忘却形のものでもよい。この場合、最近のデータに基づくモデルが重視されることになる。

[0069] 図23において、物理モデル22は、シミュレーションにより対象エンジンなどを模擬したモデルである。対象知識が十分にある場合は、対象エンジンなどを離散時間（非）線形状態空間モデル（状態方程式などで表現）で表現できるので、その中間値や出力などを推定可能となる。従って、この物理モデルに従えば、このモデルからの偏差に基づいて、異常検知を行うことも可能となる。

[0070] もちろん、物理モデルに従って、学習データのモデル（1）を修正することも可能である。或いは、逆に、学習データのモデル（1）に従って、物理モデルを修正することも可能である。物理モデルの変形として、過去の実績

としての知見を、物理モデルとして組み込むことも可能である。アラームの発生や部品交換に伴うデータの遷移を、物理モデルに組み込むことも可能である。或いは、学習データ（個々のデータや重心位置など）を、アラームの発生や部品交換に伴い、移動させてもよい。

[0071] なお、図23に対して、図1から図22までに示したように、物理モデルに対して、統計モデルを主に使うのは、データを生み出すプロセスについての理解が少ない場合に、統計モデルが有効であることによる。距離や類似性は、データの生成過程が不明瞭でも、定義できる。画像が対象である場合も、画像生成過程が不明瞭なとき、統計モデルは有効である。対象に関する知識が、すこしでも利用できる場合には、物理モデル22が活用できることになる。

[0072] 図26に、異常検知、及び異常検知後の診断を示す。同図において、設備からの時系列信号から、今までに説明した方法により、異常を検知する。同時に、保守のイベント（アラームや作業実績など）などの付帯情報を取り込み、異常を高感度に検知する。同図に示すように、早期に予兆として発見できれば、故障となって稼動停止となる前に、何らかの対策がうてることになる。そこで、この予兆に基づき、異常診断を行う。

[0073] 異常診断は、予兆を内包しているセンサを特定する現象診断と、故障を引き起こす可能性のあるパーツを特定する原因診断に分けると考えやすい。異常検知部では、異常診断部に対して、異常の有無という信号のほか、特徴量に関する情報を出力する。異常診断部は、これらの情報をもとに診断を行う。

[0074] 図27に、ハードウェア構成を示す。異常検知を実行するプロセッサに、対象とするエンジンなどのセンサデータを入力し、欠損値の修復などを行って、データベースDBに格納する。プロセッサは、観測データ、DBデータを用いて、異常検知を行い、異常検知信号を出力する。表示部では、実施例において述べたような各種表示を行う。トレンドを表示することも可能とする。

- [0075] データベースDBは、熟練エンジニアがDBを操作できる。特に、異常事例や対策事例を教示でき、格納できる。（1）学習データ（正常）、（2）異常データ、（3）対策内容が、格納される。データベースDBを、熟練エンジニアが手を加えられる構造にすることにより、洗練された、有用なデータベースができあがることになる。データ操作は、学習データ（個々のデータや重心位置など）を、アラームの発生や部品交換に伴い、移動させることにより行う。また、取得データを追加することも可能である。異常データがあれば、データの移動に、一般化ベクトル量子化などの手法も適用できる。
- [0076] 上記実施例では、エンジンなどの設備を対象に説明したが、時系列信号のたぐいならば、対象にこだわらない。人体の測定データへも適用可能である。本実施例に従えば、状態の数、遷移の回数が多くても、対応可能である。
- [0077] また、実施例にて説明した各機能、例えばクラスタリング、主成分分析や、Wavelet解析などは、必ず実施するというものでもなく、対象の信号の性質に応じて、適宜実施すればよい。
- [0078] クラスタリングも、時間軌跡のみならず、混合分布に対するEM（Expectation-Maximization）アルゴリズムやk-meansクラスタリングなどを含め、データマイニング分野における手法が使えることは言うまでもない。得られたクラスタは、これを対象に識別器を適用してもよいが、クラスタをグルーピングして、これを対象に識別器を適用してもよい。最も簡単な例は、日々の観測データが属するクラスタと、それが属するクラスタ以外に分けることである（図10の右側の特徴空間にて図示した着目データなる現データと、それより時間的に前の過去データに該当）。また、センサ信号（特微量）の選択は、ラッパー法（例えば、backward stepwise selectionにより、全ての特微量がある状態から、最も不要な特徴を一つずつ取り除く）など、既存の手法が適用可能である。
- [0079] さらに、識別器は、いくつかの識別器を準備し、それらの多数決をとも可能である。即ち、異なる識別器群を用いたアンサンブル（集団）学習が適用できる。図30にその構成例を示す。例えば、識別器1は投影距離法

、識別器2は局所部分空間法、識別器3は線形回帰法と言ったものである。

事例データに基づくものならば、任意の識別器が適用可能である。

[0080] 複数の識別器を使う理由は、識別器はそれぞれ異なる基準、異なる対象データ範囲（セグメント分けやその統合に依存）で、はずれ具合いを求めるため、その結果に微妙な違いが生まれるためである。このため、多数決をとつて安定化するか、OR（はずれ値そのもの、すなわち多値の場合は最大値検出）論理で、いずれかの識別器で異常が検知されれば異常が発生したとして出力し、異常をのこらず検知しようとするか、AND（多値の場合は最小値検出）論理で、いずれの識別器でも同時に異常が検知されれば異常が発生したとして出力し、誤検知を最小限にするといった、上位基準で識別器を構成する。もちろん、アラーム信号、部品交換などの保全情報などの情報を加味して、上記統合を行うことも可能であることは言うまでもない。

[0081] 識別器 h_1, h_2, \dots をすべて同じ識別器とし、対象データ範囲（セグメント分けやその統合に依存）を変えて学習することも可能である。例えば、パターン認識の代表的手法であるバギングやブースティングなどの手法も適用できる。この手法の適用により、異常検知に関して、より高い正解率が確保できる。ここで、バギングは、N個のデータから重複を許してデータをK個取ってきて（復元抽出）、このK個を元に1個目の識別器 h_1 を作り、またN個のデータから重複を許してデータをK個取ってきて、このK個（1個目の識別器と中身が違う）を元に2個目の学習器 h_2 を作ることを続け、異なるデータからいくつかの識別器を作り、実際に判別器として使う時は多数決を取る方法である。

[0082] ブースティング（Adaboostと呼ぶ手法）は、N個のデータに均等な重み $1/N$ をまずは割り当て、1個目の識別器 h_1 はN個全てのデータを用いて学習し、学習した後、データN個に関して正解率を調べ、それを元に信頼度 $\beta_1 (> 0)$ を求める。1個目の識別器が正解したデータの重みを $\exp(-\beta_1)$ 倍して重みを減らし、正解できなかったデータの重みを $\exp(\beta_1)$ 倍して重みを増やす。

[0083] 2個目の識別器 h_2 はN個全てのデータを用いて重み付きの学習を行い、信頼

度 $\beta_2(>0)$ を求める、データの重みを更新する。2個とも正解したデータの重みは軽く、2個とも間違えたデータの重みは重くなる。以降、これを繰返してM個の識別器を作り、実際に判別器として使う時は信頼度付き多数決を取るものである。これらの手法を、クラスタ群を対象に適用することにより、性能向上が期待できる。

- [0084] 図31に、図30に示した識別器を含んだ、異常検知全体の構成例の一例を示す。軌跡クラスタリング、特徴選択などを経て、アンサンブル学習を行い、高い識別率を達成するものである。識別器出力の統合の方法は上述した通りであるが、どのクラスタに、どの識別器を適用するかという組み合わせは、いくつか存在する。例えば、観測データとは異なるクラスタに対して局所部分空間法を適用して、異なるクラスタからのはずれ具合いを把握し（推定値も算出する）、観測データと同じクラスタに対しては回帰分析法を適用して、自クラスタからのはずれ具合いを把握する。
- [0085] そして、それらの識別器出力を統合して、異常判定を行うことができる。他のクラスタからのはずれ具合いを、投影距離法や回帰分析法により行うことも可能である。自クラスタからのはずれ具合いを投影距離法により行うこともできる。クラスタは、アラーム信号が活用できる場合には、アラーム信号の重症度レベルに応じて、重症アラーム信号が付加されていないクラスタを対象にすることもできる。
- [0086] クラスタ間の類似性を判断し、類似クラスタを統合して、これを対象にすることもできる。識別器出力の統合は、はずれ値の加算、最大・最小、OR/ANDなどのスカラー変換処理でもよいし、識別器の出力をベクトル的に、多次元として扱うこともできる。もちろん、識別器出力のスケールは、極力一致させることとする。
- [0087] 上述したクラスタとの関連の持たせ方に關して、さらに、他クラスタを対象にして第1報の異常検知を行い、自クラスタのデータが収集された時点で自クラスタを対象にして第2報の異常検知を行ってもよい。このようにして、顧客への注意喚起を促すことが可能になる。このように、本実施例は、対

象クラスタ群との関係において、信号の振舞い、behaviorに、より着目した実施例と言える。

[0088] 上述したいくつかの実施例に関する総合的効果をさらに補足する。えば、発電設備を所有している会社では、機器の保守費用削減を希望しており、保証期間中に機器を点検、部品交換を実施している。これは時間ベースの設備保全と言われている。しかし、最近は機器の状態を見て、部品交換を実施する状態ベースの保全に移行しつつある。状態保全を実施するには、機器の正常・異常データを収集する必要があり、このデータの量、質が状態保全の品質を決めてしまう。しかし、異常データの収集は、まれなケースも多く、大型の設備になるほど、異常データを収集することは困難である。従って、正常データから、はずれ値を検出することが重要となる。上述したいくつかの実施例によれば、

正常データから、異常を検知できる、

データ収集が不完全でも精度の高い異常検知が可能となる、

異常データが含まれっていても、この影響を許容できる、

といった直接的効果に加え、

ユーザにとって、現象を理解しやすい、

エンジニアの知識を活用できる

物理モデルも併用できる、

と言った副次的な効果がある。

産業上の利用可能性

[0089] プラント、設備の異常検知として利用することが出来る。

符号の説明

[0090] 1・・・多次元信号取得部、

2・・・欠損値の修正・削除部、

3・・・状態データ・知識データベース、

4・・・相関解析による無効信号の削除部、

6・・・軌跡分割クラスタリング、

- 7 . . . アラーム信号／保全情報、
- 8 . . . 各クラスタを対象にモデル化部、
- 9 . . . モデルからの偏差算出部、
- 10 . . . はづれ値検出部、
- 11 . . . 各クラスタを特徴選択でモデル化部、
- 12 . . . アラーム信号などの一定区間累積ヒストグラム、
- 13 . . . 異常特定部、
- 14 . . . Wavelet（変換）解析部、
- 15 . . . 各クラスタ軌跡散布図・相関解析部、
- 16 . . . 各クラスタ毎時間・周波数解析部、
- 17 . . . 学習データ、
- 18 . . . モデル化（1）部、
- 19 . . . プロセッサ、
- 20 . . . 表示器、
- 21 . . . データベース
- 22 . . . 物理モデル。

請求の範囲

- [請求項1] プラントまたは設備の異常を早期に検知する異常検知方法であって、
複数のセンサからデータを取得し、前記データの時間的な変化に基づいて、データ空間の軌跡を複数のクラスタに分割し、
前記分割した前記クラスタ群に対して、部分空間法でモデル化し、
はずれ値を異常候補として算出することを特徴とする異常検知方法。
- [請求項2] 前記モデル化は、前記複数のセンサから過去に得られたデータからなる学習データを対象にモデル化したリファレンスモデルを併用して、はずれ値を異常候補として算出することを特徴とする請求項1に記載の異常検知方法。
- [請求項3] 前記モデル化は、データのN個抜き ($N=0, 1, 2, \dots$) の回帰分析法、または、部分空間法を用いて行うことを特徴とする請求項1記載の異常検知方法。
- [請求項4] 前記部分空間法は、投影距離法、CLAFIC法、または、観測データの近傍を対象とする局所部分空間法であることを特徴とする請求項3記載の異常検知方法。
- [請求項5] 前記はずれ値の検出は、前記回帰分析法、前記投影距離法、前記CLAFIC法、または、観測データの近傍を対象とする前記局所部分空間法の出力を統合して行うものであることを特徴とする請求項4記載の異常検知方法。
- [請求項6] プラントまたは設備の異常を早期に検知する方法において、
複数のセンサからデータを取得し、
時間的なデータの時間的変化に基づいて、データ空間の軌跡を複数のクラスタに分割し、
分割したクラスタが分かるように、データを表示することにより、
状態を可視化することを特徴とする異常検知方法。
- [請求項7] 前記複数のセンサからの過去のデータに基づくクラスタ群に対する

異常検知により、第1報を出力し、前記複数のセンサの現在のデータに基づくクラスタ群に対する異常検知により第2報を出力することを特徴とする請求項1記載の異常検知方法。

[請求項8] プラントまたは設備の異常を早期に検知する異常検知システムにおいて、

複数のセンサからデータを取得するデータ取得部と、
前記データの時間的な変化に基づいて、データ空間の軌跡を複数のクラスタ群に分割する軌跡分割部と、

分割した前記クラスタ群に対して、部分空間でモデル化するモデル化部と、

はずれ値を異常候補として算出する異常検出部を有し、
前記軌跡分割部においては前記分割したクラスタ群に対して、部分空間でモデル化し、はずれ値を異常候補として算出することを特徴とする異常検知システム

[請求項9] 前記データ取得部において取得した前記複数のセンサからの過去のデータを蓄積し、かつ、モデル化してリファレンスマルチモデルを作成する学習データベース部を有し、前記学習データベース部における前記リファレンスマルチモデルを併用して、はずれ値を異常候補として算出することを特徴とする請求項8記載の異常検知システム。

[請求項10] 前記モデル化は、データのN個抜き ($N=0, 1, 2, \dots$) の回帰分析法、または、部分空間法を用いて行うことを特徴とする請求項8記載の異常検知システム。

[請求項11] 前記部分空間法は、投影距離法、CLAFIS法、または、観測データの近傍を対象とする局所部分空間法であることを特徴とする請求項8記載の異常検知システム。

[請求項12] 前記はずれ値の検出は、前記回帰分析法、前記投影距離法、前記CLAFIS法、または、観測データの近傍を対象とする前記局所部分空間法の出力を統合して行うものであることを特徴とする請求項8記

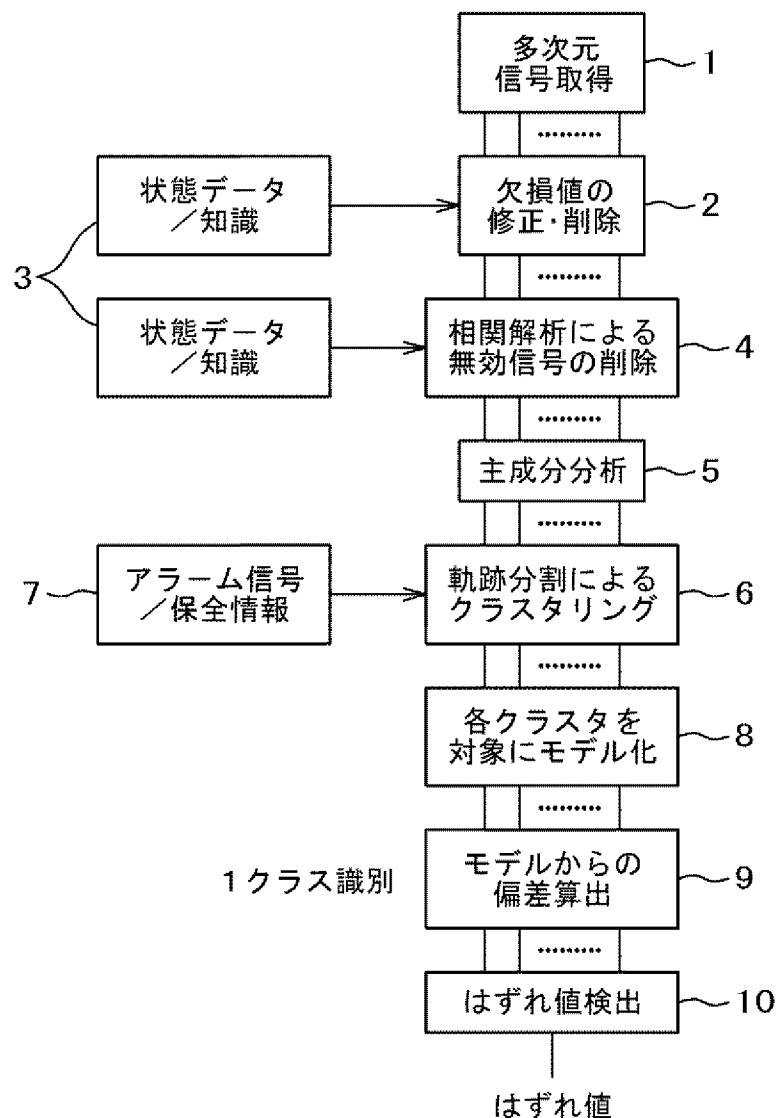
載の異常検知システム。

- [請求項13] プラントまたは設備の異常を早期に検知する異常システムにおいて、
複数のセンサからデータを取得するデータ取得部と、
前記データの時間的な変化に基づいて、データ空間の軌跡を複数の
クラスタに分割する軌跡分割部と、
前記分割したクラスタ群に対して、部分空間でモデル化するモデル
化部と、
はづれ値を異常候補として算出する異常検出部と、
データ表示部とを有し、
前記分割した複数のクラスタが分かるように、データを表示するこ
とにより、
状態を可視化することを特徴とする異常検知システム。

- [請求項14] 前記複数のセンサからの過去のデータに基づくクラスタ群に対する
異常検知により、第1報を出し、前記複数のセンサの現在のデータ
に基づくクラスタ群に対する異常検知により第2法を出力することを
特徴とする請求項13記載の異常検知システム。

[図1]

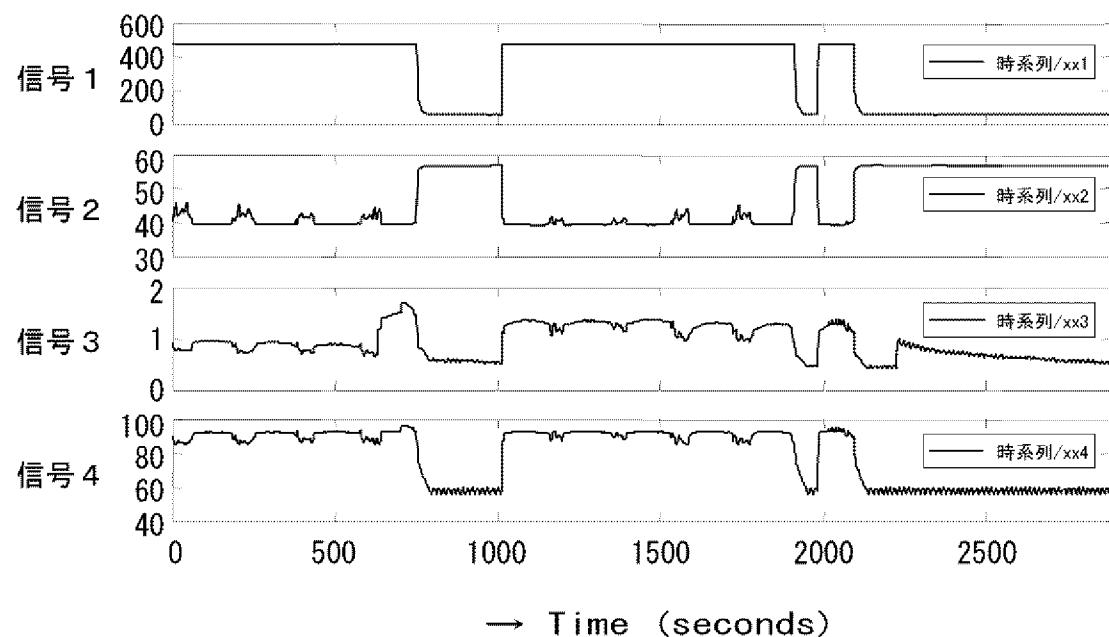
図 1



[図2]

図 2

多次元時系列信号



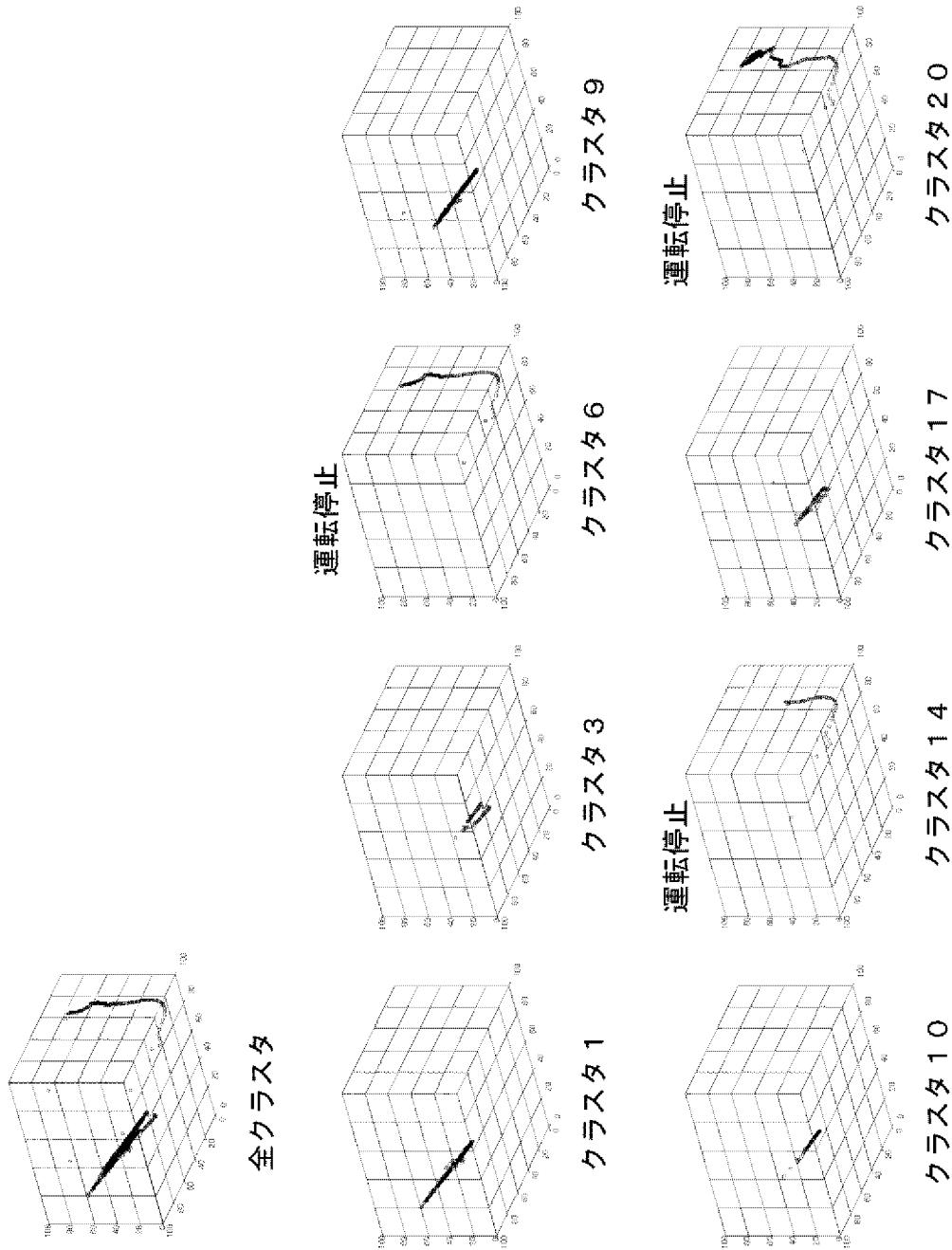
[図3]

図 3

相関行列

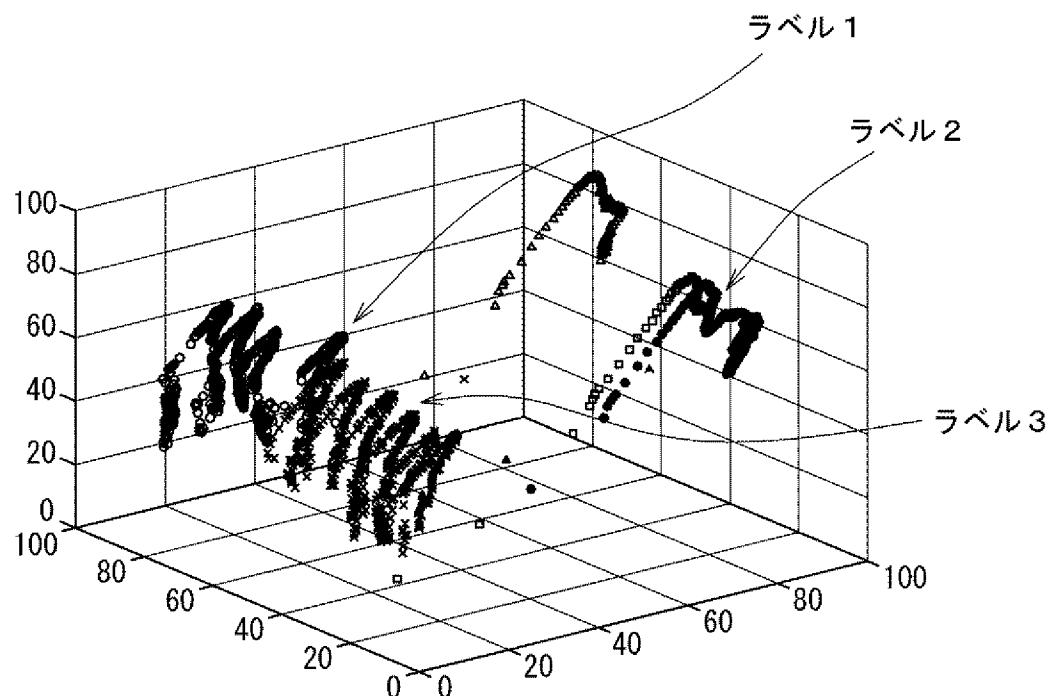
	信号 1	信号 2	信号 3	信号 4
信号 1	1.000	-0.991	0.781	0.990
信号 2	-0.991	1.000	-0.814	-0.992
信号 3	0.781	-0.814	1.000	0.816
信号 4	0.990	-0.992	0.816	1.000

[図4]



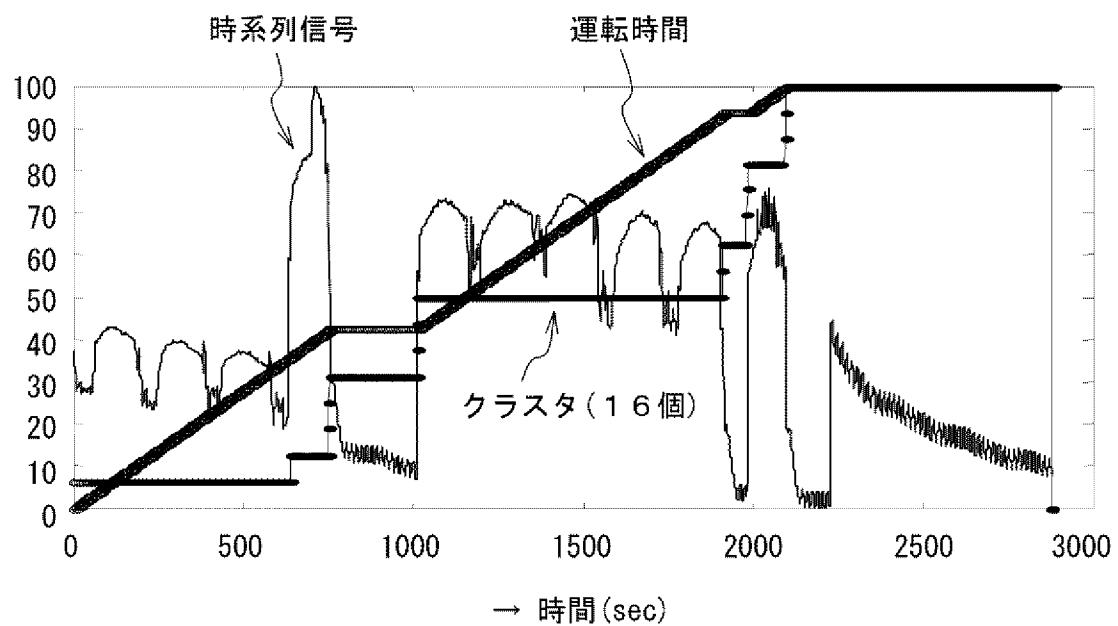
[図5]

図 5



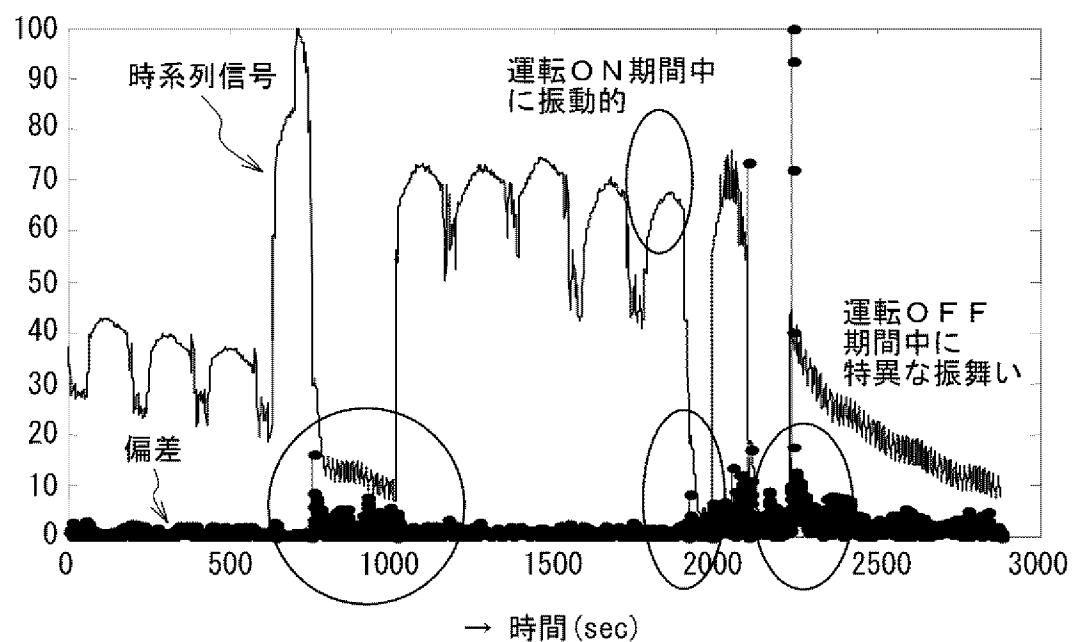
[図6]

図 6



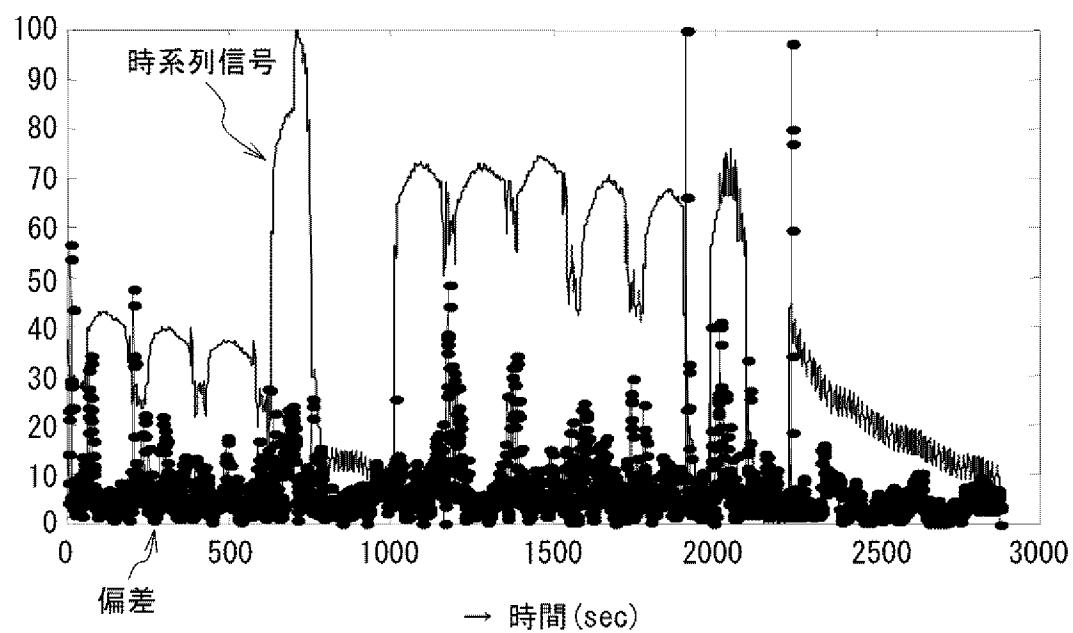
[図7]

図 7



[図8]

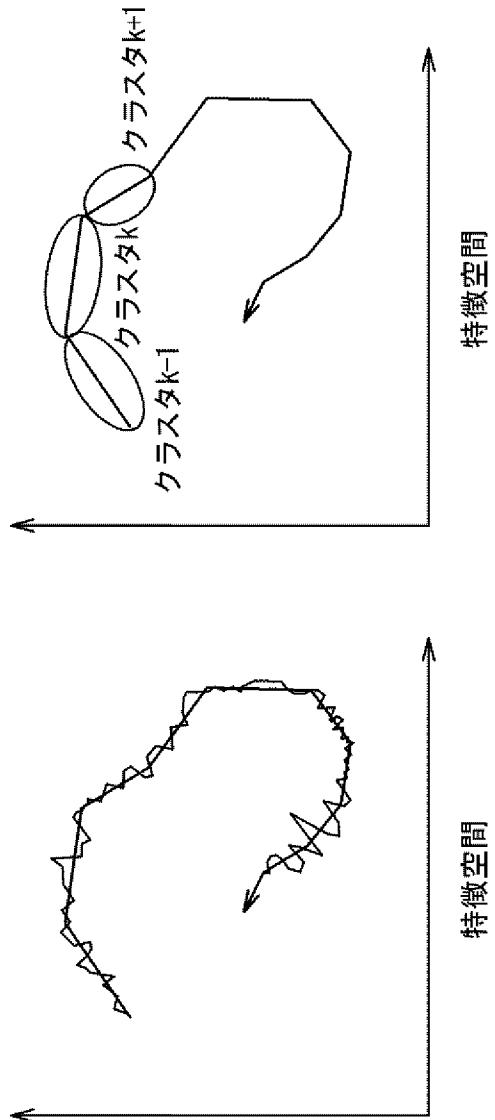
図 8



[図9]

データを単純なクラスタ毎にモデル化
データの振舞い(軌跡)を可視化

図9



[図10]

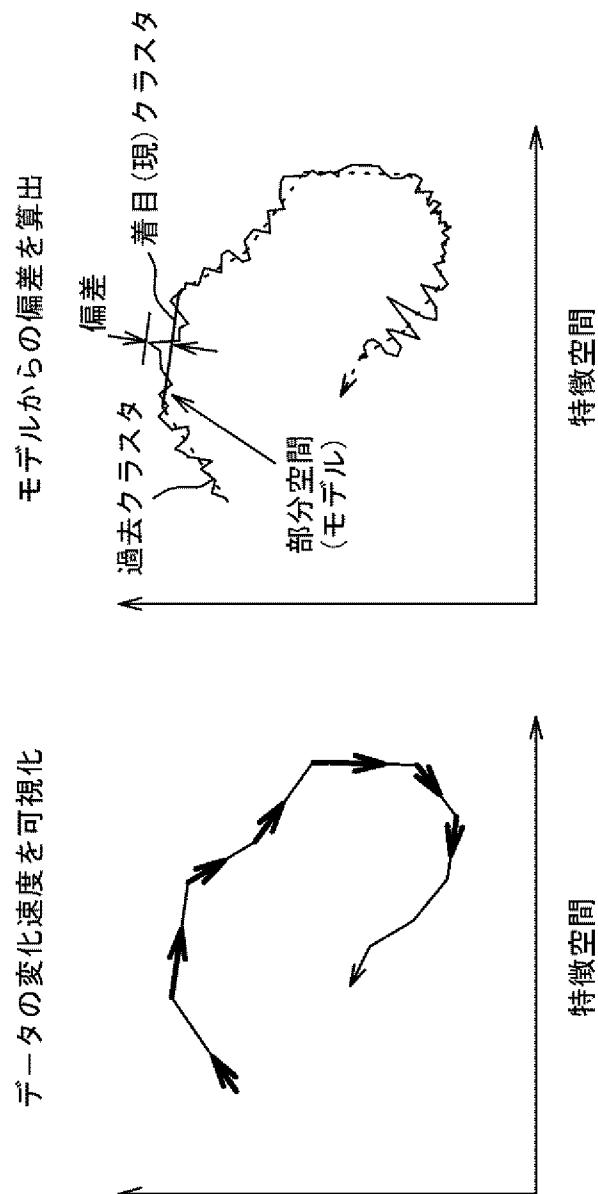


図10

[図11]

個別状態を対象
(コンパクトな空間で表現し、
見逃し防止)

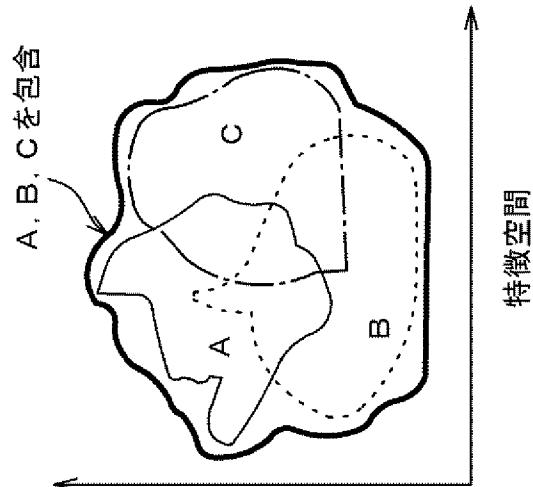
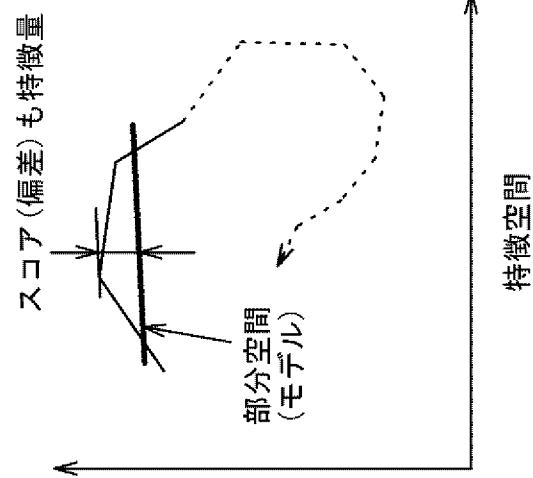


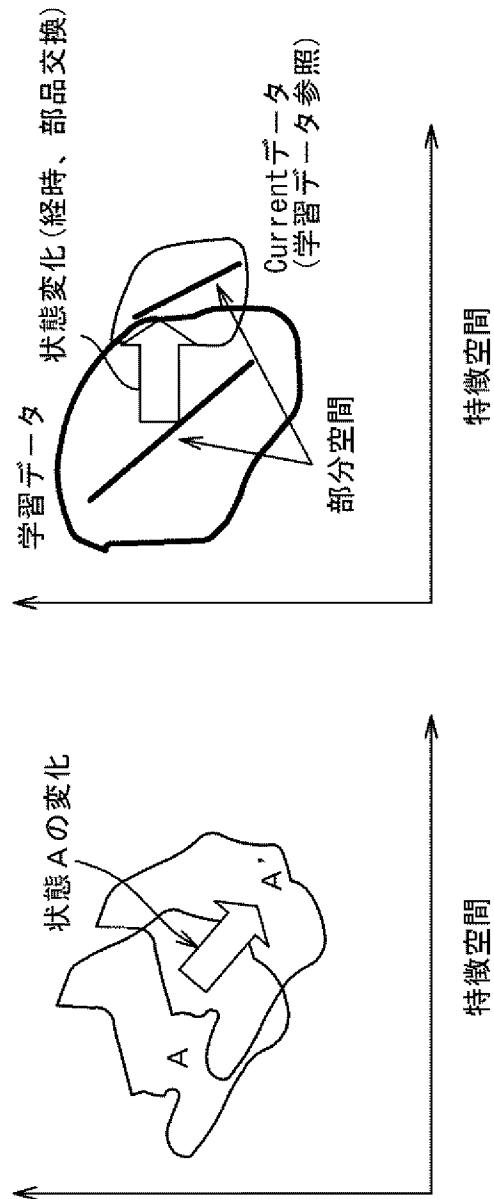
図11

クラスタのマージ
モデルからの偏差を算出



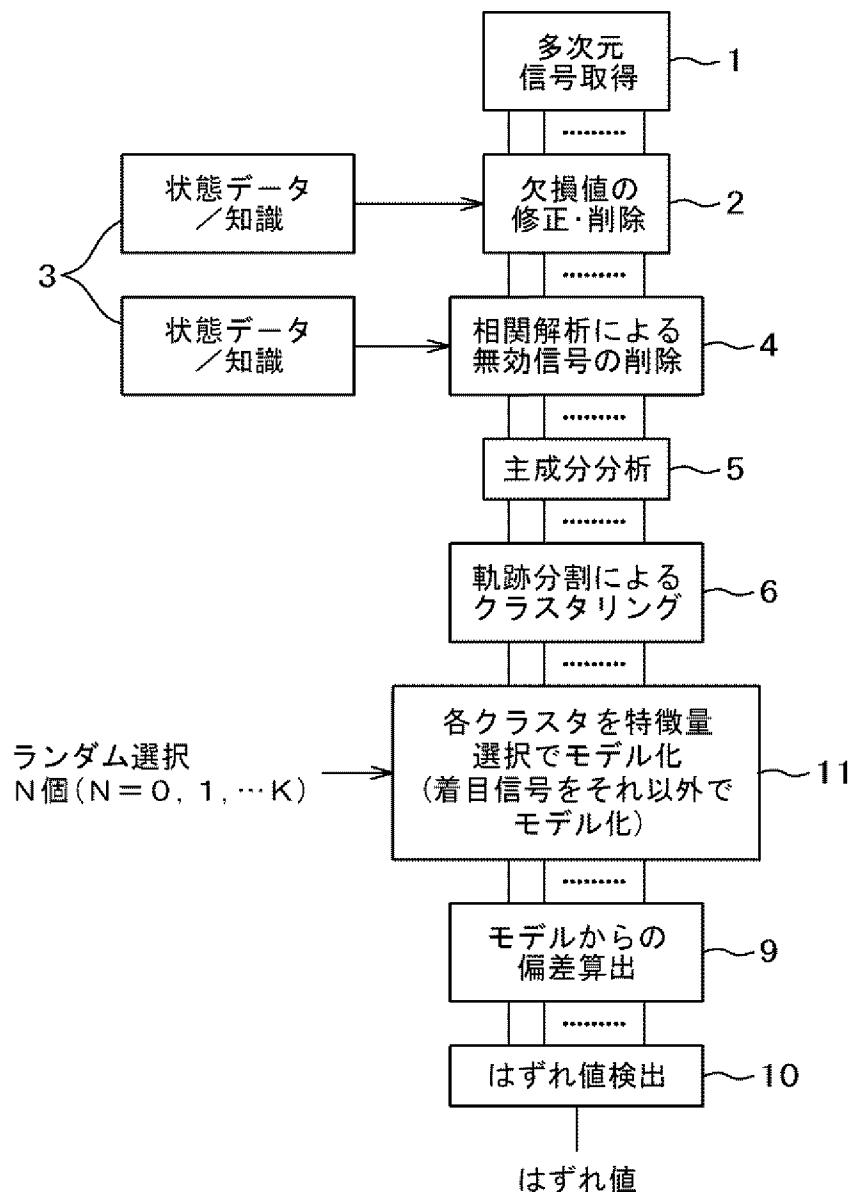
[図12]

同じ状態でも、時間で分割
状態を学習、変化をモデル化



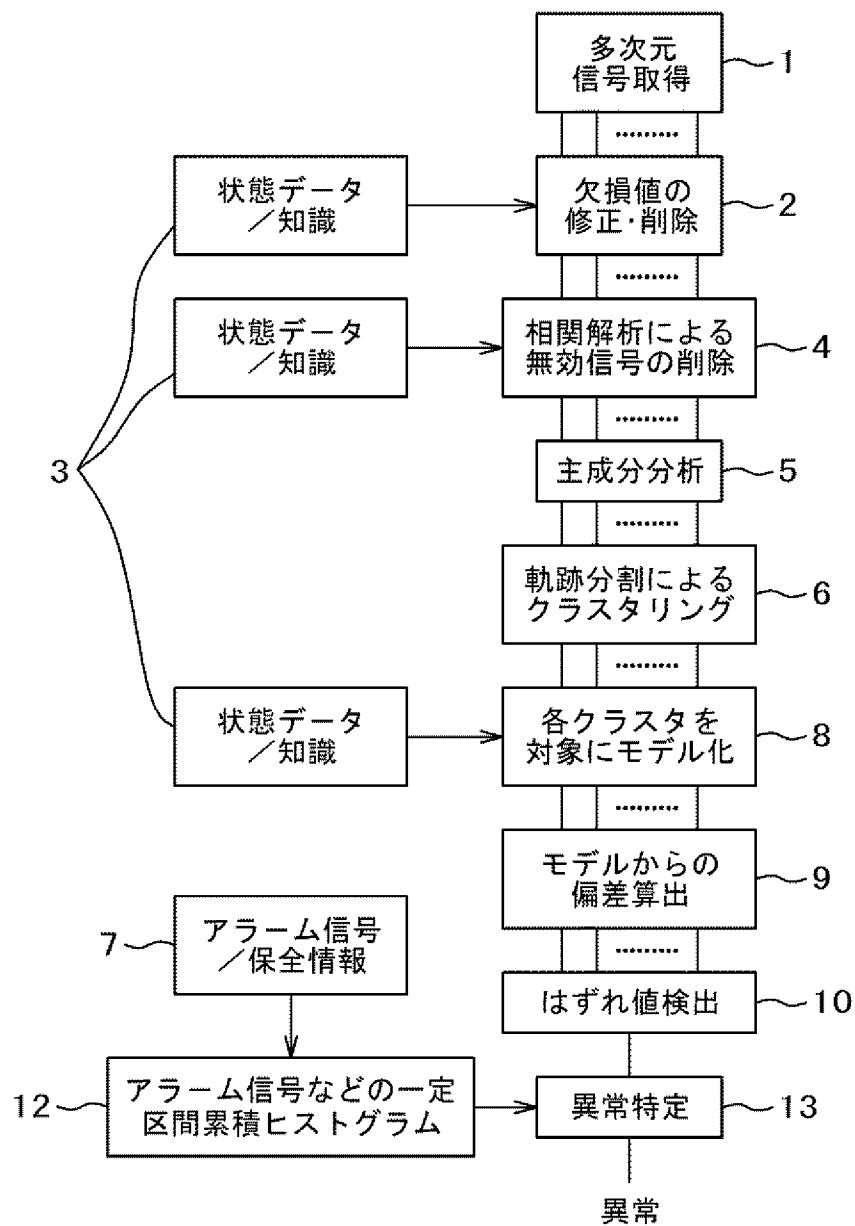
[図13]

図 13



[図14]

図 1 4

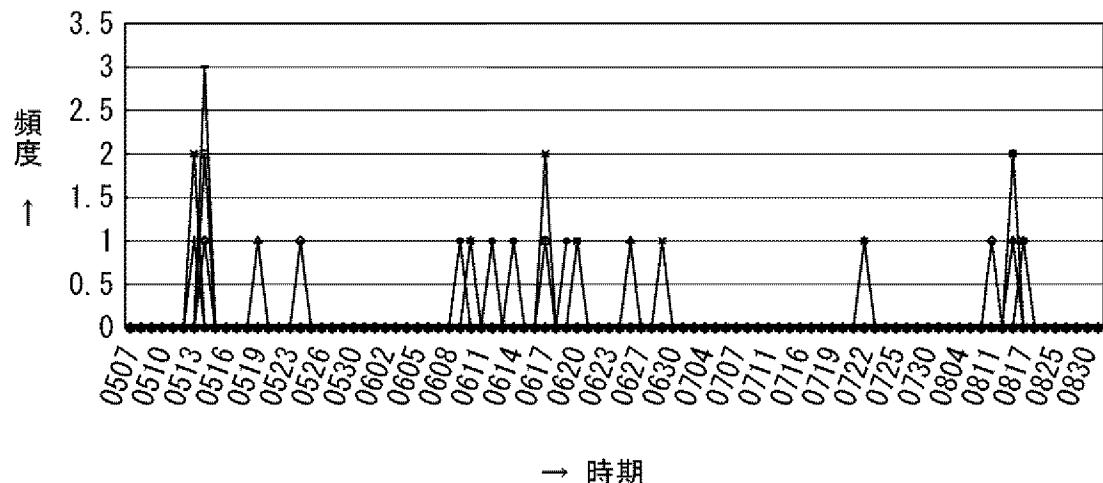


[図15]

図 15

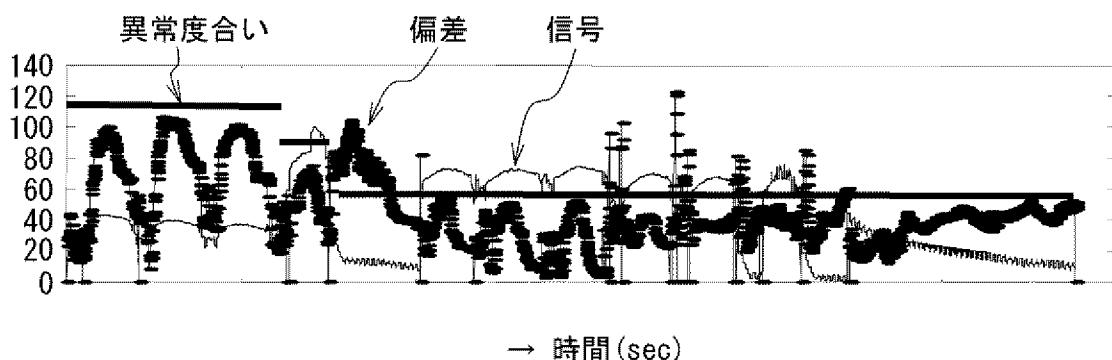
(a)

アラーム信号ヒストグラム



(b)

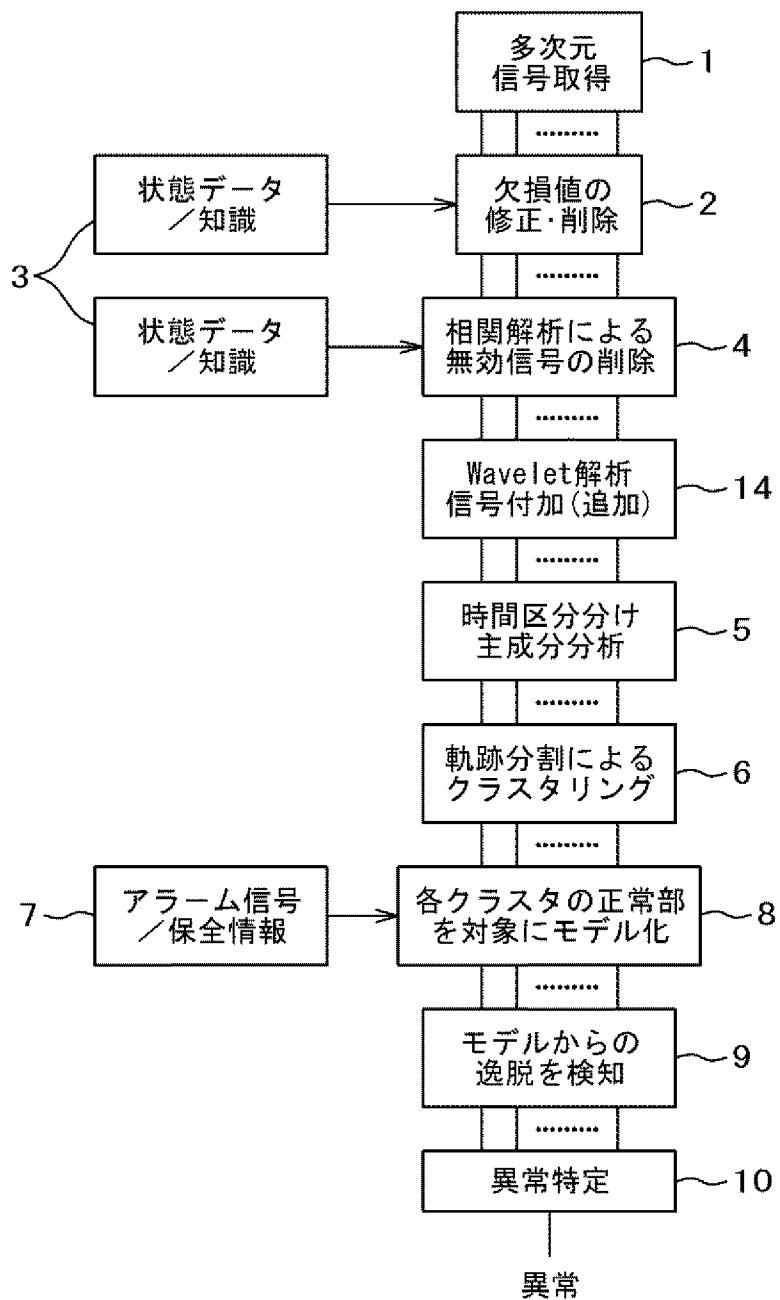
時系列信号 3、偏差、異常度合い



アラーム信号とはずれ値を組み合わせて、
異常の度合いや信頼度を付与

[図16]

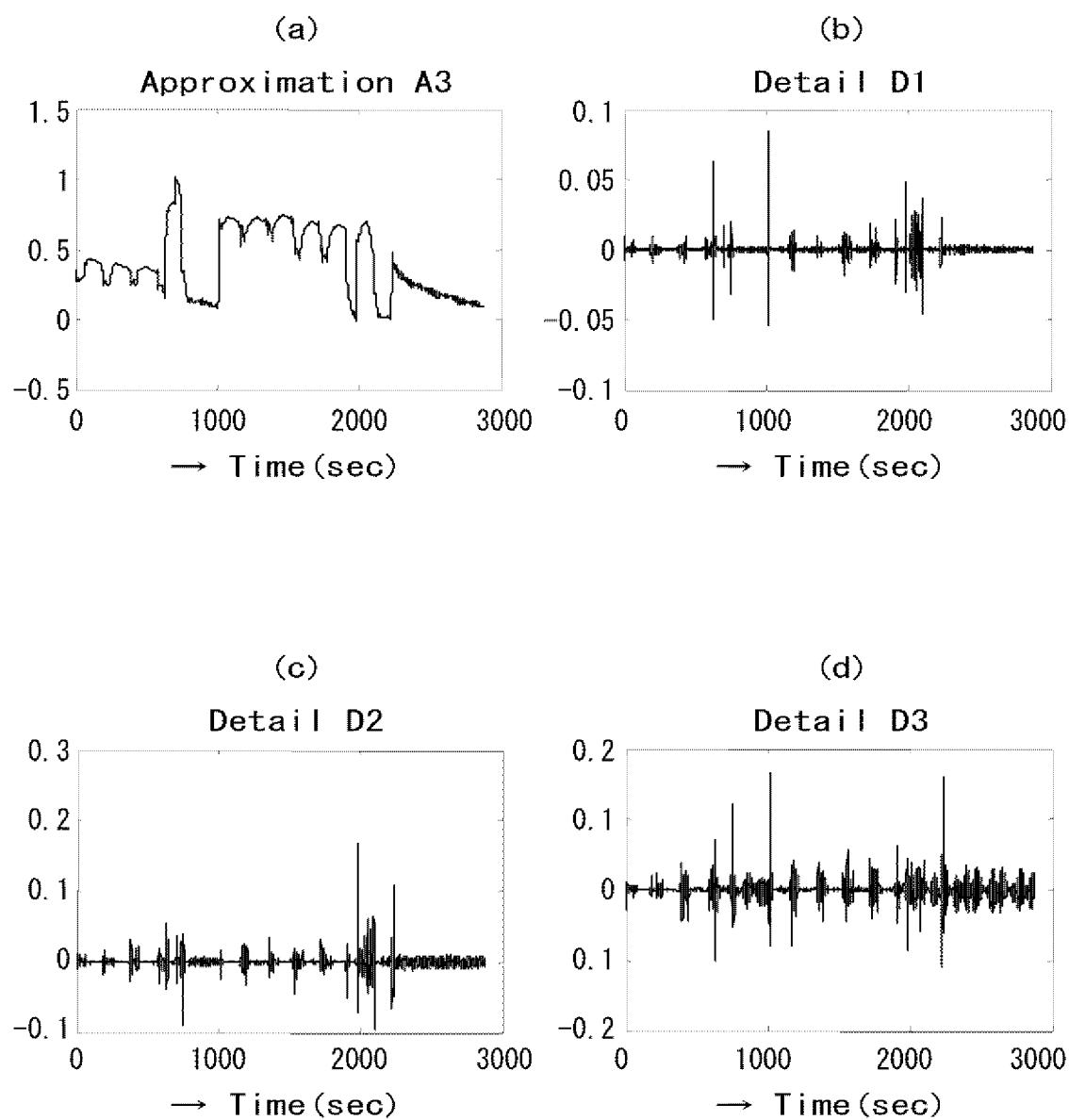
図 16



[図17]

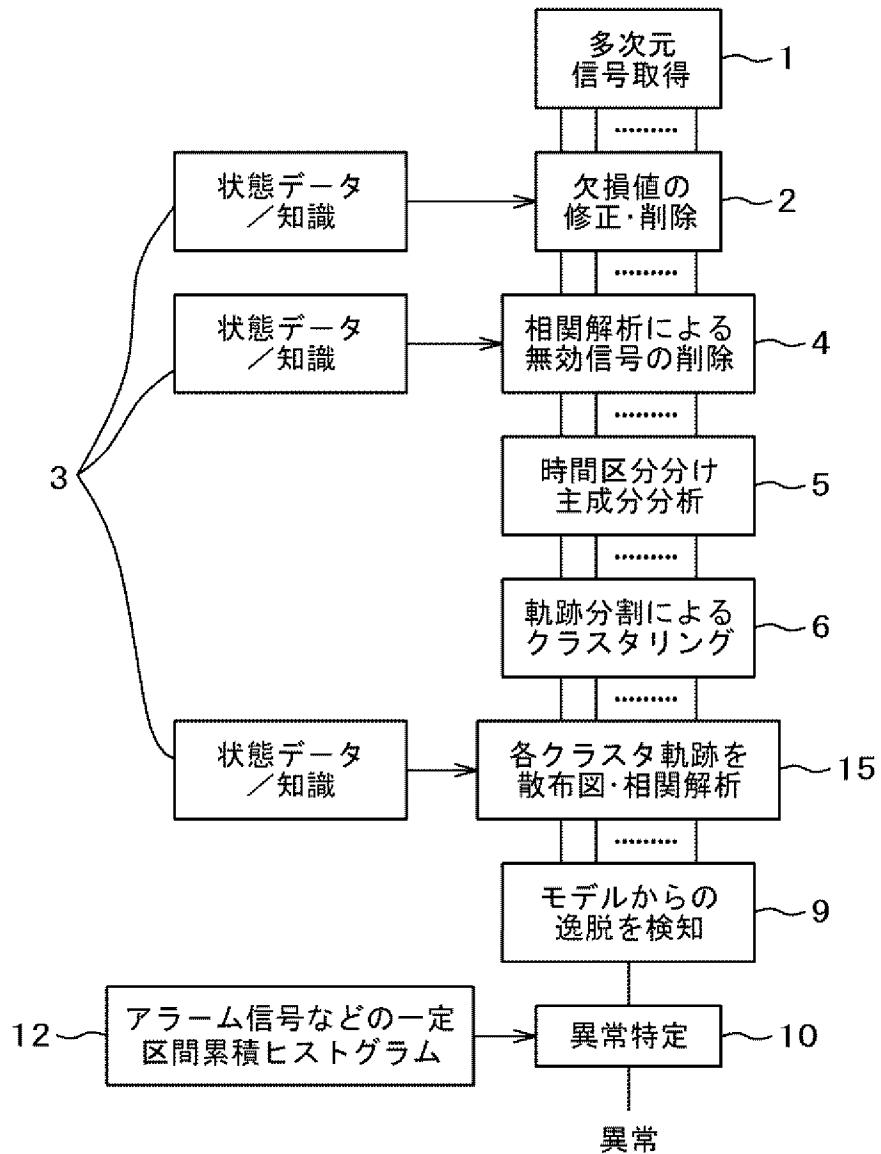
図 17

Wavelet解析



[図18]

図 18



[図19]

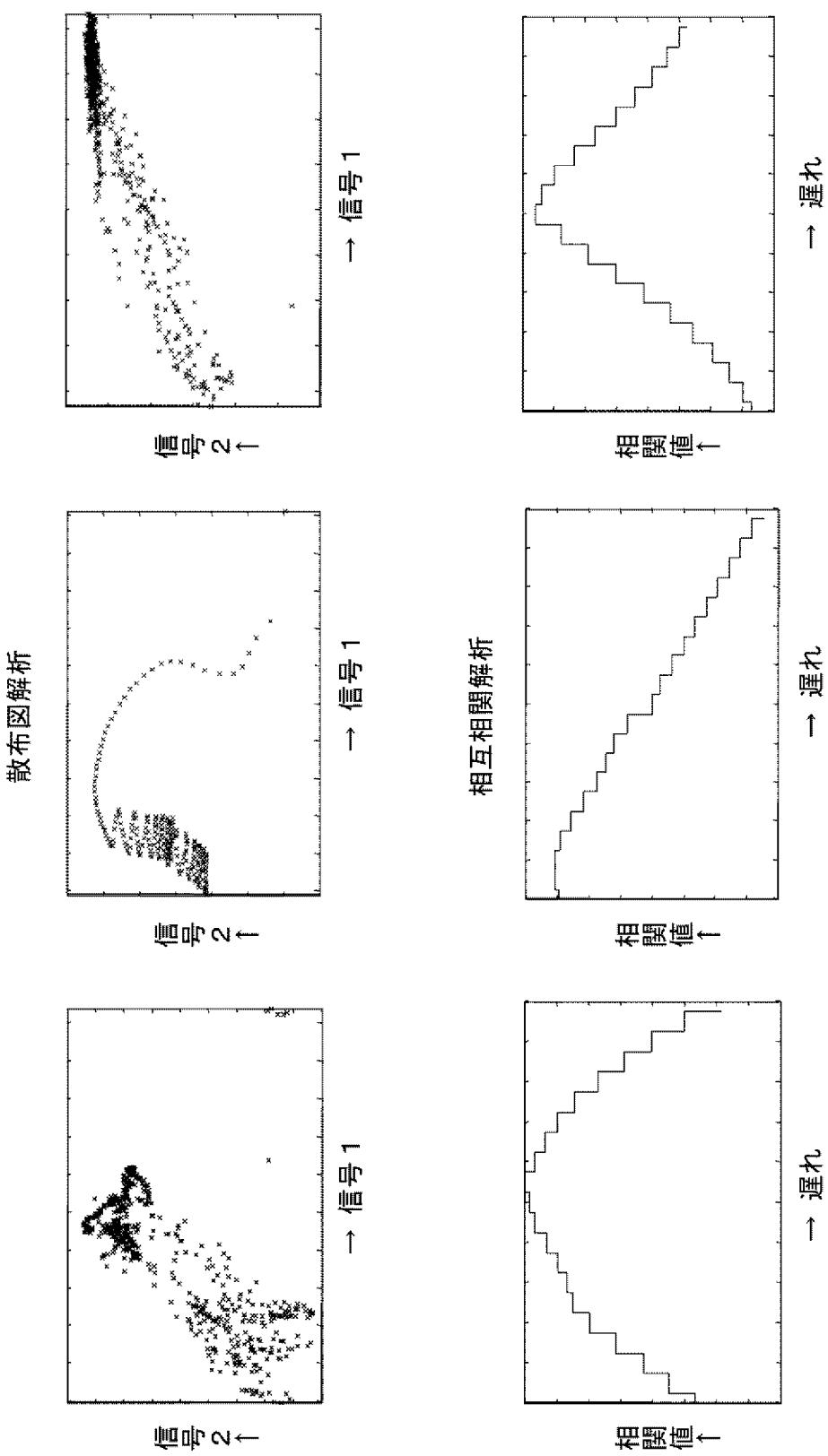
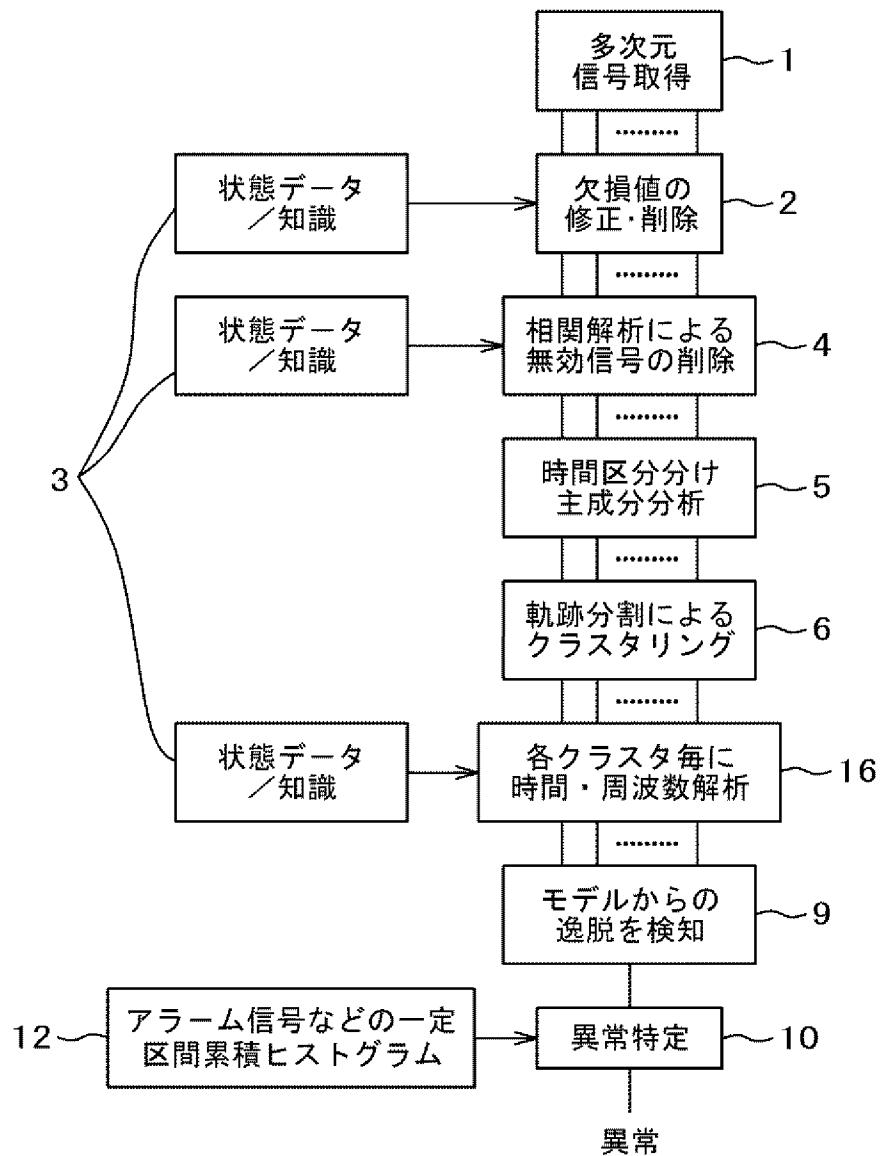


図 19

[図20]

図 2 0

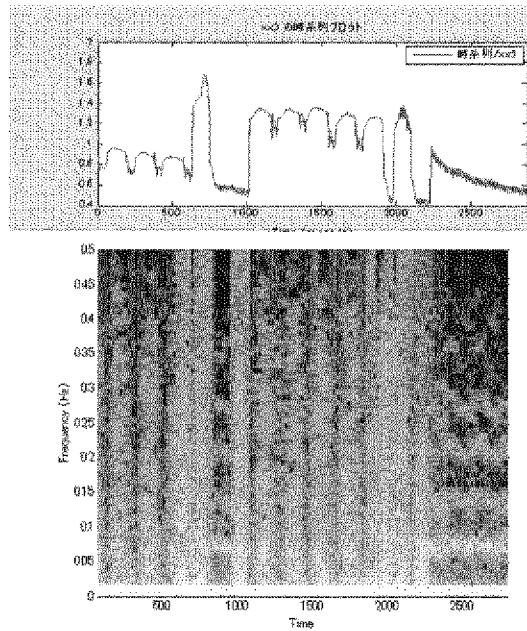


[図21]

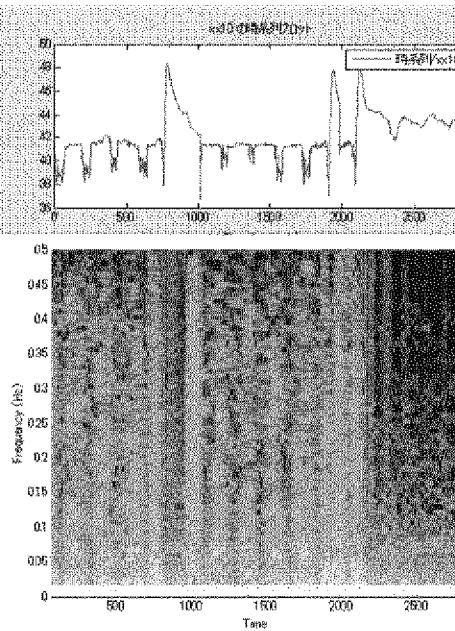
図 21

時間一周波数解析

信号3

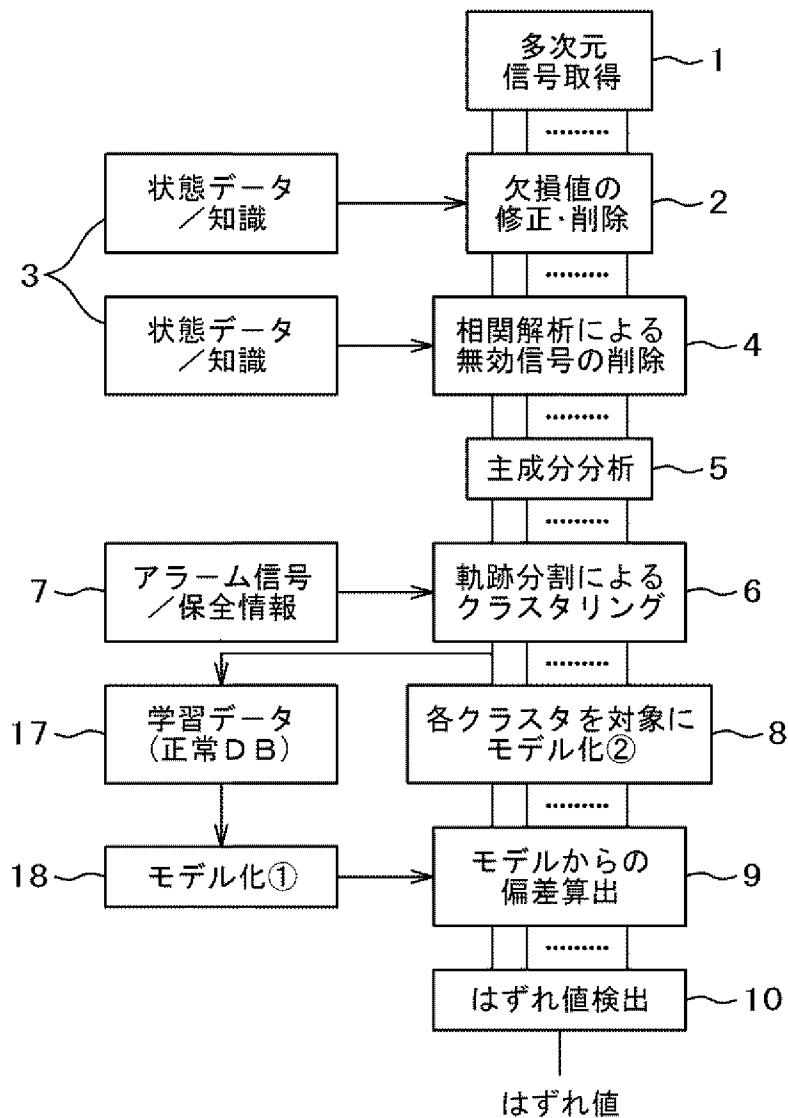


信号10



[図22]

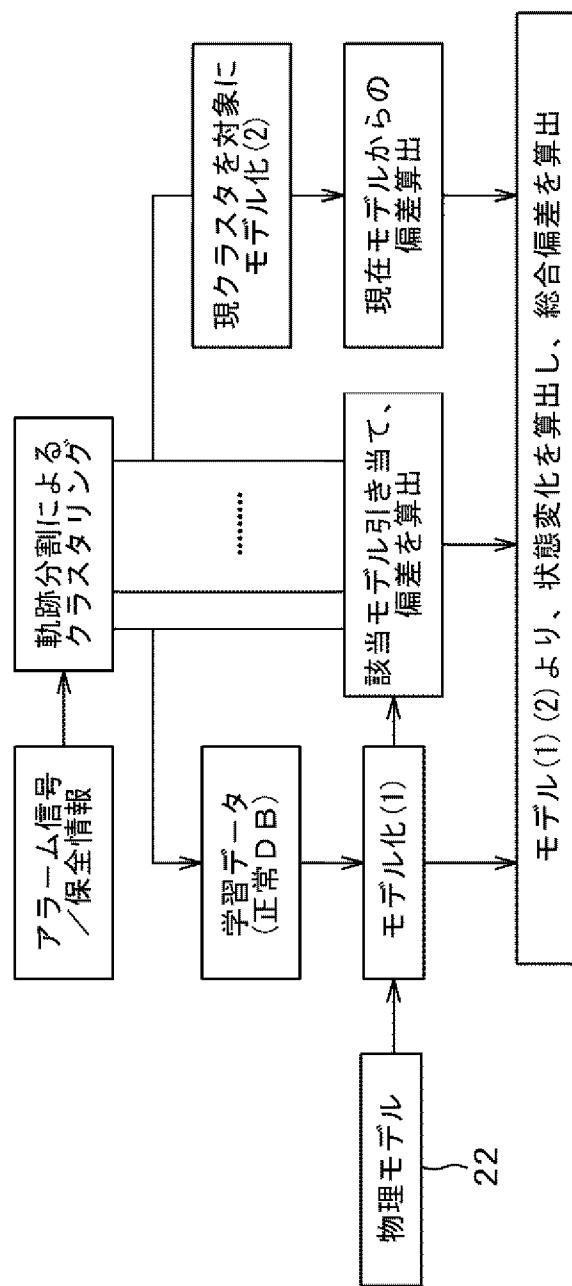
図 2 2



[図23]

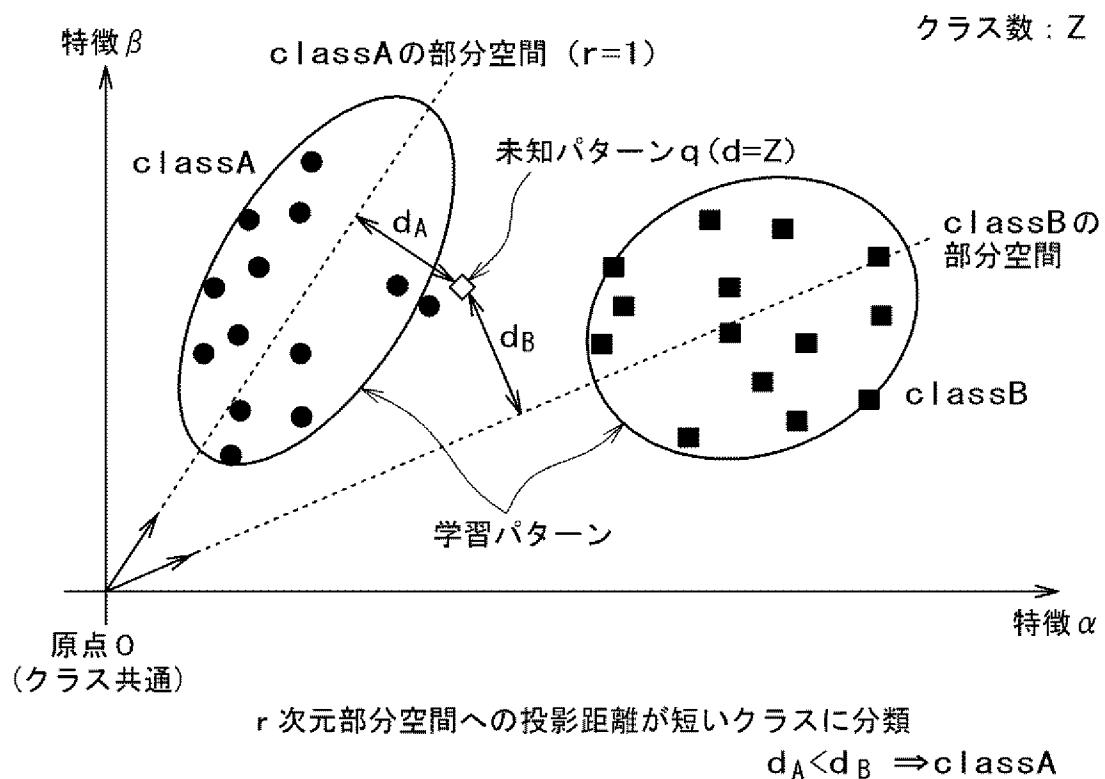
図23

1. 該当学習モデル(個別)との偏差を算出
(応答性を確保)
2. 現在モデルとの偏差を算出(個別)
3. 学習モデルと現在モデルから、状態変化を算出
4. 総合偏差を算出
5. 物理モデルをモデル化(1)に組込み
(物理モデルは、過去の実績をモデルとして組込み)



[図24]

図 2 4



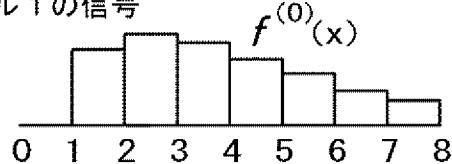
[図25]

図 2 5

Wavelet 変換

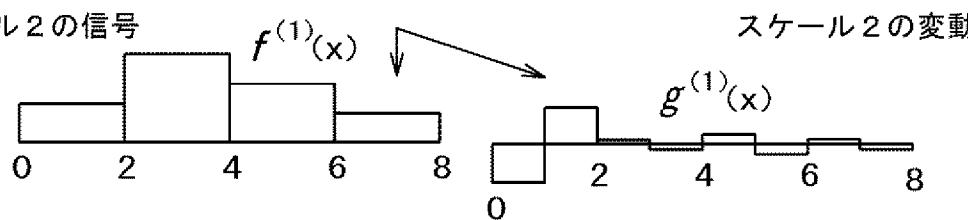
多重解像度分類

スケール 1 の信号

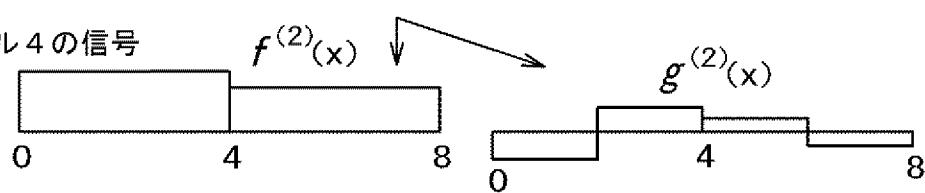


スケール : 値が一定の区間幅
スケールの逆数 : 解像度

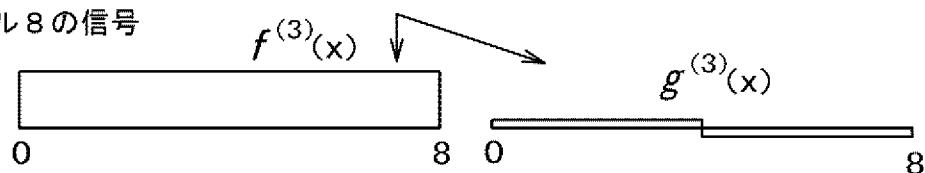
スケール 2 の信号



スケール 4 の信号

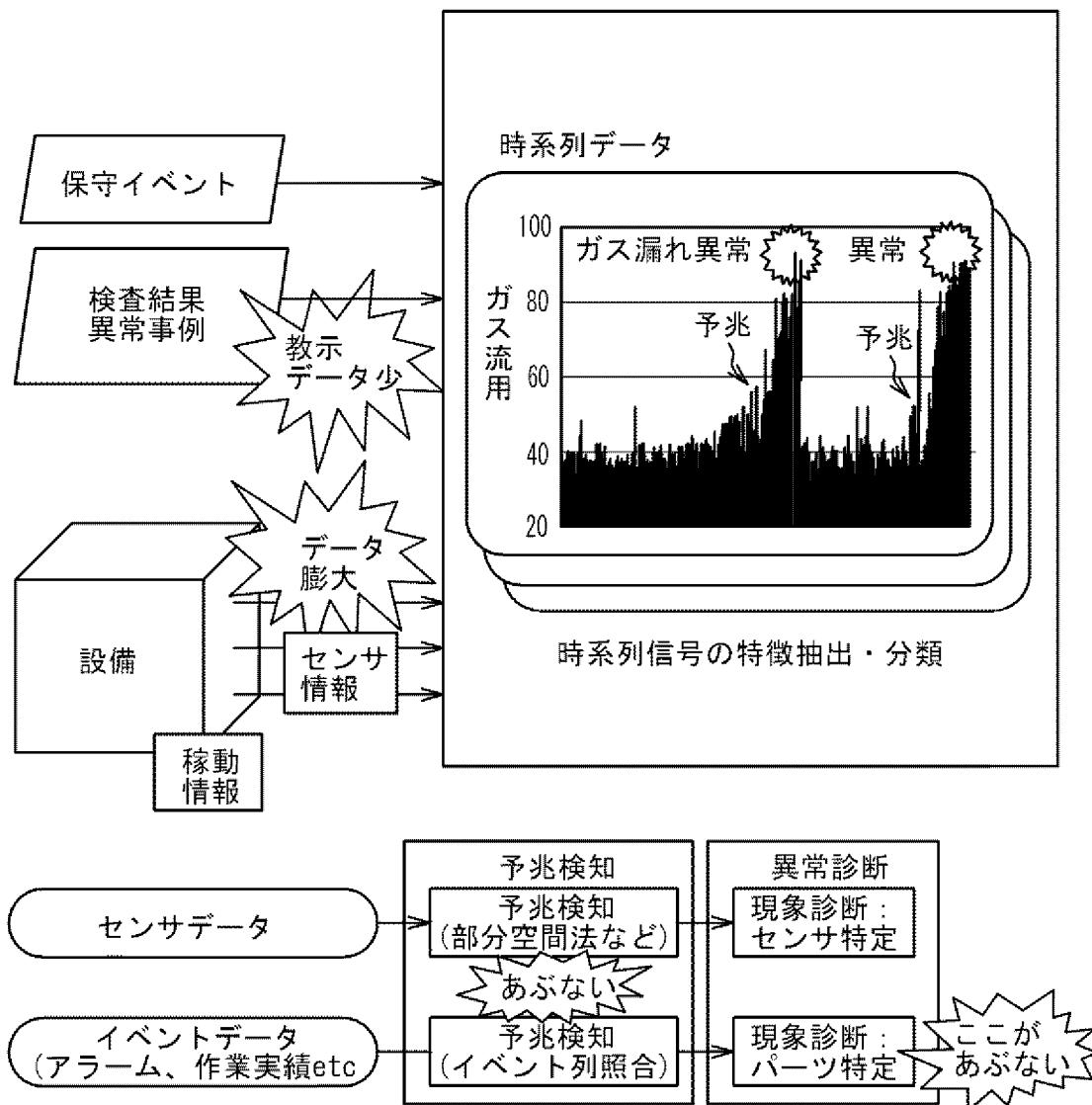


スケール 8 の信号



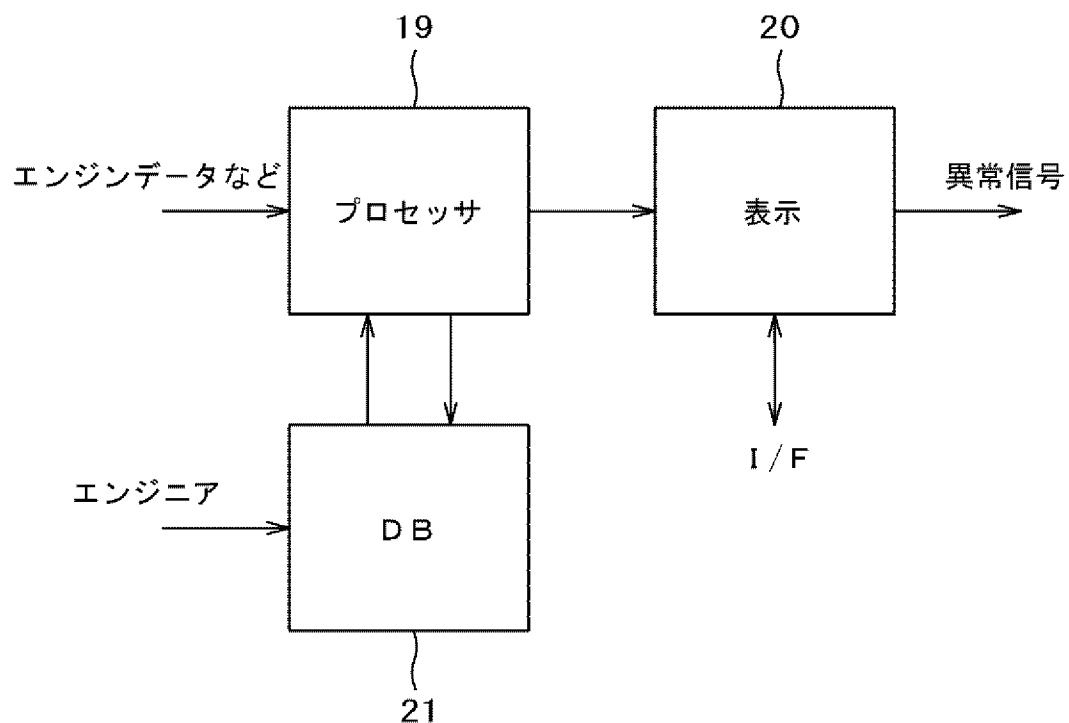
[図26]

図26



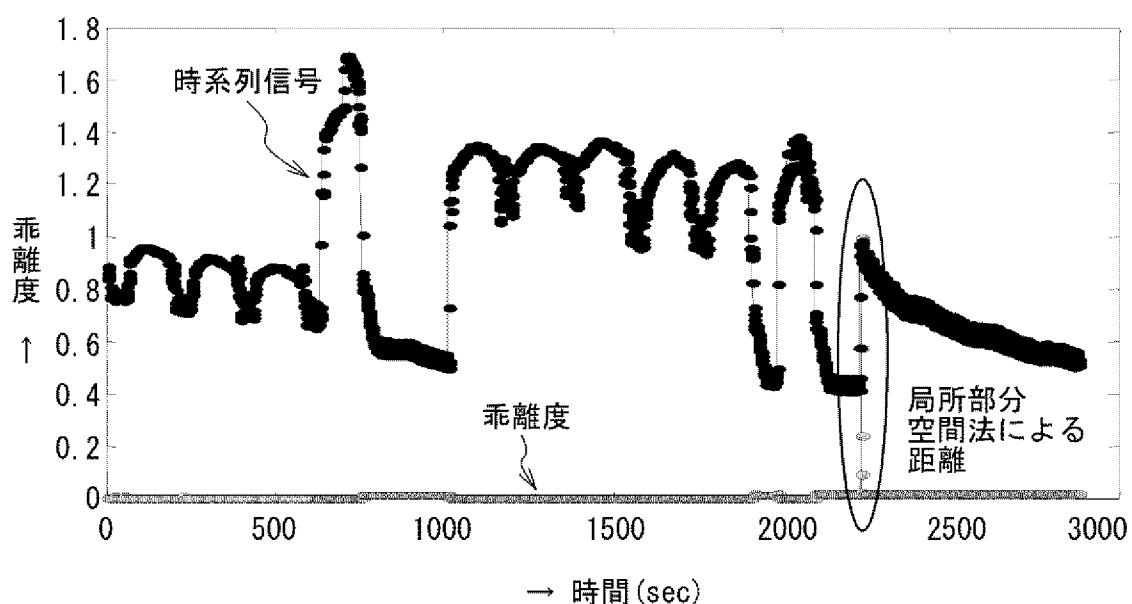
[図27]

図 2 7



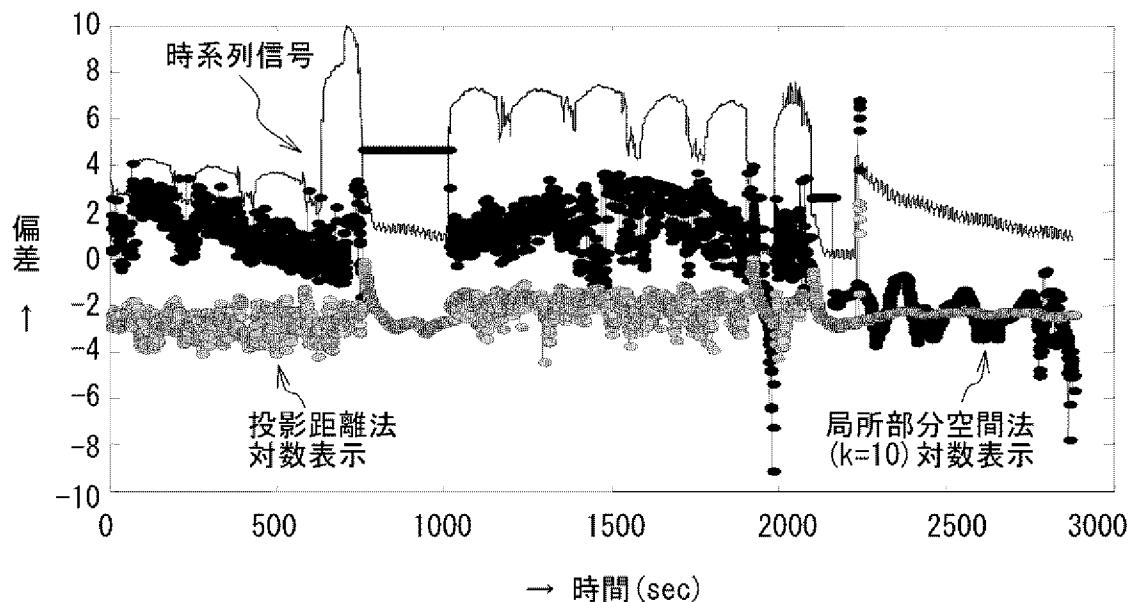
[図28]

図 2 8



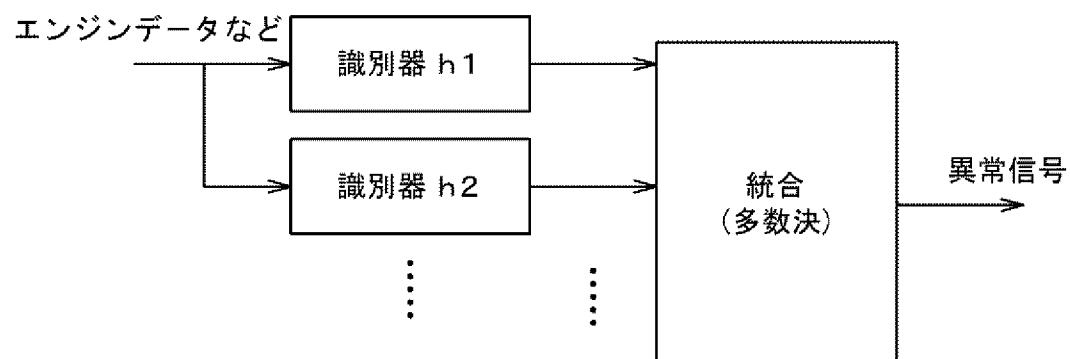
[図29]

図 2 9



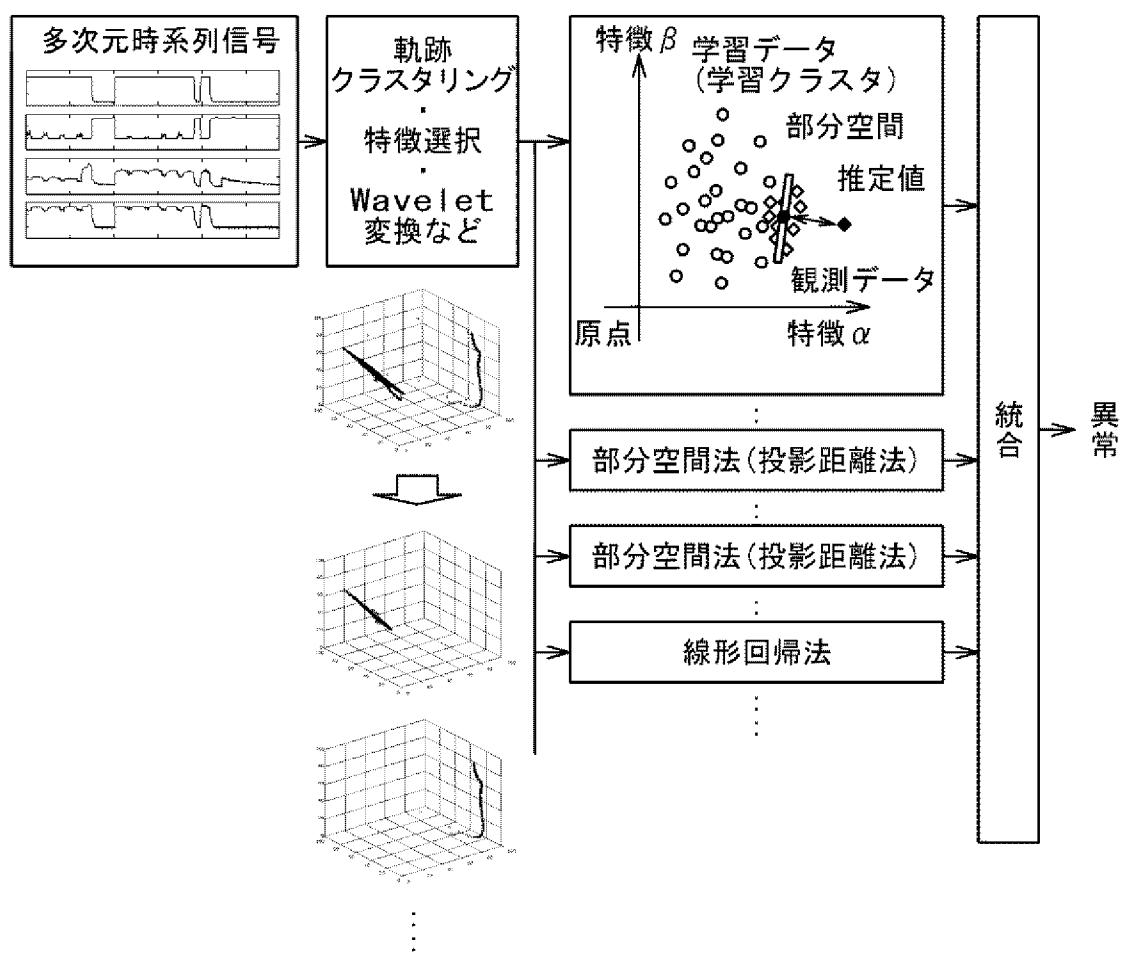
[図30]

図 3 0



[図31]

図31



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/002391

A. CLASSIFICATION OF SUBJECT MATTER
G05B23/02 (2006.01)i, G06Q50/00 (2006.01)i

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)
G05B23/02, G06Q50/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2009
Kokai Jitsuyo Shinan Koho 1971-2009 Toroku Jitsuyo Shinan Koho 1994-2009

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 2006-252259 A (NTT Data Corp.) , 21 September, 2006 (21.09.06) , Full text (Family: none)	1-14
A	JP 2006-107179 A (NTT Data Corp.) , 20 April, 2006 (20.04.06) , Full text (Family: none)	1-14
A	JP 2007-198918 A (Matsushita Electric Works, Ltd.) , 09 August, 2007 (09.08.07) , Full text (Family: none)	1-14

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
12 June, 2009 (12.06.09)

Date of mailing of the international search report
23 June, 2009 (23.06.09)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/002391

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2000-184593 A (Kyushu Electric Power Co., Ltd.) , 30 June, 2000 (30.06.00) , Full text (Family: none)	1-14
A	JP 7-28766 A (Toshiba Corp.) , 31 January, 1995 (31.01.95) , Full text (Family: none)	1-14

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G05B23/02(2006.01)i, G06Q50/00(2006.01)i

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））

Int.Cl. G05B23/02, G06Q50/00

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

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2009年
日本国実用新案登録公報	1996-2009年
日本国登録実用新案公報	1994-2009年

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）

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 2006-252259 A (株式会社エヌ・ティ・ティ・データ) 2006.09.21, 全文 (ファミリーなし)	1-14
A	JP 2006-107179 A (株式会社エヌ・ティ・ティ・データ) 2006.04.20, 全文 (ファミリーなし)	1-14
A	JP 2007-198918 A (松下電工株式会社) 2007.08.09, 全文 (ファミリーなし)	1-14

 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。

* 引用文献のカテゴリー

- 「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
 「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
 「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）
 「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

- 「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矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「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
 「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
 「&」同一パテントファミリー文献

国際調査を完了した日 12.06.2009	国際調査報告の発送日 23.06.2009
国際調査機関の名称及びあて先 日本国特許庁（ISA/JP） 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 川東 孝至 電話番号 03-3581-1101 内線 3324 3U 4135

C (続き) . 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 2000-184593 A (九州電力株式会社) 2000.06.30, 全文 (ファミリーなし)	1-14
A	JP 7-28766 A (株式会社東芝) 1995.01.31, 全文 (ファミリーなし)	1-14