

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6141235号
(P6141235)

(45) 発行日 平成29年6月7日(2017.6.7)

(24) 登録日 平成29年5月12日(2017.5.12)

(51) Int.Cl.

G O 1 M 99/00 (2011.01)

F 1

G O 1 M 99/00

Z

請求項の数 6 外国語出願 (全 9 頁)

(21) 出願番号 特願2014-124072 (P2014-124072)
 (22) 出願日 平成26年6月17日 (2014.6.17)
 (65) 公開番号 特開2015-11027 (P2015-11027A)
 (43) 公開日 平成27年1月19日 (2015.1.19)
 審査請求日 平成29年3月8日 (2017.3.8)
 (31) 優先権主張番号 13/932, 238
 (32) 優先日 平成25年7月1日 (2013.7.1)
 (33) 優先権主張国 米国 (US)

早期審査対象出願

(73) 特許権者 000006013
 三菱電機株式会社
 東京都千代田区丸の内二丁目7番3号
 (74) 代理人 100110423
 弁理士 曽我 道治
 (74) 代理人 100111648
 弁理士 梶並 順
 (74) 代理人 100122437
 弁理士 大宅 一宏
 (74) 代理人 100147566
 弁理士 上田 俊一
 (74) 代理人 100161171
 弁理士 吉田 潤一郎
 (74) 代理人 100161115
 弁理士 飯野 智史

最終頁に続く

(54) 【発明の名称】時系列データにおける異常を検出する方法

(57) 【特許請求の範囲】

【請求項 1】

時系列データにおいて異常を検出する方法であって、
前記時系列データを分割するステップであって、重なり合ってスライドする時間ウィンドウにするステップと、
前記時間ウィンドウごとの特徴を求めるステップと、
類似したウィンドウの特徴をクラスター化するステップであって、普遍的特徴を得るステップと
を含み、

前記普遍的特徴は、確率変動成分を含み、前記確率変動成分の統計は、平均、標準偏差、第1の差分の絶対値の平均、平均値交差数、前記第1の差分における正の値のパーセンテージ、前記第1の差分におけるゼロ値のパーセンテージ、および前記時間ウィンドウ内の前記時系列データの正の差分の平均ラン長、並びに前記統計の標準偏差を含み、

試験時系列データから抽出された前記普遍的特徴をトレーニング時系列データから取得された前記普遍的特徴と比較するステップであって、スコアを求め、前記普遍的特徴は、前記時系列データの軌道成分と、前記時系列データの確率変動成分とを特徴付ける、ステップと、

スコアが閾値よりも高い場合、異常を検出するステップとを含み、前記ステップは、プロセッサにおいて実行される、時系列データにおいて異常を検出する方法。

10

20

【請求項 2】

前記時系列データは、センサーによって機械から取得される、請求項 1 に記載の方法。

【請求項 3】

前記軌道成分は、平滑な曲線によって表され、前記確率変動成分は、前記平滑な曲線のまわりの不規則変動によって表される、請求項 1 に記載の方法。

【請求項 4】

前記普遍的特徴の前記軌道成分は、前記時系列データの移動平均および該平均の制限を含む、請求項 1 に記載の方法。

【請求項 5】

前記普遍的特徴の軌道成分は、前記時系列データの低周波数成分を特徴付ける、請求項 10
1 に記載の方法。

【請求項 6】

前記普遍的特徴の確率変動成分は、前記時系列データの高周波数成分を特徴付ける、請求項 1 に記載の方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、包括的には、時系列データを処理することに関し、より詳細には、機器の動作中に時系列データから異常を判断することに関する。

【背景技術】**【0002】**

複数の用途で、時系列データを生成し、解析することができる。例えば、自動機器監視は、コストの高い修理を回避することができる。これは、機器上または機器の近くのセンサーによって取得される時系列データを解析して、機器のメンテナンスまたは修理が必要であることを示す場合がある異常を検出することによって行うことができる。

【0003】

通常の確率論的かつ多変量の時系列データについて、図 1 A および図 1 B に示すように、異常を検出する従来技術による方法のほとんどは、正常動作範囲 101 外の値を単に求める。

【0004】

例えば、1つの方法は、多変量時系列をベクトル自己回帰 (AR) モデルとして局的にモデル化することができるなどを仮定する。非特許文献 1 を参照されたい。その方法は、まず、トレーニングデータの時間ウィンドウごとに、AR モデルパラメーターの分布を学習する。試験中、時間ウィンドウごとに、AR モデルパラメーターが推定され、これらのパラメーターの確率が、以前に学習された確率分布から求められる。その方法によって学習される分布は、制限的自己回帰推定を用いる。その方法は、不規則確率変動成分を有する時系列データに対し、良好に機能する。

【0005】

ほとんどの方法は、特定の用途の単一のタイプの時系列データに焦点をあてる。時系列データにおける異常検出の問題に対する、より一般的な解決を試みる方法は、ほとんどない。

【0006】

特許文献 1 「Strategies for identifying anomalies in time-series data」は、時系列データにスプラインセグメントを当てはめるモデルベースの手法である。図 1 C を参照されたい。その方法は、スプラインパラメーターにおける変化を計算し、スプライン間の L_1 距離を用いて異常を検出し、 L_2 距離を用いて正常動作を測定する。この方法は、平滑な軌道を有する時系列に対し、良好に機能する。

【先行技術文献】**【特許文献】****【0007】**

10

20

30

40

50

【特許文献 1】米国特許第 7,716,011 号

【非特許文献】

【0008】

【非特許文献 1】Bay 他「A Framework for Discovering Anomalous Regimes in Multivariate Time-Series Data with Local Models」Technical Report, Center for the Study of Language and Information, Stanford University, 2004

【発明の概要】

【0009】

本発明の実施の形態によれば、時系列を、平滑な曲線と、その平滑な曲線のまわりの不規則変動とに分解することができる。平滑な曲線は、時系列の軌道成分と呼ばれ、不規則変動は、時系列の確率変動成分と呼ばれる。従来技術による方法は、通常、これらの成分のうちの 1 つのみを含む時系列を扱う。

【0010】

実施の形態は、時系列データにおける異常を検出する方法を提供する。時系列データは、例えば、機器からセンサーによって取得することができる。本方法は、時系列データにおける異常パターンを検出する。例えば、機械からのセンサーデータは、正常時には、周期的なパターン、例えば、正弦波を生成することができる。センサーが一定の信号、すなわち、平坦なラインを生成する場合、これは、正常動作範囲から外れる値が存在しない場合であっても異常である。本発明は、時系列の確率論的成分および軌道成分の双方をモデル化する。本発明は、異常発見問題を解決し、これは、任意の時系列データとともに、良好に機能する。

【0011】

本方法は、トレーニングに用いられる時系列データを要約する 1 組の普遍的特徴を用いる。普遍的特徴は、本発明の顕著な要素である。これらの特徴は、様々なタイプの時系列データを扱うように設計されるので、普遍的であるといわれる。このため、本方法は、特徴に基づく手法である。

【0012】

本発明の実施の形態によれば、時系列データ源、例えば、機械からのセンサー出力は、確率変動成分と軌道成分との組み合わせとして特徴付けることができる。確率変動成分は、移動平均 (moving mean or average) のまわりの不規則変動を有するデータを生成する一方、軌道成分は、軌道のような平滑な曲線を生成する。

【0013】

自己回帰モデルを用いて時系列をモデル化する等の従来技術の方法は、確率論的時系列に対してのみ良好に機能する。実施の形態によって用いられる普遍的特徴は、これらの特徴が、時系列の軌道のような挙動および確率論的タイプの挙動の双方を特徴付けるという点で優れている。

【図面の簡単な説明】

【0014】

【図 1 A】従来技術による方法によって処理される時系列データである。

【図 1 B】従来技術による方法によって処理される時系列データである。

【図 1 C】従来技術による方法によって処理される時系列データである。

【図 2 A】本発明の実施の形態による時系列データにおける異常を検出する方法の流れ図である。

【図 2 B】本発明の実施の形態による、信頼区間を含む普遍的特徴の軌道成分のグラフである。

【図 3】本発明の実施の形態による、対応する普遍的特徴の軌道成分が重ねられた軌道タイプおよび確率論的タイプが混成した時系列データの値のグラフである。

【図 4】本発明の実施の形態によるトレーニング方法の流れ図である。

【図 5】本発明の実施の形態による異常を検出する方法の流れ図である。

【発明を実施するための形態】

10

20

30

40

50

【0015】**方法の概観**

図2Aに示すように、本発明の実施の形態は、時系列データにおける異常を検出する方法を提供する。例えば、データは、機器または機械からセンサーによって取得される。本方法は、いくつか例を挙げると、一般統計学、信号処理、パターン認識、計量経済学、金融サービス、天気予報、気候変動、海洋学、制御工学、通信工学、医療診断、疫学および市場分析等の他の用途において用いることができる理解すべきである。

【0016】

時系列データは、トレーニングデータ101と試験データとを含むことができる。いずれの場合においても、時系列データは、長さwの重なり合ってスライドする時間ウィンドウに分割される。サンプリングレート、ウィンドウサイズ、および重ね合わせの量は、用途に特有とすることができます、例えば、短期現象の場合、ウィンドウは、ほんの数分の一秒にすることができます。長期の挙動の場合のウィンドウは、数年または数世紀とすることができます。以下のように、ウィンドウから特徴が抽出される。これらの特徴は、様々なタイプの時系列データ、例えば、確率論的タイプ、軌道タイプ、並びに単一の特徴において軌道タイプおよび確率論的タイプの時系列データの特性を組み合わせた混成タイプを表すことができるので、普遍的である。

【0017】

トレーニング中、トレーニング用の普遍的特徴111が、トレーニングデータ、例えば、異常を一切含まないデータから抽出される(110)。

【0018】

試験中、同様にして、特徴が抽出され、特徴付けされる。次に、学習された特徴と比較することによって、試験データ102内の異常103が検出される(120)。検出は、トレーニング中に学習された特徴と、試験データからの特徴との間の距離を用いることができる。

【0019】

この方法のステップおよび本明細書において説明される他の手順は、当該技術分野において既知のメモリおよび入/出力インターフェースに接続されたプロセッサ100において実行することができる。

【0020】**軌道および確率論的特徴**

普遍的特徴は、2つの主要成分、すなわち、軌道成分および確率変動成分を含む。

【0021】**軌道成分**

軌道成分は、時系列がウィンドウ内でたどる平滑な軌道を示す。軌道成分は、ウィンドウ移動平均を用い、これは、時系列の低域通過フィルタリングに等しい。軌道成分は、各時間ステップにおけるウィンドウの平均に対する制限も含み、トレーニングデータにおける全ての類似した軌道の中で、軌道の変動が通常どの程度見られるかが規定される。これは、軌道に沿って信頼区間を配置することに等しい。

【0022】

普遍的特徴の軌道成分の例が、図2Bに示されている。軌道成分は、長さ2wのベクトルとして表され、移動平均200と、平均に対する制限201とを含む。軌道成分は、各時間ステップにおける軌道の変動を示す区間とともに、平均軌道とみなすことができる。軌道成分は、時系列データの低周波数成分を特徴付ける。

【0023】

ほとんどの従来技術による方法が、起動後にある定常状態202に達したシステムの確率論的挙動にしか関心を有しないことに留意すべきである。対照的に、本発明者らは、過渡的状態203にも関心を有する。

【0024】**確率変動成分**

10

20

30

40

50

時系列の確率変動成分は、不規則過程に起因する可能性が最も高い、時系列データにおける変動を示す。普遍的特徴の確率変動成分は、時系列データの時間ウィンドウにわたって、各統計の標準偏差とともに求められた様々な統計を含む。統計は、N要素ベクトルとして記憶される。ここで、Nは、計算される統計の数である。各統計の標準偏差が、別のN要素ベクトルとして記憶され、合計2N個の要素となる。好ましい実施の形態では、確率変動成分の統計は、

- i . 平均、
- i i . 標準偏差、
- i i i . 第1の差分、例えば

【数1】

$$\underline{z}(t) = z(t) - z(t-1)$$

10

の絶対値の平均、

- i v . 平均値交差数、
- v . 第1の差分における正の値のパーセンテージ、
- v i . 第1の差分におけるゼロ値のパーセンテージ、および
- v i i . 正の差分の平均ラン長、

を含む。

【0025】

図3は、対応する普遍的特徴の軌道成分が重ねられた混成型時系列データを示している。

20

【0026】

トレーニング

図4に示すようなトレーニング中、時系列のトレーニングデータ101から普遍的特徴411が抽出される(410)。トレーニングデータは、正常な時系列データのみを含むべきである。これは、時系列を長さwの重なり合う時間ウィンドウに分割することによって行われる。関連付けられた標準偏差を有しない普遍的特徴が、時間ウィンドウごとに求められる。

【0027】

次に、類似した普遍的特徴のクラスター421が求められる(420)。クラスターごとに、上記で説明したように、軌道の時間ステップごとの標準偏差および各統計の標準偏差431が求められ、クラスターごとの普遍的特徴が形成される(430)。次に、クラスターの普遍的特徴が収集され、時系列データを特徴付ける単一組の普遍的特徴111にされる(440)。

30

【0028】

試験

図5に示すような試験中、試験時系列データ102において、異常130が以下のように検出される。試験時系列はトレーニング中、長さwの重なり合う時間ウィンドウに分割される(510)。関連付けられた標準偏差を有しない普遍的特徴511が、時間ウィンドウごとに求められる。次に、普遍的特徴は、トレーニング用の普遍的特徴と比較され、最も類似した特徴が求められる(520)。この比較は、以下に説明される距離メトリックを用いる。

40

【0029】

距離メトリック

標準偏差を含む、クラスターのトレーニング用の普遍的特徴fv1と、標準偏差を含まない、試験データ内の单一の時間ウィンドウの普遍的特徴fv2との間の距離が、以下の式に従って求められる。

【0030】

【数2】

$$Dist(fv1, fv2) = \sum_t \max(0, |\mu_1(t) - \mu_2(t)| - 3\sigma_1(t)) \\ + \sum_{i=1}^7 \max(0, |m_1(i) - m_2(i)| - 3s_1(i))$$

【0031】

ここで、 μ_1 は、 $fv1$ の軌道成分であり、 σ_1 は、 $fv1$ の軌道成分の標準偏差のベクトルであり、 m_1 は、 $fv1$ の統計のベクトルであり、 s_1 は、 $fv1$ の統計の標準偏差のベクトルである。同様に、 μ_2 は、単一の時間ウィンドウ $fv2$ の普遍的特徴の軌道成分であり、 m_2 は、 $fv2$ の統計のベクトルである。10

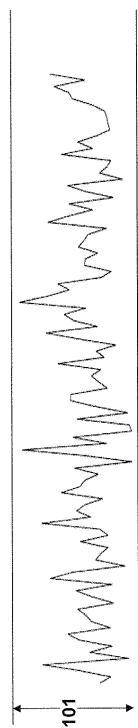
【0032】

この定式は、 $fv1$ の平均軌道の3標準偏差（3シグマ）外側からの $fv2$ の距離と、 $fv1$ の各特徴統計の3標準偏差（3シグマ）外側からの距離とを加えたものを測定する。。

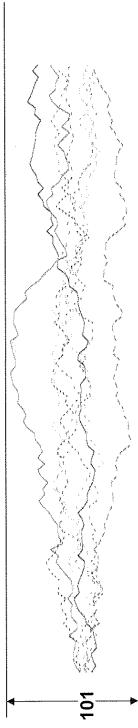
【0033】

試験用の普遍的特徴とトレーニング用の普遍的特徴との間のそのような距離の最小距離は、その時間ウィンドウの異常スコア521である。異常スコアが、ある閾値430よりも高い場合、試験時間ウィンドウは、異常動作と関連付けられる。20

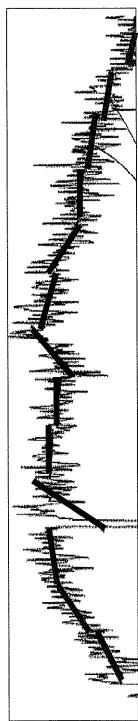
【図1A】



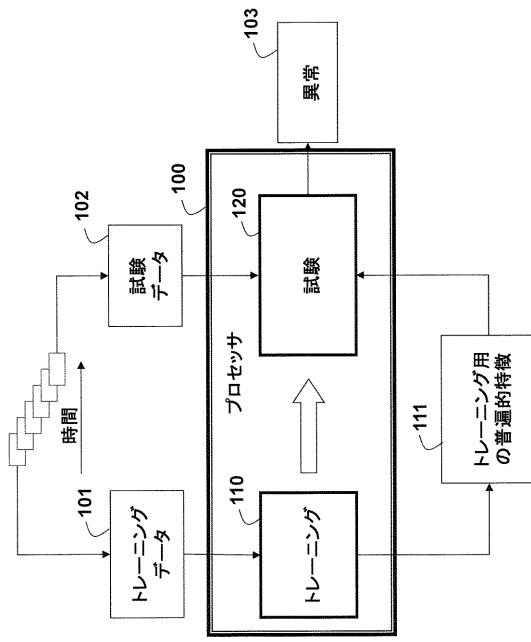
【図1B】



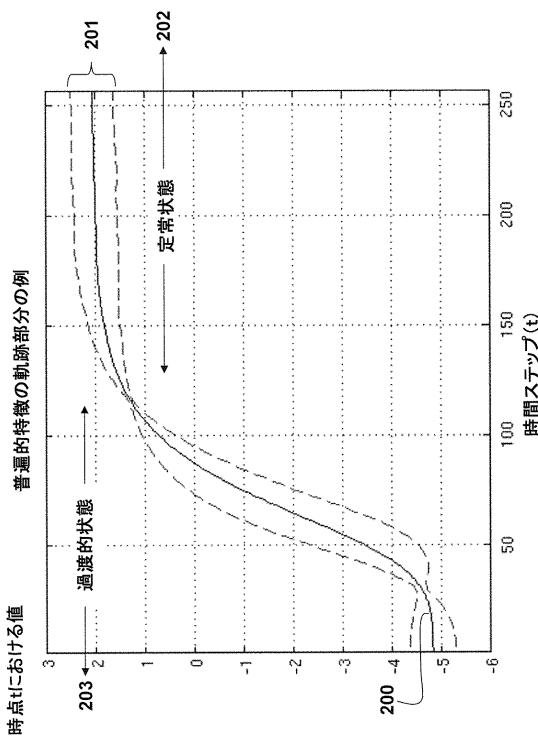
【図1C】



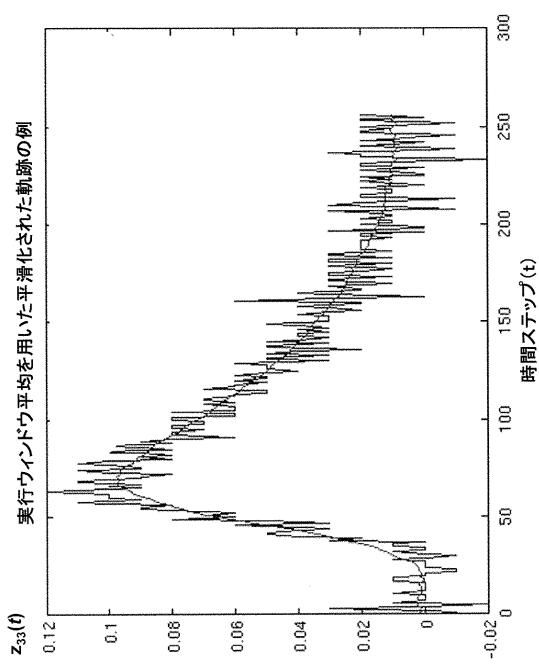
【図2A】



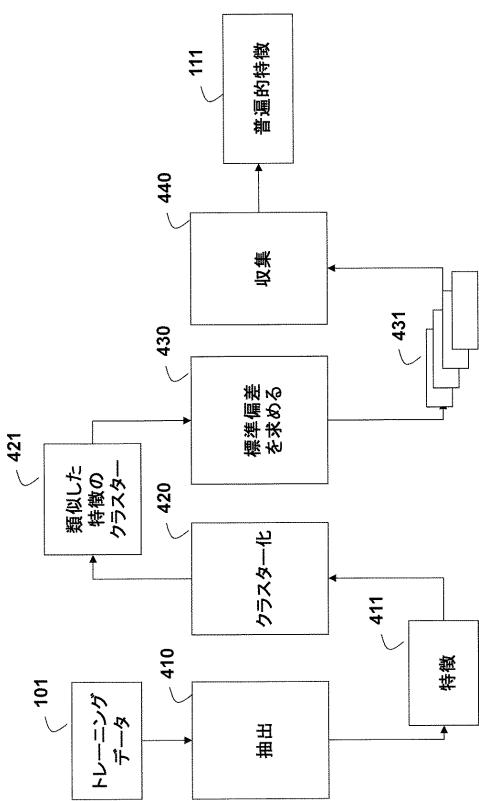
【図2B】



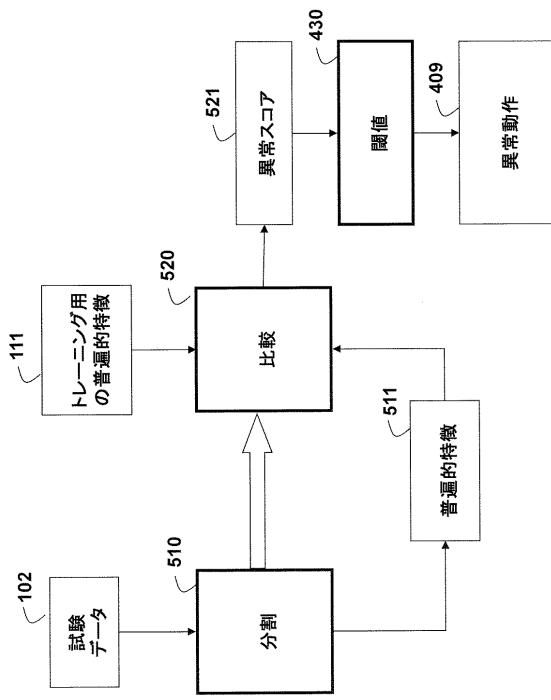
【図3】



【図4】



【図5】



フロントページの続き

(72)発明者 マイケル・ジェイ・ジョーンズ

アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウエイ 201、ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド

審査官 山口 剛

(56)参考文献 特開平08-043192(JP,A)

特開平08-159928(JP,A)

特開2002-181038(JP,A)

特開平06-129955(JP,A)

特開2013-246818(JP,A)

米国特許出願公開第2014/0149806(US,A1)

米国特許出願公開第2012/0163463(US,A1)

米国特許出願公開第2011/0185422(US,A1)

米国特許出願公開第2008/0225738(US,A1)

米国特許出願公開第2012/0136676(US,A1)

(58)調査した分野(Int.Cl., DB名)

G01M 13/00 - 13/04

G01M 99/00

G01H 1/00 - 17/00