

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7399998号
(P7399998)

(45)発行日 令和5年12月18日(2023.12.18)

(24)登録日 令和5年12月8日(2023.12.8)

(51)国際特許分類 F I
G 0 6 N 20/00 (2019.01) G 0 6 N 20/00 1 3 0

請求項の数 2 (全10頁)

(21)出願番号	特願2022-53072(P2022-53072)	(73)特許権者	000005326 本田技研工業株式会社 東京都港区南青山二丁目1番1号
(22)出願日	令和4年3月29日(2022.3.29)	(74)代理人	100095566 弁理士 高橋 友雄
(65)公開番号	特開2023-146080(P2023-146080 A)	(74)代理人	100114775 弁理士 高岡 亮一
(43)公開日	令和5年10月12日(2023.10.12)	(74)代理人	100121511 弁理士 小田 直
審査請求日	令和4年11月29日(2022.11.29)	(72)発明者	軽部 俊和 東京都港区南青山2丁目1番1号 本田 技研工業株式会社内
		審査官	円子 英紀

最終頁に続く

(54)【発明の名称】 教師データ収集装置

(57)【特許請求の範囲】

【請求項1】

検査対象物が異常品とされる、当該検査対象物の外観画像である不良品データを、所定の学習モデルの学習に用いられる教師データとして収集する教師データ収集装置であって、前記不良品データを、非エキスパートの選定によって取得する非エキスパート不良品データ取得手段と、

前記不良品データを、エキスパートの選定によって取得するエキスパート不良品データ取得手段と、

前記非エキスパート不良品データ取得手段によって取得された前記不良品データを、非エキスパートデータとして格納する非エキスパートデータ格納部と、

前記エキスパート不良品データ取得手段によって取得された前記不良品データを、エキスパートデータとして格納するエキスパートデータ格納部と、

前記非エキスパートデータ格納部に格納されている全ての非エキスパートデータの特徴量の標準偏差を、第1標準偏差として算出する第1標準偏差算出手段と、

前記エキスパートデータ格納部に格納されている全てのエキスパートデータの特徴量の標準偏差を、第2標準偏差として算出する第2標準偏差算出手段と、

前記非エキスパートデータ格納部に格納されている全ての非エキスパートデータの特徴量の順位和を、第1順位和として算出する第1順位和算出手段と、

前記エキスパートデータ格納部に格納されている全てのエキスパートデータの特徴量の順位和を、第2順位和として算出する第2順位和算出手段と、

前記第 1 及び第 2 標準偏差並びに前記第 1 及び第 2 順位和に基づいて、前記エキスパート不良品データ取得手段による前記不良品データの取得の継続及び終了を決定する決定手段と、

を備えていることを特徴とする教師データ収集装置。

【請求項 2】

前記第 1 標準偏差を σ_1 、前記第 2 標準偏差を σ_2 、前記第 1 順位和を S_1 、前記第 2 順位和を S_2 とした場合に、前記決定手段は、下式 (1) によって算出される判定係数 $Vrank$ が所定値以上のときに、前記エキスパート不良品データ取得手段による前記不良品データの取得の終了を決定することを特徴とする請求項 1 に記載の教師データ収集装置。

【数 1】

$$Vrank = 1 - \frac{1}{\left(\frac{\beta}{\alpha}\right)^2 \times \left(\frac{\theta}{\xi}\right)^2} \cdot \cdot \cdot (1)$$

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ニューラルネットワークを用いた機械学習機能を有する検査装置などの学習に用いられる教師データを収集するための教師データ収集装置に関する。

20

【背景技術】

【0002】

近年、ニューラルネットワークを用いた機械学習機能を有する検査装置により、各種の工業製品や部品などの検査対象物について、正常品（良品）であるか、異常品（不良品）であるかを判定する検品作業の自動化技術の開発が進んでいる。上記のような検査装置では、良品と不良品に分類された検査対象物の外観の画像データを多数読み込ませることにより、学習が行われる。そして、分類基準を学習した検査装置により、カメラで撮影した新規の検査対象物を、良品と不良品に分類することが可能になる。

【0003】

上述したように、検査装置の学習には、良品及び不良品の画像データが用いられ、検査精度を高めるためには、良品及び不良品のいずれも、多数の画像データが必要である。しかし、工業製品などの製造現場では一般に、できるだけ不良品を出さないように製造が行われるため、良品の数は多いものの、不良品の数は非常に少ない。そのため、比較的容易に収集可能な良品の画像データ（以下「良品データ」という）に比べて、不良品の画像データ（以下「不良品データ」という）の収集が困難である。

30

【0004】

また、不良品データには、熟練者や経験年数の長い作業員などのエキスパートによって選定されるデータ（以下「エキスパートデータ」という）と、新人や経験年数の短い作業員などの非エキスパートによって選定されるデータ（以下「非エキスパートデータ」という）がある。上記のようなエキスパートデータ及び非エキスパートデータを学習させることにより、学習モデルを生成するものとして、例えば特許文献 1 に記載のものが知られている。

40

【0005】

この特許文献 1 では、「エキスパートデータ」を、知識、経験が豊かな専門家によってラベル付けが行われ、ラベル付けの精度（信頼性）が高い教師データであると規定する一方、「非エキスパートデータ」を、ラベル付けは行われているが、その精度（信頼性）が不明確な教師データであると規定している。また、非エキスパートデータには、エキスパートデータを参照して、信頼度が付与されるようになっている。そして、エキスパートデータと、信頼度付非エキスパートデータを用いて学習することにより、分類モデルとしての学習モデルを生成している。

50

【先行技術文献】

【特許文献】

【0006】

【文献】特開2009-110064号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

上述した従来の学習モデルの生成方法では、非エキスパートデータ自体の質についてはある程度の評価がなされているものの、エキスパートデータ自体の質や量について適切に評価されているとは言えない。そのため、上記のエキスパートデータと信頼度付非エキスパートデータを用いて生成された学習モデルにおいて、エキスパートデータ自体の質や量が十分でない場合には、学習モデルの分類精度が低くなってしまふ。また、エキスパートデータの収集には通常、非エキスパートデータの収集に比べてコストがかかるため、エキスパートデータの量が必要以上に多い場合には、学習モデルの生成コストの上昇を招いてしまふ。

10

【0008】

本発明は、以上のような課題を解決するためになされたものであり、エキスパートデータをできる限り少なくしながら、高い分類精度を有する学習モデルを生成するための教師データを収集することができる教師データ収集装置を提供することを目的とする。

【課題を解決するための手段】

20

【0009】

上記の目的を達成するために、請求項1に係る発明は、検査対象物が異常品とされる、検査対象物の外観画像である不良品データを、所定の学習モデルの学習に用いられる教師データとして収集する教師データ収集装置であって、不良品データを、非エキスパートの選定によって取得する非エキスパート不良品データ取得手段と、不良品データを、エキスパートの選定によって取得するエキスパート不良品データ取得手段と、非エキスパート不良品データ取得手段によって取得された不良品データを、非エキスパートデータとして格納する非エキスパートデータ格納部と、エキスパート不良品データ取得手段によって取得された不良品データを、エキスパートデータとして格納するエキスパートデータ格納部と、非エキスパートデータ格納部に格納されている全ての非エキスパートデータの特徴量の標準偏差を、第1標準偏差として算出する第1標準偏差算出手段と、エキスパートデータ格納部に格納されている全てのエキスパートデータの特徴量の標準偏差を、第2標準偏差として算出する第2標準偏差算出手段と、非エキスパートデータ格納部に格納されている全ての非エキスパートデータの特徴量の順位和を、第1順位和として算出する第1順位和算出手段と、エキスパートデータ格納部に格納されている全てのエキスパートデータの特徴量の順位和を、第2順位和として算出する第2順位和算出手段と、第1及び第2標準偏差並びに第1及び第2順位和に基づいて、エキスパート不良品データ取得手段による不良品データの取得の継続及び終了を決定する決定手段と、を備えていることを特徴とする。

30

【0010】

この構成によれば、教師データ収集装置によって、検査対象物の不良品データを教師データとして収集する場合、非エキスパート不良品データ取得手段により、非エキスパートの選定によって不良品データが取得され、その不良品データが非エキスパートデータとして、非エキスパートデータ格納部に格納される。また、エキスパート不良品データ取得手段により、エキスパートの選定によって不良品データが取得され、その不良品データがエキスパートデータとして、エキスパートデータ格納部に格納される。

40

【0011】

上記の非エキスパートは、新人や検品作業の経験年数の短い作業員などであり、そのような非エキスパートが選定する非エキスパートデータは、比較的発生しやすい不良品画像である。そのため、収集された大量の非エキスパートデータについて、各不良品画像から特徴を抽出して得られる値、すなわち特徴量に変換した場合には、平均値の付近に偏った

50

分布状態となり、標準偏差が小さくなる。本発明では、非エキスパートデータ格納部に格納されている全ての非エキスパートデータの特徴量の標準偏差を、第1標準偏差として算出する。

【0012】

一方、エキスパートは、熟練者や検品作業の経験年数の長い作業者などであり、そのようなエキスパートが選定するエキスパートデータは、比較的発生しやすいものから数年に一度発生するものまで、あらゆる不良品画像である。そのため、収集された多数のエキスパートデータの特徴量に変換した場合には、非エキスパートデータと異なり、全体的に満遍なく広がった分布状態となり、標準偏差が大きくなる。本発明では、エキスパートデータ格納部に格納されている全てのエキスパートデータの特徴量の標準偏差を、第2標準偏差として算出する。

10

【0013】

また、収集された大量の非エキスパートデータについて、例えばほとんどのものが平均値の付近に偏在していても、1つのデータが大きく飛び抜けた外れ値がある場合、標準偏差が大きくなってしまふ。一般に、学習モデルの学習に最適な良質の教師データは、平均から外れ値まで満遍なくデータが散らばっている状態であることが好ましい。そのため、全ての非エキスパートデータの特徴量について、値の小さいものから大きいものまで順に順位を付け、その順位和を、第1順位和として算出する。一方、全てのエキスパートデータの特徴量についての順位和を、第2順位和として算出する。

【0014】

20

以上のように算出された第1及び第2標準偏差、並びに第1及び第2順位和に基づき、決定手段により、エキスパート不良品データ取得手段による不良品データの取得の継続及び終了を決定する。具体的には、例えば、第1及び第2標準偏差による比（標準偏差比）が、所定の基準値以上のときには、エキスパートデータとして、あらゆる不良品画像を網羅した教師データが収集できたと判断することができる。また、第1及び第2順位和による比（順位和比）が、所定の基準値以上のときには、エキスパートデータとして、非エキスパートデータにおいて外れ値とされるデータを十分に含む教師データが収集できたと判断することができる。したがって、上記の標準偏差比及び順位和比を用い、エキスパートによって選定される不良品データの取得の継続及び終了を決定することにより、エキスパートデータをできる限り少なくしながら、高い分類精度を有する学習モデルを生成するための教師データを収集することができる。

30

【0015】

請求項2に係る発明は、請求項1に記載の教師データ収集装置において、第1標準偏差を、第2標準偏差を、第1順位和を、第2順位和をとした場合に、決定手段は、下式(1)によって算出される判定係数 $Vrank$ が所定値以上のときに、前記エキスパート不良品データ取得手段による前記不良品データの取得の終了を決定することを特徴とする。

【数1】

$$Vrank = 1 - \frac{1}{\left(\frac{\beta}{\alpha}\right)^2 \times \left(\frac{\theta}{\xi}\right)^2} \quad \dots (1)$$

40

【0016】

この構成によれば、第1標準偏差と第2標準偏差による標準偏差比（ β/α ）、及び第1順位和と第2順位和による順位和比（ θ/ξ ）を用いて算出される判定係数 $Vrank$ は、標準偏差比及び順位和比が大きくなると、値が1に近づく。この判定係数 $Vrank$ が所定値以上になったときに、教師データとして収集した非エキスパートデータ及びエキスパートデータが、データ全体としての平均付近に偏在せず、かつ適切に分散した状態であることを判定することが可能となる。

50

【図面の簡単な説明】

【0017】

【図1】本発明の一実施形態による教師データ収集装置によって収集された教師データが学習に利用される検査システムの概要を説明するための図である。

【図2】本発明の一実施形態による教師データ収集装置を示すブロック図である。

【図3】教師データ収集装置による教師データの収集処理を示すフローチャートである。

【図4】分類モデルの生成処理を示すフローチャートである。

【発明を実施するための形態】

【0018】

以下、図面を参照しながら、本発明の好ましい実施形態を詳細に説明する。図1は、後述する教師データ収集装置11によって収集された多数の不良品画像のデータ（不良品データ）、及び多数の良品画像のデータ（良品データ）を用いて学習された学習モデルを備えた検査システムを示している。この検査システム1は、例えば車両部品の製造工場などに設置され、製造された車両部品（例えばシリンダブロック）が正常品（良品）であるか、異常品（不良品）であるかを、車両部品の外観を検査することによって自動で判別するものである。以下、検査すべき車両部品を、「検査対象物」というものとする。

10

【0019】

図1に示すように、検査システム1は、検査対象物Gを所定方向に所定速度で搬送する搬送機2と、検査対象物Gが所定の検査位置に到達したときに、その検査対象物Gの良否を判定する検査装置3とを備えている。なお、図示は省略するが、検査装置3によって不良品であると判定された検査対象物Gは、搬送機2から取り除かれたり、不良品専用の格納場所に搬送されたりするようになっている。

20

【0020】

検査装置3は、主にコンピュータから成る情報処理装置で構成されており、制御部4、画像取得部5、記憶部6、学習部7、入力部8、出力部9及びカメラ10を備えている。

【0021】

制御部4は、CPUを備えており、検査装置3の上記各部5～9及びカメラ10などを制御する。画像取得部5は、カメラ10で撮影された検査対象物Gの外観画像を、デジタルデータとして取得する。記憶部6は、ROM及びRAMを有しており、検査装置3の制御で使用される各種のプログラムが記憶されているとともに、各種データが記憶される。学習部7は、検査対象物Gの良否を判別するための基準が学習された学習モデルを有している。入力部8は、作業者によって操作されるキーボードやマウスを有するとともに、外部からデータや信号が入力可能に構成されている。出力部9は、検査対象物Gの判定結果が表示されるディスプレイなどの表示器を有している。

30

【0022】

図2は、本発明の一実施形態による教師データ収集装置11を示している。この教師データ収集装置11は、検査対象物Gの検品作業を行う作業者によって操作されることにより、不良品データを収集するためのものである。教師データ収集装置11は、前述した検査装置3と同様、コンピュータから成る情報処理装置で構成されており、不良品画像取得部12（非エキスパート不良品データ取得手段、エキスパート不良品データ取得手段）と、非エキスパートデータ格納部13、エキスパートデータ格納部14、特徴量変換部15、標準偏差算出部16（第1標準偏差算出手段、第2標準偏差算出手段）、順位和算出部17（第1順位和算出手段、第2順位和算出手段）、判定係数 $Vrank$ 算出部18、及び不良品画像取得の継続・終了決定部19（決定手段）を備えている。

40

【0023】

不良品画像取得部12は、前述した検査装置3のカメラ10と同様のカメラによって撮影された検査対象物Gの外観画像について、作業者によって不良品であると判定されたものを、不良品データとして取得する。

【0024】

非エキスパートデータ格納部13は、非エキスパート（新人や検品作業の経験年数の短

50

い作業員)によって選定された不良品データ(非エキスパートデータ)を格納する。一方、エキスパートデータ格納部14は、エキスパート(熟練者や検品作業の経験年数の長い作業員)によって選定された不良品データ(エキスパートデータ)を格納する。

【0025】

特徴量変換部15は、不良品データを所定の特徴量に変換する。具体的には、例えばSIFT(Scale-Invariant Feature Transform)やCNN(Convolution Neural Network)を用いて、不良品データを特徴量に変換する。

【0026】

標準偏差算出部は、上記の非エキスパートデータ格納部13及びエキスパートデータ格納部14にそれぞれ格納された多数の不良品データについて、特徴量による標準偏差を算出するものである。また、順位和算出部17は、上記の多数の不良品データについて、特徴量による順位和を算出するものである。

10

【0027】

判定係数Vrank算出部18は、後述するように、非エキスパートデータの標準偏差及び順位和と、エキスパートデータの標準偏差及び順位和の4つをパラメータとして、不良品画像取得の継続又は終了を判定するための判定係数Vrankを算出する。

【0028】

不良品画像取得の継続・終了決定部19は、算出された判定係数Vrankに応じて、不良品画像の取得を継続するか、あるいは終了するかを決定する。

【0029】

図3は、上述した教師データ収集装置11による教師データとしての不良品画像の収集処理を示している。本処理ではまず、ステップ1(「S1」と図示。以下同じ)において、不良品画像を取得する。

20

【0030】

次いで、ステップ2において、取得した不良品画像が、非エキスパートによる選定か否かを判定する。例えば、教師データ収集装置11を非エキスパートが操作しているときには、取得された不良品画像は非エキスパートによる選定であり、教師データ収集装置11をエキスパートが操作しているときには、取得された不良品画像はエキスパートによる選定であると判定される。上記ステップ2の判定結果がYESのときには、取得画像を非エキスパートデータ格納部13に格納する(ステップ3)。なお、非エキスパートが不良品画像を選定する場合、選定した不良品画像が今まで見たことのないものであるときには、その不良品画像は外れ値であるとし、エキスパートデータとの条件を揃えるために、非エキスパートデータ格納部13には格納されないようになっている。

30

【0031】

次いで、非エキスパートデータ格納部13の全画像を、特徴量変換部15によって特徴量に変換する(ステップ4)。なお、非エキスパートデータ格納部13に今回格納された不良品画像以外のものは、あらかじめ特徴量にすでに変換されており、したがって、ステップ4では、今回取得された不良品画像のみを、特徴量に変換する。そして、非エキスパートデータ格納部13に格納されている全画像の特徴量に基づき、標準偏差算出部16及び順位和算出部17によって、標準偏差(第1標準偏差)及び順位和(第1順位和)をそれぞれ算出する(ステップ5)。なお、

40

【0032】

一方、前記ステップ2の判別結果がNOで、今回取得した不良品画像がエキスパートによって選定されたものであるときには、取得画像をエキスパートデータ格納部14に格納する(ステップ6)。次いで、前記ステップ4と同様にして、エキスパートデータ格納部14の全画像を特徴量に変換する(ステップ7)。そして、エキスパートデータ格納部14に格納されている全画像の特徴量に基づき、標準偏差算出部16及び順位和算出部17によって、標準偏差(第2標準偏差)及び順位和(第2順位和)をそれぞれ算出する(ステップ8)。

【0033】

50

前記ステップ5又はステップ8の終了後、判定係数Vrank算出部18により、ステップ5及び8で算出された4つのパラメータ、すなわち、標準偏差、標準偏差、順位和及び順位和を用い、不良品画像取得の継続・終了決定部19により、下式(1)を用いて、判定係数Vrankを算出する。

【数1】

$$Vrank = 1 - \frac{1}{\left(\frac{\beta}{\alpha}\right)^2 \times \left(\frac{\theta}{\xi}\right)^2} \dots (1)$$

10

【0034】

上式(1)の標準偏差及びによる標準偏差比(/)、及び順位和及びによる順位和比(/)を用いて算出される判定係数Vrankは、標準偏差比及び順位和が大きくなると、値が1に近づく。この判定係数Vrankが後述する所定値VREF以上になったときに、教師データとして収集した非エキスパートデータ及びエキスパートデータが、データ全体としての平均付近に偏在せず、かつ適切に分散した状態であることを判定することが可能となる。

【0035】

上記ステップ10において、判定係数Vrankが所定値VREF以上であるか否かを判別する。この判別結果がNOで、判定係数Vrankが所定値VREFよりも小さいときは、エキスパートデータがまだ十分に揃っていないとして、ステップ1に戻り、不良品画像取得を継続する。

20

【0036】

一方、ステップ10の判別結果がYESで、Vrank > VREFのときには、エキスパートデータが十分揃っており、エキスパートによる不良品画像の収集を完了すべきであるとして、エキスパートデータ取得完了フラグF_COMPを「1」にセットし(ステップ11)、教師データ収集処理を終了する。なお、上記フラグF_COMPに「1」がセットされることにより、教師データ収集装置11において、図示しない表示部などにより、教師データの収集が完了したことが報知される。

【0037】

図4は、前述した検査システム1において、検査装置3の学習部7に搭載される学習モデルとしての分類モデルの生成処理を示している。本処理では、教師データ収集装置11のエキスパートデータ格納部14から全てのデータ(エキスパートデータ)を出力するとともに(ステップ21)、非エキスパートデータ格納部13から全て、あるいはエキスパートデータの数よりも多い所定数のデータ(非エキスパートデータ)を出力する(ステップ22)。次いで、出力されたエキスパートデータと非エキスパートデータを結合する(ステップ23)。これにより、分類モデルの生成に利用する多数の不良品教師データが作成される。

30

【0038】

そして、作成された多数の不良品教師データと、すでに収集されている多数の良品データ(良品教師データ)を用いて、分類モデルを学習する(ステップ24)。これにより、分類精度の高い分類モデルを得ることができ、検査システム1において、検査対象物Gの良否を精度良く判定することができる。

40

【0039】

以上詳述したように、本実施形態によれば、非エキスパートデータの特徴量に基づく標準偏差及び順位和、並びにエキスパートデータの特徴量に基づく標準偏差及び順位和を用い、エキスパートによって選定される不良品画像の取得の継続及び終了を決定することにより、エキスパートデータをできる限り少なくしながら、高い分類精度を有する学習モデルを生成するための教師データを収集することができる。

【0040】

50

なお、本発明は、説明した上記実施形態に限定されることなく、種々の態様で実施することができる。例えば、実施形態では、判定係数 $Vrank$ を用いて、エキスパートによる不良品画像の収集の継続・終了を決定したが、第1標準偏差及び第2標準偏差比の標準偏差比(/)、並びに第1順位和及び第2順位和の順位和比(/)を、それぞれ所定の基準値と比較することにより、エキスパートによる不良品画像の収集の継続・終了を決定するようにしてもよい。また、実施形態で示した教師データ収集装置11の細部の構成などは、あくまで例示であり、本発明の趣旨の範囲内で適宜、変更することができる。

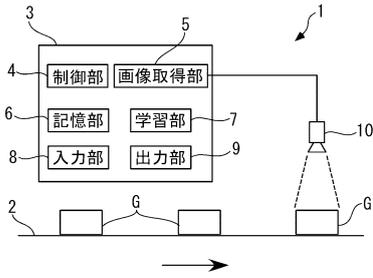
【符号の説明】

【0041】

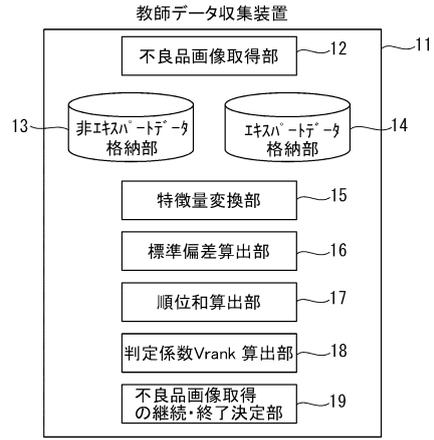
1	検査システム	
2	搬送機	
3	検査装置	
4	制御部	
5	画像取得部	
6	記憶部	
7	学習部	
8	入力部	
9	出力部	
10	カメラ	10
11	教師データ収集装置	
12	不良品画像取得部(非エキスパート不良品データ取得手段、エキスパート不良品データ取得手段)	
13	非エキスパートデータ格納部	
14	エキスパートデータ格納部	
15	特徴量変換部	
16	標準偏差算出部(第1標準偏差算出手段、第2標準偏差算出手段)	
17	順位和算出部(第1順位和算出手段、第2順位和算出手段)	
18	判定係数 $Vrank$ 算出部	
19	不良品画像取得の継続・終了決定部(決定手段)	30
G	検査対象物	

【図面】

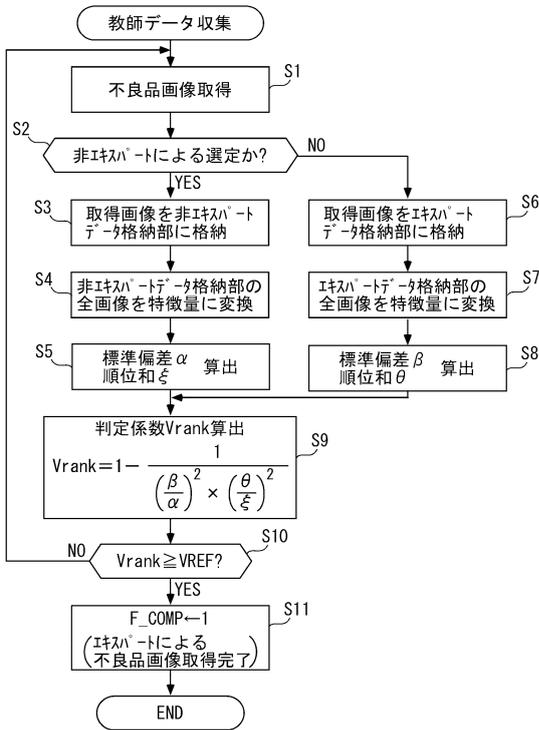
【図 1】



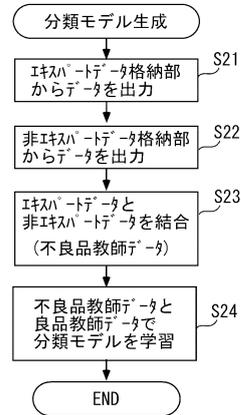
【図 2】



【図 3】



【図 4】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開2009-282686(JP,A)
特開2009-110064(JP,A)
特開2017-211689(JP,A)
特開2014-059754(JP,A)
国際公開第2021/181520(WO,A1)
国際公開第2020/065808(WO,A1)
特開2009-135649(JP,A)
- (58)調査した分野 (Int.Cl., DB名)
G06N 3/00-99/00