

(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
2020年1月30日(30.01.2020)



(10) 国際公開番号

WO 2020/021586 A1

(51) 国際特許分類:
G06F 11/36 (2006.01)

(21) 国際出願番号: PCT/JP2018/027420

(22) 国際出願日: 2018年7月23日(23.07.2018)

(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:三菱電機株式会社(MITSUBISHI ELECTRIC CORPORATION) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 Tokyo (JP).

(72) 発明者:千田 伸男(CHIDA Nobuo); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP). 川上 大介

(KAWAKAMI Daisuke); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP). 青木 利晃(AOKI Toshiaki); 〒9231292 石川県能美市旭台一丁目1番地 北陸先端科学技術大学院大学内 Ishikawa (JP).

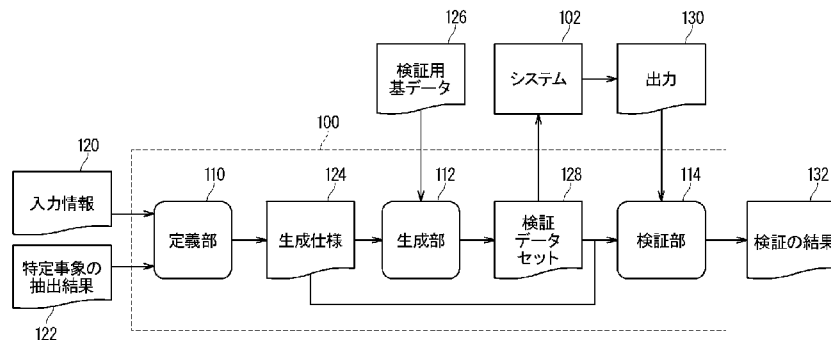
(74) 代理人:吉竹 英俊, 外(YOSHITAKE Hidetoshi et al.); 〒5400001 大阪府大阪市中央区域見1丁目4番70号住友生命OBPプラザビル10階 Osaka (JP).

(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, 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,

(54) Title: SYSTEM VERIFYING DEVICE AND SYSTEM VERIFYING METHOD

(54) 発明の名称: システム検証装置及びシステム検証方法

【図1】



102... SYSTEM
110... DEFINITION UNIT
112... GENERATION UNIT
114... VERIFICATION UNIT
120... INPUT INFORMATION
122... EXTRACTION RESULT OF SPECIFIC EVENT
124... GENERATION SPECIFICATION
126... BASE DATA FOR VERIFICATION
128... VERIFICATION DATA SET
130... OUTPUT
132... RESULT OF VERIFICATION

(57) Abstract: In the present invention, systematic verification is performed as to whether or not a system for which the basis of output is unclear complies with a specification. Generation specification is defined on the basis of the extraction result of a specific event that would cause an undesirable output to be generated by the system. A verification data set is generated on the basis of the generation specification. Verification as to whether or not the system complies with the specification is performed on the basis of the verification data set and the output generated by the system when the

HN, HR, HU, ID, IL, IN, IR, IS, 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, SM, ST, SV, SY, TH, 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, 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))

verification data set is inputted.

(57) 要約: 出力の根拠が不明確であるシステムが仕様を満たしているか否かについて系統的に検証を行う。システムにより望ましくない出力が生成される原因となる特定事象の抽出結果に基づいて生成仕様が定義される。生成仕様に基づいて検証データセットが生成される。検証データセット、及び検証データセットがシステムに入力された場合にシステムにより生成される出力に基づいてシステムが仕様を満たしているか否かについて検証が行われる。

明 細 書

発明の名称： システム検証装置及びシステム検証方法

技術分野

[0001] 本発明は、システムが仕様を満たしているか否かについて検証を行うシステム検証装置及びシステム検証方法に関する。

背景技術

[0002] 近年注目を集めている人工知能は、機能仕様を定義することが難しい問題に対して高い性能を有する。例えば、ニューラルネットワークを活用する手書き数字の認識アルゴリズムは、従来の手書き数字の認識アルゴリズムと比較して、高い認識率を有する。しかし、人工知能は、人工知能により生成される出力の根拠が不明確であるという特徴を有する。このため、人工知能を備えるシステムが満たすべき機能仕様を詳細に定義することは難しい。また、システムが満たすべき機能仕様を詳細に定義することが可能である場合は、そもそもシステムが人工機能を備える必要はない。

[0003] 特許文献1に記載された知能ロボットにおいては、仮想空間における部品の位置及び姿勢がランダムに変えられ、複数の部品が仮想空間に山積みされる（段落0072及び0074）。また、山積みされた部品を含む仮想空間の画像が生成される（段落0074）。生成された仮想空間の画像は、検証用の画像として使用される。

先行技術文献

特許文献

[0004] 特許文献1：特開2017-102755号公報

発明の概要

発明が解決しようとする課題

[0005] 人工知能は、先述したように、それにより生成される出力の根拠が不明確であるという特徴を有する。このため、人工知能を備えるシステムが仕様を満たしているか否かについての検証を、発生しうる多数の事象について系統

的に行うことは、困難である。特許文献 1 に記載された技術も、ランダムに生成された検証用の画像が使用されるため、この問題を解決することはできない。

[0006] また、この問題は、人工知能を備えるシステムに固有の問題ではなく、システムにより生成される出力の根拠が不明確であるという特徴を有するシステムに共通する問題である。

[0007] 本発明は、この問題に鑑みてなされる。本発明が解決しようとする課題は、出力の根拠が不明確であるシステムが仕様を満たしているか否かについて系統的に検証を行うことである。

課題を解決するための手段

[0008] 本発明は、システム検証装置に向けられる。

[0009] システム検証装置は、定義部、生成部及び検証部を備える。

[0010] 定義部は、システムにより望ましくない出力が生成される原因となる特定事象の抽出結果に基づいて生成仕様を定義する。

[0011] 生成部は、生成仕様に基づいて検証データセットを生成する。

[0012] 検証部は、検証データセット、及び検証データセットがシステムに入力された場合にシステムにより生成される出力に基づいてシステムが仕様を満たしているか否かについて検証を行う。

[0013] 本発明は、システム検証方法にも向けられる。

発明の効果

[0014] 本発明によれば、特定事象の抽出結果を反映する検証データセットを用いてシステムが仕様を満たしているか否かについて検証が行われる。このため、システムが仕様を満たしているか否かについて、特定事象の抽出結果を得るために行われた系統的な抽出を反映する系統的な検証が行われる。これにより、生成する出力の根拠が不明確であるシステムが仕様を満たしているか否かについて系統的に検証を行うことができる。

[0015] この発明の目的、特徴、局面及び利点は、以下の詳細な説明と添付図面とによって、より明白となる。

図面の簡単な説明

[0016] [図1]実施の形態 1 のシステム検証装置を図示するブロック図である。

[図2]実施の形態 1 及び 2 のシステム検証装置に入力される入力情報の例を図示する図である。

[図3]実施の形態 1 及び 2 のシステム検証装置に入力される特定事象の抽出結果を図示する図である。

[図4]通常のフォルトツリー解析 (F T A) により生成される木構造の例を図示する図である。

[図5]実施の形態 1 及び 2 のシステム検証装置に入力される特定事象の抽出結果に含まれる木構造の例を図示する図である。

[図6]実施の形態 1 及び 2 のシステム検証装置に入力される特定事象の抽出結果の例を図示する図である。

[図7]実施の形態 1 及び 2 のシステム検証装置に備えられる定義部の動作の流れを図示するフローチャートである。

[図8]実施の形態 1 及び 2 のシステム検証装置に備えられる定義部により定義される生成仕様の例を図示する図である。

[図9]実施の形態 1 及び 2 の検証システムにおいてベルヌーイ試行の区間推定が用いられる場合に仮定される認識率の分布を表すグラフを図示する図である。

[図10]実施の形態 1 及び 2 のシステム検証装置に入力される検証用基データの例を図示する図である。

[図11]実施の形態 1 及び 2 のシステム検証装置に備えられる生成部の動作の流れを図示するフローチャートである。

[図12]実施の形態 1 及び 2 のシステム検証装置に備えられる生成部の動作の流れを図示するフローチャートである。

[図13]実施の形態 1 及び 2 のシステム検証装置に備えられる生成部により生成される検証データセットの例を図示する図である。

[図14]実施の形態 1 及び 2 のシステム検証装置により検証されるシステムに

より生成される出力の例を図示する図である。

[図15]実施の形態1のシステム検証装置に備えられる検証部の動作の流れを図示するフローチャートである。

[図16]実施の形態1及び2のシステム検証装置に備えられる検証部により出力される検証の結果の例を図示する図である。

[図17]実施の形態1のシステム検証装置に入力される特定事象の抽出結果の別例を図示する図である。

[図18]実施の形態2のシステム検証装置を図示するブロック図である。

[図19]実施の形態2のシステム検証装置に備えられる検証部及び定義部の動作の流れを図示するフローチャートである。

[図20]実施の形態2のシステム検証装置に備えられる検証部及び定義部の動作の流れを図示するフローチャートである。

[図21]実施の形態2の検証システムにおいて算出される認識率の分布の例を表すグラフを図示する図である。

[図22]実施の形態2のシステム検証装置に備えられる定義部により定義される生成仕様の例を図示する図である。

[図23]実施の形態2のシステム検証装置に備えられる生成部により生成される検証データセットの例を図示する図である。

[図24]実施の形態2のシステム検証装置に備えられる検証部により出力される検証の結果の例を図示する図である。

発明を実施するための形態

[0017] 1 実施の形態1

1. 1 システム検証装置の概略

図1は、実施の形態1のシステム検証装置を図示するブロック図である。

[0018] 図1に図示されるシステム検証装置100は、システム102を検証する。システム検証装置100は、画像により表現される手書き数字を認識する画像認識システムを検証するのに適する。このため、以下では、システム102が当該画像認識システムである場合が説明される。しかし、システム1

０２は、当該画像認識システム以外の画像認識システムであってもよく、画像認識システム以外の認識システムであってもよく、認識システム以外のシステムであってもよい。例えば、システム１０２は、制御システム等であってもよい。システム１０２は、人工知能を備えるシステムである。しかし、システム１０２は、人工知能を備えないシステムであってもよい。システム１０２は、具体的には、人工知能を備える自動運転システム、ロボットシステム等である。

- [0019] システム検証装置１００は、定義部１１０、生成部１１２及び検証部１１４を備える。システム検証装置１００がこれらの要素以外の要素を備えてもよい。
- [0020] システム検証装置１００は、コンピュータを備える。コンピュータは、プロセッサ及びメモリを備える。プロセッサは、中央演算処理装置（ＣＰＵ）等である。メモリは、リードオンリーメモリ（ＲＯＭ）、ランダムアクセスメモリ（ＲＡＭ）等である。定義部１１０、生成部１１２及び検証部１１４は、メモリに記憶されたプログラムをプロセッサが実行することにより構成される。定義部１１０、生成部１１２及び検証部１１４の全部又は一部が、プログラムを実行しないハードウェアにより構成されてもよい。
- [0021] 定義部１１０は、入力された入力情報１２０、及び特定事象の抽出結果１２２に基づいて生成仕様１２４を定義する。また、定義部１１０は、定義した生成仕様１２４を出力する。
- [0022] 入力情報１２０は、システム１０２に入力される検証データセット１２８に含まれる複数の検証データの属性を含む。
- [0023] 特定事象は、システム１０２により望ましくない出力が生成される原因となる事象である。望ましくない出力は、システム１０２に期待される出力に反する出力である。
- [0024] 生成仕様１２４は、検証データセット１２８を生成するための仕様である。
- [0025] 生成部１１２は、入力された生成仕様１２４及び検証用基データ１２６に

基づいて検証データセット 128 を生成する。また、生成部 112 は、生成した検証データセット 128 を出力する。

[0026] 検証データセット 128 は、システム 102 が仕様を満たしているか否かについて検証を行うためのデータセットである。

[0027] システム 102 には、検証データセット 128 が入力される。システム 102 は、検証データセット 128 が入力された場合に出力 130 を生成する。検証データセット 128 は、複数の検証データを含む。出力 130 は、複数の応答を含む。複数の応答は、それぞれ複数の検証データに対する応答である。複数の検証データは、それぞれ複数の画像である。複数の応答は、それぞれ複数の認識結果である。

[0028] 検証部 114 は、入力された生成仕様 124、検証データセット 128 及び出力 130 に基づいてシステム 102 が仕様を満たしているか否かについて検証を行う。また、検証部 114 は、検証の結果 132 を出力する。

[0029] システム検証装置 100 によれば、特定事象の抽出結果 122 を反映する検証データセット 128 を用いてシステム 102 が仕様を満たしているか否かについて検証が行われる。このため、システム 102 が仕様を満たしているか否かについて、特定事象の抽出結果 122 を得るために行われた系統的な抽出を反映する系統的な検証が行われる。これにより、出力 130 の根拠が不明確であるシステム 102 が仕様を満たしているか否かについて系統的に検証を行うことができる。

[0030] 1. 2 入力情報の例

図 2 は、実施の形態 1 のシステム検証装置に入力される入力情報の例を図示する図である。

[0031] 図 2 に図示される入力情報 120 P は、図 1 に図示される入力情報 120 の例である。以下では、図 1 に図示される入力情報 120 が図 2 に図示される入力情報 120 P であるとして説明が進められる。

[0032] 入力情報 120 P は、画像サイズ 150 及びビット数 152 を含む。画像サイズ 150 及びビット数 152 は、それぞれ検証データセット 128 に含

まれる複数の検証データの画像サイズ及びビット数である。入力情報120 Pが、画像サイズ150及びビット数152以外の画像の属性を含んでもよく、画像以外のデータの属性を含んでもよい。

[0033] 検証データセット128に含まれる複数の検証データは、MNISTと呼ばれるデータセットに含まれる、複数の手書き数字をそれぞれ表現する複数の画像を加工することにより作成される。このため、画像サイズ150は、MNISTに含まれる複数の画像の画像サイズである「 28×28 (pixel)」に一致する。また、ビット数152は、MNISTに含まれる複数の画像のビット数である「8 (bit)」に一致する。

[0034] 1. 3 木構造

図3は、実施の形態1のシステム検証装置に入力される特定事象の抽出結果を図示する図である。

[0035] 特定事象の抽出結果122は、フォルトツリー解析(FTA)により生成される。このため、特定事象の抽出結果122は、図3に図示されるように、木構造200を含む。

[0036] 木構造200は、ノード210A、210B、210C、210D及び210Eを含む。

[0037] ノード210A、210B、210C、210D及び210Eは、特定事象A、特定基本事象B、特定事象C、特定基本事象D及び特定基本事象Eをそれぞれ表す。

[0038] 木構造200は、分岐のゲート220及び222をさらに含む。

[0039] 分岐のゲート220は、上位ノード210Aに下位ノード210B及び210Cを連結する。上位ノード210Aにより表される特定事象A、並びに下位ノード210B及び210Cによりそれぞれ表される特定基本事象B及び特定事象Cにおいては、特定事象Aが上位事象である。また、特定基本事象B及び特定事象Cが下位事象である。特定基本事象B及び特定事象Cは、特定事象Aが発生する原因となる。

[0040] 特定事象Aは、特定基本事象B及び特定事象Cの組み合わせにより表され

る。特定事象Aは、特定基本事象B及び特定事象Cのいずれかが発生した場合に発生する。このため、特定事象Aは、特定基本事象B及び特定事象Cの論理和により表される。したがって、分岐のゲート220は、ORゲートである。

[0041] 分岐のゲート222は、上位ノード210Cに下位ノード210D及び210Eを連結する。上位ノード210Cにより表される特定事象C、並びに下位ノード210D及び210Eによりそれぞれ表される特定基本事象D及びEにおいては、特定事象Cが上位事象である。また、特定基本事象D及びEが下位事象である。特定基本事象D及びEは、特定事象Cが発生する原因となる。

[0042] 特定事象Cは、特定基本事象D及びEの組み合わせにより表される。特定事象Cは、特定基本事象D及びEのすべてが発生した場合に発生する。このため、特定事象Cは、特定基本事象D及びEの論理積により表される。したがって、分岐のゲート222は、ANDゲートである。

[0043] ノード210A、210B、210C、210D及び210Eは、特定事象A及びCをそれぞれ表す情報となる内部ノード210A及び210Cを含む。

[0044] ノード210A、210B、210C、210D及び210Eは、特定事象Aが発生する原因となる特定基本事象B、D及びEをそれぞれ表す情報となる葉ノード210B、210D及び210Eを含む。葉ノード210B、210D及び210Eは、上位ノードになることはできない。

[0045] 木構造200の構造が図3に図示される構造から他の構造に変更される場合は、特定事象、特定基本事象、ノード、上位ノード、下位ノード、内部ノード、葉ノード、分岐のゲート、ORゲート及びANDゲートの数も変化する。

[0046] 1. 4 フォルトツリー解析

図4は、通常のFTAにより生成される木構造の例を図示する図である。

図5は、実施の形態1のシステム検証装置に入力される特定事象の抽出結果

に含まれる木構造の例を図示する図である。

[0047] 通常のF T Aにおいては、上位事象が発生する原因となる下位事象が専ら分析される。これに対して、木構造200を生成するために行われるF T Aにおいては、上位事象が発生する原因となる下位事象が分析され、上位事象及び下位事象に含まれる各事象に属するデータセットがさらに分析される。このため、木構造200を生成するために行われるF T Aにおいては、複数の事象にそれぞれ属する複数のデータセットの共通集合が空集合である場合は、複数の事象の論理積により表される事象に属するデータセットを作成することができず、当該事象を表すノードを作成することができない。

[0048] したがって、図4に図示される、通常のF T Aにより生成される木構造250においては、ANDゲート260が、「インクがにじむ」という上位事象を表す上位ノード270に、「文字が太くなる」及び「文字が薄くなる」という下位事象をそれぞれ表す下位ノード272及び274を連結することができる。これに対して、図5に図示される、実施の形態1のシステム検証装置100に入力される特定事象の抽出結果122に含まれる木構造300においては、ORゲート310が、「インクがにじむ」という上位事象をあらわす上位ノード320に、「文字が太くなるかつ薄くなる」という下位事象を表す下位ノード322を連結しなければならない。このような相違が生じるのは、「文字が太くなる」及び「文字が薄くなる」という2個の事象にそれぞれ属する2個のデータセットの共通集合が空集合であるため、「文字が太くなる」及び「文字が薄くなる」という2個の事象の論理積により表される事象に属するデータセットを作成することができず、当該事象を表すノードを作成することができないためである。

[0049] 1. 5 発生確率、誤認識率、基本誤認識率及び加工条件

特定事象の抽出結果122は、図3に図示されるように、発生確率 p_{f_A} 及び p_{f_C} を含む。

[0050] 発生確率 p_{f_A} 及び p_{f_C} は、それぞれ特定事象A及びCについて定義される。発生確率 p_{f_A} 及び p_{f_C} は、それぞれ特定事象A及びCが発生する確率

である。

[0051] 特定事象の抽出結果 1 2 2 は、誤認識率 p_{e_A} 及び p_{e_C} をさらに含む。

[0052] 誤認識率 p_{e_A} 及び p_{e_C} は、それぞれ特定事象 A 及び C について定義される。誤認識率 p_{e_A} 及び p_{e_C} は、それぞれ特定事象 A 及び C が発生した場合に特定事象 A 及び C が原因となってシステム 1 0 2 が手書き数字を誤認識する確率である。

[0053] 特定事象の抽出結果 1 2 2 は、発生確率 p_{f_B} 、 p_{f_D} 及び p_{f_E} をさらに含む。

[0054] 発生確率 p_{f_B} 、 p_{f_D} 及び p_{f_E} は、それぞれ特定基本事象 B、D 及び E について定義される。発生確率 p_{f_B} 、 p_{f_D} 及び p_{f_E} は、それぞれ特定基本事象 B、D 及び E が発生する確率である。

[0055] 特定事象の抽出結果 1 2 2 は、基本誤認識率 p_{b_B} 、 p_{b_D} 及び p_{b_E} をさらに含む。

[0056] 基本誤認識率 p_{b_B} 、 p_{b_D} 及び p_{b_E} は、それぞれ特定基本事象 B、D 及び E について定義される。基本誤認識率 p_{b_B} 、 p_{b_D} 及び p_{b_E} は、それぞれ特定基本事象 B、D 及び E が単独で発生した場合に特定基本事象 B、D 及び E が原因となってシステム 1 0 2 が手書き数字を誤認識する確率である。

[0057] 特定事象の抽出結果 1 2 2 は、加工条件 P_{C_B} 、 P_{C_D} 及び P_{C_E} をさらに含む。

[0058] 加工条件 P_{C_B} 、 P_{C_D} 及び P_{C_E} は、それぞれ特定基本事象 B、D 及び E について定義される。加工条件 P_{C_B} 、 P_{C_D} 及び P_{C_E} は、それぞれ特定基本事象 B、D 及び E に属するデータセットを生成するための加工条件であり、それぞれ特定基本事象 B、D 及び E について検証データセットを生成するために検証用基データ 1 2 6 を加工する際にしたがうべき加工条件である。

[0059] 1. 6 データセットの間の関係

全体のデータ U、並びに特定事象 A、特定基本事象 B 及び特定事象 C にそれぞれ属するデータセット D_A 、 D_B 及び D_C は、式 (1) 及び式 (2) により表される関係を有する。

[0060] [数1]

$$U \supset D_A = D_B \cup D_C \quad \dots(1)$$

$$D_B \cap D_C = \emptyset \quad \dots(2)$$

[0061] 式(1)は、データセット D_A が全体のデータ U に含まれることを示し、データセット D_A がデータセット D_B 及び D_C の和集合であることを示す。式(2)は、データセット D_B 及び D_C の積集合が空集合であることを示す。

[0062] 特定事象 C 、特定基本事象 D 及び特定基本事象 E にそれぞれ属するデータセット D_C 、 D_D 及び D_E は、式(3)により表される関係を有する。

[0063] [数2]

$$D_C = D_D \cap D_E \quad \dots(3)$$

[0064] 式(3)は、データセット D_C がデータセット D_D 及び D_E の積集合であることを示す。

[0065] 特定事象 A 、特定基本事象 B 、特定事象 C 、特定基本事象 D 及び特定基本事象 E にそれぞれ属する誤認識データセット E_A 、 E_B 、 E_C 、 E_D 及び E_E 、並びにデータセット D_A 、 D_B 、 D_C 、 D_D 及び D_E は、式(4)により表される関係を有する。ただし、 X は、 A 、 B 、 C 、 D 及び E の各々である。

[0066] [数3]

$$E_X \subset D_X \quad \dots(4)$$

[0067] 式(4)は、誤認識データセット E_X がデータセット D_X に含まれることを示す。誤認識データセット E_A 、 E_B 、 E_C 、 E_D 及び E_E に含まれる誤認識データにより表現される手書き数字は、システム102に誤認識される。

[0068] 誤認識データセット E_A 、 E_B 及び E_C は、式(5)により表される関係を有する。

[0069] [数4]

$$E_A = E_B \cup E_C \quad \dots(5)$$

[0070] 式(5)は、誤認識データセット E_A が誤認識データセット E_B 及び E_C の和集合であることを示す。

[0071] 誤認識データセット E_C 、 E_D 及び E_E は、式(6)により表される関係を有

する。

[0072] [数5]

$$E_C = E_D \cap E_E \quad \dots(6)$$

[0073] 式(6)は、誤認識データセット E_C が誤認識データセット E_D 及び E_E の和集合であることを示す。

[0074] 発生確率 pf_x 、誤認識率 pe_x 及び基本誤認識率 pb_x は、それぞれ式(7)、式(8)及び式(9)により定義される。

[0075] [数6]

$$pf_x = \frac{|D_x|}{|U|} \quad \dots(7)$$

$$pe_x = \frac{|E_x|}{|U|} \quad \dots(8)$$

$$pb_x = \frac{|E_x|}{|D_x|} \quad \dots(9)$$

[0076] 基本誤認識率 pb_x を求める際の母数 $|D_x|$ が発生確率 pf_x 及び誤認識率 pe_x を求める際の母数 $|U|$ と相違するのは、基本誤認識率 pb_x を求める際の母集団が発生確率 pf_x 及び誤認識率 pe_x を求める際の母集団と相違するためである。

[0077] 上位ノード 210A、並びにORゲート220により上位ノード210Aに連結される下位ノード210B及び210Cに着目した場合は、発生確率 pf_A 、 pf_B 及び pf_C 、誤認識率 pe_A 及び pe_C 、並びに基本誤認識率 pb_B は、式(10)及び式(11)により表される関係を有する。

[0078] [数7]

$$pf_A = pf_B + pf_C \quad \dots(10)$$

$$pe_A = pf_B \times pb_B + pe_C \quad \dots(11)$$

[0079] 上位ノード210C、並びにANDゲート222により上位ノード210Cに連結される下位ノード210D及び210Eに着目した場合は、発生確率 pf_C 、 pf_D 及び pf_E 、誤認識率 pe_C 、並びに基本誤認識率 pb_D 及び pb_E は、式(12)及び式(13)により表される関係を有する。

[0080] [数8]

$$pf_c = \frac{|D_c|}{|U|} = \frac{|D_D \cap D_E|}{|U|} \leq \min\left(\frac{|D_D|}{|U|}, \frac{|D_E|}{|U|}\right) = \min(pf_D, pf_E) \quad \dots(12)$$

$$pe_c \leq \min(pf_D \times pb_D, pf_E \times pb_E) \quad \dots(13)$$

[0081] 式(12)は、発生確率 pf_c が、発生確率 pf_D 及び pf_E のうちの最小値以下であることを示す不等式になっている。式(13)は、誤認識率 pe_c が、発生確率 pf_D と基本誤認識率 pb_D との積、及び発生確率 pf_E と基本誤認識率 pb_E との積のうちの最小値以下であることを示す不等式となっている。式(12)及び式(13)が不等式となっているのは、データセット D_D 及び D_E の積集合に含まれるデータの数、正確には不明であるものの、データセット D_D に含まれるデータの数、及びデータセット D_E に含まれるデータの数のうちの最小値以下であるためである。

[0082] 1. 7 特定事象の抽出結果の例

図6は、実施の形態1のシステム検証装置に入力される特定事象の抽出結果の例を図示する図である。

[0083] 図6に図示される特定事象の抽出結果122Pは、図1に図示される特定事象の抽出結果122の例である。以下では、図1に図示される特定事象の抽出結果122が図6に図示される特定事象の抽出結果122Pであるとして説明が進められる。

[0084] 特定事象の抽出結果122Pは、木構造350を含む。

[0085] 木構造350は、ノード360A、360B、360C、360D、360E、360F、360G、360H等のノードを含む。

[0086] ノード360Aは、「ノイズによって手書き数字を誤認識する」という特定事象Aを表す。ノード360Bは、「インク切れ」という特定事象Bを表す。ノード360Cは、「紙の配置を誤る」という特定事象Cを表す。ノード360Dは、「紙が汚れている」という特定事象Dを表す。ノード360Eは、「文字全体が薄くなる」という特定基本事象Eを表す。ノード360Fは、「文字全体が傾いている」という特定基本事象Fを表す。ノード36

OGは、「ドット足し」という特定基本事象Gを表す。ノード360Hは、「ドット欠け」という特定基本事象Hを表す。「ノイズによって手書き数字を誤認識する」、「インク切れ」、「紙の配置を誤る」、「紙が汚れている」、「文字全体が薄くなる」、「文字全体が傾いている」、「ドット足し」及び「ドット欠け」は、それぞれノード360A, 360B, 360C, 360D, 360E, 360F, 360G及び360Hのノード名となっている。

[0087] 特定事象の抽出結果122Pは、ORゲート370, 372, 374, 376等の分岐のゲートをさらに含む。

[0088] ORゲート370は、上位ノード360Aに、下位ノード360B, 360C及び360Dを連結する。ORゲート372は、上位ノード360Bに、下位ノード360Eを連結する。ORゲート374は、上位ノード360Cに、下位ノード360Fを連結する。ORゲート376は、上位ノード360Dに、下位ノード360G及び360Hを連結する。

[0089] 特定事象の抽出結果122Pは、発生確率 p_{fA} , p_{fB} , p_{fC} 及び p_{fD} をさらに含む。発生確率 p_{fA} , p_{fB} , p_{fC} 及び p_{fD} は、それぞれ特定事象A, B, C及びDについて定義される。

[0090] 特定事象の抽出結果122Pは、誤認識率 p_{eA} , p_{eB} , p_{eC} 及び p_{eD} をさらに含む。誤認識率 p_{eA} , p_{eB} , p_{eC} 及び p_{eD} は、それぞれ特定事象A, B, C及びDについて定義される。

[0091] 特定事象の抽出結果122Pは、発生確率 p_{fE} , p_{fF} , p_{fG} 及び p_{fH} をさらに含む。発生確率 p_{fE} , p_{fF} , p_{fG} 及び p_{fH} は、それぞれ特定基本事象E, F, G及びHについて定義される。

[0092] 特定事象の抽出結果122Pは、基本誤認識率 p_{bE} , p_{bF} , p_{bG} 及び p_{bH} をさらに含む。基本誤認識率 p_{bE} , p_{bF} , p_{bG} 及び p_{bH} は、それぞれ特定基本事象E, F, G及びHについて定義される。

[0093] 特定事象の抽出結果122Pは、加工条件 PC_E , PC_F , PC_G 及び PC_H をさらに含む。加工条件 PC_E , PC_F , PC_G 及び PC_H は、それぞれ特定基

本事象E, F, G及びHについて定義される。

[0094] 全体のデータU、上位事象である特定事象Aに属するデータセット D_A 、及び下位事象である特定事象B, C, D, ..., Xにそれぞれ属するデータセット $D_B, D_C, D_D, \dots, D_X$ は、式(14)及び式(15)により表される関係を有する。

[0095] [数9]

$$U \supset D_A = D_B \cup D_C \cup D_D \cup \dots \cup D_X \quad \dots(14)$$

$$D_B \cap D_C \cap D_D \cap \dots \cap D_X = \emptyset \quad \dots(15)$$

[0096] 式(14)は、データセット D_A が全体のデータUに含まれることを示し、データセット D_A がデータセット $D_B, D_C, D_D, \dots, D_X$ の和集合であることを示す。式(15)は、データセット $D_B, D_C, D_D, \dots, D_X$ の積集合が空集合であることを示す。

[0097] 上位事象である特定事象Aに属する誤認識データセット E_A 、並びに下位事象である特定事象B, C, D, ..., Xにそれぞれ属する誤認識データセット $E_B, E_C, E_D, \dots, E_X$ は、式(16)及び式(17)により表される関係を有する。

[0098] [数10]

$$E_A = E_B \cup E_C \cup E_D \cup \dots \cup E_X \quad \dots(16)$$

$$E_B \cap E_C \cap E_D \cap \dots \cap E_X = \emptyset \quad \dots(17)$$

[0099] 式(16)は、誤認識データセット E_A が誤認識データセット $E_B, E_C, E_D, \dots, E_X$ の和集合であることを示す。式(17)は、誤認識データセット $E_B, E_C, E_D, \dots, E_X$ の積集合が空集合であることを示す。誤認識データセット E_A は、全体のデータUに含まれる。

[0100] 特定事象の抽出結果122Pが生成される際には、複数の手書き数字をそれぞれ表す複数の画像の全体においてノイズを含む画像の発生確率が50%であると仮定される。また、手書き数字の認識に関する目標値が「手書き数字を95%以上正しく認識すること」であると定義される。これにより、図6に図示されるように、発生確率 p_{fA} が0.5に設定され、誤認識率 p_{eA}

が $0.05 * 0.50$ に設定される。

[0101] また、上位事象について定義された発生確率が少なくともひとつの下位事象についてそれぞれ定義された少なくともひとつの発生確率に等分に分配される。これにより、図6に図示されるように、発生確率 pf_B , pf_C , pf_D , $\dots$, pf_X が $0.5/X$ に設定され、誤認識率 pe_B , pe_C , pe_D , $\dots$, pe_X が $0.05 * 0.5/X$ に設定される。また、発生確率 pf_E が $0.5/X$ に設定され、基本誤認識率 pb_E が 0.05 に設定される。また、発生確率 pf_F が $0.5/X$ に設定され、基本誤認識率 pb_F が 0.05 に設定される。また、発生確率 pf_G 及び pf_H が $0.5/X/2$ に設定され、基本誤認識率 pb_F 及び pb_H が 0.05 に設定される。

[0102] 上位ノード360A、並びにORゲート370により上位ノード360Aに連結される下位ノード360B, 300C, 300D, $\dots$ に着目した場合は、発生確率 pf_A , pf_B , pf_C , pf_D , $\dots$, pf_X 、及び誤認識率 pe_A , pe_B , pe_C , pe_D , $\dots$, pe_X は、式(18)及び式(19)により表される関係を有する。

[0103] [数11]

$$pf_A = pf_B + pf_C + pf_D + \dots + pf_X = 0.50 \quad \dots(18)$$

$$pe_A = pe_B + pe_C + pe_D + \dots + pe_X = 0.05 \times 0.50 \quad \dots(19)$$

[0104] 上位ノード360B、及びORゲート372により上位ノード360Bに連結される下位ノード360Eに着目した場合は、発生確率 pf_B 及び pf_E 、誤認識率 pe_B 、並びに基本誤認識率 pb_E は、式(20)及び式(21)により表される関係を有する。

[0105] [数12]

$$pf_B = pf_E = 0.50 \div X \quad \dots(20)$$

$$pe_B = pb_E \times pf_E = 0.05 \times 0.5 \div X \quad \dots(21)$$

[0106] 上位ノード360C、及びORゲート374により上位ノード360Cに連結される下位ノード360Fに着目した場合は、発生確率 pf_C 及び pf_F 、誤認識率 pe_C 、並びに基本誤認識率 pb_F は、式(22)及び式(23)

により表される関係を有する。

[0107] [数13]

$$pf_C = pf_F = 0.50 \div X \quad \dots (22)$$

$$pe_C = pb_F \times pf_F = 0.05 \times 0.50 \div X \quad \dots (23)$$

[0108] 上位ノード360D、並びにORゲート376により上位ノード360Dに連結される下位ノード360G及び360Hに着目した場合は、発生確率 pf_D 、 pf_G 及び pf_H 、誤認識率 pe_D 、並びに基本誤認識率 pb_G 及び pb_H は、式(24)及び式(25)により表される関係を有する。

[0109] [数14]

$$pf_D = pf_G + pf_H = 0.50 \div X \quad \dots (24)$$

$$pe_D = pb_G \times pf_G + pb_H \times pf_H = 0.05 \times 0.50 \div X \quad \dots (25)$$

[0110] 加工条件 PC_E 、 PC_F 、 PC_G 及び PC_H は、それぞれ基本誤認識率 pb_E 、 pb_F 、 pb_G 及び pb_H に応じて設定される。加工条件 PC_E 、 PC_F 、 PC_G 及び PC_H は、それぞれ「文字領域のみ69%から50%まで1%ずつ薄くする」、「文字領域のみ時計回りに10°から49°まで1°ずつ回転させる」、「文字領域外に1×1から3×3までのノイズを1個から10個まで加える」及び「文字領域内に1×1から3×3までのノイズを1個から10個まで加える」という加工条件に設定される。

[0111] 1. 8 定義部の動作

図7は、実施の形態1のシステム検証装置に備えられる定義部の動作の流れを図示するフローチャートである。図8は、実施の形態1のシステム検証装置に備えられる定義部により定義される生成仕様の例を図示する図である。

[0112] 定義部110は、図7に図示されるステップS101において、入力された入力情報120Pを読み込む。

[0113] 定義部110は、続くステップS102において、入力された特定事象の抽出結果122Pを読み込む。

[0114] 定義部110は、続くステップS103からS109までのループにおい

て、特定基本事象E， F， G及びHのすべてについてそれぞれ図8に図示される生成仕様1 2 4 E， 1 2 4 F， 1 2 4 G及び1 2 4 Hを定義するまで、特定基本事象E， F， G及びHに含まれるひとつの特定基本事象Xについて生成仕様を定義することを繰り返す。生成仕様1 2 4 E， 1 2 4 F， 1 2 4 G及び1 2 4 Hの各々は、図1に図示される生成仕様1 2 4の例である。定義部1 1 0は、ループ内のステップS 1 0 4からS 1 0 8までにおいて、ひとつの特定基本事象Xについて生成仕様を定義する。

[0115] 定義部1 1 0は、続くステップS 1 1 0において、定義した生成仕様1 2 4 E， 1 2 4 F， 1 2 4 G及び1 2 4 Hを出力する。

[0116] 続いて、ステップS 1 0 4からS 1 0 8までにおいて行われる、特定基本事象Xについての生成仕様の定義が説明される。

[0117] 定義部1 1 0は、ステップS 1 0 4において、特定基本事象Xについて認識仕様を決定し、決定した認識仕様を特定基本事象Xについて定義する生成仕様を含める。定義部1 1 0は、ステップS 1 0 4を繰り返し実行することにより、図8に図示される認識仕様4 0 0 E， 4 0 0 F， 4 0 0 G及び4 0 0 Hを、それぞれ生成仕様1 2 4 E， 1 2 4 F， 1 2 4 G及び1 2 4 Hに含める。

[0118] 定義部1 1 0は、特定基本事象Xが発生した場合でも特定基本事象Xについて生成された出力が期待される出力であること、すなわち特定基本事象Xが発生した場合でも手書き数字を正しく認識することを、特定基本事象Xについて決定する認識仕様にする。これにより、「「文字全体が薄くなる」という特定基本事象が発生した場合でも手書き数字を正しく認識すること」が認識仕様4 0 0 Eにされ、「「文字全体が傾いている」という特定基本事象が発生した場合でも手書き数字を正しく認識すること」が認識仕様4 0 0 Fにされ、「「ドット足し」という特定基本事象が発生した場合でも手書き数字を正しく認識すること」が認識仕様4 0 0 Gにされ、「「ドット欠け」という特定基本事象が発生した場合でも手書き数字を正しく認識すること」が認識仕様4 4 0 Hにされる。

- [0119] 定義部 110 は、続くステップ S105 において、特定基本事象検 X について加工条件を決定し、決定した加工条件を特定基本事象検 X について定義する生成仕様に含める。定義部 110 は、ステップ S105 を繰り返し実行することにより、図 8 に図示される加工条件 402E, 402F, 402G 及び 402H を、それぞれ生成仕様 124E, 124F, 124G 及び 124H に含める。
- [0120] 定義部 110 は、特定事象の抽出結果 122P に含まれる、特定基本事象 X について定義された加工条件 PC_x を、特定基本事象 X について決定する加工条件にする。これにより、「文字領域のみ 69% から 50% まで 1% ずつ薄くする」という加工条件 PC_E が加工条件 402E にされ、「文字領域のみ時計回りに 10° から 49° まで 1° ずつ回転させる」という加工条件 PC_F が加工条件 402F にされ、「文字領域外に 1×1 から 3×3 までのノイズを 1 個から 10 個まで加える」という加工条件 PC_G が加工条件 402G にされ、「文字領域内に 1×1 から 3×3 までのノイズを 1 個から 10 個まで加える」という加工条件 PC_H が加工条件 402H にされる。
- [0121] 定義部 110 は、続くステップ S106 において、特定基本事象検 X について検証方法を決定し、決定した検証方法を特定基本事象検 X について定義する生成仕様に含める。定義部 110 は、ステップ S106 を繰り返し実行することにより、図 8 に図示される検証方法 404E, 404F, 404G 及び 404H を、それぞれ生成仕様 124E, 124F, 124G 及び 124H に含める。
- [0122] 定義部 110 は、特定基本事象 X について定義された加工条件 PC_x にしたがって検証用基データ 126 を加工することにより作成可能な検証データの数 N_x を決定する。作成可能データ数 N_x は、検証用基データ 126 に含まれる画像の数 N 、及び特定基本事象 X について定義された加工条件 PC_x にしたがって検証用基データ 126 に含まれるひとつの画像を加工することにより作成可能な作成可能データ数 M_x を用いて、式 (26) により決定される。

[0123] [数15]

$$N_X = N \times M_X \quad \cdots (26)$$

[0124] 以下では、作成可能データ数 N_E 、 N_F 、 N_G 及び N_H がそれぞれ200000、400000、1400000及び1500000であるとする。

[0125] また、定義部110は、求めた生成可能データ数 N_X が最大データ数 N_{MAX} 以下である場合は、全数検証を特定基本事象 X について決定する検証方法にする。また、定義部110は、求めた生成可能データ数 N_X が最大データ数 N_{MAX} より多い場合は、統計的検証を特定基本事象 X について決定する検証方法にする。全数検証が検証方法にされた場合は、作成可能な検証データの全部が検証データセットに含められる。統計的検証が検証方法にされた場合は、作成可能な検証データからランダムに選択された検証データが検証データセットに含められる。最大データ数 N_{MAX} は、例えばRAMの容量に基づいて決定される。

[0126] 以下では、最大データ数 N_{MAX} が700000であるとする。このため、検証方法404E、404F、404G及び404Hは、それぞれ全数検証、全数検証、統計的検証及び統計的検証である。

[0127] 定義部110は、続くステップS107において、特定基本事象 X について生成される検証データセットに含める検証データの数を示すデータ数 N'_X を決定し、決定したデータ数 N'_X を特定基本事象 X について定義する生成仕様に含める。定義部110は、ステップS107を繰り返し実行することにより、図8に図示されるデータ数406E、406F、406G及び406H（データ数 N'_E 、 N'_F 、 N'_G 及び N'_H ）を、それぞれ生成仕様124E、124F、124G及び124Hに含める。

[0128] 定義部110は、特定基本事象 X について決定した検証方法に応じて、特定基本事象 X について決定するデータ数 N'_X を変化させる。

[0129] 定義部110は、特定基本事象 X について決定した検証方法が全数検証である場合は、特定基本事象 X について決定した生成可能データ数 N_X を、特定基本事象 X について決定するデータ数 N'_X にする。このため、データ数40

6 E 及び 4 0 6 F (データ数 N'_E 及び N'_F) は、それぞれ 2 0 0 0 0 0 及び 4 0 0 0 0 0 である。

[0130] 一方、定義部 1 1 0 は、特定基本事象 X について決定した検証方法が統計的検証である場合は、特定基本事象 X について決定した生成可能データ数 N_X より少ないデータ数を、特定基本事象 X について決定するデータ数 N'_X にする。このため、データ数 4 0 6 G 及び 4 0 6 H (データ数 N'_G 及び N'_H) は、それぞれ 1 4 0 0 0 0 0 及び 1 5 0 0 0 0 0 より少ないデータ数である。

[0131] 特定基本事象 X について決定された検証方法が統計的検証である場合、すなわち作成可能データ数 N_X が最大データ数 N_{MAX} より多い場合は、決定されるデータ数 N'_X は、作成可能な検証データの全部を検証データセットに含めた場合に行われる検証の結果を区間推定により設定信頼度以上で推定できるデータ数である。

[0132] 特定基本事象 X について決定された検証方法が統計的検証である場合にデータ数 N'_X が決定されるときには、ベルヌーイ試行の区間推定が用いられる。

[0133] ベルヌーイ試行の区間推定が用いられる場合は、データ数 n は、チェルノフヘフディング境界 (Chernoff-Hoeffding bound) から、定義部 1 1 0 の内部パラメータ ϵ 及び δ を用いて、式 (27) により求められる。内部パラメータ ϵ は、誤差である。

[0134] [数16]

$$n \geq \frac{1}{2\epsilon^2} \ln \left(\frac{2}{\delta} \right) \quad \dots (27)$$

[0135] 図 9 は、実施の形態 1 の検証システムにおいてベルヌーイ試行の区間推定が用いられる場合に仮定される認識率の分布を表すグラフを図示する図である。

[0136] ベルヌーイ試行の区間推定が用いられる場合は、 N_X 個の検証データを含む母集団から p 個の検証データを含む標本を取り出し、取り出した標本におけ

る認識率を求めることが、 q 回繰り返される。これにより、 q 個の認識率が求められる。また、求められた q 個の認識率の分布が、図9に図示されるグラフ450により表される正規分布であると仮定される。この場合は、95%から100%までの認識率の区間が信頼区間 $1 - \delta$ であるとする、認識率は $(1 - \delta) \times 100\%$ の確率で信頼区間に含まれるといえる。このことを検証するためには、式(27)により求められるデータ数 n より多いデータ数の検証データにおいて認識率を求め、求めた認識率が97.5%以上であることを確かめればよい。以下では、 ε が0.025であるとされ、 δ が0.01であるとされる。このため、データ数406G及び406H（データ数 N'_G 及び N'_H ）は、4239以上である。以下では、データ数406G及び406H（データ数 N'_G 及び N'_H ）が4240であるとする。

[0137] このように統計的手法を用いてデータ数 N'_x を決定することにより、検証の結果132に対する信頼性を定量的に評価することができる。

[0138] 定義部110は、続くステップS108において、特定基本事象Xについて目標値を決定し、決定した目標値を特定基本事象Xについて定義する生成仕様に含める。定義部110は、ステップS108を繰り返し実行することにより、図8に図示される目標値408E、408F、408G及び408Hを、それぞれ生成仕様124E、124F、124G及び124Hに含める。

[0139] 定義部110は、特定基本事象Xについて生成された検証データセットがシステム102に入力された場合にシステム102が正しく認識をすることが求められる手書き数字を表現する検証データの数を、特定基本事象Xについて決定する目標値 NT_x にする。

[0140] 特定基本事象Xについて決定した検証方法が全数検証である場合は、目標値 NT_x は、基本誤認識率 pb_x 、及び作成可能データ数 N_x を用いて、式(28)により決定される。

[0141] [数17]

$$NT_x = (1 - pb_x) \times N_x \quad \cdots (28)$$

[0142] 特定基本事象E及びFについて決定された検証方法は、全数検証である。

また、基本誤認識率 pb_E 及び pb_F は、0.05である。また、作成可能データ数 N_E 及び N_F は、それぞれ200000及び400000である。このため、目標値 408_E 及び 408_F （目標値 NT_E 及び NT_F ）は、それぞれ190000及び380000である。

[0143] 特定基本事象Xについて決定された検証方法が統計的検証である場合は、目標値 NT_X は、基本誤認識率 pb_X 、誤差 ϵ 及びデータ数 N'_X を用いて、式(29)により求められる。

[0144] [数18]

$$NT_X = ((1 - pb_X) + \epsilon) \times N'_X \quad \cdots (29)$$

[0145] 特定基本事象G及びHについて決定された検証方法は、統計的検証である。また、基本誤認識率 pb_g 及び pb_h は、0.05である。また、誤差 ϵ は、0.025である。また、データ数 N'_G 及び N'_H は、それぞれ4240である。このため、目標値 408_G 及び 408_H （目標値 NT_G 及び NT_H ）は、4134である。

[0146] 1.9 検証用基データ

図10は、実施の形態1のシステム検証装置に入力される検証用基データの例を図示する図である。

[0147] 図10に図示される検証用基データ126Pは、図1に図示される検証用基データ126の例である。以下では、図1に図示される検証用基データ126が図10に図示される検証用基データ126Pであるとして説明が進められる。

[0148] 検証用基データ126Pは、複数の画像500を含む。複数の画像500は、MNISTに含まれる複数の画像である。

[0149] MNISTは、60000枚の学習用の画像、及び10000枚のテスト用の画像を含む。複数の画像500は、MNISTに含まれる10000枚のテスト用の画像である。複数の画像500が、60000枚の学習用の画像のうちの、システム102の学習フェーズにおいて使用されなかった学習

用の画像であってもよい。

[0150] 検証用基データ 1 2 6 P は、複数のデータ番号 5 0 2 をさらに含む。複数のデータ番号 5 0 2 は、それぞれ複数の画像 5 0 0 に対応づけられる。

[0151] 検証用基データ 1 2 6 P は、複数の教師ラベル 5 0 4 をさらに含む。複数の教師ラベル 5 0 4 は、それぞれ複数の画像 5 0 0 に対応づけられる。

[0152] 1. 1 0 生成部の動作

図 1 1 及び図 1 2 は、実施の形態 1 のシステム検証装置に備えられる生成部の動作の流れを図示するフローチャートである。図 1 3 は、実施の形態 1 のシステム検証装置に備えられる生成部により生成される検証データセットの例を図示する図である。

[0153] 生成部 1 1 2 は、図 1 1 に図示されるステップ S 1 2 1 において、定義された生成仕様 1 2 4 E, 1 2 4 F, 1 2 4 G 及び 1 2 4 H を読み込む。

[0154] 生成部 1 1 2 は、続くステップ S 1 2 2 から S 1 3 4 までのループにおいて、特定基本事象 E, F, G 及び H のすべてについてそれぞれ図 1 3 に図示される検証データセット 1 2 8 E, 1 2 8 F, 1 2 8 G 及び 1 2 8 H を生成するまで、特定基本事象 E, F, G 及び H に含まれるひとつの特定基本事象 X について検証データセットを生成することを繰り返す。検証データセット 1 2 8 E, 1 2 8 F, 1 2 8 G 及び 1 2 8 H の各々は、図 1 に図示される検証データセット 1 2 8 の例である。生成部 1 1 2 は、ループ内のステップ S 1 2 3 から S 1 3 3 までにおいて、ひとつの特定基本事象 X について検証データセットを生成する。

[0155] 生成部 1 1 2 は、続くステップ S 1 3 5 において、生成した検証データセット 1 2 8 E, 1 2 8 F, 1 2 8 G 及び 1 2 8 H を出力する。

[0156] 続いて、ステップ S 1 2 3 から S 1 3 3 までにおいて行われる、特定基本事象 X についての検証データセットの生成が説明される。

[0157] 生成部 1 1 2 は、ステップ S 1 2 3 において、特定基本事象 X について定義された生成仕様に含まれる検証方法が、全数検証であるか否かを判定する。

- [0158] 生成部 112 は、検証方法が全数検証であると判定した場合は、ステップ S124 から S128 までのループにおいて、作成した検証データの数 N'_x （作成可能データ数 N_x ）に達するまで、当該生成仕様に含まれる加工条件に含まれるひとつの加工方法にしたがって検証用基データ 126P に含まれるひとつの画像を加工することを繰り返す。生成部 112 は、ループ内のステップ S125 から S127 までにおいて、当該生成仕様に含まれる加工条件に含まれるひとつの加工方法にしたがって検証用基データ 126P に含まれるひとつの画像を加工する。生成部 112 は、ステップ S125 から S127 までを繰り返し実行することにより、当該生成仕様に含まれる加工条件にしたがって検証用基データ 126P を加工することにより特定基本事象 X について N'_x 個の検証データを含む検証データセットを生成する。
- [0159] 生成部 112 は、ステップ S125 において、検証用基データ 126P に含まれる複数の画像 500 からひとつの画像を順番に選択する。ひとつの画像の選択は、画像に対応づけられたデータ番号の順番で行われる。
- [0160] 生成部 112 は、続くステップ S126 において、特定基本事象 X について定義された生成仕様に含まれる加工条件に含まれる複数の加工方法からひとつの加工方法を順番に選択する。
- [0161] 生成部 112 は、続くステップ S127 において、選択したひとつの加工方法にしたがって選択したひとつの画像を加工する。これにより、生成部 112 は、加工された画像である検証データを作成し、作成した検証データを特定基本事象 X について生成する検証データセットに含める。
- [0162] 検証方法 404E は、全数検証である。また、加工条件 402E は、「文字領域のみ 69% から 50% まで 1% ずつ薄くする」という加工条件である。このため、生成部 112 は、まず、文字領域のみ 69% 薄くするという加工方法にしたがって、検証用基データ 126P に含まれる、データ番号「No. 1」が対応づけられた画像を加工する。また、生成部 112 は、データ番号「No. 1」、及び加工された画像である検証データを検証データセット 128E

に含める。生成部 112 は、続いて、文字領域のみ 68% 薄くするという加工方法にしたがって、検証用基データ 126 P に含まれる、データ番号「No. 1」が対応づけられた画像を加工する。また、生成部 112 は、データ番号「No. 2」、及び加工された画像である検証データを検証データセット 128 E に含める。生成部 112 は、作成した検証データの数データ数 406 E（データ数 N'_E すなわち作成可能データ数 N_E ）に達するまで、同様の処理を繰り返す。

[0163] 検証方法 404 F は、全数検証である。また、加工条件 402 F は、「文字領域のみ時計回りに 10° から 49° まで 1° ずつ回転させる」という加工条件である。このため、生成部 112 は、まず、文字領域のみ時計回りに 10° 回転させるという加工方法にしたがって、検証用基データ 126 P に含まれる、データ番号「No. 1」が対応づけられた画像を加工する。また、生成部 112 は、データ番号「No. 1」、及び加工された画像である検証データを検証データセット 128 F に含める。生成部 112 は、続いて、文字領域のみ時計回りに 11° 回転させるという加工方法にしたがって、検証用基データ 126 P に含まれる、データ番号「No. 1」が対応づけられた画像を加工する。また、生成部 112 は、データ番号「No. 2」、及び加工された画像である検証データを検証データセット 128 F に含める。生成部 112 は、作成した検証データの数データ数 406 F（データ数 N'_F すなわち作成可能データ数 N_F ）に達するまで、同様の処理を繰り返す。

[0164] 一方、生成部 112 は、検証方法が全数検証でないと判定した場合は、ステップ S129 から S133 までのループにおいて、作成した検証データの数データ数 N'_X が特定基本事象 X について定義された生成仕様に含まれるデータ数 N'_X に達するまで、当該生成仕様に含まれる加工条件に含まれるひとつの加工方法にしたがって検証用基データ 126 P に含まれるひとつの画像を加工することを繰り返す。生成部 112 は、ループ内のステップ S130 から S132 までにおいて、当該生成仕様に含まれる加工条件に含まれるひとつの加工方法にしたがって検証用基データ 126 P に含まれるひとつの画像を加工する

。生成部 112 は、ステップ S130 から S132 までを繰り返し実行することにより、当該生成仕様に含まれる加工条件にしたがって検証用基データ 126P を加工することにより特定基本事象 X について N'_x 個の加工された画像を含む検証データセットを生成する。

[0165] 生成部 112 は、ステップ S130 において、検証用基データ 126P に含まれる複数の画像 500 からひとつの画像をランダムに選択する。

[0166] 生成部 112 は、続くステップ S131 において、特定基本事象 X について定義された生成仕様に含まれる加工条件に含まれる複数の加工方法からひとつの加工方法をランダムに選択する。

[0167] 生成部 112 は、続くステップ S132 において、選択したひとつの加工方法にしたがって選択したひとつの画像を加工する。

[0168] 生成部 112 は、ステップ S130 から S132 までを実行することにより、作成可能データ数 N_x が最大データ数 N_{MAX} より多い場合に、作成可能な検証データからランダムに選択された検証データを作成し、作成した検証データを検証データセットに含める。これにより、検証データセットに含まれる複数の検証データの偏りを抑制することができる。

[0169] 検証方法 404G は、統計的検証である。また、加工条件 402G は、「文字領域外に 1×1 から 3×3 までのノイズを 1 個から 10 個まで加える」という加工条件である。加工条件 402G は、複数の加工方法を含む。複数の加工方法の各々は、ノイズの大きさとノイズの数との組を含む。このため、生成部 112 は、まず、加工条件 402G に含まれる複数の加工方法からランダムに選択した第 1 の加工方法にしたがって検証用基データ 126P に含まれる複数の画像 500 からランダムに選択した、データ番号「No. 5897」が対応づけられた画像を加工する。また、生成部 112 は、データ番号「No. 1」、及び加工された画像である検証データを検証データセット 128G に含める。生成部 112 は、続いて、当該加工条件 402G に含まれる複数の加工方法からランダムに選択した第 2 の加工方法にしたがって検証用基データ 126P に含まれる複数の画像 500 からランダムに選択した、データ番号

「No.256」が対応づけられた画像を加工する。また、生成部112は、データ番号「No.2」、及び加工された画像である検証データを検証データセット128Gに含める。生成部112は、作成した検証データの数データ数406G（データ数 N'_G ）に達するまで、同様の処理を繰り返す。

[0170] 検証方法404Hは、統計的検証である。また、加工条件402Hは、「文字領域内に1×1から3×3までのノイズを1個から10個まで加える」という加工条件である。加工条件402Hは、複数の加工方法を含む。複数の加工方法の各々は、ノイズの大きさとノイズの数との組を含む。このため、生成部112は、まず、加工条件402Hに含まれる複数の加工方法からランダムに選択した第1の加工方法にしたがって検証用基データ126Pに含まれる複数の画像500からランダムに選択した、データ番号「No.1854」が対応づけられた画像を加工する。また、生成部112は、データ番号「No.1」、及び加工された画像である検証データを検証データセット128Hに含める。生成部112は、続いて、当該加工条件402Hに含まれる複数の加工方法からランダムに選択した第2の加工方法にしたがって検証用基データ126Pに含まれる複数の画像500からランダムに選択した、データ番号「No.9543」が対応づけられた画像を加工する。また、生成部112は、データ番号「No.2」、及び加工された画像である検証データを検証データセット128Hに含める。生成部112は、作成した検証データの数データ数406H（データ数 N'_H ）に達するまで、同様の処理を繰り返す。

[0171] 1. 11 システム

図14は、実施の形態1のシステム検証装置により検証されるシステムにより生成される出力の例を図示する図である。

[0172] 図14に図示される出力130Pは、図1に図示される出力130の例である。以下では、図1に図示される出力130が図14に図示される出力130Pであるとして説明が進められる。

[0173] 出力130Pは、複数のデータ番号600を含む。

[0174] 出力130Pは、複数の認識結果602をさらに含む。複数の認識結果6

02は、それぞれ複数のデータ番号600に対応づけられる。複数の認識結果602に含まれる各認識結果は、各認識結果に対応づけられたデータ番号と同じデータ番号に対応づけられた画像により表現される手書き数字の認識結果を表す。

[0175] 複数の認識結果602に含まれる各認識結果は、10個の確率610を含む。10個の確率610は、それぞれ「0」から「9」までの数字に対応づけられる。10個の確率610に含まれる各確率は、各認識結果に対応づけられたデータ番号と同じデータ番号に対応づけられた画像により表現される手書き数字が、各確率に対応づけられた数字である確率を表す。

[0176] 図14によれば、例えば、データ番号「No.1」に対応づけられた画像により表現される手書き数字が「0」である確率は、100%であり、当該手書き数字が「1」から「9」までの数字である確率は、0.00%である。また、データ番号「No.2」に対応づけられた画像により表現される手書き数字が「3」である確率は、98.98%であり、当該手書き数字が「0」である確率は、0.01%であり、当該手書き数字が「1」である確率は、0.03%であり、当該手書き数字が「2」である確率は、0.01%であり、当該手書き数字が「4」である確率は、0.00%であり、当該手書き数字が「5」である確率は、0.05%であり、当該手書き数字が「6」である確率は、0.00%であり、当該手書き数字が「7」である確率は、0.00%であり、当該手書き数字が「8」である確率は、0.51%であり、当該手書き数字が「9」である確率は、0.42%である。

[0177] 1.12 検証部の動作

図15は、実施の形態1のシステム検証装置に備えられる検証部の動作の流れを図示するフローチャートである。図16は、実施の形態1のシステム検証装置に備えられる検証部により出力される検証の結果の例を図示する図である。

[0178] 検証部114は、図15に図示されるステップS141からS146までのループにおいて、特定基本事象E、F、G及びHのすべてについてシステ

ム 1 0 2 が仕様を満たしているか否かについて検証を行うまで、特定基本事象 E, F, G 及び H に含まれるひとつの特定基本事象 X についてシステム 1 0 2 が仕様を満たしているか否かについて検証を行うことを繰り返す。検証部 1 1 4 は、ループ内のステップ S 1 4 2 から S 1 4 5 までにおいて、ひとつの特定基本事象 X についてシステム 1 0 2 が仕様を満たしているか否かについて検証を行う。

[0179] 検証部 1 1 4 は、続くステップ S 1 4 7 において、特定事象 E, F, G 及び H について行った検証の結果をそれぞれ示す検証の結果 1 3 2 E, 1 3 2 F, 1 3 2 G 及び 1 3 2 H を出力する。検証の結果 1 3 2 E, 1 3 2 F, 1 3 2 G 及び 1 3 2 H の各々は、図 1 に図示される検証の結果 1 3 2 の例である。

[0180] 続いて、ステップ S 1 4 2 から S 1 4 5 までにおいて行われる、特定基本事象 X についての、システム 1 0 2 が仕様を満たしているか否かについての検証が説明される。

[0181] 検証部 1 1 4 は、ステップ S 1 4 2 において、特定基本事象 X について定義された生成仕様を読み込む。

[0182] 検証部 1 1 4 は、続くステップ S 1 4 3 において、特定基本事象 X について生成された検証データセットがシステム 1 0 2 に入力された際にシステム 1 0 2 により生成された出力 1 3 0 P を読み込む。

[0183] 検証部 1 1 4 は、続くステップ S 1 4 4 において、特定基本事象 X について生成された検証データセットを読み込む。

[0184] 検証部 1 1 4 は、続くステップ S 1 4 5 において、特定基本事象 X についてシステム 1 0 2 が仕様を満たしているか否かについて検証を行う。

[0185] 検証部 1 1 4 は、読み込んだ出力 1 3 0 P が読み込んだ生成仕様に含まれる目標値 $N T_x$ を達成している場合は、システム 1 0 2 が仕様を満たしていると判定する。また、検証部 1 1 4 は、システム 1 0 2 が仕様を満たしていると判定した場合は、特定基本事象 X について出力する検証の結果に、システム 1 0 2 が仕様を満たしていることを示す情報「○」を含める。一方、検証

部 1 1 4 は、読み込んだ出力 1 3 0 P が読み込んだ生成仕様に含まれる目標値 NT_x を達成していない場合は、システム 1 0 2 が仕様を満たしていないと判定する。また、検証部 1 1 4 は、システム 1 0 2 が仕様を満たしていないと判定した場合は、特定基本事象 X について出力する検証の結果に、システム 1 0 2 が仕様を満たしていることを示す情報「×」を含める。

[0186] 検証部 1 1 4 は、読み込んだ出力 1 3 0 P に含まれる複数の認識結果 6 0 2 がそれぞれ複数の教師ラベルと一致するか否かを判定する。検証部 1 1 4 は、認識結果に含まれる最大の確率に対応づけられた数字が、認識結果が対応づけられたデータ番号と同じデータ番号に対応づけられた教師ラベルと一致する場合は、認識結果が教師ラベルと一致すると判定し、当該データ番号に対応づけられた画像により表現される手書き数字が正しく認識されたと判定する。一方、検証部 1 1 4 は、当該数字が当該教師ラベルと一致しない場合は、当該認識結果が当該教師ラベルと一致しないと判定し、当該手書き数字が誤認識されたと判定する。また、検証部 1 1 4 は、教師ラベルと一致する認識結果の数を示す認識数を算出する。また、検証部 1 1 4 は、算出した認識数を出力する。検証部 1 1 4 は、ステップ S 1 4 5 を繰り返し実行することにより、特定基本事象 E, F, G 及び H についてそれぞれ認識数 6 5 2 E, 6 5 2 F, 6 5 2 G 及び 6 5 2 H を算出し、算出した認識数 6 5 2 E, 6 5 2 F, 6 5 2 G 及び 6 5 2 H を出力する。また、検証部 1 1 4 は、算出した認識数が目標値 NT_x 以上である場合は、システム 1 0 2 が仕様を満たしていると判定する。一方、検証部 1 1 4 は、算出した認識数が目標値 NT_x より少ない場合は、システム 1 0 2 が仕様を満たしていないと判定する。

[0187] 以下では、複数の認識数 6 5 2 E, 6 5 2 F, 6 5 2 G 及び 6 5 2 H が、それぞれ 1 9 5 2 4 1, 3 9 5 6 7 5, 4 1 8 6 及び 4 0 6 0 であるとする。

[0188] 認識数 6 5 2 E, 6 5 2 F 及び 6 5 2 G は、それぞれ 1 9 5 2 4 1, 3 9 5 6 7 5 及び 4 1 8 6 である。また、目標値 4 0 8 E, 4 0 8 F 及び 4 0 8 G は、それぞれ 1 9 0 0 0 0, 3 8 0 0 0 0 及び 4 1 3 4 である。このため

、特定基本事象E、F及びGについて生成された出力は、それぞれ目標値408E、408F及び408Gを達成している。したがって、検証部114は、システム102が仕様を満たしていることを示す情報「○」を検証の結果132E、132F及び132Gに含める。一方、認識数652Hは、4060である。また、目標値408Hは、4134である。このため、特定基本事象Hについて生成された出力は、目標値408Hを達成していない。したがって、検証部114は、システム102が仕様を満たしていないことを示す情報「×」を検証の結果132Hに含める。

[0189] 1. 13 特定事象の抽出結果の別例

図17は、実施の形態1のシステム検証装置に入力される特定事象の抽出結果の別例を図示する図である。

[0190] 図17に図示される特定事象の抽出結果122Qは、図1に図示されるシステム102が人工知能を備える自動運転システムである場合の図1に図示される特定事象の抽出結果122の例となっている。

[0191] 特定事象の抽出結果122Qは、「カメラが誤認識する」という特定事象Aを表すノード700A、「雪」という特定事象Bを表すノード700B、「雨」という特定事象Cを表すノード700C、「西日（逆行）」という特定事象Dを表すノード700D、「画像の一部が白飛びする」という特定基本事象Eを表すノード700E等の自動運転システムを検証するために必要な事象を表すノードを含む点で、図6に図示される特定事象の抽出結果122Pと相違する。

[0192] この他、特定事象の抽出結果122Qは、分岐のノード710、712、714及び716を含む。

[0193] 2 実施の形態2

図18は、実施の形態2のシステム検証装置を図示するブロック図である。

[0194] 図2は、実施の形態2のシステム検証装置に入力される入力情報の例を図示する図でもある。図3は、実施の形態2のシステム検証装置に入力される

特定事象の抽出結果を図示する図でもある。図5は、実施の形態2のシステム検証装置に入力される特定事象の抽出結果に含まれる木構造の例を図示する図でもある。図6は、実施の形態2のシステム検証装置に入力される特定事象の抽出結果の例を図示する図でもある。図7は、実施の形態2のシステム検証装置に備えられる定義部の動作の流れを図示するフローチャートでもある。図8は、実施の形態2のシステム検証装置に備えられる定義部により定義される生成仕様の例を図示する図でもある。図9は、実施の形態2の検証システムにおいてベルヌーイ試行の区間推定が用いられる場合に仮定される認識率の分布を表すグラフを図示する図でもある。図10は、実施の形態2のシステム検証装置に入力される検証用基データの例を図示する図でもある。図11及び図12は、実施の形態2のシステム検証装置に備えられる生成部の動作の流れを図示するフローチャートでもある。図13は、実施の形態2のシステム検証装置に備えられる生成部により生成される検証データセットの例を図示する図でもある。図14は、実施の形態2のシステム検証装置により検証されるシステムにより生成される出力の例を図示する図でもある。図16は、実施の形態2のシステム検証装置に備えられる検証部により出力される検証の結果を図示する図でもある。

[0195] 図18に図示される実施の形態2のシステム検証装置750は、特定基本事象Xについて決定された検証方法が統計的検証であり、特定基本事象Xについて行われた検証においてシステム102が仕様を満たしていないと判定された場合に、データ数 N'_x が増加させられる点で、実施の形態1のシステム検証装置100と相違する。

[0196] 以下では、上記の相違に関連して実施の形態2のシステム検証装置750において採用された構成が説明される。説明されない点については、実施の形態1のシステム検証装置100において採用された構成が、実施の形態2のシステム検証装置750においても採用される。

[0197] システム検証装置750に備えられる検証部114は、特定基本事象Xについて定義された生成仕様に含まれる検証方法が統計的検証であり、特定基

本事象Xについて行った検証においてシステム102が仕様を満たしていないと判定した場合に、特定基本事象Xについて行った検証においてシステム102が仕様を満たしていないことを示す情報を定義部110に入力する。

[0198] システム検証装置750に備えられる定義部110は、特定基本事象Xについて行った検証においてシステム102が仕様を満たしていないことを示す情報が入力された場合に、特定基本事象Xについて定義する生成仕様に含まれるデータ数 N'_x を増加させる。

[0199] 図19及び図20は、実施の形態2のシステム検証装置に備えられる検証部及び定義部の動作の流れを図示するフローチャートである。

[0200] 検証部114は、図19に図示されるステップS161からS171までのループにおいて、特定基本事象E、F、G及びHのすべてについて更新要求のための処理を行うまで、特定基本事象E、F、G及びHに含まれるひとつの特定基本事象Xについて更新要求のための処理を行うことを繰り返す。検証部114は、ループ内のステップS162からS170までにおいて、ひとつの特定基本事象Xについて更新要求のための処理を行う。

[0201] 検証部114は、ステップS162からS165までにおいて、それぞれ図15に図示されるステップS142からS146までにおいて行われる処理と同様の処理を行う。

[0202] 検証部114は、続くステップS166において、特定基本事象Xについて定義された生成仕様に含まれる検証方法が統計的手法であるか否かを判定する。検証部114は、検証方法が統計的手法であると判定した場合は、ステップS167においてさらに処理を行う。検証部114は、検証方法が全数検証であると判定した場合は、さらに処理を行うことなく更新要求のための処理を終了する。

[0203] 検証部114は、続くステップS167において、特定基本事象Xについて算出された認識数が特定基本事象Xについて定義された生成仕様に含まれる目標値 N_{Tx} より少ないか否かを判定する。これにより、検証部114は、特定基本事象Xについて行われた検証においてシステム102が仕様を満た

しているか否かを判定する。検証部 114 は、認識数が目標値 NT_x より少ない場合は、ステップ S 168 においてさらに処理を行う。検証部 114 は、認識数が目標値 NT_x 以上である場合は、さらに処理を行うことなく更新要求のための処理を終了する。

[0204] 検証部 114 は、ステップ S 168 において、特定基本事象 X について算出された認識数が、特定基本事象 X について算出された許容認識数 NP_x 以上であるか否かを判定する。許容認識数は、目標値 NT_x より少ない数であり、式 (30) により表される。以下では、 ζ が 0.02 であるとする。

[0205] [数19]

$$NP_x = (1 - pb_x - \zeta) \times N'_x \quad \dots (30)$$

[0206] これにより、検証部 114 は、特定基本事象 X について生成された出力が、目標値 NT_x により定義される仕様より緩やかな、許容認識数 NP_x により定義される許容条件を満たすか否かを判定する。

[0207] 検証部 114 は、認識数が許容認識数 NP_x 以上である場合は、ステップ S 169 においてさらに処理を行う。検証部 114 は、認識数が許容認識数 NP_x より小さい場合は、さらに処理を行うことなく更新要求のための処理を終了する。

[0208] 検証部 114 は、ステップ S 169 において、区間推定において考慮される誤差 ε から設定値 η を減じることにより得られる更新後の誤差 ε が 0 より大きいかな否かを判定する。更新後の誤差 ε は、式 (31) により算出される。以下では、 η が 0.005 であるとする。

[0209] [数20]

$$\varepsilon = \varepsilon - \eta \quad \dots (31)$$

[0210] 検証部 114 は、更新後の誤差が 0 より大きい場合は、ステップ S 170 において誤差 ε を更新することを定義部 110 に要求してから更新要求のための処理を終了する。一方、検証部 114 は、更新後の誤差が 0 以下である場合は、さらに処理を行うことなく更新要求のための処理を終了する。

[0211] したがって、更新要求のための処理においては、下記の条件 1 から 4 まで

が満たされる場合に、誤差 ε を更新することが要求される。

[0212] 条件 1 : 検証方法が統計的検証である。

[0213] 条件 2 : 認識数が目標値 $N T_x$ より少ない。

[0214] 条件 3 : 認識数が許容認識数 $N P_x$ 以上である。

[0215] 条件 4 : 更新後の誤差 ε が 0 より大きい。

[0216] 例えば、図 8 に図示されるように目標値 $408H$ が 4134 であり、図 16 に図示されるように認識数 $652H$ が 4060 である場合は、許容認識数 $N P_H$ が 3994 となり、更新後の誤差 ε が 0.020 となる。したがって、条件 1 から 4 までは満たされ、誤差 ε を更新することが要求される。

[0217] ステップ S 172 においては、特定基本事象 E, F, G 及び H のすべてについて誤差 ε を更新することが要求されなかったか否かが判定される。

[0218] 特定基本事象 E, F, G 及び H のすべてについて誤差 ε を更新することが要求されなかったと判定された場合は、検証部 114 が、ステップ S 173 において、検証の結果 $132E$, $132F$, $132G$ 及び $132H$ を出力する。

[0219] 一方、特定基本事象 E, F, G 及び H のいずれかについて誤差 ε を更新することが要求されたと判定された場合は、定義部 110 が、ステップ S 174 において、誤差 ε を更新することが要求された特定基本事象について誤差 ε を更新し、続くステップ S 175 において、誤差 ε を更新することが要求された特定基本事象について更新後の誤差 ε を用いて生成仕様を新たに定義する。これにより、定義部 110 は、システム 102 が仕様を満たさないが許容条件を満たし、誤差 ε から設定値 η を減じることにより得られる更新後の誤差 ε が 0 より大きい場合に、誤差 ε を更新後の誤差 ε に更新することによりデータ数 N'_x を増加させる。

[0220] 図 21 は、実施の形態 2 の検証システムにおいて算出される認識率の分布の例を表すグラフを図示する図である。

[0221] 図 21 に図示されるグラフ 800 には、誤差 ε が 0.025 である場合の認識率の分布 810、及び誤差 ε が 0.020 である場合の認識率の分布 8

1 2 が表される。認識率の分布 8 1 0 及び 8 1 2 のいずれから、認識率が 9 5 % であることを 9 9 % 以上の信頼度で区間推定により推定することができる。誤差 ε が 0. 0 2 0 である場合は、検証データセットに含めるデータ数 n は、式 (2 7) により、6 6 2 3 以上となる。したがって、データ数 N'_{H} は、6 6 2 3 である。また、式 (2 9) より、更新後の誤差 ε に基づいて定義される生成仕様に含まれるデータ数 N'_{H} は、6 4 2 5 であり、6 6 2 3 より増加させられている。

[0222] 図 2 2 は、実施の形態 2 のシステム検証装置に備えられる定義部により定義される生成仕様の例を図示する図である。図 2 3 は、実施の形態 2 のシステム検証装置に備えられる生成部により生成される検証データセットの例を図示する図である。図 2 4 は、実施の形態 2 のシステム検証装置に備えられる検証部により出力される検証の結果の例を図示する図である。

[0223] 図 2 2 に図示される生成仕様 1 2 4 S は、更新後の誤差 ε に基づいて定義された更新後の生成仕様である。図 2 3 に図示される検証データセット 1 2 8 S は、更新後の生成仕様 1 2 4 S に基づいて生成された更新後の検証データセットである。図 2 4 に図示される検証の結果 1 3 2 S は、更新後の検証データセット 1 2 8 S がシステム 1 0 2 に入力された場合に検証部 1 1 4 により出力される検証の結果である。検証の結果 1 3 2 S は、誤差 ε の更新により、システム 1 0 2 が仕様を満たすことを示す情報「○」を含む。このような検証の結果 1 3 2 S が得られるのは、認識数 6 5 2 S が 6 5 3 3 であり、認識数 6 5 2 S が目標値 4 0 8 S 以上であるためである。

[0224] なお、本発明は、その発明の範囲内において、実施の形態を適宜、変形、省略することが可能である。

[0225] この発明は詳細に説明されたが、上記した説明は、すべての局面において、例示であって、この発明がそれに限定されるものではない。例示されていない無数の変形例が、この発明の範囲から外れることなく想定され得るものと解される。

符号の説明

[0226] 100, 750 システム検証装置、102 システム、110 定義部、112 生成部、114 検証部、120, 120P 入力情報、122, 122P, 122Q 特定事象の抽出結果、124, 124E, 124F, 124G, 124H, 124S 生成仕様、126, 126P 検証用基データ、128, 128E, 128F, 128G, 128H, 128S 検証データセット、130, 130P 出力、132, 132E, 132F, 132G, 132H, 132S 検証の結果、200, 250, 300, 350 木構造、210A, 210B, 210C, 210D, 210E, 270, 272, 274, 320, 322, 360A, 360B, 360C, 360D, 360E, 360F, 360G, 360H, 700A, 700B, 700C, 700D, 700E ノード、220, 222, 260, 310, 370, 372, 374, 376, 710, 712, 714, 716 分岐のゲート。

請求の範囲

- [請求項1] システムにより望ましくない出力が生成される原因となる特定事象の抽出結果に基づいて生成仕様を定義する定義部と、
前記生成仕様に基づいて検証データセットを生成する生成部と、
前記検証データセット、及び前記検証データセットが前記システムに入力された場合に前記システムにより生成される出力に基づいて前記システムが仕様を満たしているか否かについて検証を行う検証部と、
を備えるシステム検証装置。
- [請求項2] 前記生成仕様は、目標値を含み、
前記検証部は、前記出力が前記目標値を達成している場合は、前記システムが前記仕様を満たしていると判定し、前記出力が前記目標値を達成していない場合は、前記システムが前記仕様を満たしていないと判定する
請求項1のシステム検証装置。
- [請求項3] 前記抽出結果は、前記特定事象が発生する原因となる複数の特定基本事象を表す情報を含み、
前記定義部は、前記複数の特定基本事象に含まれる各特定基本事象について前記生成仕様を定義する
請求項1又は2のシステム検証装置。
- [請求項4] 前記抽出結果は、木構造を含み、
前記木構造は、
複数の事象をそれぞれ表す複数のノードと、
前記複数のノードに含まれ上位事象を表す上位ノードに、前記複数のノードに含まれ前記上位事象が発生する原因となる下位事象を表す下位ノードを連結する分岐のゲートと、
を含み、
前記複数のノードは、

前記特定事象を表す内部ノードと、
前記複数の特定基本事象をそれぞれ表す複数の葉ノードと、
を含む
請求項3のシステム検証装置。

[請求項5] 前記抽出結果は、前記各特定基本事象について定義された加工条件をさらに含み、
前記定義部は、前記加工条件を前記生成仕様に含め、
前記生成部は、前記加工条件にしたがって検証用基データを加工することにより前記検証データセットを生成する
請求項3又は4のシステム検証装置。

[請求項6] 前記システムは、認識システムであり、
前記抽出結果は、前記各特定基本事象について定義された基本誤認識率をさらに含み、
前記定義部は、前記基本誤認識率を用いて目標値を決定し、前記目標値を前記生成仕様に含め、
前記検証部は、前記出力が前記目標値を達成している場合は、前記認識システムが前記仕様を満たしていると判定し、前記出力が前記目標値を達成していない場合は、前記システムが前記仕様を満たしていないと判定する
請求項3から5までのいずれかのシステム検証装置。

[請求項7] 前記システムは、認識システムであり、
前記抽出結果は、
前記特定事象について定義された発生確率と、
前記特定事象について定義された誤認識率と、
前記各特定基本事象について定義された発生確率と、
をさらに含む
請求項3から6までのいずれかのシステム検証装置。

[請求項8] 前記定義部は、作成可能な検証データの数を示す作成可能データ数

を決定し、前記作成可能データ数が最大データ数より多い場合は、前記作成可能な検証データの全部を前記検証データセットに含めたときに行われる検証の結果を区間推定により設定信頼度以上で推定可能な検証データの数を示すデータ数を決定し、前記データ数を前記生成仕様に含める

請求項 1 から 7 までのいずれかのシステム検証装置。

[請求項9] 前記生成部は、前記作成可能データ数が前記最大データ数より多い場合は、前記作成可能な検証データからランダムに選択された検証データを前記検証データセットに含める

請求項 8 のシステム検証装置。

[請求項10] 前記生成仕様は、複数の加工方法を含み、

前記生成部は、検証用基データに含まれる複数のデータからデータをランダムに選択し、前記複数の加工方法から加工方法をランダムに選択し、前記加工方法にしたがって前記データを加工することにより前記検証データセットを作成する

請求項 9 のシステム検証装置。

[請求項11] 前記システムは、認識システムであり、

前記出力は、複数の認識結果を含み、

前記検証部は、前記複数の認識結果がそれぞれ複数の教師ラベルと一致するか否かを判定し、教師ラベルと一致する認識結果の数を示す認識数を算出し、前記認識数が目標値以上である場合は、前記認識システムが前記仕様を満たしていると判定し、前記認識数が前記目標値より少ない場合は、前記認識システムが前記仕様を満たしていないと判定する

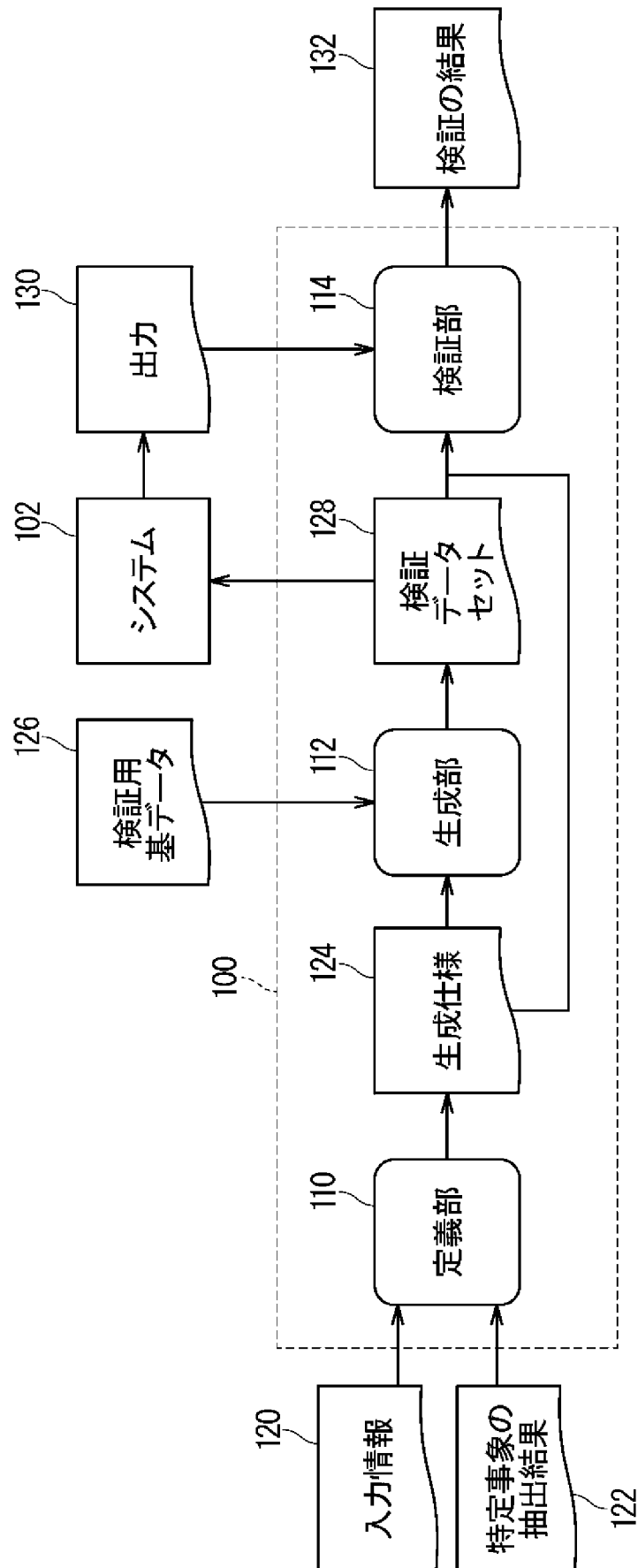
請求項 1 から 10 までのいずれかのシステム検証装置。

[請求項12] 前記定義部は、前記システムが前記仕様を満たさないが前記仕様より緩やかな許容条件を満たす場合に、前記データ数を増加させる

請求項 8 から 10 までのいずれかのシステム検証装置。

- [請求項13] 前記区間推定は、誤差を考慮し、
前記定義部は、前記システムが前記仕様を満たさないが前記仕様より緩やかな許容条件を満たし、前記誤差から設定値を減じることにより得られる更新後の誤差が0より大きい場合に、前記誤差を前記更新後の誤差に更新することにより前記データ数を増加させる
請求項12のシステム検証装置。
- [請求項14] 前記定義部は、前記検証データセットに含まれる検証データの属性を含む入力情報に基づいて前記生成仕様を定義する
請求項1から13までのいずれかに記載のシステム検証装置。
- [請求項15] システムにより望ましくない出力が生成される原因となる特定事象の抽出結果に基づいて生成仕様を定義する工程と、
前記生成仕様に基づいて検証データセットを生成する工程と、
前記検証データセット、及び前記検証データセットが前記システムに入力された場合に前記システムにより生成される出力に基づいて前記システムが仕様を満たしているか否かについて検証を行う工程と、
を備えるシステム検証方法。

[図1]

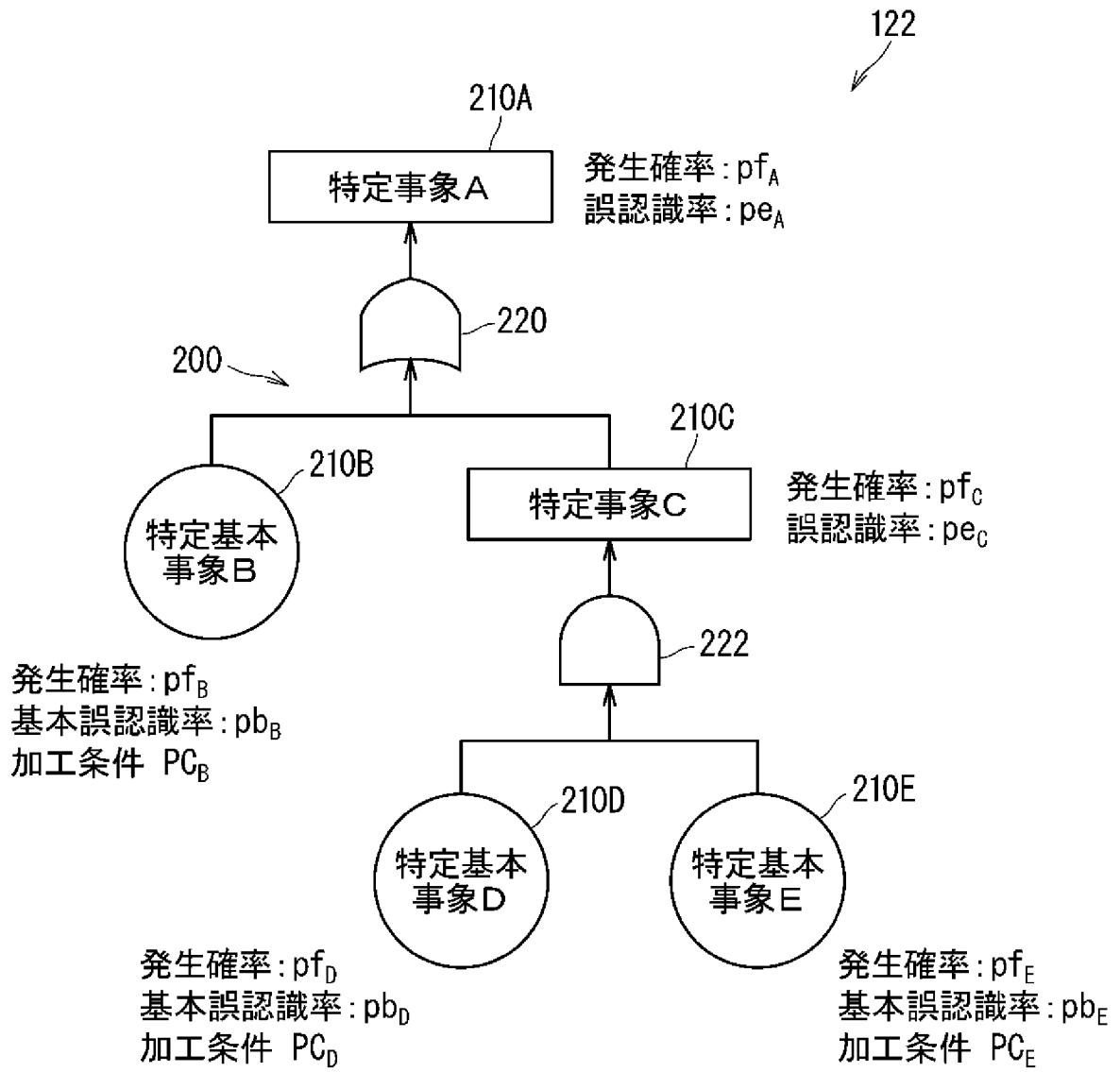


[図2]

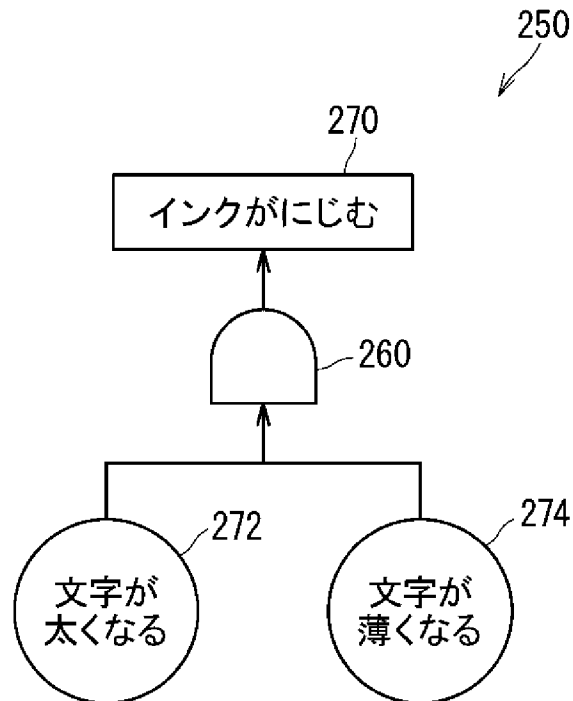
120P

画像サイズ	28×28(pixel) ← 150
ビット数	8(bit) ← 152

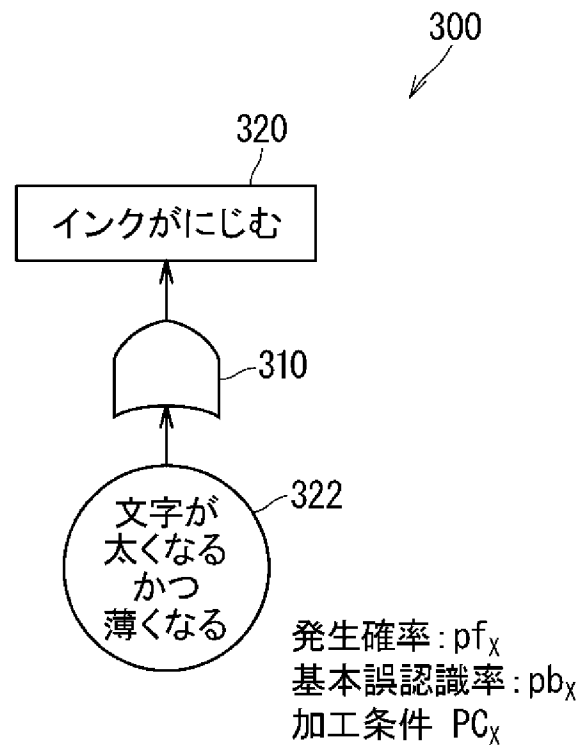
[図3]



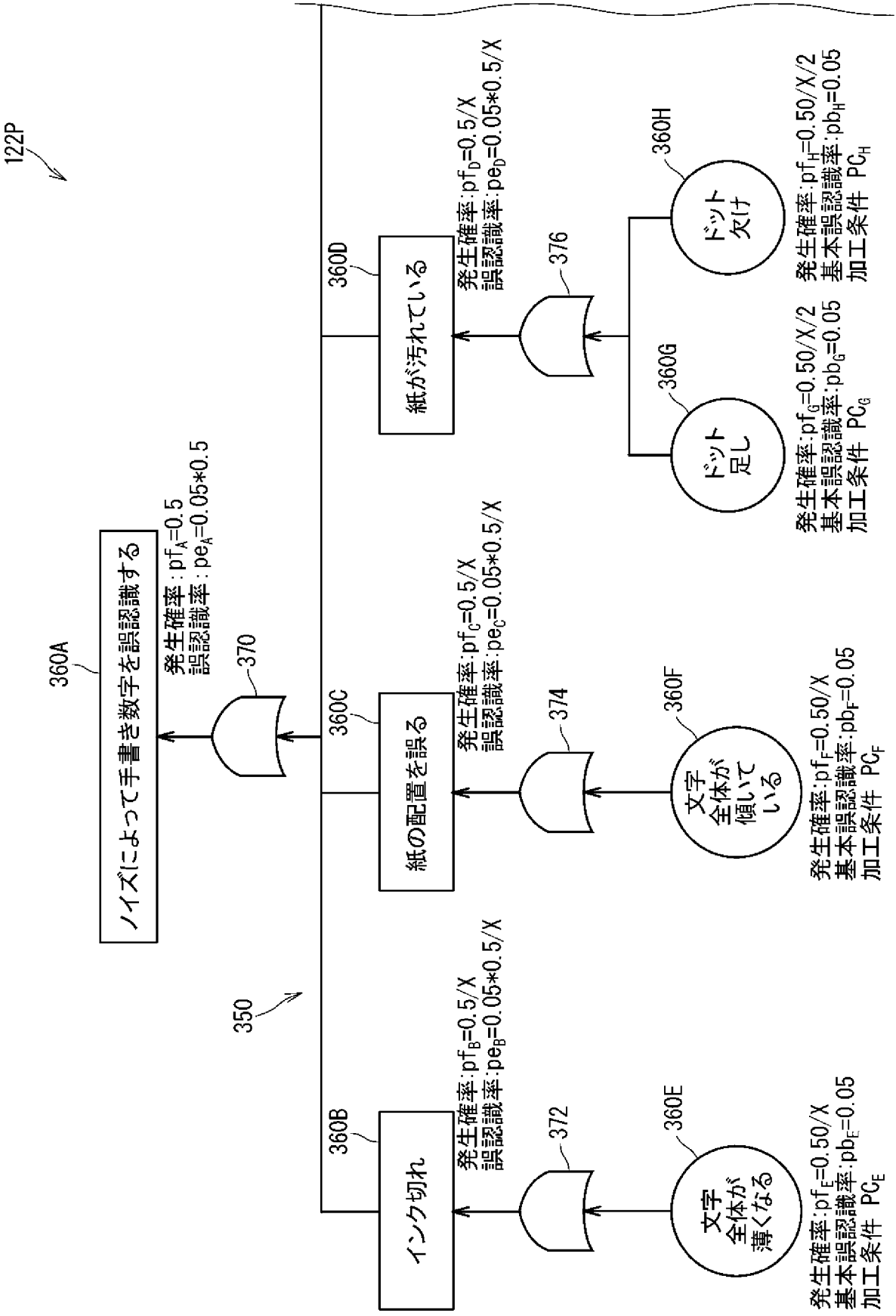
[図4]



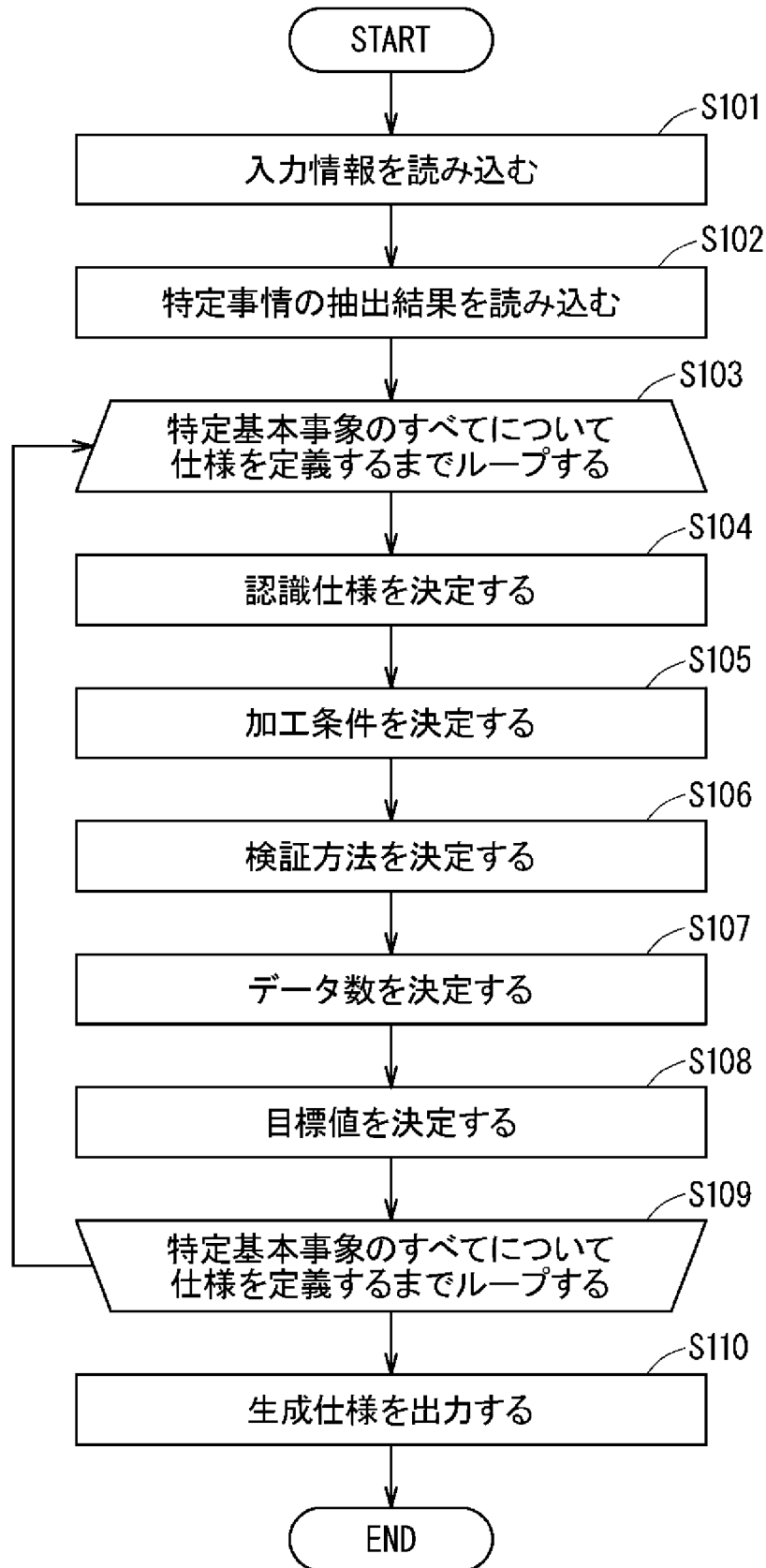
[図5]



[図6]



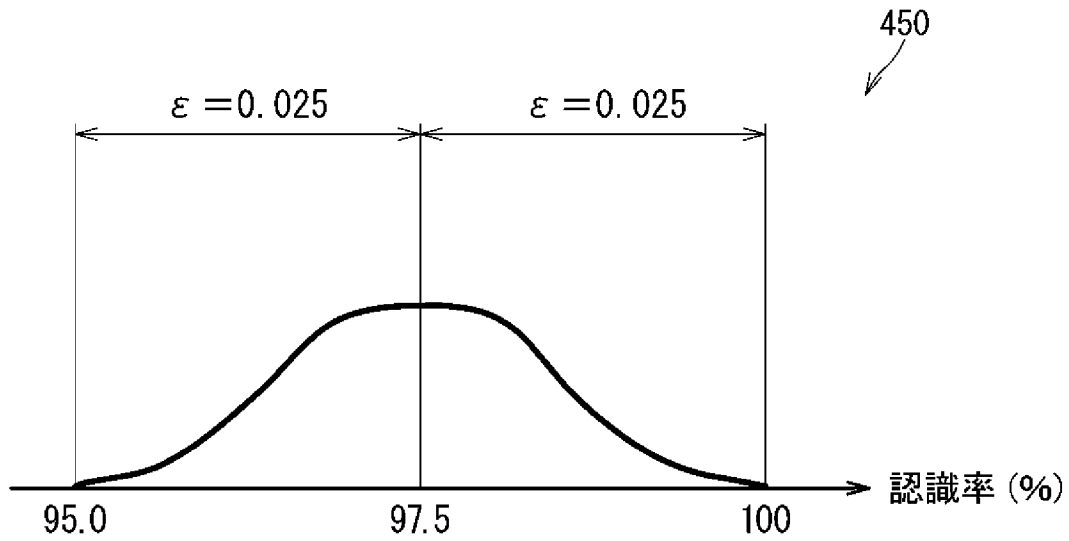
[図7]



[図8]

(a)	124E	<table> <tr> <td>認識仕様</td><td>「文字全体が薄くなる」という特定基本事象が発生した場合でも手書き数字を正しく認識すること</td><td>400E</td></tr> <tr> <td>加工条件</td><td>文字領域のみ69%から50%まで1%ずつ薄くする</td><td>402E</td></tr> <tr> <td>検証方法</td><td>全数検証</td><td>404E</td></tr> <tr> <td>目標値</td><td>190000</td><td>408E</td></tr> <tr> <td>データ数</td><td>200000</td><td>406E</td></tr> </table>	認識仕様	「文字全体が薄くなる」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400E	加工条件	文字領域のみ69%から50%まで1%ずつ薄くする	402E	検証方法	全数検証	404E	目標値	190000	408E	データ数	200000	406E
認識仕様	「文字全体が薄くなる」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400E															
加工条件	文字領域のみ69%から50%まで1%ずつ薄くする	402E															
検証方法	全数検証	404E															
目標値	190000	408E															
データ数	200000	406E															
(b)	124F	<table> <tr> <td>認識仕様</td><td>「文字全体が傾いている」という特定基本事象が発生した場合でも手書き数字を正しく認識すること</td><td>400F</td></tr> <tr> <td>加工条件</td><td>文字領域のみ時計回りに10°から49°まで1°ずつ回転させる</td><td>402F</td></tr> <tr> <td>検証方法</td><td>全数検証</td><td>404F</td></tr> <tr> <td>目標値</td><td>380000</td><td>408F</td></tr> <tr> <td>データ数</td><td>400000</td><td>406F</td></tr> </table>	認識仕様	「文字全体が傾いている」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400F	加工条件	文字領域のみ時計回りに10°から49°まで1°ずつ回転させる	402F	検証方法	全数検証	404F	目標値	380000	408F	データ数	400000	406F
認識仕様	「文字全体が傾いている」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400F															
加工条件	文字領域のみ時計回りに10°から49°まで1°ずつ回転させる	402F															
検証方法	全数検証	404F															
目標値	380000	408F															
データ数	400000	406F															
(c)	124G	<table> <tr> <td>認識仕様</td><td>「ドット足し」という特定基本事象が発生した場合でも手書き数字を正しく認識すること</td><td>400G</td></tr> <tr> <td>加工条件</td><td>文字領域外に1×1から3×3までのノイズを1個から10個まで加える</td><td>402G</td></tr> <tr> <td>検証方法</td><td>統計的検証</td><td>404G</td></tr> <tr> <td>目標値</td><td>4134</td><td>408G</td></tr> <tr> <td>データ数</td><td>4240</td><td>406G</td></tr> </table>	認識仕様	「ドット足し」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400G	加工条件	文字領域外に1×1から3×3までのノイズを1個から10個まで加える	402G	検証方法	統計的検証	404G	目標値	4134	408G	データ数	4240	406G
認識仕様	「ドット足し」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400G															
加工条件	文字領域外に1×1から3×3までのノイズを1個から10個まで加える	402G															
検証方法	統計的検証	404G															
目標値	4134	408G															
データ数	4240	406G															
(d)	124H	<table> <tr> <td>認識仕様</td><td>「ドット欠け」という特定基本事象が発生した場合でも手書き数字を正しく認識すること</td><td>400H</td></tr> <tr> <td>加工条件</td><td>文字領域内に1×1から3×3までのノイズを1個から10個まで加える</td><td>402H</td></tr> <tr> <td>検証方法</td><td>統計的検証</td><td>404H</td></tr> <tr> <td>目標値</td><td>4134</td><td>408H</td></tr> <tr> <td>データ数</td><td>4240</td><td>406H</td></tr> </table>	認識仕様	「ドット欠け」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400H	加工条件	文字領域内に1×1から3×3までのノイズを1個から10個まで加える	402H	検証方法	統計的検証	404H	目標値	4134	408H	データ数	4240	406H
認識仕様	「ドット欠け」という特定基本事象が発生した場合でも手書き数字を正しく認識すること	400H															
加工条件	文字領域内に1×1から3×3までのノイズを1個から10個まで加える	402H															
検証方法	統計的検証	404H															
目標値	4134	408H															
データ数	4240	406H															

[図9]

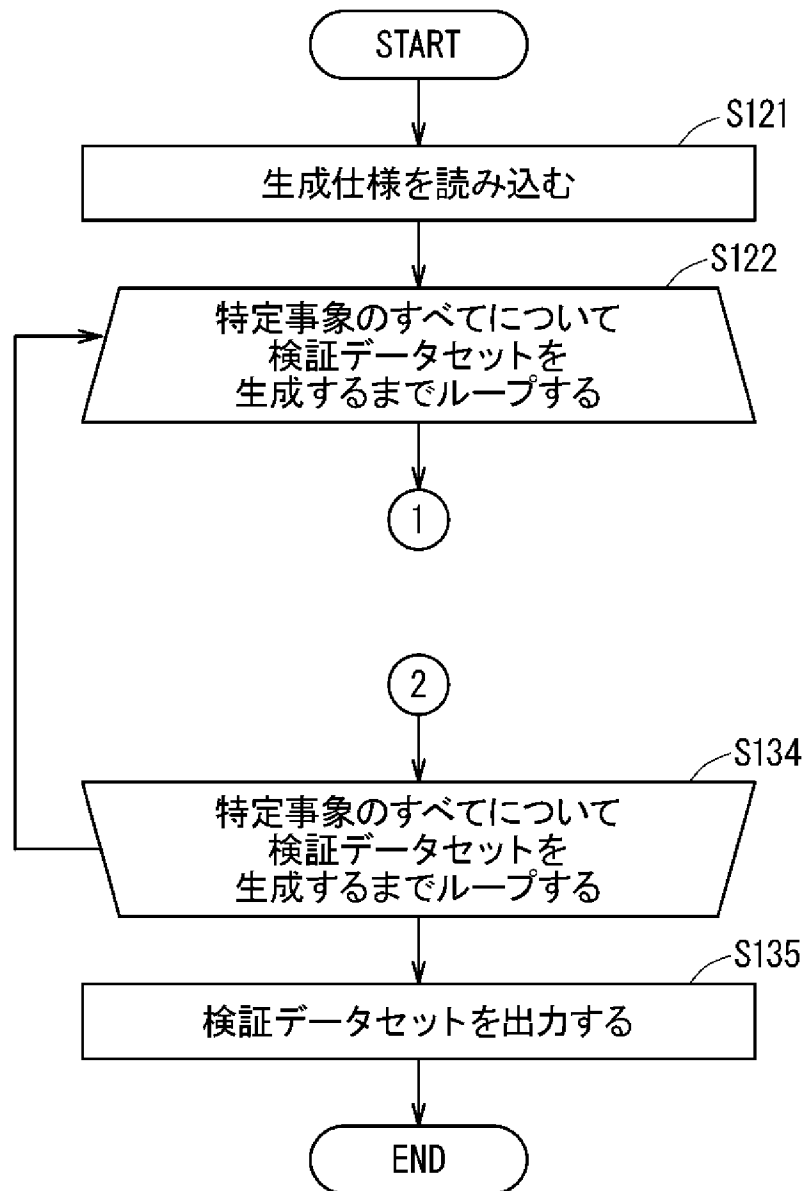


[図10]

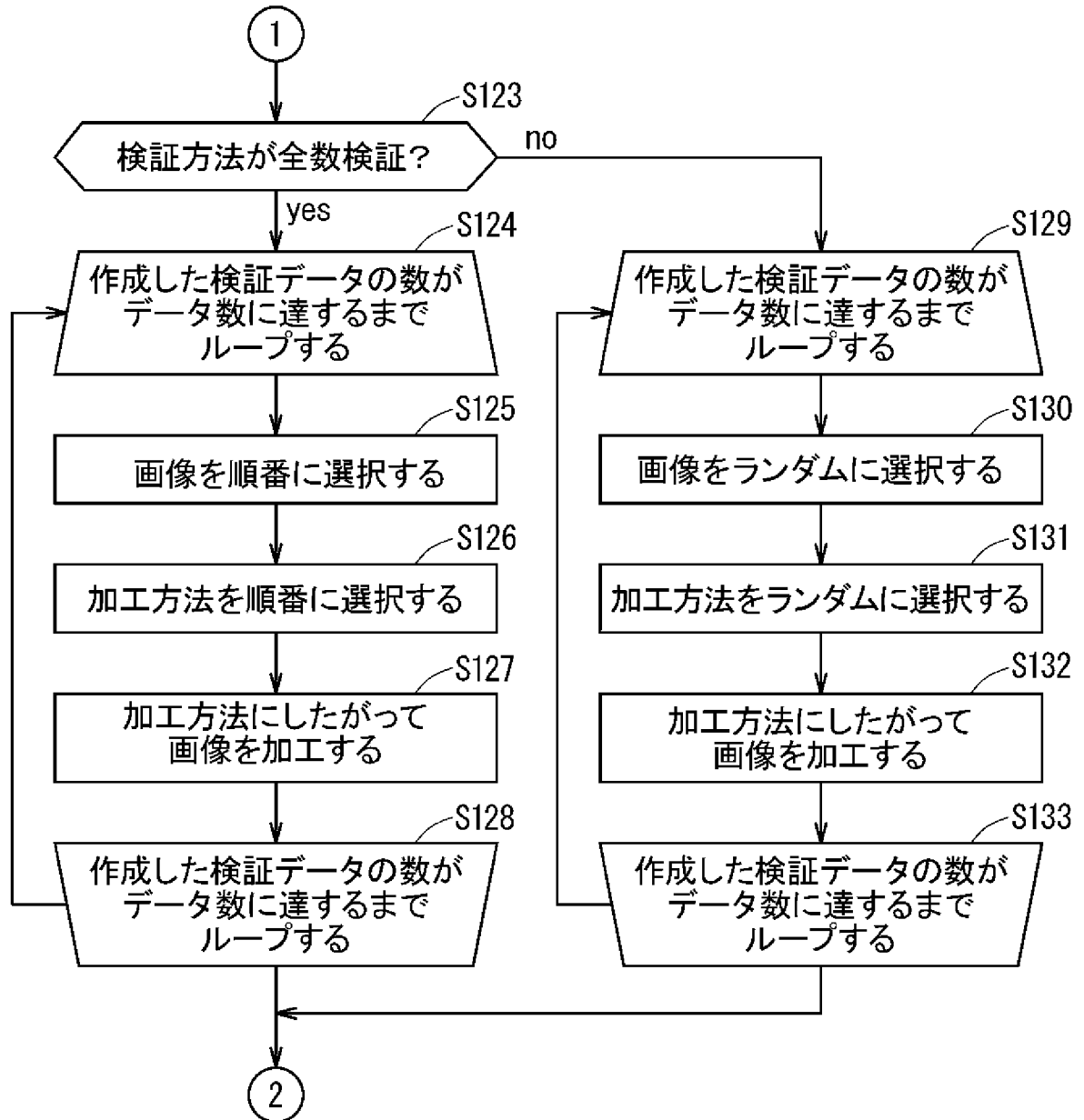
Figure 10 shows a table with 6 columns. The first column is labeled "No." and contains the values 1, 2, 3, ..., 10000. The second column is labeled "教師ラベル" (Teacher Label) and contains the values 3, 5, 502, 8, ..., 4. The third column contains handwritten digits 3, 5, 8, ..., 4. The fourth column contains the number of samples for each digit: 500, 504, 502, 8, ..., 4. The fifth column contains the number of samples for each digit: 500, 504, 502, 8, ..., 4.

No.	1	2	3		10000
教師ラベル	3	5	502 8		4
	3	5	504 8		4
			500		

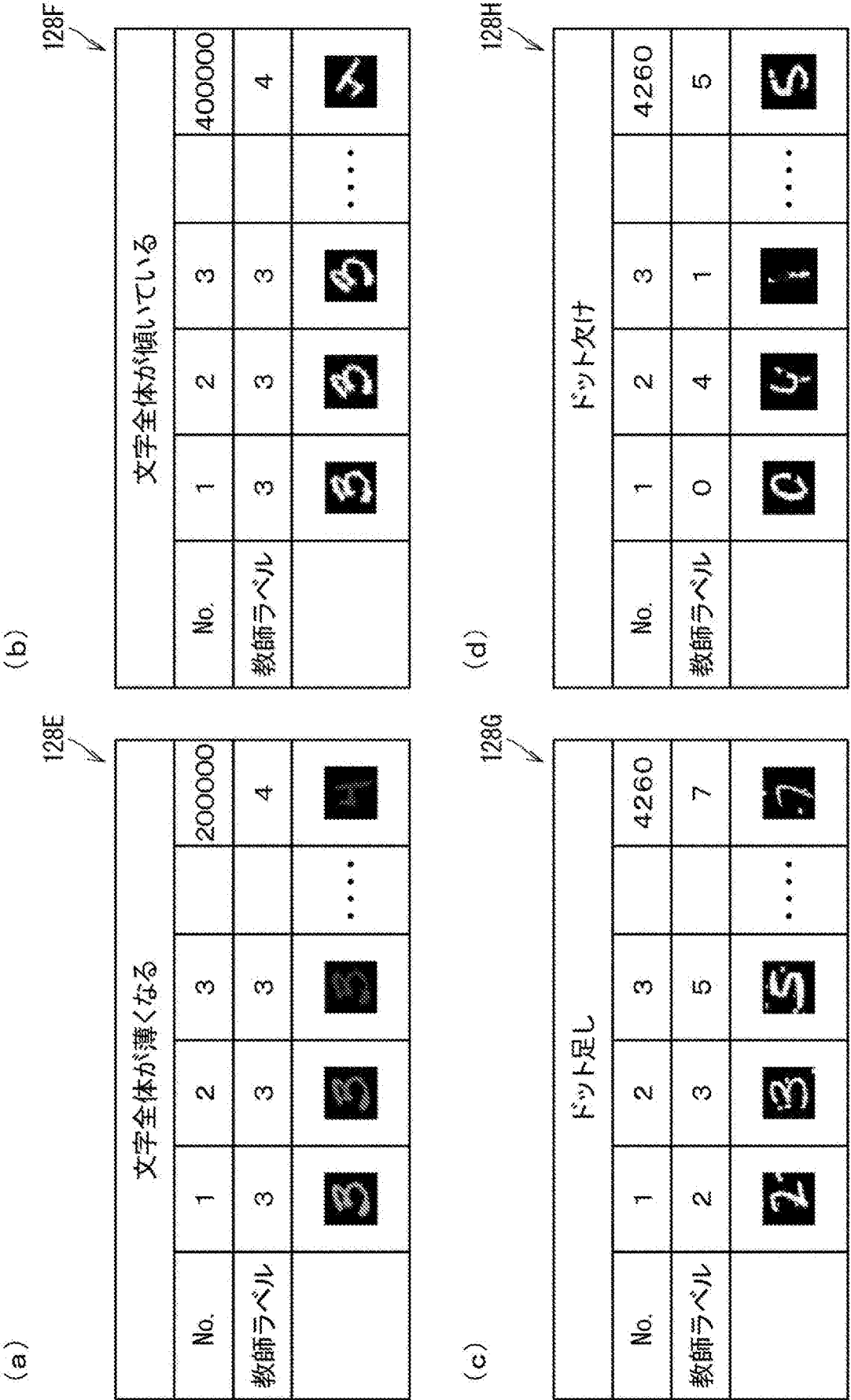
[図11]



[図12]



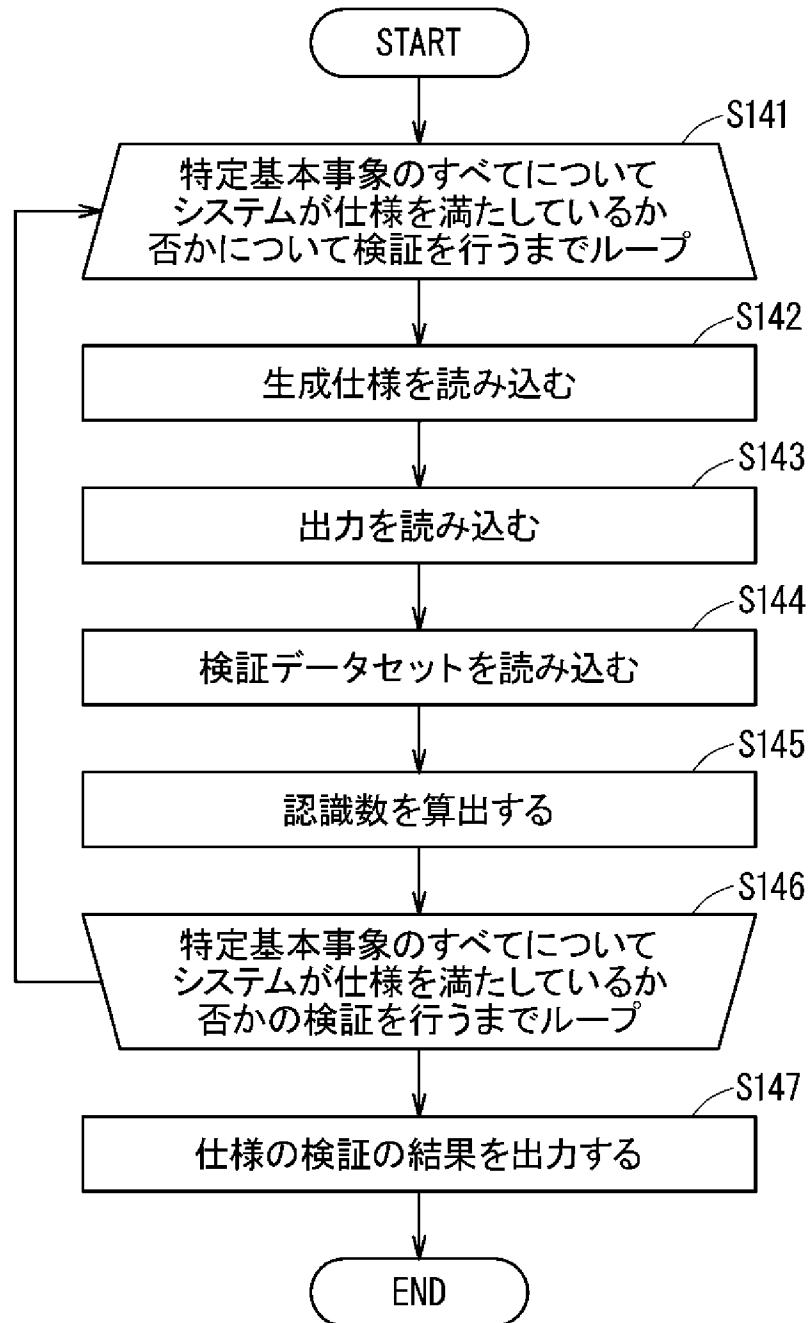
[図13]



130P

[illegible]

[図15]



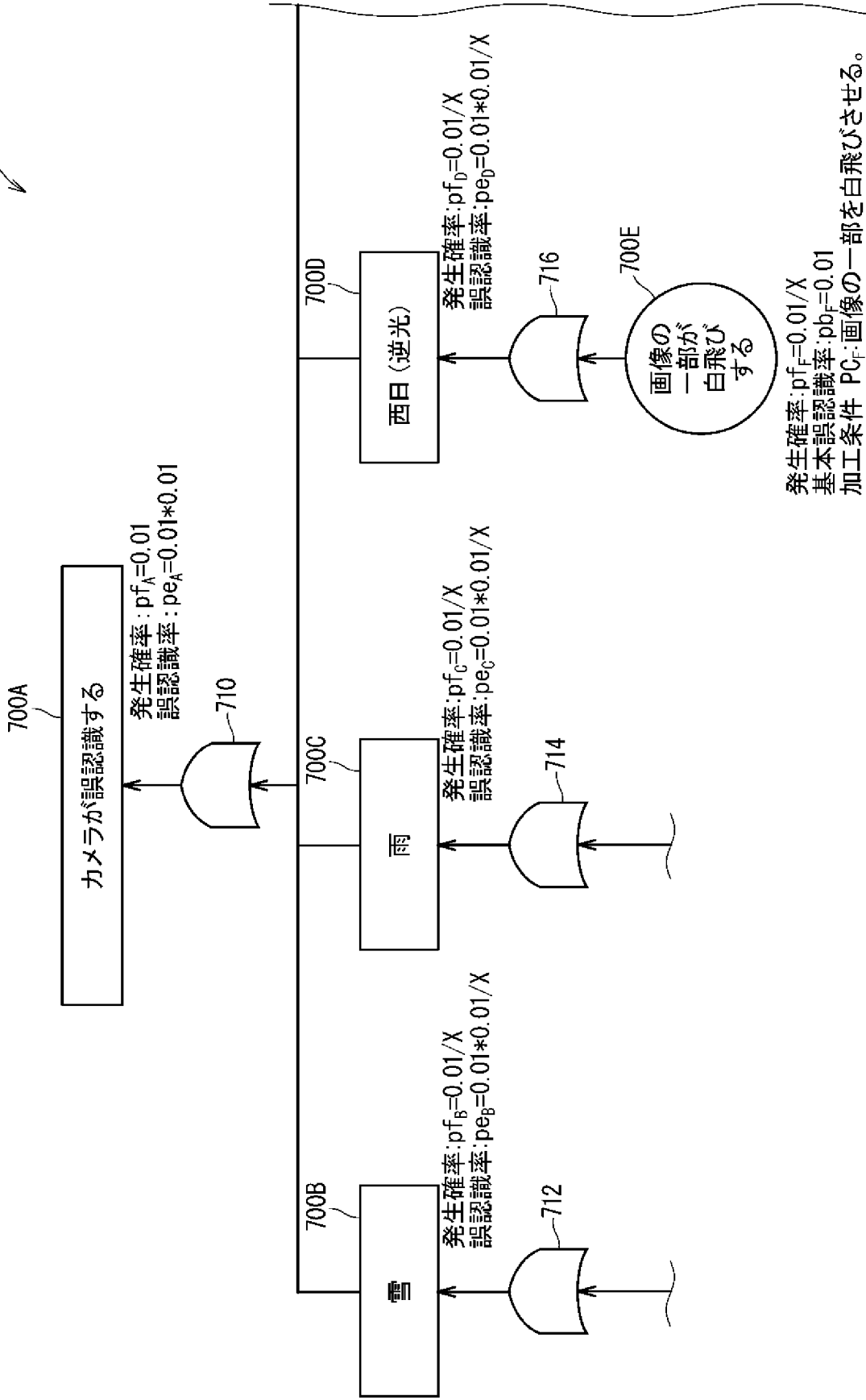
[図16]

特定基本事象	認識数	検証の結果
文字全体が薄くなる	195241	○ ← 132E
文字全体が傾いている	395675	○ ← 132F
ドット足し	4186	○ ← 132G
ドット欠け	4060	× ← 132H

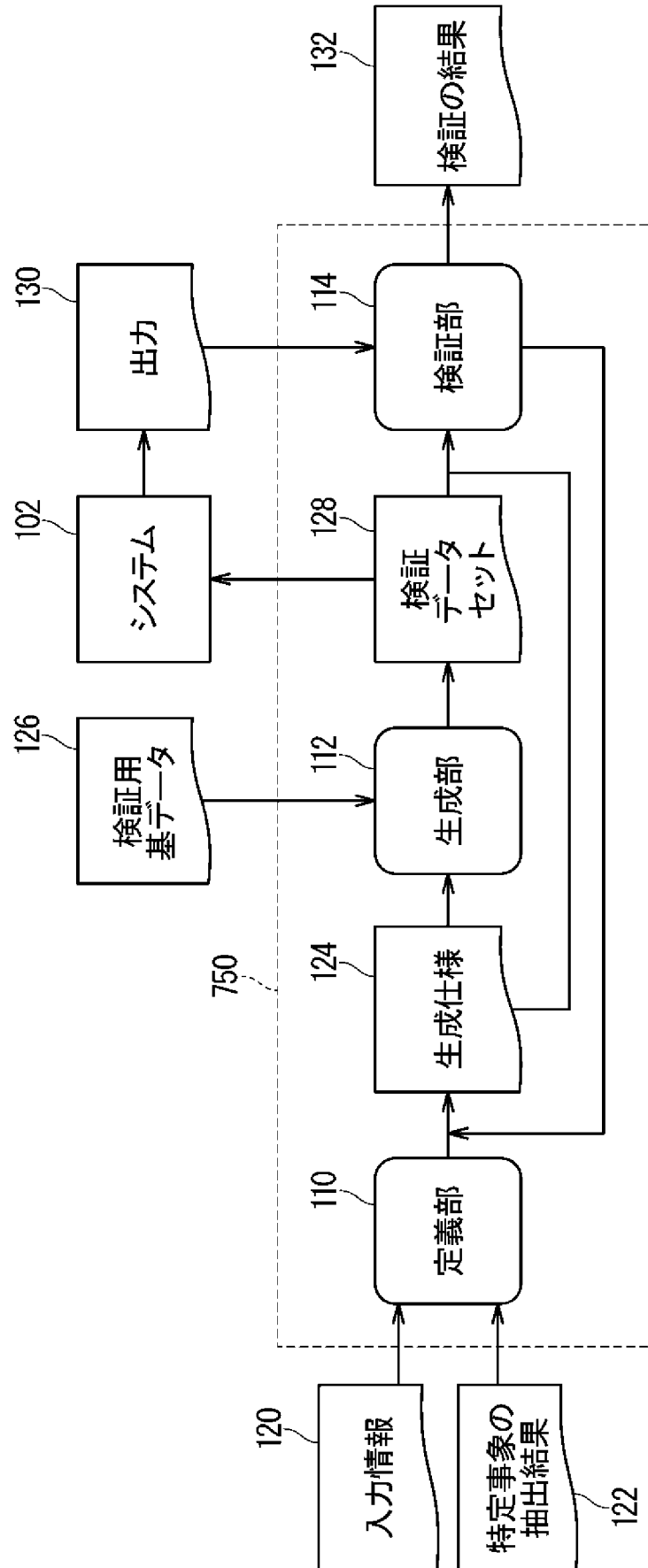
111
 652E
 652F
 652G
 652H

[図17]

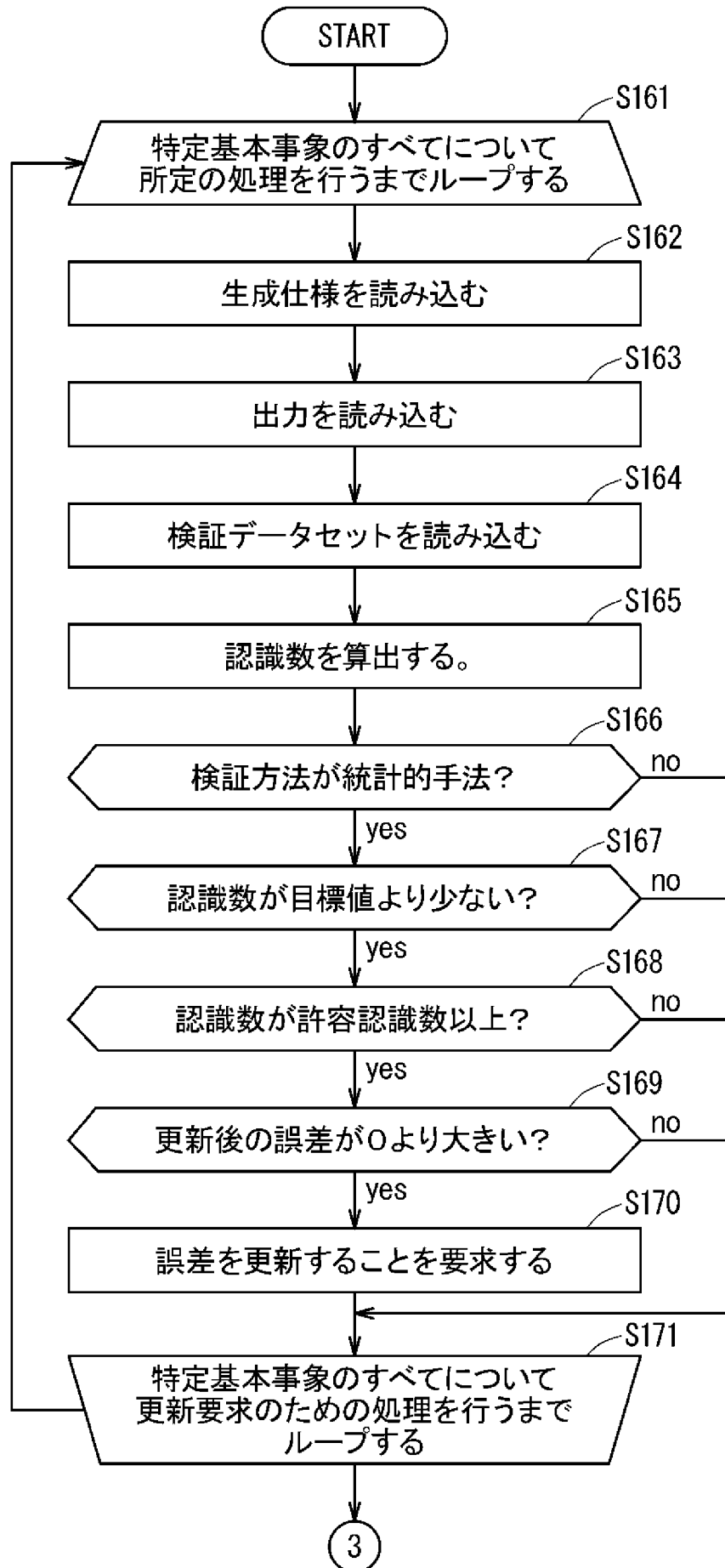
1220



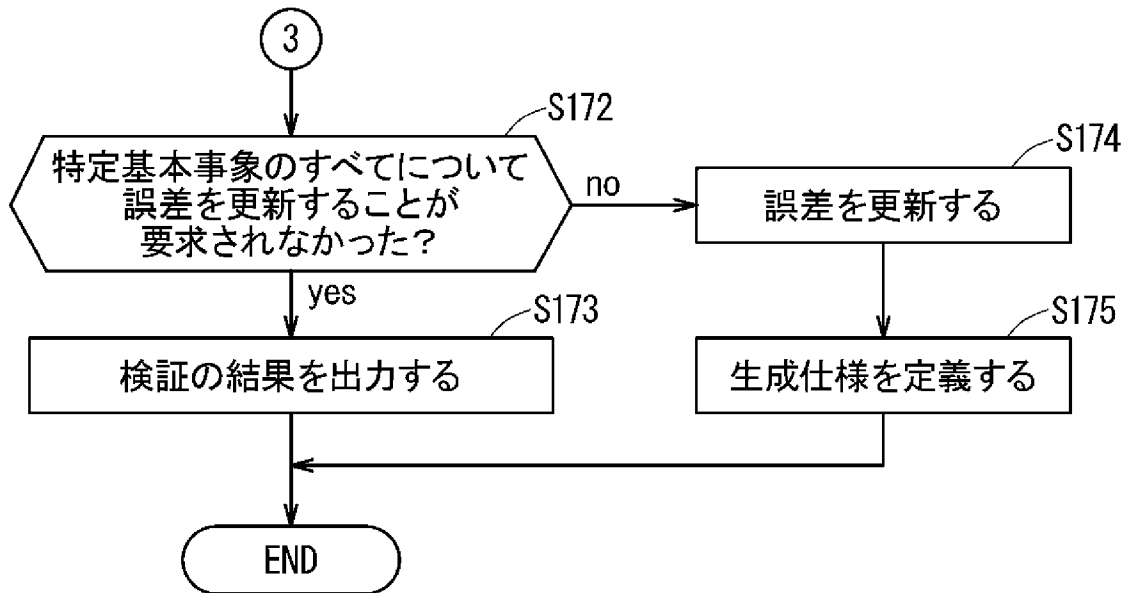
[図18]



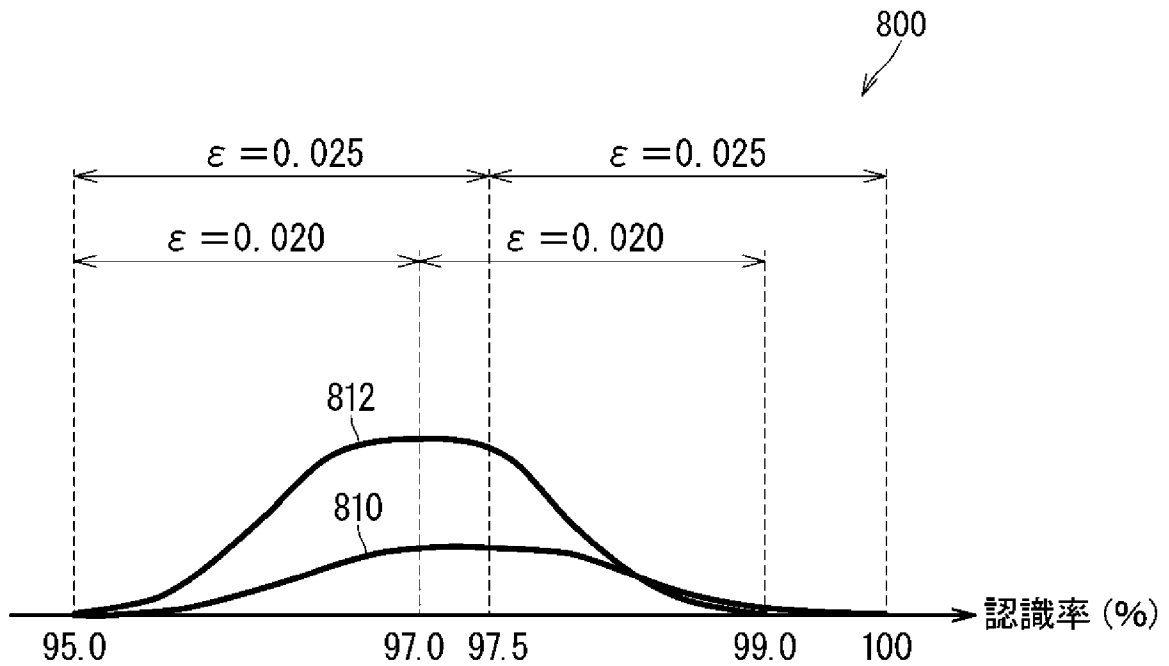
[図19]



[図20]



[図21]



[図22]

124S

認識仕様	「ドット欠け」という特定基本事象が発生した場合でも手書き数字を正しく認識すること
加工条件	文字領域内に 1×1から3×3までのノイズを1個から10個まで加える
検証方法	統計的検証
ε	0.020
目標値	6425
データ数	6623

408S

[図23]

128S

ドット欠け					
No.	1	2	3		6623
教師ラベル	2	8	9		6
					

[図24]

特定基本事象	認識数	検証の結果
文字全体が薄くなる	195241	○
文字全体が傾いている	395675	○
ドット足し	4186	○
ドット欠け	6533	○

652S 132S

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/027420

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06F11/36 (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)

Int.Cl. G06F11/36

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 1971-2018

Registered utility model specifications of Japan 1996-2018

Published registered utility model applications of Japan 1994-2018

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.
X A	JP 2001-184232 A (TOSHIBA CORP.) 06 July 2001, paragraphs [0002], [0005], [0008], [0042], [0046]-[0047] (Family: none)	1, 3-4, 14-15 2, 5-13
A	JP 2016-115278 A (FUJITSU LTD.) 23 June 2016, entire text (Family: none)	5-13
A	JP 2017-187918 A (NIPPON TELEGRAPH AND TELEPHONE CORP.) 12 October 2017, entire text (Family: none)	5-13

☐ 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

"I" 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 October 2018 (12.10.2018)

Date of mailing of the international search report
23 October 2018 (23.10.2018)

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

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F11/36(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. G06F11/36

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	1971-2018年
日本国実用新案登録公報	1996-2018年
日本国登録実用新案公報	1994-2018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
X A	JP 2001-184232 A（株式会社東芝）2001.07.06, 段落[0002], [0005], [0008], [0042], [0046]-[0047]（ファミリーなし）	1, 3-4, 14-15 2, 5-13
A	JP 2016-115278 A（富士通株式会社）2016.06.23, 全文（ファミリーなし）	5-13
A	JP 2017-187918 A（日本電信電話株式会社）2017.10.12, 全文（ファミリーなし）	8-13

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

12.10.2018

国際調査報告の発送日

23.10.2018

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

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）

小林 哲雄

5B

5881

電話番号 03-3581-1101 内線 3545