

(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3月5日(05.03.2020)



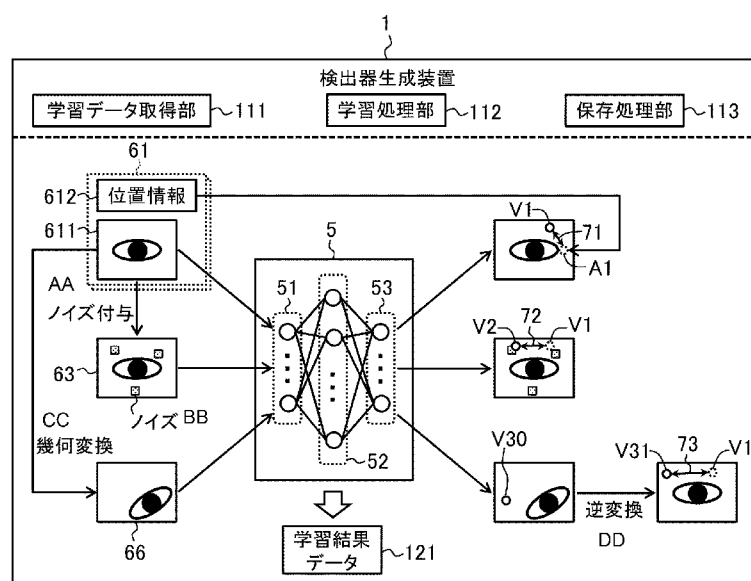
(10) 国際公開番号

WO 2020/044630 A1

- (51) 国際特許分類:
G06T 7/00 (2017.01)
- (21) 国際出願番号: PCT/JP2019/010187
- (22) 国際出願日: 2019年3月13日(13.03.2019)
- (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 2018-162683 2018年8月31日(31.08.2018) JP
- (71) 出願人: オムロン株式会社 (OMRON CORPORATION) [JP/JP]; 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町801番地 Kyoto (JP).
- (72) 発明者: 長江 成典 (NAGAE, Shigenori); 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町801番地 オムロン株式会社内 Kyoto (JP). 竹内 倭 (TAKEUCHI, Yamato); 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町801番地 オムロン株式会社内 Kyoto (JP).
- (74) 代理人: 立花 顕治, 外 (TACHIBANA, Kenji et al.); 〒5300005 大阪府大阪市北区中之島6-2-40 中之島インテス21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,

(54) Title: DETECTOR GENERATION DEVICE, MONITORING DEVICE, DETECTOR GENERATION METHOD, AND DETECTOR GENERATION PROGRAM

(54) 発明の名称: 検出器生成装置、モニタリング装置、検出器生成方法及び検出器生成プログラム



- | | | | |
|-----|--------------------------------|-----|--------------------------|
| 1 | Detector generation device | 612 | Position information |
| 111 | Learning data acquisition unit | AA | Apply noise |
| 112 | Learning processing unit | BB | Noise |
| 113 | Save processing unit | CC | Geometric transformation |
| 121 | Learning result data | DD | Reverse transformation |

(57) Abstract: Provided is technology for generating a detector that is robust against the effects of noise and that makes it possible to detect an object with a higher degree of precision. A detector generation device according to one aspect of the present invention acquires a plurality of learning data sets each comprising a combination of a first face image and position information indicating the positions of the organs of the face, then uses the acquired plurality of learning data sets to subject a detector to machine learning. Subjecting the detector to machine learning includes using the detector to

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, 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:

- 国際調査報告 (条約第21条(3))

calculate the total value of first errors between a specified first detection position and a position indicated by the position information in each first face image, generating each of second face images by applying noise to each of the first face images, using the detector to calculate the total value of second errors between a specified second detection position and each first detection position in each of the second face images, and training the detector so that the total value of each of the errors decreases.

(57) 要約: ノイズの影響にロバストで、対象物をより高精度に検出可能な検出器を生成するための技術を提供する。本発明の一側面に係る検出器生成装置は、第1顔画像及び顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得し、取得した複数の学習データセットを利用して、検出器の機械学習を実施する。機械学習を実施することは、検出器を利用して各第1顔画像において特定される第1検出位置と位置情報により示される位置との第1誤差の合計値を算出すること、各第1顔画像にノイズを付与することで各第2顔画像を生成すること、検出器を利用して各第2顔画像において特定される第2検出位置と各第1検出位置との第2誤差の合計値を算出すること、及び各誤差の合計値が小さくなるように検出器を訓練することを含む。

明 細 書

発明の名称：

検出器生成装置、モニタリング装置、検出器生成方法及び検出器生成プログラム

技術分野

[0001] 本発明は、検出器生成装置、モニタリング装置、検出器生成方法及び検出器生成プログラムに関する。

背景技術

[0002] 近年、脇見等に起因する自動車の交通事故を防止するために、運転者の状態を監視する技術の開発が進んでいる。また、自動車の自動運転の実現に向けた動きが加速している。自動運転は、システムにより自動車の操舵を制御するものであるが、システムに代わって運転者が運転しなければならない場面もあり得ることから、自動運転中であっても、運転者が運転操作を行える状態にあるか否かを監視する必要性があるとされている。この自動運転中に運転者の状態を監視する必要性があることは、国連欧州経済委員会（UNECE）の政府間会合（WP29）においても確認されている。この点からも、運転者の状態を監視する技術の開発が進められている。

[0003] 運転者の状態を監視する技術の一つとして、視線方向等の運転者の顔の状態を推定するために、運転者を撮影することで得られた画像から運転者の顔の器官を検出する技術がある。顔の器官を検出することができれば、検出した器官の位置関係に基づいて、運転者の状態を推定することができる。例えば、顔画像から各目を検出し、各目の位置を特定することができれば、特定した各目の位置関係に基づいて、運転者の視線方向を推定することができる。

[0004] 人物の器官を検出する技術として、例えば、特許文献1では、畳み込みニューラルネットワーク（Convolutional Neural Networks）を利用して、顔の器官を検出するための装置が提案されている。具体的には、この装置は、畳

み込みニューラルネットワークを使用して、顔の器官に関連する特徴点の候補を顔画像データから抽出し、抽出した各候補に対して幾何学的な補正処理を実行することで、各特徴点の位置を決定する。これにより、当該装置は、決定した各特徴点の位置に基づいて、対象人物を認識することができる。

先行技術文献

特許文献

[0005] 特許文献1：特開2012-243285号公報

発明の概要

発明が解決しようとする課題

[0006] 従来、特許文献1で例示されるように、顔の器官の位置を検出するために、ニューラルネットワーク等の学習モデルにより構成された検出器を採用することがある。この検出器は、機械学習により、入力された顔画像から顔の器官の位置を検出することができるようになる。検出器の機械学習を実施するためには、顔の写る顔画像及び顔の器官の位置を示す位置情報が用意される。機械学習では、用意した顔画像が入力されると、位置情報に対応する出力値を出力するように検出器のパラメータが調整される。この機械学習により、顔画像から顔の器官の位置を検出する能力を獲得した学習済みの検出器が生成される。

[0007] 本件発明者らは、このような学習済みの検出器を用いる従来の方法では、図1に示すような問題点が生じることを見出した。図1は、従来の方法により顔の器官の位置を検出する場面の一例を模式的に例示する。図1では、顔の器官の一例として目（具体的には、目尻）を検出する場면을例示する。機械学習により、検出器は、同一の顔画像に対して同一の出力値を出力するように、換言すると、同一の顔画像に対して同一の位置を対象の器官の検出位置として回答する能力を獲得する。したがって、運転者が顔の器官を動かさなければ、本来、検出器から回答される器官の検出位置は一定のはずである。

[0008] しかしながら、動画像等の連続的に取得される複数の画像には、撮影装置の性能、撮影環境等に応じたノイズがランダムに含まれ得る。このランダムなノイズの影響によって、完全に同一の画像を得ることは難しく、各画像は微妙に異なり得る。この各画像の微妙な相違に起因して、運転者が顔の器官を動かしていないにも関わらず、検出器から回答される器官の検出位置にぶれが生じてしまう。すなわち、図1の示すように、器官（目尻）を動かしていないにも関わらず、各顔画像に含まれるノイズの影響によって、各顔画像間で器官の検出位置が変動してしまうという問題点が生じ得ることを本件発明者らは見出した。

[0009] 運転者が顔の器官を動かしていないにも関わらず、検出器による器官の検出位置にぶれが生じてしまうと、この検出結果の変動に起因して、運転者の状態を推定する精度が低くなってしまう。例えば、上記の視線方向を推定する例では、運転者が視線を変えていないにも関わらず、各目の検出位置がぶれてしまうことにより、運転者の視線方向の推定結果が変動してしまうことになる。これにより、例えば、運転者が前方を注視しているにも関わらず運転者が脇見をしていると認識してしまう等のように、システムが運転者の状態を誤認識してしまう可能性が生じる。

[0010] なお、この課題は、上記のような運転者の顔の器官を検出する場面ではなく、例えば、運転者以外の対象者の顔の器官を検出する場面、対象者の顔の器官以外の対象物を検出する場面等の、対象物を検出するあらゆる場面で生じ得る。例えば、生産ラインを流れる製品、製品に形成された傷等の対象物を検出する場合に、従来の検出器を用いた方法では、対象物が動いていないにも関わらず、複数の画像間で対象物の検出位置がぶれてしまうという問題点が生じ得る。

[0011] 本発明は、一側面では、このような実情を鑑みてなされたものであり、その目的は、ノイズの影響にロバストで、対象物をより高精度に検出可能な検出器を生成するための技術を提供することである。

課題を解決するための手段

[0012] 本発明は、上述した課題を解決するために、以下の構成を採用する。

[0013] すなわち、本発明の一側面に係る検出器生成装置は、車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、を備える。そして、前記検出器の機械学習を実施することは、前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記器官の第1検出位置と当該第1顔画像に組み合わせられた前記位置情報により示される前記位置との第1誤差の合計値を算出するステップと、前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1顔画像に対して生成された前記第2顔画像に対して得られた前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、を含む。

[0014] 上記構成に係る検出器生成装置は、第1顔画像及び位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得し、取得した複数の学習データセットを利用して、検出器の機械学習を実施する。第1顔画像は訓練データ（入力データ）に対応し、位置情報は正解データ（教師データ）に対応する。検出器は、顔画像の入力を受け付け、入力された顔画像に写る顔の器官の位置を検出した結果（例えば、器官の座標値）に対応する出力

値を出力するように構成される。

[0015] 上記構成に係る検出器生成装置は、機械学習の過程において、第1顔画像を検出器に入力することで第1出力値を取得し、取得した第1出力値に基づいて特定される器官の第1検出位置と位置情報により示される位置との第1誤差の合計値を算出する。この第1誤差の合計値は、検出器により第1顔画像から器官を検出した結果と位置情報により指定される器官の位置の正解との間の損失に相当する。そこで、上記構成に係る検出器生成装置は、この第1誤差の合計値が小さくなるように検出器を訓練する。これにより、位置情報により示される正解と一致するように、第1顔画像から器官を検出する能力を習得した検出器が構築される。

[0016] 加えて、上記構成に係る検出器生成装置は、第1顔画像にノイズを付与することで第2顔画像を生成し、生成した第2顔画像を検出器に入力することで第2出力値を取得する。そして、上記構成に係る検出器生成装置は、上記第1検出位置と取得した第2出力値に基づいて特定される器官の第2検出位置との第2誤差の合計値を算出する。この第2誤差の合計値は、検出器による器官の検出結果のブレであって、ノイズの有無によって生じるブレの損失に相当する。そこで、上記構成に係る検出器生成装置は、上記第1誤差の合計値と共に、第2誤差の合計値が小さくなるように検出器を訓練する。これにより、器官の検出結果がノイズによる影響を受け難い検出器を構築することができる。

[0017] したがって、上記構成に係る検出器生成装置によれば、位置情報により示される正解と一致するように第1顔画像から器官を検出し、かつその検出結果がノイズによる影響を受け難い、ように訓練された検出器を構築することができる。よって、ノイズの影響にロバストで、対象物（上記一側面では、顔の器官）をより高精度に検出可能な検出器を生成することができる。

[0018] なお、検出器は、機械学習を実施可能な学習モデルにより構成される。このような検出器は、例えば、ニューラルネットワークにより構成されてよい。顔画像から検出される器官の種類は、顔に含まれる何らかの部位であれば

特に限定されなくともよく、実施の形態に応じて適宜選択されてよい。検出の対象となる器官は、例えば、目、鼻、口、眉、顎及びこれらの組み合わせから選択されてよい。また、検出の対象となる器官には、顔の輪郭が含まれてもよい。

[0019] 上記一側面に係る検出器生成装置において、前記検出器の機械学習を実施することは、前記各学習データセットの前記第1顔画像に所定の幾何変換を適用することで、前記各学習データセットの前記第1顔画像に対して第3顔画像を生成するステップと、前記各第3顔画像を前記検出器に入力することで、前記各第3顔画像に対する第3出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1顔画像に対して生成された前記第3顔画像に対して得られた前記第3出力値に基づいて特定される前記器官の第3検出位置との第3誤差の合計値を算出するステップと、を更に含んでもよく、前記訓練するステップでは、前記学習処理部は、前記第1誤差の合計値、前記第2誤差の合計値及び前記第3誤差の合計値が小さくなるように前記検出器を訓練してもよい。

[0020] 本件発明者らは、上記画像に生じるランダムなノイズの影響の他、次のような理由によっても、器官の検出位置が変動してしまうという問題点が生じ得ることを見出した。すなわち、機械学習の正解データとして利用される位置情報にはノイズが含まれ得る。つまり、位置情報は、常に一定の基準に従って対象の器官の位置を正確に示すとは限らないという不安定性を有している。例えば、位置情報が人間により入力される場合に、位置情報を入力する人物が異なれば、器官の存在する位置として指定される位置は異なり得る。また、同一の人物が位置情報を入力したとしても、同一の顔画像に対して完全に一致する位置を複数回指定するのは困難である。

[0021] 具体例として、上記図1に例示される目尻の位置を位置情報として入力する場面を想定する。この場面において、複数の異なる人物が顔画像に対して目尻の位置を指定する場合に、各人物の間には、目尻の位置の定義に対する

解釈に微妙な差が存在し得る。そのため、同一の顔画像に対して各人物が目尻の位置を指定したとしても、各人物により指定される目尻の位置は微妙に相違し得る。また、同一の人物により目尻の位置を指定したとしても、完全に一致する位置を複数回指定するのは困難である。そのため、同一の顔画像に対して同一の人物が目尻の位置を指定したとしても、各回に指定される目尻の位置は微妙に相違し得る。

[0022] したがって、複数の異なる人物が位置情報を入力する場合は勿論のこと、同一の人物が位置情報を入力する場合であっても、入力された位置情報は、常に一定の基準に従って対象の器官の位置を正確に示すとは限らない。すなわち、位置情報により示される器官の存在する位置は変動し得る。このようなノイズを含む位置情報を機械学習の正解データとして利用した場合、検出器は、顔画像に対して与えられた器官の位置だけではなく、このノイズをも再現するように学習してしまう。これに起因して、運転者が顔の器官を動かしていないにもかかわらず、学習済みの検出器から回答される器官の検出位置にぶれが生じてしまう。すなわち、上記図1に例示されるように、器官（目尻）を動かしていないにもかかわらず、機械学習に利用した位置情報に含まれるノイズを学習した影響によっても、各顔画像間で器官の検出位置が変動してしまうという問題点が生じ得る。

[0023] これに対して、当該構成に係る検出器生成装置は、第1顔画像に所定の幾何変換を適用することで第3顔画像を生成し、生成した第3顔画像を検出器に入力することで第3出力値を取得する。そして、当該構成に係る検出器生成装置は、上記第1検出位置と取得した第3出力値に基づいて特定される器官の第3検出位置との第3誤差の合計値を算出する。

[0024] 各第1顔画像に付与される位置情報が、常に一定の基準に従って対象の器官（図1の例では、目尻）の位置を正確に示していた場合には、第1誤差に基づく機械学習により構築される検出器は、一定の基準に対応する対象の器官の何らかの特徴を捉えて、当該対象の器官の位置を正確に検出することができるようになる。そのため、この検出器は、第1顔画像を幾何変換するこ

とで生成された第3顔画像に対しても、その特徴に基づいて、対象の器官の位置を正確に検出することができる。

[0025] 一方、各第1顔画像に付与される位置情報が一定の基準に従って対象の器官の位置を示すものではない場合、第1誤差に基づく機械学習により構築される検出器は、上記の特徴を正確には捉えることができず、対象の器官の検出する位置がぶれてしまう。そのため、この検出器は、第1顔画像を幾何変換することで生成された第3顔画像に対して、対象の器官の位置を正確に検出することができない可能性が高い。なお、各第1顔画像に付与される位置情報が一定の基準に従って対象の器官の位置を示すものではないケースとは、例えば、各第1顔画像において顔の位置がずれているにも関わらず、各顔画像に対して一定の位置を示す位置情報が各第1顔画像に付与されるケースである。

[0026] 上記第3誤差の合計値は、検出器による器官の検出結果のブレであって、このような顔画像における幾何的なノイズの影響によって生じるブレの損失に相当する。そこで、当該構成に係る検出器生成装置は、上記第1誤差及び第2誤差それぞれの合計値と共に、第3誤差の合計値が小さくなるように検出器を訓練する。これにより、画像ノイズ及び位置情報に含まれる幾何的なノイズによる影響にロバストな検出器を構築することができる。

[0027] なお、所定の幾何変換（幾何学的な変換）は、例えば、平行移動、回転移動、反転、拡大、縮小又はこれらの組み合わせであってよい。具体例として、所定の幾何変換は、平行移動及び回転移動を伴う変換であってよい。これにより、平行移動及び回転移動による影響にロバストな検出器を構築することができる。

[0028] 上記一側面に係る検出器生成装置において、前記各学習データセットの前記位置情報は、当該位置情報に組み合わせられる前記第1顔画像を、機械学習を実施済みの他の検出器に入力することで、当該他の検出器から得られる出力値に基づいて与えられてよい。当該構成によれば、他の検出器による検出結果を位置情報として利用するため、上記人為的なノイズが位置情報に含

まれるのを防止することができる。これにより、位置情報に含まれ得るノイズによる影響にロバストな検出器を構築することができる。

[0029] なお、他の検出器は、機械学習により最終的な検出器を得る過程で生成される暫定的な検出器であってよい。つまり、検出器生成装置は、学習データセットを利用した機械学習による検出器の構築する処理と、構築された検出器により位置情報を更新する処理とを交互に繰り返し実行してもよい。上記人為的なノイズによる影響にロバストな検出器を構築するために、この検出器を構築する処理と位置情報を更新する処理とを交互に実行する部分のみを抽出して、新たな形態に係る検出器生成装置が構成されてもよい。

[0030] 例えば、本発明の一側面に係る検出器生成装置は、車両を運転する被験者の顔の写る顔画像、及び前記顔画像に写る前記顔の器官の位置を示す第1位置情報の組み合わせによりそれぞれ構成される複数の第1学習データセットを取得する学習データ取得部と、前記複数の第1学習データセットを利用して、第1検出器の機械学習を実施することで、前記各第1学習データセットの前記顔画像が入力されると、入力された前記顔画像に組み合わせられた前記第1位置情報に対応する出力値を出力する第1検出器を構築する学習処理部と、を備える。そして、前記学習データ取得部は、前記各第1学習データセットの前記顔画像、及び構築された前記第1検出器に当該顔画像を入力することで、当該第1検出器から得られる出力値に基づいて与えられる第2位置情報の組み合わせによりそれぞれ構成される複数の第2学習データセットを更に取得する。また、前記学習処理部は、前記複数の第2学習データセットを利用して、第2検出器の機械学習を実施することで、前記各第2学習データセットの前記顔画像が入力されると、入力された前記顔画像に組み合わせられた前記第2位置情報に対応する出力値を出力する第2検出器を更に構築する。なお、第1検出器は、上記「他の検出器」、すなわち、暫定的な検出器の一例である。また、第2学習データセットの取得及び第2検出器の構築を繰り返す場合、あるステップで構築される第2検出器が、次のステップの第1検出器と利用される。そして、最後に構築される第2検出器が、上記

「検出器」、すなわち、最終的な検出器の一例である。

[0031] また、本発明の一側面に係るモニタリング装置は、車両を運転する運転者の顔の写る顔画像を取得するデータ取得部と、上記いずれかの形態に係る検出器生成装置による機械学習を実施済みの前記検出器に取得した前記顔画像を入力することで、前記運転者の前記顔の器官を検出した結果に対応する出力値を前記検出器から取得する検出部と、前記運転者の顔の器官を検出した結果に関する情報を出力する出力部と、を備える。当該構成に係るモニタリング装置によれば、ノイズの影響にロバストな検出器により、運転者の顔の器官を高精度に検出することができる。したがって、例えば、運転操作中における運転者の状態等を精度よく認識することができる。

[0032] また、上記各形態に係る検出器生成装置及びモニタリング装置は、車両の運転者の顔の器官を検出する場面だけではなく、例えば、生産ラインの作業者の顔の器官を検出する場面等、運転者以外の対象者の顔の器官を検出するあらゆる場面に適用されてよい。更に、上記各形態に係る検出器生成装置及びモニタリング装置は、対象者の顔の器官を検出する場面だけではなく、例えば、生産ラインを流れる製品の何らかの特徴（例えば、傷）の位置を検出する場面等、人物以外の対象物の写る画像から当該対象物の位置を検出するあらゆる場面に適用されてよい。

[0033] 例えば、本発明の一側面に係る検出器生成装置は、被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、を備える。そして、前記検出器の機械学習を実施することは、前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記器官の第1検出位置と当該第1顔画像に組み合わせられた前記位置情報により示さ

れる前記位置との第1誤差の合計値を算出するステップと、前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1顔画像に対して生成された前記第2顔画像に対して得られた前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、を含む。

[0034] また、例えば、本発明の一側面に係る検出器生成装置は、対象物の写る第1画像、及び前記第1画像に写る前記対象物の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、を備える。そして、前記検出器の機械学習を実施することは、前記各学習データセットの前記第1画像を前記検出器に入力することで、前記各学習データセットの前記第1画像に対する第1出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1画像に対して得られた前記第1出力値に基づいて特定される前記対象物の第1検出位置と当該第1画像に組み合わせられた前記位置情報により示される前記位置との第1誤差の合計値を算出するステップと、前記各学習データセットの前記第1画像にノイズを付与することで、前記各学習データセットの前記第1画像に対して第2画像を生成するステップと、前記各第2画像を前記検出器に入力することで、前記各第2画像に対する第2出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1画像に対して生成された前記第2画像に対して得られた前記第2出力値に基づいて特定される前記対象物の第2検出位置との第2誤差の合計値を算出するス

テップと、前記第 1 誤差の合計値及び前記第 2 誤差の合計値が小さくなるように前記検出器を訓練するステップと、を含む。

[0035] 例えば、本発明の一側面に係る検出器生成装置は、被験者の顔の写る顔画像、及び前記顔画像に写る前記顔の器官の位置を示す第 1 位置情報の組み合わせによりそれぞれ構成される複数の第 1 学習データセットを取得する学習データ取得部と、前記複数の第 1 学習データセットを利用して、第 1 検出器の機械学習を実施することで、前記各第 1 学習データセットの前記顔画像が入力されると、入力された前記顔画像に組み合わせられた前記第 1 位置情報に対応する出力値を出力する第 1 検出器を構築する学習処理部と、を備える。そして、前記学習データ取得部は、前記各第 1 学習データセットの前記顔画像、及び構築された前記第 1 検出器に当該顔画像を入力することで、当該第 1 検出器から得られる出力値に基づいて与えられる第 2 位置情報の組み合わせによりそれぞれ構成される複数の第 2 学習データセットを更に取得する。また、前記学習処理部は、前記複数の第 2 学習データセットを利用して、第 2 検出器の機械学習を実施することで、前記各第 2 学習データセットの前記顔画像が入力されると、入力された前記顔画像に組み合わせられた前記第 2 位置情報に対応する出力値を出力する第 2 検出器を更に構築する。

[0036] また、例えば、本発明の一側面に係る検出器生成装置は、対象物の写る画像、及び前記画像に写る前記対象物の位置を示す第 1 位置情報の組み合わせによりそれぞれ構成される複数の第 1 学習データセットを取得する学習データ取得部と、前記複数の第 1 学習データセットを利用して、第 1 検出器の機械学習を実施することで、前記各第 1 学習データセットの前記画像が入力されると、入力された前記画像に組み合わせられた前記第 1 位置情報に対応する出力値を出力する第 1 検出器を構築する学習処理部と、を備える。そして、前記学習データ取得部は、前記各第 1 学習データセットの前記画像、及び構築された前記第 1 検出器に当該画像を入力することで、当該第 1 検出器から得られる出力値に基づいて与えられる第 2 位置情報の組み合わせによりそれぞれ構成される複数の第 2 学習データセットを更に取得する。また、前記

学習処理部は、前記複数の第2学習データセットを利用して、第2検出器の機械学習を実施することで、前記各第2学習データセットの前記画像が入力されると、入力された前記画像に組み合わせられた前記第2位置情報に対応する出力値を出力する第2検出器を更に構築する。

[0037] なお、検出器を訓練することは、検出器のパラメータを調節することである。検出器のパラメータは、何らかの対象物を画像から検出する（出力値を得る）ための演算処理に利用される。検出器がニューラルネットワークにより構成される場合、パラメータは、例えば、各ニューロン間の結合の重み、各ニューロンの閾値等である。

[0038] 対象物の種類は、画像に写り得るものであれば特に限定されなくてよく、実施の形態に応じて適宜選択されてよい。人間の写る画像が第1画像として取得される場合、対象物は、例えば、顔の器官、顔そのもの、顔以外の身体部位等であってよい。また、生産ラインで生産される製品の写る画像が第1画像として取得される場合、対象物は、例えば、製品そのもの、製品に含まれる何らかの特徴（例えば、傷）等であってよい。

[0039] 各検出位置は、各画像に対して検出器が対象物を検出した位置である。第2誤差は、ノイズ画像（第2顔画像、第2画像）に対する対象物の検出位置と元の画像（第1顔画像、第1画像）に対する対象物の検出位置との差分に相当する。第3誤差は、幾何変換画像（第3顔画像、第3画像）に対する対象物の検出位置と元の画像（第1顔画像、第1画像）に対する対象物の検出位置との相対的な差分に相当する。この第3誤差は、幾何変換画像から検出した位置に幾何変換の逆変換を適用することで得られた位置と元の画像から検出した位置との差分を算出することで導出されてよい。あるいは、第3誤差は、幾何変換画像から検出した位置と元の画像から検出した位置に幾何変換を適用することで得られた位置との差分を算出することで導出されてもよい。つまり、「第1出力値に基づいて特定される器官の第1検出位置と第3出力値に基づいて特定される器官の第3検出位置との第3誤差」を計算することは、幾何変換画像から検出した位置に幾何変換の逆変換を適用すること

で得られた位置と元の画像から検出した位置との差分を算出すること、及び幾何変換画像から検出した位置と元の画像から検出した位置に幾何変換を適用することで得られた位置との差分を算出することを含んでよい。検出器は、「学習器」と称されてもよい。また、機械学習後の検出器は、「学習済みの学習器」と称されてもよい。位置情報の形式は、特に限定されなくてもよく、実施の形態に応じて適宜決定されてよい。位置情報は、例えば、対象物の座標値により構成されてもよいし、対象物の領域を示す情報により構成されてもよい。

[0040] 上記各形態に係る検出器生成装置及びモニタリング装置それぞれの別の態様として、本発明の一側面は、以上の各構成を実現する情報処理方法であってもよいし、プログラムであってもよいし、このようなプログラムを記憶した、コンピュータ等が読み取り可能な記憶媒体であってもよい。ここで、コンピュータ等が読み取り可能な記憶媒体とは、プログラム等の情報を、電氣的、磁氣的、光学的、機械的、又は、化学的作用によって蓄積する媒体である。また、本発明の一側面に係る検出システムは、上記いずれかの形態に係る検出器生成装置及びモニタリング装置により構成されてもよい。

[0041] 例えば、本発明の一側面に係る検出器生成方法は、コンピュータが、車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得するステップと、前記複数の学習データセットを利用して、検出器の機械学習を実施するステップと、を実行する、情報処理方法である。そして、前記検出器の機械学習を実施するステップは、前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記器官の第1検出位置と当該第1顔画像に組み合わせられた前記位置情報により示される前記位置との第1誤差の合計値を算出するステップと、前記各学習データセットの前記第1顔画

像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1顔画像に対して生成された前記第2顔画像に対して得られた前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、を含む。

[0042] また、例えば、本発明の一側面に係る検出器生成プログラムは、コンピュータに、車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得するステップと、前記複数の学習データセットを利用して、検出器の機械学習を実施するステップと、を実行させるためのプログラムである。そして、前記検出器の機械学習を実施するステップは、前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記器官の第1検出位置と当該第1顔画像に組み合わせられた前記位置情報により示される前記位置との第1誤差の合計値を算出するステップと、前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、前記各学習データセットの前記第1顔画像に対して得られた前記第1出力値に基づいて特定される前記第1検出位置と当該第1顔画像に対して生成された前記第2顔画像に対して得られた前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の

合計値を算出するステップと、前記第 1 誤差の合計値及び前記第 2 誤差の合計値が小さくなるように前記検出器を訓練するステップと、を含む。

発明の効果

[0043] 本発明によれば、ノイズの影響にロバストで、対象物をより高精度に検出可能な検出器を生成することができる。

図面の簡単な説明

[0044] [図1]図 1 は、従来の検出器に生じる問題点を説明するための図である。

[図2]図 2 は、本発明が適用される場面の一例を模式的に例示する。

[図3]図 3 は、実施の形態に係る検出器生成装置のハードウェア構成の一例を模式的に例示する。

[図4]図 4 は、実施の形態に係るモニタリング装置のハードウェア構成の一例を模式的に例示する。

[図5]図 5 は、実施の形態に係る検出器生成装置のソフトウェア構成の一例を模式的に例示する。

[図6]図 6 は、実施の形態に係るモニタリング装置のソフトウェア構成の一例を模式的に例示する。

[図7]図 7 は、実施の形態に係る検出器生成装置の処理手順の一例を例示する。

[図8A]図 8 A は、実施の形態に係る機械学習の処理手順の一例を例示する。

[図8B]図 8 B は、実施の形態に係る機械学習の処理手順の一例を例示する。

[図9]図 9 は、機械学習を繰り返し実施する場面の一例を模式的に例示する。

[図10]図 1 0 は、実施の形態に係るモニタリング装置の処理手順の一例を例示する。

[図11]図 1 1 は、変形例に係る検出器生成装置のソフトウェア構成の一例を模式的に例示する。

[図12]図 1 2 は、本発明が適用される場面のその他の例を模式的に例示する。

[図13]図 1 3 は、本発明が適用される場面のその他の例を模式的に例示する。

。

[図14]図14は、変形例に係る検出器生成装置のソフトウェア構成の一例を模式的に例示する。

[図15]図15は、変形例に係る検出器生成装置のソフトウェア構成の一例を模式的に例示する。

発明を実施するための形態

[0045] 以下、本発明の一側面に係る実施の形態（以下、「本実施形態」とも表記する）を、図面に基づいて説明する。ただし、以下で説明する本実施形態は、あらゆる点において本発明の例示に過ぎない。本発明の範囲を逸脱することなく種々の改良や変形を行うことができることは言うまでもない。つまり、本発明の実施にあたって、実施形態に応じた具体的構成が適宜採用されてもよい。なお、本実施形態において登場するデータを自然言語により説明しているが、より具体的には、コンピュータが認識可能な疑似言語、コマンド、パラメータ、マシン語等で指定される。

[0046] §1 適用例

まず、図2を用いて、本発明が適用される場面の一例について説明する。図2は、本実施形態に係る検出システム100の適用場面の一例を模式的に例示する。図2の例では、車両を運転する運転者の顔を撮影し、得られる顔画像から顔の器官を検出する場面を想定している。顔画像は、本発明の「画像」の一例であり、顔の器官は、本発明の「対象物」の一例である。しかしながら、本発明の適用対象は、このような例に限定されなくてもよく、何らかの対象物を画像から検出するあらゆる場面に適用可能である。

[0047] 図2に示されるとおり、本実施形態に係る検出システム100は、ネットワークを介して互いに接続される検出器生成装置1及びモニタリング装置2を備えており、顔の器官を検出するための検出器を生成し、生成した検出器により運転者の顔の器官を検出するように構成される。検出器生成装置1及びモニタリング装置2の間のネットワークの種類は、例えば、インターネット、無線通信網、移動通信網、電話網、専用網等から適宜選択されてよい。

[0048] 本実施形態に係る検出器生成装置 1 は、機械学習を実施することにより、顔画像から顔の器官を検出するための検出器 5 を構築するように構成されたコンピュータである。具体的には、検出器生成装置 1 は、車両を運転する被験者 T の顔の写る第 1 顔画像 6 1 1、及び第 1 顔画像 6 1 1 に写る顔の器官の位置 A 1 を示す位置情報 6 1 2 の組み合わせによりそれぞれ構成される複数の学習データセット 6 1 を取得する。第 1 顔画像 6 1 1 は、例えば、被験者 T の顔を撮影可能な車両内の位置に設置されたカメラ 3 1 により得られる。第 1 顔画像 6 1 1 は、機械学習における訓練データ（入力データ）に対応し、位置情報 6 1 2 は、正解データ（教師データ）に対応する。検出器生成装置 1 は、このような複数の学習データセット 6 1 を利用して、検出器 5 の機械学習を実施する。

[0049] 具体的に、機械学習の処理として、検出器生成装置 1 は、まず、各学習データセット 6 1 の第 1 顔画像 6 1 1 を検出器 5 に入力することで、各学習データセット 6 1 の第 1 顔画像 6 1 1 に対する第 1 出力値を検出器 5 から取得する。検出器 5 は、顔画像の入力を受け付け、入力された顔画像に写る顔の器官の位置を検出した結果に対応する出力値を出力するように構成される。つまり、第 1 出力値は、第 1 顔画像 6 1 1 に写る被験者 T の顔の器官の位置を検出器 5 が検出した結果に対応する。このような検出器 5 は、機械学習を実施可能な学習モデルにより構成される。本実施形態では、検出器 5 は、後述するニューラルネットワークにより構成される。

[0050] 次に、検出器生成装置 1 は、各学習データセット 6 1 の第 1 顔画像 6 1 1 に対して得られた第 1 出力値に基づいて特定される器官の第 1 検出位置 V 1 と当該第 1 顔画像 6 1 1 に組み合わせられた位置情報 6 1 2 により示される位置 A 1 との第 1 誤差 7 1 の合計値を算出する。つまり、検出器生成装置 1 は、各学習データセット 6 1 について、第 1 顔画像 6 1 1 に対して得られた第 1 検出位置 V 1、及び対応する位置情報 6 1 2 により示される位置 A 1 の差分を第 1 誤差 7 1 として算出し、算出した各第 1 誤差 7 1 を足し合わせることで、第 1 誤差 7 1 の合計値を算出する。

[0051] また、検出器生成装置 1 は、各学習データセット 6 1 の第 1 顔画像 6 1 1 にノイズを付与することで、各学習データセット 6 1 の第 1 顔画像 6 1 1 に対して第 2 顔画像 6 3 を生成する。続いて、検出器生成装置 1 は、生成した各第 2 顔画像 6 3 を検出器 5 に入力することで、各第 2 顔画像 6 3 に対する第 2 出力値を検出器 5 から取得する。更に、検出器生成装置 1 は、各学習データセット 6 1 の第 1 顔画像 6 1 1 に対して得られた第 1 出力値に基づいて特定される第 1 検出位置 V 1 と当該第 1 顔画像 6 1 1 に対して生成された第 2 顔画像 6 3 に対して得られた第 2 出力値に基づいて特定される器官の第 2 検出位置 V 2 との第 2 誤差 7 2 の合計値を算出する。つまり、検出器生成装置 1 は、各学習データセット 6 1 について、第 1 顔画像 6 1 1 に対して得られた第 1 検出位置 V 1、及び対応する第 2 顔画像 6 3 に対して得られた第 2 検出位置 V 2 の差分を第 2 誤差 7 2 として算出し、算出した各第 2 誤差 7 2 を足し合わせることで、第 2 誤差 7 2 の合計値を算出する。

[0052] そして、検出器生成装置 1 は、第 1 誤差 7 1 の合計値及び第 2 誤差 7 2 の合計値が小さくなるように検出器 5 を訓練する。検出器 5 を訓練することは、詳細には、検出器 5 のパラメータを調整することである。検出器 5 のパラメータは、何らかの対象物を画像から検出する（出力値を得る）ための演算処理に利用されるものである。本実施形態では、検出器 5 は後述するニューラルネットワークにより構成される。そのため、このパラメータは、例えば、各ニューロン間の結合の重み、各ニューロンの閾値等である。この検出器 5 のパラメータの調整には、例えば、誤差逆伝播法等の公知の学習アルゴリズムが用いられてよい。

[0053] 第 1 誤差 7 1 の合計値は、検出器 5 により第 1 顔画像 6 1 1 から器官を検出した結果と位置情報 6 1 2 により指定される器官の位置の正解との間の損失に相当する。また、第 2 誤差 7 2 の合計値は、検出器 5 による器官の検出結果のブレであって、第 1 顔画像 6 1 1 におけるノイズの有無によって生じるブレの損失に相当する。そのため、上記機械学習により、検出器生成装置 1 は、位置情報 6 1 2 により示される正解と一致するように第 1 顔画像 6 1

1 から器官を検出し、かつその検出結果がノイズによる影響を受け難い、ように訓練された（学習済みの）検出器 5 を構築（生成）することができる。

[0054] 一方、本実施形態に係るモニタリング装置 2 は、検出器生成装置 1 により構築された学習済みの検出器 5 を利用して、車両を運転する運転者 D の顔の器官を顔画像から検出するように構成されたコンピュータである。具体的には、モニタリング装置 2 は、車両を運転する運転者 D の顔の写る顔画像を取得する。顔画像は、例えば、運転者 D の顔を撮影可能な車両内の位置に設置されたカメラ 4 1 により得られる。

[0055] 続いて、モニタリング装置 2 は、検出器生成装置 1 による機械学習を実施済みの検出器 5 に取得した顔画像を入力することで、運転者 D の顔の器官を検出した結果に対応する出力値を検出器 5 から取得する。そして、モニタリング装置 2 は、運転者 D の顔の器官を検出した結果に関する情報を出力する。例えば、モニタリング装置 2 は、検出した顔の器官の位置に基づいて、視線方向等の運転者 D の状態を推定することができる。そこで、モニタリング装置 2 は、検出結果に関する情報として、検出結果に基づいて運転者 D の状態を推定した結果を出力してもよい。これにより、モニタリング装置 2 は、運転者 D の状態を監視することができる。

[0056] 上記のとおり、検出器 5 は、ノイズの影響にロバストに構築される。そのため、モニタリング装置 2 は、この検出器 5 を利用することで、運転者 D の顔の器官を高精度に検出することができる。したがって、モニタリング装置 2 は、例えば、運転操作中における運転者 D の視線方向等の状態を精度よく推定することができる。

[0057] なお、図 2 の例では、検出器生成装置 1 及びモニタリング装置 2 は別個のコンピュータである。しかしながら、検出システム 100 の構成は、このような例に限定されなくてもよい。検出器生成装置 1 及びモニタリング装置 2 は一体のコンピュータで構成されてもよい。また、検出器生成装置 1 及びモニタリング装置 2 はそれぞれ複数台のコンピュータにより構成されてもよい。更に、検出器生成装置 1 及びモニタリング装置 2 はネットワークに接続さ

れていなくてもよい。この場合、検出器生成装置 1 及びモニタリング装置 2 の間のデータのやりとりは、不揮発メモリ等の記憶媒体を介して行われてもよい。

[0058] また、説明の便宜のため、図 2 では、検出の対象となる顔の器官の一例として、目（詳細には、目尻）を例示している。しかしながら、検出の対象となる顔の器官は、目に限定されなくてもよい。検出の対象となる顔の器官は、例えば、目、鼻、口、眉、顎及びこれらの組み合わせから選択されてよい。また、検出の対象となる器官には、顔の輪郭が含まれてもよい。後述する図 5 及び図 11 においても同様である。

[0059] § 2 構成例

[ハードウェア構成]

<検出器生成装置>

次に、図 3 を用いて、本実施形態に係る検出器生成装置 1 のハードウェア構成の一例について説明する。図 3 は、本実施形態に係る検出器生成装置 1 のハードウェアの一例を模式的に例示する。

[0060] 図 3 に示されるとおり、本実施形態に係る検出器生成装置 1 は、制御部 11、記憶部 12、通信インタフェース 13、入力装置 14、出力装置 15、及びドライブ 16 が電氣的に接続されたコンピュータである。なお、図 3 では、通信インタフェースを「通信 I/F」と記載している。

[0061] 制御部 11 は、ハードウェアプロセッサである CPU (Central Processing Unit)、RAM (Random Access Memory)、ROM (Read Only Memory) 等を含み、プログラム及び各種データに基づいて情報処理を実行するように構成される。記憶部 12 は、メモリの一例であり、例えば、ハードディスクドライブ、ソリッドステートドライブ等で構成される。本実施形態では、記憶部 12 は、検出器生成プログラム 81、複数の学習データセット 61、学習結果データ 121 等の各種情報を記憶する。

[0062] 検出器生成プログラム 81 は、検出器生成装置 1 に、後述する機械学習の情報処理（図 7、図 8A、図 8B）を実行させ、学習済みの検出器 5 を構築

させるためのプログラムである。検出器生成プログラム 8 1 は、この情報処理の一連の命令を含む。各学習データセット 6 1 は、上記第 1 顔画像 6 1 1 及び位置情報 6 1 2 により構成される。学習結果データ 1 2 1 は、各学習データセット 6 1 を利用した機械学習により構築された学習済みの検出器 5 の設定を行うためのデータである。学習結果データ 1 2 1 は、検出器生成プログラム 8 1 の実行結果として生成される。詳細は後述する。

[0063] 通信インタフェース 1 3 は、例えば、有線 LAN (Local Area Network) モジュール、無線 LAN モジュール等であり、ネットワークを介した有線又は無線通信を行うためのインタフェースである。検出器生成装置 1 は、この通信インタフェース 1 3 を利用することで、ネットワークを介したデータ通信を他の情報処理装置（例えば、モニタリング装置 2）と行うことができる。

[0064] 入力装置 1 4 は、例えば、マウス、キーボード等の入力を行うための装置である。また、出力装置 1 5 は、例えば、ディスプレイ、スピーカ等の出力を行うための装置である。オペレータは、入力装置 1 4 及び出力装置 1 5 を利用することで、検出器生成装置 1 を操作することができる。

[0065] ドライブ 1 6 は、例えば、CD ドライブ、DVD ドライブ等であり、記憶媒体 9 1 に記憶されたプログラムを読み込むためのドライブ装置である。ドライブ 1 6 の種類は、記憶媒体 9 1 の種類に応じて適宜選択されてよい。上記検出器生成プログラム 8 1 及び複数の学習データセット 6 1 の少なくともいずれかは、この記憶媒体 9 1 に記憶されていてもよい。

[0066] 記憶媒体 9 1 は、コンピュータその他装置、機械等が、記録されたプログラム等の情報を読み取り可能なように、当該プログラム等の情報を、電氣的、磁氣的、光学的、機械的又は化学的作用によって蓄積する媒体である。検出器生成装置 1 は、この記憶媒体 9 1 から、上記検出器生成プログラム 8 1 及び複数の学習データセット 6 1 の少なくともいずれかを取得してもよい。

[0067] ここで、図 3 では、記憶媒体 9 1 の一例として、CD、DVD 等のディスク型の記憶媒体を例示している。しかしながら、記憶媒体 9 1 の種類は、デ

ディスク型に限定される訳ではなく、ディスク型以外であってもよい。ディスク型以外の記憶媒体として、例えば、フラッシュメモリ等の半導体メモリを挙げることができる。

[0068] なお、検出器生成装置 1 の具体的なハードウェア構成に関して、実施形態に応じて、適宜、構成要素の省略、置換及び追加が可能である。例えば、制御部 11 は、複数のハードウェアプロセッサを含んでもよい。ハードウェアプロセッサは、マイクロプロセッサ、FPGA (field-programmable gate array)、DSP (digital signal processor) 等で構成されてよい。記憶部 12 は、制御部 11 に含まれる RAM 及び ROM により構成されてもよい。通信インタフェース 13、入力装置 14、出力装置 15 及びドライブ 16 の少なくともいずれかは省略されてもよい。検出器生成装置 1 は、カメラ 31 と接続するための外部インタフェースを更に備えてもよい。この外部インタフェースは、モニタリング装置 2 の後述する外部インタフェース 24 と同様に構成されてよい。検出器生成装置 1 は、複数台のコンピュータで構成されてもよい。この場合、各コンピュータのハードウェア構成は、一致していてもよいし、一致していなくてもよい。また、検出器生成装置 1 は、提供されるサービス専用に設計された情報処理装置の他、汎用のサーバ装置、PC (Personal Computer) 等であってもよい。

[0069] <モニタリング装置>

次に、図 4 を用いて、本実施形態に係るモニタリング装置 2 のハードウェア構成の一例について説明する。図 4 は、本実施形態に係るモニタリング装置 2 のハードウェア構成の一例を模式的に例示する。

[0070] 図 4 に示されるとおり、本実施形態に係るモニタリング装置 2 は、制御部 21、記憶部 22、通信インタフェース 23、外部インタフェース 24、入力装置 25、出力装置 26、及びドライブ 27 が電氣的に接続されたコンピュータである。なお、図 4 では、通信インタフェース及び外部インタフェースをそれぞれ「通信 I/F」及び「外部 I/F」と記載している。

[0071] モニタリング装置 2 の制御部 21～通信インタフェース 23 及び入力装置

25～ドライブ27はそれぞれ、上記検出器生成装置1の制御部11～ドライブ16それぞれと同様に構成されてよい。すなわち、制御部21は、ハードウェアプロセッサであるCPU、RAM、ROM等を含み、プログラム及びデータに基づいて各種情報処理を実行するように構成される。記憶部22は、例えば、ハードディスクドライブ、ソリッドステートドライブ等で構成される。記憶部22は、モニタリングプログラム82、学習結果データ121等の各種情報を記憶する。

[0072] モニタリングプログラム82は、学習済みの検出器5を利用して、運転者Dの状態を監視する後述の情報処理（図10）をモニタリング装置2に実行させるためのプログラムである。モニタリングプログラム82は、この情報処理の一連の命令を含む。詳細は後述する。

[0073] 通信インタフェース23は、例えば、有線LANモジュール、無線LANモジュール等であり、ネットワークを介した有線又は無線通信を行うためのインタフェースである。モニタリング装置2は、この通信インタフェース23を利用することで、ネットワークを介したデータ通信を他の情報処理装置（例えば、検出器生成装置1）と行うことができる。

[0074] 外部インタフェース24は、例えば、USB（Universal Serial Bus）ポート、専用ポート等であり、外部装置と接続するためのインタフェースである。外部インタフェース24の種類及び数は、接続される外部装置の種類及び数に応じて適宜選択されてよい。本実施形態では、モニタリング装置2は、外部インタフェース24を介して、カメラ41に接続される。

[0075] カメラ41は、運転者Dの顔を撮影することで、顔画像を取得するのに利用される。カメラ41の種類及び配置場所は、特に限定されなくてもよく、実施の形態に応じて適宜決定されてよい。カメラ41には、例えば、デジタルカメラ、ビデオカメラ等の公知のカメラが利用されてよい。また、カメラ41は、例えば、少なくとも運転者Dの上半身を撮影範囲とするように、運転席の前方上方に配置されてよい。なお、カメラ41が通信インタフェースを備える場合、モニタリング装置2は、外部インタフェース24ではなく、

通信インタフェース 23 を介して、カメラ 41 に接続されてもよい。

[0076] 入力装置 25 は、例えば、マウス、キーボード等の入力を行うための装置である。また、出力装置 26 は、例えば、ディスプレイ、スピーカ等の出力を行うための装置である。運転者 D 等のオペレータは、入力装置 25 及び出力装置 26 を利用することで、モニタリング装置 2 を操作することができる。

[0077] ドライブ 27 は、例えば、CD ドライブ、DVD ドライブ等であり、記憶媒体 92 に記憶されたプログラムを読み込むためのドライブ装置である。上記モニタリングプログラム 82 及び学習結果データ 121 のうちの少なくともいずれかは、記憶媒体 92 に記憶されていてもよい。また、モニタリング装置 2 は、記憶媒体 92 から、上記モニタリングプログラム 82 及び学習結果データ 121 のうちの少なくともいずれかを取得してもよい。

[0078] なお、モニタリング装置 2 の具体的なハードウェア構成に関して、上記検出器生成装置 1 と同様に、実施の形態に応じて、適宜、構成要素の省略、置換及び追加が可能である。例えば、制御部 21 は、複数のハードウェアプロセッサを含んでもよい。ハードウェアプロセッサは、マイクロプロセッサ、FPGA、DSP 等で構成されてよい。記憶部 22 は、制御部 21 に含まれる RAM 及び ROM により構成されてもよい。通信インタフェース 23、外部インタフェース 24、入力装置 25、出力装置 26 及びドライブ 27 の少なくともいずれかは省略されてもよい。モニタリング装置 2 は、複数台のコンピュータで構成されてもよい。この場合、各コンピュータのハードウェア構成は、一致していてもよいし、一致していなくてもよい。また、モニタリング装置 2 は、提供されるサービス専用に設計された情報処理装置の他、汎用のサーバ装置、汎用のデスクトップ PC、ノート PC、タブレット PC、スマートフォンを含む携帯電話等が用いられてよい。

[0079] [ソフトウェア構成]

<検出器生成装置>

次に、図 5 を用いて、本実施形態に係る検出器生成装置 1 のソフトウェア

構成の一例について説明する。図5は、本実施形態に係る検出器生成装置1のソフトウェア構成の一例を模式的に例示する。

[0080] 検出器生成装置1の制御部11は、記憶部12に記憶された検出器生成プログラム81をRAMに展開する。そして、制御部11は、RAMに展開された検出器生成プログラム81をCPUにより解釈及び実行して、各構成要素を制御する。これによって、図5に示されるとおり、本実施形態に係る検出器生成装置1は、学習データ取得部111、学習処理部112、及び保存処理部113をソフトウェアモジュールとして備えるコンピュータとして動作する。すなわち、本実施形態では、検出器生成装置1の各ソフトウェアモジュールは、制御部11（CPU）により実現される。

[0081] 学習データ取得部111は、車両を運転する被験者Tの顔の写る第1顔画像611、及び第1顔画像611に写る顔の器官の位置を示す位置情報612の組み合わせによりそれぞれ構成された複数の学習データセット61を取得する。対象となる器官は、例えば、目、鼻、口、眉、顎及びこれらの組み合わせから選択されてよい。また、対象となる器官には、顔の輪郭が含まれてもよい。

[0082] 学習処理部112は、複数の学習データセット61を利用して、検出器5の機械学習を実施する。具体的には、機械学習の処理として、学習処理部112は、まず、各学習データセット61の第1顔画像611を検出器5に入力することで、各学習データセット61の第1顔画像611に対する第1出力値を検出器5から取得する。続いて、学習処理部112は、各学習データセット61の第1顔画像611に対して得られた第1出力値に基づいて特定される器官の第1検出位置V1と当該第1顔画像611に組み合わせられた位置情報612により示される位置A1との第1誤差71の合計値を算出する。

[0083] また、学習処理部112は、各学習データセット61の第1顔画像611にノイズを付与することで、各学習データセット61の第1顔画像611に対して第2顔画像63を生成する。続いて、学習処理部112は、生成した

各第2顔画像63を検出器5に入力することで、各第2顔画像63に対する第2出力値を検出器5から取得する。そして、学習処理部112は、各学習データセット61の第1顔画像611に対して得られた第1出力値に基づいて特定される第1検出位置V1と当該第1顔画像611に対して生成された第2顔画像63に対して得られた第2出力値に基づいて特定される器官の第2検出位置V2との第2誤差72の合計値を算出する。

[0084] 更に、本実施形態では、学習処理部112は、各学習データセット61の第1顔画像611に所定の幾何変換を適用することで、各学習データセット61の第1顔画像611に対して第3顔画像66を生成する。所定の幾何変換（幾何的な変換）は、例えば、平行移動、回転移動、反転、拡大、縮小又はこれらの組み合わせであってよい。具体例として、所定の幾何変換は、平行移動及び回転移動を伴う変換であってよい。図5の例では、平行移動及び回転移動を伴う変換を第1顔画像611に適用した場面を例示している。続いて、学習処理部112は、生成した各第3顔画像66を検出器5に入力することで、各第3顔画像66に対する第3出力値を検出器5から取得する。そして、学習処理部112は、各学習データセット61の第1顔画像611に対して得られた第1出力値に基づいて特定される第1検出位置と当該第1顔画像611に対して生成された第3顔画像66に対して得られた第3出力値に基づいて特定される器官の第3検出位置との第3誤差73の合計値を算出する。

[0085] ここで、第3顔画像66は、第1顔画像611に幾何変換を適用することで得られるため、第3顔画像66の座標軸と第1顔画像611の座標軸とは幾何変換の分だけずれている。そのため、第1顔画像611に対する器官の検出結果と第3顔画像66に対する器官の検出結果とを比較するため、学習処理部112は、第1顔画像611の座標軸と第3顔画像66の座標軸とを一致させる変換処理を実行する。

[0086] 図5の例では、まず、学習処理部112は、検出器5から取得した第3出力値に基づいて、第3顔画像66の座標軸における器官の第3検出位置V3

0を特定する。次に、学習処理部112は、特定した検出位置V30に所定の幾何変換の逆変換を適用し、第3顔画像66の座標軸を第1顔画像611の座標軸に一致させて、第1顔画像611の座標軸における第3検出位置V31を特定する。そして、学習処理部112は、第3検出位置V31と第1検出位置V1との差分を第3誤差73として算出し、各学習データセット61についての第3誤差73を足し合わせることで、第3誤差73の合計値を算出する。

[0087] しかしながら、第1検出位置と第3検出位置との第3誤差の合計値を算出する方法は、このような例に限定されなくてもよい。学習処理部112は、第1顔画像611の座標軸における第1検出位置V1に所定の逆変換を適用することで、第1顔画像611の座標軸を第3顔画像66の座標軸に一致させてもよい。そして、学習処理部112は、第3顔画像66の座標軸における第1検出位置と第3検出位置V30との差分を第3誤差73として算出してもよい。

[0088] 以上の各処理によって、第1誤差71、第2誤差72、及び第3誤差73それぞれの合計値が算出される。学習処理部112は、第1誤差71の合計値、第2誤差72の合計値、及び第3誤差73の合計値が小さくなるように検出器5を訓練する。これにより、学習処理部112は、運転者の顔の器官を検出するための学習済みの検出器5を構築する。

[0089] (検出器)

次に、検出器5の構成の一例について説明する。図5に示されるとおり、本実施形態に係る検出器5は、ニューラルネットワークにより構成されている。具体的には、検出器5は、いわゆる深層学習に用いられる多層構造のニューラルネットワークにより構成されており、入力層51、中間層（隠れ層）52、及び出力層53を備えている。

[0090] なお、図5の例では、検出器5を構成するニューラルネットワークは、1層の中間層52を備えており、入力層51の出力が中間層52に入力され、中間層52の出力が出力層53に入力されている。ただし、検出器5の構成

は、このような例に限定されなくてもよい。中間層 5 2 の数は、1 層に限られなくてもよい。検出器 5 は、2 層以上の中間層 5 2 を備えてもよい。

[0091] 各層 5 1 ～ 5 3 は、1 又は複数のニューロンを備えている。例えば、入力層 5 1 のニューロンの数は、第 1 顔画像 6 1 1 の画素数に応じて設定されてよい。中間層 5 2 のニューロンの数は、実施の形態に応じて適宜設定されてよい。また、出力層 5 3 のニューロンの数は、検出する器官の数、位置表現等に応じて設定されてよい。

[0092] 隣接する層のニューロン同士は適宜結合され、各結合には重み（結合荷重）が設定されている。図 5 の例では、各ニューロンは、隣接する層の全てのニューロンと結合されている。しかしながら、ニューロンの結合は、このような例に限定されなくてもよく、実施の形態に応じて適宜設定されてよい。

[0093] 各ニューロンには閾値が設定されており、基本的には、各入力と各重みとの積の和が閾値を超えているか否かによって各ニューロンの出力が決定される。つまり、これらの各ニューロン間の結合の重み及び各ニューロンの閾値は、演算処理に利用される検出器 5 のパラメータの一例である。学習処理部 1 1 2 は、各顔画像（6 1 1、6 3、6 6）を入力層 5 1 に入力し、これらのパラメータを利用して、検出器 5 の演算処理を実行する。

[0094] この演算処理の結果、学習処理部 1 1 2 は、各顔画像（6 1 1、6 3、6 6）に対する器官の検出結果として、出力層 5 3 から各出力値を取得する。続いて、学習処理部 1 1 2 は、上記のとおり、取得した各出力値に基づいて、各誤差 7 1 ～ 7 3 の合計値を算出する。そして、学習処理部 1 1 2 は、上記訓練処理として、各誤差 7 1 ～ 7 3 の合計値が小さくなるように、検出器 5 のパラメータを調節する。これにより、運転者の顔の器官を検出するための学習済みの検出器 5 が構築される。

[0095] 保存処理部 1 1 3 は、構築された学習済みの検出器 5 の構成（例えば、ニューラルネットワークの層数、各層におけるニューロンの個数、ニューロン同士の結合関係、各ニューロンの伝達関数）、及び演算パラメータ（例えば、各ニューロン間の結合の重み、各ニューロンの閾値）を示す情報を学習結

果データ１２１として記憶部１２に保存する。

[0096] <モニタリング装置>

次に、図６を用いて、本実施形態に係るモニタリング装置２のソフトウェア構成の一例について説明する。図６は、本実施形態に係るモニタリング装置２のソフトウェア構成の一例を模式的に例示する。

[0097] モニタリング装置２の制御部２１は、記憶部２２に記憶されたモニタリングプログラム８２をＲＡＭに展開する。そして、制御部２１は、ＲＡＭに展開されたモニタリングプログラム８２をＣＰＵにより解釈及び実行して、各構成要素を制御する。これによって、図６に示されるとおり、本実施形態に係るモニタリング装置２は、データ取得部２１１、検出部２１２、及び出力部２１３をソフトウェアモジュールとして備えるコンピュータとして動作する。すなわち、本実施形態では、モニタリング装置２の各ソフトウェアモジュールも、上記検出器生成装置１と同様に、制御部２１（ＣＰＵ）により実現される。

[0098] データ取得部２１１は、車両を運転する運転者Ｄの顔の写る顔画像２２１を取得する。例えば、データ取得部２１１は、カメラ４１により運転者Ｄの顔を撮影することで、顔画像２２１を取得する。検出部２１２は、学習結果データ１２１を保持することで、学習済みの検出器５を含んでいる。具体的には、検出部２１２は、学習結果データ１２１を参照し、学習済みの検出器５の設定を行う。そして、検出部２１２は、取得された顔画像２２１を検出器５に入力することで、運転者Ｄの顔の器官を顔画像２２１より検出した結果に対応する出力値を検出器５から取得する。出力部２１３は、運転者Ｄの顔の器官を検出した結果に関する情報を出力する。

[0099] <その他>

検出器生成装置１及びモニタリング装置２の各ソフトウェアモジュールに関しては後述する動作例で詳細に説明する。なお、本実施形態では、検出器生成装置１及びモニタリング装置２の各ソフトウェアモジュールがいずれも汎用のＣＰＵによって実現される例について説明している。しかしながら、

以上のソフトウェアモジュールの一部又は全部が、１又は複数の専用のプロセッサにより実現されてもよい。また、検出器生成装置１及びモニタリング装置２それぞれのソフトウェア構成に関して、実施形態に応じて、適宜、ソフトウェアモジュールの省略、置換及び追加が行われてもよい。

[0100] § 3 動作例

[検出器生成装置]

次に、図７を用いて、検出器生成装置１の動作例について説明する。図７は、本実施形態に係る検出器生成装置１の処理手順の一例を示すフローチャートである。以下で説明する処理手順は、本発明の「検出器生成方法」の一例である。ただし、以下で説明する処理手順は一例に過ぎず、各処理は可能な限り変更されてよい。また、以下で説明する処理手順について、実施の形態に応じて、適宜、ステップの省略、置換、及び追加が可能である。

[0101] (ステップＳ１０１)

ステップＳ１０１では、制御部１１は、学習データ取得部１１１として動作し、第１顔画像６１１及び位置情報６１２の組み合わせによりそれぞれ構成された複数の学習データセット６１を取得する。

[0102] 複数件の学習データセット６１を取得する方法は、特に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。例えば、カメラ３１を搭載した車両及び被験者Ｔを用意し、車両を運転する被験者Ｔをカメラにより様々な条件で撮影することで、被験者Ｔの顔の写る複数の第１顔画像６１１を取得することができる。用意する車両及び被験者Ｔの数は、実施の形態に応じて適宜決定されてよい。そして、得られた各第１顔画像６１１に対して、各第１顔画像６１１に写る対象の器官の位置を示す位置情報６１２を組み合わせることで、各学習データセット６１を生成することができる。対象の器官は、例えば、目、鼻、口、眉、顎及びこれらの組み合わせから選択されてよい。また、対象の器官には、顔の輪郭が含まれてもよい。

[0103] この各学習データセット６１は、自動的に生成されてもよいし、手動的に

生成されてもよい。また、各学習データセット 6 1 の生成は、検出器生成装置 1 により行われてもよいし、検出器生成装置 1 以外の他のコンピュータにより行われてもよい。

[0104] 各学習データセット 6 1 が検出器生成装置 1 により生成される場合、制御部 1 1 は、外部インタフェース、ネットワーク、記憶媒体 9 1 等を介して、各第 1 顔画像 6 1 1 をカメラ 3 1 から取得する。次に、制御部 1 1 は、各第 1 顔画像 6 1 1 に付与する位置情報 6 1 2 を生成する。位置情報 6 1 2 を手動的に生成する場合、制御部 1 1 は、オペレータによる入力装置 1 4 を介する器官の位置の指定を受け付ける。続いて、制御部 1 1 は、オペレータによる入力装置 1 4 の操作に応じて、指定された器官の位置を示す位置情報 6 1 2 を生成する。他方、位置情報 6 1 2 を自動的に生成する場合、制御部 1 1 は、各第 1 顔画像 6 1 1 から対象の器官の位置を検出する。この検出には、エッジ検出、パターンマッチング等の公知の画像処理方法が用いられてもよいし、後述する暫定的に生成された検出器が用いられてもよい。続いて、制御部 1 1 は、検出された器官の位置を示す位置情報 6 1 2 を生成する。そして、制御部 1 1 は、生成した位置情報 6 1 2 を第 1 顔画像 6 1 1 に組み合わせることで、各学習データセット 6 1 を生成する。これにより、制御部 1 1 は、複数の学習データセット 6 1 を取得してもよい。

[0105] 一方、各学習データセット 6 1 が他のコンピュータにより生成される場合、本ステップ S 1 0 1 では、制御部 1 1 は、ネットワーク、記憶媒体 9 1 等を介して、他のコンピュータにより生成された複数の学習データセット 6 1 を取得してもよい。他のコンピュータでは、上記検出器生成装置 1 と同様の方法で、各学習データセット 6 1 が生成されてよい。

[0106] なお、第 1 顔画像 6 1 1 及び位置情報 6 1 2 のデータ形式は、実施の形態に応じて適宜選択されてよい。位置情報 6 1 2 は、座標値等により器官の位置を点で示してもよいし、器官の位置を領域で示してもよい。また、取得する学習データセット 6 1 の件数は、特に限定されなくてもよく、検出器 5 の機械学習を実施可能な程度に適宜決定されてよい。複数の学習データセット

61を取得すると、制御部11は、次のステップS102に処理を進める。

[0107] (ステップS102)

ステップS102では、制御部11は、学習処理部112として動作し、ステップS101により取得した複数の学習データセット61を利用して、検出器5の機械学習を実施する。機械学習の処理手順は後述する。これにより、学習済みの検出器5を構築すると、制御部11は、次のステップS103に処理を進める。

[0108] (ステップS103)

ステップS103では、制御部11は、保存処理部113として動作し、機械学習により構築された検出器5の構成及び演算パラメータを示す情報を学習結果データ121として生成する。そして、制御部11は、生成した学習結果データ121を記憶部12に保存する。これにより、制御部11は、本動作例に係る処理を終了する。

[0109] なお、学習結果データ121の保存先は、記憶部12に限られなくてもよい。制御部11は、例えば、NAS (Network Attached Storage) 等のデータサーバに学習結果データ121を格納してもよい。また、学習済みの検出器5を構築した後、制御部11は、生成した学習結果データ121を任意のタイミングでモニタリング装置2に転送してもよい。

[0110] モニタリング装置2は、検出器生成装置1から転送を受け付けることで学習結果データ121を取得してもよいし、検出器生成装置1又はデータサーバにアクセスすることで学習結果データ121を取得してもよい。また、学習結果データ121は、モニタリング装置2に予め組み込まれてもよい。

[0111] <機械学習の処理>

次に、図8A及び図8Bを用いて、上記ステップS102の処理の一例を詳細に説明する。図8A及び図8Bは、本実施形態に係る検出器生成装置1による機械学習の処理手順の一例を示すフローチャートである。本実施形態に係るステップS102の処理は、以下のステップS201～S210の処理を含む。ただし、以下で説明する処理手順は一例に過ぎず、各処理は可能

な限り変更されてよい。また、以下で説明する処理手順について、実施の形態に応じて、適宜、ステップの省略、置換、及び追加が可能である。

[0112] ステップS201の処理を開始する前に、制御部11は、機械学習を実施する対象となるニューラルネットワーク（学習前の検出器5）を用意する。用意するニューラルネットワークの構成及び各パラメータは、テンプレートにより与えられてもよいし、オペレータの入力により与えられてもよい。また、再学習を実施する場合には、制御部11は、再学習の対象となる学習結果データ121に基づいて、ニューラルネットワークを用意してもよい。これにより、学習前の検出器5の用意が完了すると、制御部11は、ステップS201の処理を開始する。

[0113] （ステップS201）

ステップS201では、制御部11は、各学習データセット61の第1顔画像611を検出器5に入力し、検出器5の演算処理を実行する。具体的には、制御部11は、各第1顔画像611を検出器5の入力層51に入力し、入力側から順に各層51～53に含まれる各ニューロンの発火判定を行う。これにより、制御部11は、各第1顔画像611に対する第1出力値を出力層53から取得する。各第1出力値は、現段階における検出器5が各第1顔画像611から器官の位置を検出した結果に対応する。各第1出力値を取得すると、制御部11は、次のステップS202に処理を進める。

[0114] なお、検出器5の出力値は、器官を検出した位置を示し得るものであれば、検出器5の出力値の形式は、特に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。例えば、検出器5の出力値は、座標値等のように、器官の位置を直接的に示すものであってもよい。また、例えば、検出器5の出力値は器官の位置を間接的に示し、検出器5の出力値に何らかの演算処理を適用することにより、器官の位置を示す情報が得られてもよい。いずれの形式であっても、検出器5の出力値は同様に扱うことができる。そのため、以下では、説明の便宜上、検出器5の出力値は、器官の位置を直接示すものとして取り扱うことにする。

[0115] (ステップS202)

ステップS202では、制御部11は、各第1出力値に基づいて特定される器官の第1検出位置V1と対応する位置情報612により示される位置A1との第1誤差71の合計値を算出する。具体的には、制御部11は、各学習データセット61について、各第1顔画像611に対して得られた第1検出位置V1と対応する位置情報612により示される位置A1との差分を第1誤差71として算出する。そして、制御部11は、算出した各第1誤差71を足し合わせることで、第1誤差71の合計値を算出する。

[0116] なお、第1検出位置V1と位置A1との差分は、器官の位置を指定する形式に応じて適宜算出されてよい。例えば、器官の位置が点で指定される場合、制御部11は、第1検出位置V1と位置A1との間の距離（ノルム）を差分として算出してもよい。また、例えば、器官の位置が領域で指定される場合、制御部11は、IoU（Intersection Over Union）等の指標を利用して、第1検出位置V1と位置A1との差分を算出してもよい。第1誤差71の合計値を算出すると、制御部11は、次のステップS203に処理を進める。

[0117] (ステップS203)

ステップS203では、制御部11は、各学習データセット61の第1顔画像611にノイズを付与することで、各第2顔画像63を生成する。各第1顔画像611に付与するノイズの種類は、特に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。ノイズには、例えば、ガウシアンノイズ、ポアソンノイズ、一様乱数によるノイズ等が用いられてよい。例えば、制御部11は、ランダムなガウシアンノイズを各第1顔画像611に付与してもよい。また、例えば、制御部11は、局所平均フィルタ等のフィルタを各第1顔画像611に適用することで、各第1顔画像611にノイズを付与してもよい。各第1顔画像611に付与するノイズの種類は、同じであってもよいし、異なってもよい。また、例えば、制御部11は、敵対的な例（Adversarial example）を形成するように、検出器5による検出が変化し

やすい位置にノイズを付与してもよい。検出が変化しやすいか否かは、例えば、ノイズを付与したことによる検出位置の変化量と閾値との比較に基づいて判定することができる。各第2顔画像63を生成すると、制御部11は、次のステップS204に処理を進める。

[0118] (ステップS204)

ステップS204では、制御部11は、生成した各第2顔画像63を検出器5に入力し、検出器5の演算処理を実行する。この検出器5の演算処理は、上記ステップS201と同様であってよい。これにより、制御部11は、各第2顔画像63に対する第2出力値を検出器5の出力層53から取得する。各第2出力値は、現段階における検出器5が各第2顔画像63から器官の位置を検出した結果に対応する。各第2出力値を取得すると、制御部11は、次のステップS205に処理を進める。

[0119] (ステップS205)

ステップS205では、制御部11は、各第1出力値に基づいて特定される器官の第1検出位置V1と対応する第2顔画像63に対して得られた第2出力値に基づいて特定される器官の第2検出位置V2との第2誤差72の合計値を算出する。具体的には、制御部11は、各学習データセット61について、各第1顔画像611に対して得られた第1検出位置V1と対応する第2顔画像63に対して得られた第2検出位置V2との差分を第2誤差72として算出する。この差分を算出する方法は、上記ステップS202と同様であってよい。そして、制御部11は、算出した各第2誤差72を足し合わせることで、第2誤差72の合計値を算出する。第2誤差72の合計値を算出すると、制御部11は、次のステップS206に処理を進める。

[0120] (ステップS206)

ステップS206では、制御部11は、各学習データセット61の第1顔画像611に所定の幾何変換を適用することで、各第3顔画像66を生成する。各第1顔画像611に適用する幾何変換は、例えば、平行移動、回転移動、反転、拡大、縮小又はこれらの組み合わせであってよい。具体例として

、各第1顔画像611に適用する幾何変換は、平行移動及び回転移動を伴う変換であってよい。各第3顔画像66に適用する幾何変換は、同じであってもよいし、異なってもよい。制御部11は、幾何変換のパラメータをランダムに変更しながら、各第3顔画像66を生成してもよい。幾何変換のパラメータは、例えば、各変換量（平行移動量、回転量等）である。各第3顔画像66を生成すると、制御部11は、次のステップS207に処理を進める。

[0121] (ステップS207)

ステップS207では、制御部11は、生成した各第3顔画像66を検出器5に入力し、検出器5の演算処理を実行する。この検出器5の演算処理は、上記ステップS201と同様であってよい。これにより、制御部11は、各第3顔画像66に対する第3出力値を検出器5の出力層53から取得する。各第3出力値は、現段階における検出器5が各第3顔画像66から器官の位置を検出した結果に対応する。各第3出力値を取得すると、制御部11は、次のステップS208に処理を進める。

[0122] (ステップS208)

ステップS208では、制御部11は、各第1出力値に基づいて特定される器官の第1検出位置と対応する第3顔画像66に対して得られた第3出力値に基づいて特定される器官の第3検出位置との第3誤差73の合計値を算出する。具体的には、制御部11は、各第1顔画像611に対して得られた第3出力値から導出される第3検出位置V30に幾何変換の逆変換を適用して、第3検出位置V31を算出する。そして、制御部11は、各学習データセット61について、各第1顔画像611に対して得られた第1検出位置V1と対応する第3顔画像66に対して得られた第3検出位置V31との差分を第3誤差73として算出する。この差分を算出する方法は、上記ステップS202と同様であってよい。また、上記のとおり、制御部11は、第1検出位置V1に幾何変換を適用することで導出される位置と第3検出位置V30との差分を第3誤差73として算出してもよい。そして、制御部11は、

算出した各第3誤差73を足し合わせることで、第3誤差73の合計値を算出する。第3誤差73の合計値を算出すると、制御部11は、次のステップS209に処理を進める。

[0123] (ステップS209)

ステップS209では、制御部11は、算出した各誤差71～73の合計値が閾値以下であるか否かを判定する。閾値は、顔画像から器官を適切に検出するように検出器5が十分に訓練されたか否かを判定するための基準となる。この閾値は、適宜設定されてよい。

[0124] 各誤差71～73の合計値が閾値以下であることは、換言すると、各誤差71～73の合計値が十分に小さいことは、顔画像から器官を適切に検出可能に検出器5が十分に訓練されたことを示す。そのため、各誤差71～73の合計値が閾値以下であると判定した場合には、制御部11は、本動作例に係る機械学習の処理を終了し（すなわち、ステップS102の処理を完了し）、上記ステップS103に処理を進める。

[0125] 一方、各誤差71～73の合計値が閾値を超えていることは、換言すると、各誤差71～73の合計値が大きいことは、検出器5がまだ十分には訓練されていないことを示す。そのため、各誤差71～73の合計値が閾値を超えていると判定した場合には、制御部11は、次のステップS210に処理を進める。

[0126] (ステップS210)

ステップS210では、制御部11は、各誤差71～73の合計値が小さくなるように検出器5の訓練を行う。具体的には、制御部11は、各誤差71～73の合計値が小さくなるように、検出器5のパラメータを調整する。このパラメータの調整には、誤差逆伝播法等の公知の方法が採用されてよい。すなわち、制御部11は、各誤差71～73の合計値を利用して、出力層53側から順に、各ニューロン間の結合の重み、各ニューロンの閾値等の各パラメータの誤差を算出する。そして、制御部11は、算出した誤差に基づいて、各パラメータの値を更新する。各パラメータの更新幅を定める学習率

は、適宜設定されてよい。

[0127] 各層 5 1 ~ 5 3 のパラメータの値の更新が完了すると、制御部 1 1 は、ステップ S 2 0 1 から処理を繰り返す。制御部 1 1 は、ステップ S 2 0 1 ~ S 2 1 0 の処理を繰り返すことで、各誤差 7 1 ~ 7 3 の合計値が十分に小さくなるように、検出器 5 のパラメータを最適化する。なお、学習時間を短縮化するために、このパラメータの最適化には、確率的勾配降下法等の公知の方法が採用されてもよい。これにより、検出器 5 は、各誤差 7 1 ~ 7 3 の合計値が小さくなるように訓練される。すなわち、ステップ S 2 0 9 において、各誤差 7 1 ~ 7 3 の合計値が閾値以下であると判定されるようになり、本動作例に係る機械学習の処理が完了する。

[0128] なお、上記処理手順は一例に過ぎず、各処理は可能な限り変更されてよい。例えば、ステップ S 2 0 2 の処理がステップ S 2 0 1 の後に実行され、ステップ S 2 0 4 の処理がステップ S 2 0 3 の後に実行され、ステップ S 2 0 5 の処理がステップ S 2 0 1 及び S 2 0 4 の後に実行され、ステップ S 2 0 7 の処理がステップ S 2 0 6 の後に実行され、ステップ S 2 0 8 の処理がステップ S 2 0 1 及び S 2 0 7 の後に実行されるのであれば、ステップ S 2 0 1 ~ S 2 0 8 の処理の順序は適宜入れ替えられてよい。

[0129] <繰り返し学習>

検出器生成装置 1 は、上記ステップ S 1 0 1 ~ S 1 0 3 の処理を繰り返してもよい。この繰り返しの過程で、暫定的に生成される検出器 5 を利用して、各学習データセット 6 1 に含まれる位置情報 6 1 2 を更新してもよい。すなわち、検出器生成装置 1 は、ステップ S 1 0 2 により、各学習データセット 6 1 を利用して学習済みの検出器 5 を生成した後、生成された学習済みの検出器 5 に各第 1 顔画像 6 1 1 を入力し、各第 1 顔画像 6 1 1 から器官の位置を検出してもよい。そして、検出器生成装置 1 は、元の位置（すなわち、学習済みの検出器 5 を生成する際に位置情報 6 1 2 により示されていた位置）から学習済みの検出器 5 により検出された位置を示すように、各学習データセット 6 1 の位置情報 6 1 2 を更新してもよい。

[0130] 図9を用いて、この繰り返し学習について詳細に説明する。図9は、機械学習を繰り返し実施する場面の一例を模式的に例示する。図9に示されるとおり、ステップS101～S103を最初に実施する際に、ステップS101により、制御部11は、第1顔画像611及び位置情報622aによりそれぞれ構成された複数の学習データセット62aを取得したと想定する。位置情報622aは、上記のとおり、オペレータの指定により与えられてもよいし、公知の画像処理方法により与えられてもよいし、他の検出器から得られる出力値に基づいて与えられてもよい。

[0131] 1回目の学習処理では、制御部11は、得られた各学習データセット62aを利用して、ステップS102の機械学習の処理を実行する。これにより、制御部11は、各学習データセット62aを与えられると、位置情報622aにより示される正解と一致するように第1顔画像611から器官を検出するように訓練された学習済みの検出器50aを暫定的に生成する。制御部11は、ステップS103の処理を実行することで、この暫定的に生成した学習済みの検出器50aの学習結果データを所定の記憶領域に保存してもよい。学習済みの検出器50aが生成されることで、1回目の学習処理は終了である。

[0132] 2回目の学習処理では、制御部11は、暫定的に生成された学習済みの検出器50aを利用して、各学習データセット62aの位置情報622aを更新する。すなわち、制御部11は、1回目の学習処理で生成された学習済みの検出器50aに各第1顔画像611を入力し、検出器50aの演算処理を実行する。これにより、制御部11は、各第1顔画像611より器官の位置を検出した結果に対応する出力値を学習済みの検出器50aから取得する。

[0133] 制御部11は、得られた各出力値に基づいて、学習済みの検出器50aによって各第1顔画像611から対象の器官を検出した位置を特定する。そして、制御部11は、各学習データセット62aについて、特定した検出位置を示す位置情報622bを生成し、生成した位置情報622bに位置情報622aを置き換える。これにより、ステップS101では、制御部11は、

第1顔画像611及び位置情報622bの組み合わせにより構成された新たな各学習データセット62bを取得する。

[0134] つまり、2回目の学習処理では、位置情報622bは、位置情報622bに組み合わせられる第1顔画像611を、機械学習を実施済みの検出器50aに入力することで、当該検出器50aから得られる出力値に基づいて与えられる。この2回目の学習処理の場面では、学習済みの検出器50aは、本発明の「他の検出器」及び「第1検出器」の一例である。各学習データセット62aは、「第1学習データセット」の一例である。各学習データセット62bは、「第2学習データセット」の一例である。

[0135] 2回目の学習処理では、制御部11は、この各学習データセット62bを利用して、ステップS102の機械学習の処理を実行する。これにより、制御部11は、各学習データセット62bを与えられると、位置情報622bにより示される正解と一致するように第1顔画像611から器官を検出するように訓練された学習済みの検出器50bを生成する。1回目の学習処理と同様に、制御部11は、ステップS103の処理を実行することで、この生成した学習済みの検出器50bの学習結果データを所定の記憶領域に保存してもよい。学習済みの検出器50bが生成されることで、2回目の学習処理は終了する。

[0136] 以上のように、検出器生成装置1は、各学習データセットを利用した機械学習により学習済みの検出器を生成する処理、及び生成された学習済みの検出器を利用して位置情報を更新する処理を交互に繰り返し実行してもよい。検出器生成装置1が、検出器を生成する処理及び位置情報を更新する処理を交互に繰り返し実行した場合、最終的に取得される各学習データセットが上記各学習データセット61に対応し、最終的に生成される学習済みの検出器が上記学習済みの検出器5に対応する。位置情報612は、最終の1回前の学習処理により生成された学習済みの検出器に第1顔画像611を入力することで、当該学習済みの検出器から得られる出力値に基づいて与えられる。

[0137] この検出器の生成処理及び位置情報の更新処理の繰り返しにより、暫定的

に生成された検出器による検出結果を位置情報 6 1 2 として利用するため、人為的なノイズが位置情報 6 1 2 に含まれるのを防止することができる。また、初期段階における位置情報 6 2 2 a が、オペレータの入力により与えられることで人為的なノイズを含んでいたとしても、この繰り返しの過程において、暫定的に生成された検出器を利用した位置情報の更新により、位置情報に含まれる人為的なノイズを低減することができる。そのため、位置情報 6 2 2 a に含まれ得る人為的なノイズによる影響にロバストで、対象者の顔の器官を顔画像から検出可能な検出器 5 を構築することができる。なお、図 9 では、機械学習の処理を 3 回以上繰り返す場面を例示している。しかしながら、機械学習の処理を繰り返す回数は、このような例に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。繰り返しの回数は、2 回であってもよい。

[0138] [モニタリング装置]

次に、図 1 0 を用いて、モニタリング装置 2 の動作例について説明する。図 1 0 は、本実施形態に係るモニタリング装置 2 の処理手順の一例を示すフローチャートである。ただし、以下で説明する処理手順は一例に過ぎず、各処理は可能な限り変更されてよい。また、以下で説明する処理手順について、実施の形態に応じて、適宜、ステップの省略、置換、及び追加が可能である。

[0139] (ステップ S 3 0 1)

ステップ S 3 0 1 では、制御部 2 1 は、データ取得部 2 1 1 として動作し、車両を運転する運転者 D の顔の写る顔画像 2 2 1 を取得する。本実施形態では、モニタリング装置 2 は、外部インタフェース 2 4 を介してカメラ 4 1 に接続されている。そのため、制御部 2 1 は、カメラ 4 1 から顔画像 2 2 1 を取得する。この顔画像 2 2 1 は、動画データであってもよいし、静止画像データであってもよい。顔画像 2 2 1 を取得すると、制御部 2 1 は、次のステップ S 3 0 2 に処理を進める。

[0140] ただし、顔画像 2 2 1 を取得する経路は、このような例に限定されなくて

もよく、実施の形態に応じて適宜選択されてよい。例えば、モニタリング装置 2 とは異なる他のコンピュータが、カメラ 4 1 に接続されていてもよい。この場合、制御部 2 1 は、他のコンピュータから顔画像 2 2 1 の送信を受け付けることで、顔画像 2 2 1 を取得してもよい。

[0141] (ステップ S 3 0 2)

ステップ S 3 0 2 では、制御部 2 1 は、検出部 2 1 2 として動作し、学習結果データ 1 2 1 を参照して、学習済みの検出器 5 の設定を行う。続いて、制御部 2 1 は、取得した顔画像 2 2 1 を学習済みの検出器 5 に入力し、検出器 5 の演算処理を実行する。この検出器 5 の演算処理は、上記ステップ S 2 0 1 と同様であってよい。これにより、制御部 2 1 は、運転者 D の顔の器官を顔画像 2 2 1 より検出した結果に対応する出力値を検出器 5 から取得する。出力値を取得すると、制御部 2 1 は、次のステップ S 3 0 3 に処理を進める。

[0142] (ステップ S 3 0 3)

ステップ S 3 0 3 では、制御部 2 1 は、出力部 2 1 3 として動作し、運転者 D の顔の器官を検出した結果に関する情報を出力する。出力先及び出力する情報の内容はそれぞれ、実施の形態に応じて適宜決定されてよい。例えば、制御部 2 1 は、ステップ S 3 0 2 により運転者 D の顔の器官を顔画像 2 2 1 から検出した結果をそのまま出力装置 2 6 を介して出力してもよい。

[0143] また、例えば、制御部 2 1 は、運転者 D の顔の器官を顔画像 2 2 1 から検出した結果に基づいて、運転者 D の状態を推定する等の何らかの情報処理を実行してもよい。そして、制御部 2 1 は、その情報処理を実行した結果を出力してもよい。顔の器官として目を検出した場合、情報処理の一例として、制御部 2 1 は、検出した各目の位置関係に基づいて、運転者 D の視線方向を推定してもよい。そして、制御部 2 1 は、視線方向を推定した結果に関連する情報を出力してもよい。例えば、視線方向が車両の前方を向いていない場合に、制御部 2 1 は、出力装置 2 6 を介して、視線方向を前方に向けるように促す警告を出力してもよい。また、情報処理の他の例として、制御部 2 1

は、検出した各器官の位置関係に基づいて、運転者Dの表情を推定してもよい。そして、制御部21は、表情を推定した結果に関連する情報を出力してもよい。例えば、運転者Dの表情が疲れている場合に、制御部21は、出力装置26を介して、停車を促すメッセージを出力してもよい。

[0144] 情報の出力が完了すると、制御部21は、本動作例に係る処理を終了する。なお、運転者Dが車両の運転操作を行っている間、制御部21は、ステップS301～S303の一連の処理を継続的に繰り返し実行してもよい。これにより、モニタリング装置2は、運転者Dの顔の器官の位置を顔画像から検出することを介して、当該運転者Dの状態を監視することができる。

[0145] [特徴]

以上のように、本実施形態に係る検出器生成装置1は、上記ステップS201～S210の処理により、各誤差71～73の合計値が小さくなるように検出器5を訓練する。各誤差71～73のうち第1誤差71の合計値は、検出器5により第1顔画像611から器官を検出した結果と位置情報612により指定される器官の位置の正解との間の損失に相当する。また、第2誤差72の合計値は、検出器5による器官の検出結果のブレであって、第1顔画像611におけるノイズの有無によって生じるブレの損失に相当する。そのため、検出器生成装置1は、上記ステップS201～S210の機械学習の処理により、位置情報612により示される正解と一致するように第1顔画像611から器官を検出し、かつその検出結果がノイズによる影響を受け難い、ように訓練された（学習済みの）検出器5を構築（生成）することができる。

[0146] 加えて、第3誤差73の合計値は、検出器5による器官の検出結果のブレであって、位置情報612に含まれる幾何学的なノイズの影響によって生じるブレの損失に相当する。そのため、検出器生成装置1は、上記ステップS201～S210の機械学習の処理により、位置情報に含まれる幾何的なノイズによる影響を受け難い、ように更に訓練された検出器5を構築することができる。

[0147] § 4 変形例

以上、本発明の実施の形態を詳細に説明してきたが、前述までの説明はあらゆる点において本発明の例示に過ぎない。本発明の範囲を逸脱することなく種々の改良や変形を行うことができることは言うまでもない。例えば、以下のような変更が可能である。なお、以下では、上記実施形態と同様の構成要素に関しては同様の符号を用い、上記実施形態と同様の点については、適宜説明を省略した。以下の変形例は適宜組み合わせ可能である。

[0148] <4. 1>

上記実施形態では、検出器5を構成するニューラルネットワークは、全結合ニューラルネットワークである。しかしながら、検出器5を構成するニューラルネットワークの構造及び種類は、このような例に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。例えば、検出器5には、畳み込みニューラルネットワーク、再帰型ニューラルネットワーク等が利用されてよい。

[0149] <4. 2>

上記実施形態では、検出器5は、ニューラルネットワークにより構成されている。しかしながら、画像を入力として利用可能であれば、検出器5を構成する学習モデルの種類は、ニューラルネットワークに限られなくてよく、実施の形態に応じて適宜選択されてよい。また、機械学習の方法は、利用する学習モデルによって適宜選択されてよい。例えば、検出する器官の位置の初期値を与え、その位置を更新することで対象の顔画像に写る器官の位置を適切に検出できるように検出器5を訓練してもよい。学習用の顔画像に与えられた正解データの平均値を器官の位置の初期値として利用することができる。また、位置の更新は、特徴量を用いて行うことができる。この特徴量には、例えば、Haar-like特徴量、輝度差等が用いられてよい。学習モデルには、例えば、サポートベクタマシン、線形回帰モデル、ランダムフォレスト等が用いられてもよい。この場合、検出器5のパラメータは、例えば、重みベクトル、変換ベクトル、各分岐の閾値等である。

[0150] <4. 3>

上記実施形態では、学習結果データ121は、学習済みのニューラルネットワーク（検出器5）の構成を示す情報を含んでいる。しかしながら、学習結果データ121の構成は、このような例に限定されなくてもよく、学習済みの検出器5の設定に利用可能であれば、実施の形態に応じて適宜決定されてよい。例えば、利用するニューラルネットワークの構成が各装置で共通化されている場合、学習結果データ121は、学習済みのニューラルネットワークの構成を示す情報を含んでいなくてもよい。

[0151] <4. 4>

上記実施形態では、検出器生成装置1は、ステップS102の機械学習により、各誤差71～73の合計値が小さくなるように、検出器5のパラメータを調整している。しかしながら、機械学習の処理は、このような例に限定されなくてもよい。例えば、各誤差71～73のうち第3誤差73は省略されてもよい。

[0152] 図11は、本変形例に係る検出器生成装置1Aのソフトウェア構成の一例を模式的に例示する。検出器生成装置1Aのハードウェア構成は、上記実施形態に係る検出器生成装置1と同様である。また、図11に示されるとおり、検出器生成装置1Aのソフトウェア構成も、上記実施形態に係る検出器生成装置1と同様である。

[0153] 検出器生成装置1Aは、第3誤差73を導出する処理を省略する点を除き、上記実施形態に係る検出器生成装置1と同様に動作する。すなわち、検出器生成装置1Aの制御部は、上記検出器生成装置1と同様に、ステップS101及びS103の処理を実行する。また、制御部は、学習処理部112として動作し、ステップS102の処理を実行する。本変形例では、このステップS102の機械学習の処理において、制御部は、上記ステップS206～S208の処理を省略する。つまり、制御部は、上記検出器生成装置1と同様に、上記ステップS201～S205の処理を実行した後に、ステップS209の処理を実行する。そして、ステップS209では、制御部は、第

1 誤差 7 1 の合計値及び第 2 誤差 7 2 の合計値が閾値以下であるか否かを判定する。

[0154] 第 1 誤差 7 1 の合計値及び第 2 誤差 7 2 の合計値が閾値以下であると判定した場合には、制御部は、機械学習の処理を終了し、上記ステップ S 1 0 3 に処理を進める。一方、第 1 誤差 7 1 の合計値及び第 2 誤差 7 2 の合計値が閾値を超えていると判定した場合には、制御部は、次のステップ S 2 1 0 に処理を進める。ステップ S 2 1 0 では、制御部は、各誤差 7 1 ~ 7 2 の合計値が小さくなるように検出器 5 の訓練を行う。検出器 5 の訓練、すなわち、検出器 5 のパラメータを調整する方法は、上記実施形態と同様であってよい。これにより、本変形例に係る検出器生成装置 1 A は、上記実施形態と同様に、位置情報 6 1 2 により示される正解と一致するように第 1 顔画像 6 1 1 から器官を検出し、かつその検出結果がノイズによる影響を受け難い、ように訓練された（学習済みの）検出器 5 を構築（生成）することができる。

[0155] < 4. 5 >

上記実施形態では、車両を運転する運転者の顔を撮影し、得られる顔画像から顔の器官の位置を検出する場面に本発明を適用した例を示している。しかしながら、本発明の適用可能な範囲は、このような運転者の顔画像から器官の位置を検出する場面に限られなくてもよい。本発明は、運転者以外の対象者の顔の器官を顔画像から検出する場面に広く適用可能である。

[0156] 図 1 2 は、本発明が適用される他の場面の一例を模式的に例示する。具体的に、図 1 2 は、生産現場で作業する作業員 U を撮影し、得られる顔画像から作業員 U の顔の器官を検出する場面に本発明を適用した例を示している。本変形例に係る検出器生成装置 1 B のハードウェア構成及びソフトウェア構成は、上記実施形態に係る検出器生成装置 1 と同様である。本変形例に係るモニタリング装置 2 B のハードウェア構成及びソフトウェア構成は、上記実施形態に係るモニタリング装置 2 と同様である。本変形例に係る検出器生成装置 1 B 及びモニタリング装置 2 B は、運転者以外の対象者の顔画像を取り扱う点を除いて、上記実施形態に係る検出器生成装置 1 及びモニタリング装

置 2 と同様に動作する。

[0157] すなわち、ステップ S 1 0 1 では、検出器生成装置 1 B の制御部は、学習データ取得部として動作し、被験者の顔の写る第 1 顔画像及び第 1 顔画像に写る顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する。被験者の属性が異なる点を除き、本変形例に係るステップ S 1 0 1 の処理は、上記実施形態と同様である。次のステップ S 1 0 2 では、制御部は、学習処理部として動作し、取得した複数の学習データセットを利用して、検出器の機械学習を実施する。本変形例に係るステップ S 1 0 2 の処理は、上記実施形態と同様であってもよいし、上記<4. 4>と同様であってもよい。次のステップ S 1 0 3 では、制御部は、保存処理部として動作し、機械学習により構築された検出器の構成及び演算パラメータを示す情報を学習結果データとして保存する。これにより、本変形例に係る検出器生成装置 1 B は、運転者以外の対象者の顔の写る顔画像から器官の位置を検出する能力を習得した検出器を生成することができる。

[0158] 一方、モニタリング装置 2 B は、検出器生成装置 1 B により構築された検出器を利用して、生産ラインでロボット装置 R と共に作業を行う作業者 U をカメラ 4 1 により撮影し、作業者 U について得られた顔画像から顔の器官の位置を検出する。ステップ S 3 0 1 では、モニタリング装置 2 B の制御部は、データ取得部として動作し、ロボット装置 R と共に作業を行う作業者 U の顔の写る顔画像をカメラ 4 1 から取得する。次のステップ S 3 0 2 では、制御部は、検出部として動作し、検出器生成装置 1 B により構築された学習済みの検出器に得られた顔画像を入力することで、作業者 U の顔の器官を顔画像より検出した結果に対応する出力値を検出器から取得する。次のステップ S 3 0 3 では、制御部は、作業者 U の顔の器官を検出した結果に関する情報を出力する。

[0159] 本変形例に係るステップ S 3 0 3 において、出力する情報の内容は、上記実施形態と同様に、実施の形態に応じて適宜決定されてよい。例えば、顔の

器官として目を検出した場合、制御部は、検出した各目の位置関係に基づいて、作業者Uの視線方向を推定し、視線方向を推定した結果に関連する情報を出力してもよい。この情報の一例として、ロボット装置Rとの作業に無関係な方向を視線方向が向いている場合に、制御部は、視線方向を作業に関係のある方向に向けるように促す警告を出力してもよい。また、制御部は、作業者Uの視線方向に適合する動作を実施する指示をロボット装置Rに出力してもよい。これにより、本変形例に係るモニタリング装置2Bは、検出器生成装置1Bにより生成された検出器を利用して、作業者Uの顔を撮影することで得られた顔画像から顔の器官の位置を検出することができる。

[0160] <4. 6>

上記実施形態及び変形例では、顔の器官の位置を顔画像から検出する場面に適用した例を示している。しかしながら、本発明の適用範囲は、このような顔の器官の位置を顔画像から検出する場面に限られなくてもよい。本発明は、何らかの対象物の位置を画像から検出するあらゆる場面に広く適用可能である。

[0161] 図13は、本発明が適用される他の場面の一例を模式的に例示する。具体的に、図13は、生産ラインRCを流れる製品UCを撮影し、製品UCに欠陥が存在する場合に、得られる画像から製品UCの欠陥の位置を検出する場面に本発明を適用した例を示している。図14は、本変形例に係る検出器生成装置1Cのソフトウェア構成の一例を模式的に例示する。

[0162] 本変形例に係る製品UCの欠陥は、本発明の「対象物」の一例である。なお、「対象物」は、このような製品の欠陥に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。人間の写る画像が得られる場合、「対象物」は、顔の器官（上記実施形態）の他、例えば、顔そのもの、顔以外の身体部位等であってよい。また、本変形例と同様に製品の写る画像が得られる場合、「対象物」は、例えば、製品そのもの、製品に含まれる何らかの特徴等であってよい。

[0163] 本変形例に係る検出器生成装置1Cのハードウェア構成は、上記実施形態

に係る検出器生成装置１と同様である。図１４に示されるとおり、本変形例に係る検出器生成装置１Ｃのソフトウェア構成も、上記実施形態に係る検出器生成装置１と同様である。また、本変形例に係るモニタリング装置２Ｃのハードウェア構成及びソフトウェア構成も、上記実施形態に係るモニタリング装置２と同様である。本変形例に係る検出器生成装置１Ｃ及びモニタリング装置２Ｃは、製品の写る画像を取り扱う点を除いて、上記実施形態に係る検出器生成装置１及びモニタリング装置２と同様に動作する。

[0164] すなわち、ステップＳ１０１では、検出器生成装置１Ｃの制御部は、学習データ取得部１１１として動作し、欠陥を含む製品の写る第１画像６１１Ｃ、及び第１画像６１１Ｃに写る欠陥の位置を示す位置情報６１２Ｃの組み合わせによりそれぞれ構成される複数の学習データセット６１Ｃを取得する。撮影の対象が異なる点を除き、本変形例に係るステップＳ１０１の処理は、上記実施形態と同様である。

[0165] 次のステップＳ１０２では、制御部は、学習処理部１１２として動作し、取得した複数の学習データセット６１Ｃを利用して、検出器５の機械学習を実施する。

[0166] ステップＳ２０１では、制御部は、各学習データセット６１Ｃの第１画像６１１Ｃを検出器５に入力し、検出器５の演算処理を実行する。これにより、制御部は、各第１画像に対する第１出力値を検出器５から取得する。ステップＳ２０２では、制御部は、各第１画像に対して得られた第１出力値に基づいて特定される製品の欠陥の第１検出位置Ｖ１Ｃと対応する位置情報６１２Ｃにより示される位置Ａ１Ｃとの第１誤差７１Ｃの合計値を算出する。

[0167] ステップＳ２０３では、制御部は、各第１画像６１１Ｃにノイズを付与することで、各第１画像６１１Ｃに対して第２画像６３Ｃを生成する。ステップＳ２０４では、制御部は、生成した各第２画像６３Ｃを検出器５に入力し、検出器５の演算処理を実行する。これにより、制御部は、各第２画像６３Ｃに対する第２出力値を検出器５から取得する。ステップＳ２０５では、制御部は、各第１出力値に基づいて特定される第１検出位置Ｖ１Ｃと対応する

第2画像63Cに対して得られた第2出力値に基づいて特定される第2検出位置V2Cとの第2誤差72Cの合計値を算出する。

[0168] ステップS206では、制御部は、各第1画像611Cに所定の幾何変換を適用することで、各第1画像611Cに対して第3画像66Cを生成する。ステップS207では、制御部は、生成した各第3画像66Cを検出器5に入力し、検出器5の演算処理を実行する。これにより、制御部は、各第3画像66Cに対する第3出力値を検出器5から取得する。ステップS208では、制御部は、各第1出力値に基づいて特定される第1検出位置と対応する第3画像66Cに対して得られた第3出力値に基づいて特定される第3検出位置との第3誤差73Cの合計値を算出する。

[0169] なお、本変形例に係るステップS208では、上記実施形態と同様に、制御部は、各第1画像611Cに対して得られた第1検出位置V1Cと対応する第3画像66Cに対して導出された第3検出位置V31Cとの差分を第3誤差73Cとして算出してもよい。また、制御部は、各第1画像611Cに対して得られた第1検出位置V1Cに所定の幾何変換を適用することで導出される位置と第3検出位置V30Cとの差分を第3誤差73Cとして算出してもよい。

[0170] ステップS209では、制御部は、算出した各誤差71C～73Cの合計値が閾値以下であるか否かを判定する。各誤差71C～73Cの合計値が閾値以下であると判定した場合、制御部は、機械学習の処理を終了し、次のステップS103に処理を進める。一方、各誤差71C～73Cの合計値が閾値を超えていると判定した場合、制御部は、次のステップS210に処理を進める。ステップS210では、制御部は、各誤差71C～73Cの合計値が小さくなるように検出器5の訓練を行う。検出器5の各層51～53のパラメータの値の更新が完了すると、制御部は、ステップS201から処理を繰り返す。

[0171] なお、上記本変形例に係るステップS102の処理手順は、上記実施形態と同様である。しかしながら、本変形例に係るステップS102の処理手順

は、このような例に限定されなくてもよく、例えば、上記<4. 4>と同様であってもよい。この場合、本変形例に係るステップS 1 0 2の処理において、上記ステップS 2 0 6～S 2 0 8の処理は省略されてよい。ステップS 2 0 9では、制御部は、算出した各誤差7 1 C～7 2 Cの合計値が閾値以下であるか否かを判定してもよい。各誤差7 1 C～7 2 Cの合計値が閾値以下であると判定した場合に、制御部は、次のステップS 1 0 3に処理を進めてもよい。一方、各誤差7 1 C～7 2 Cの合計値が閾値を超えていると判定した場合、制御部は、次のステップS 2 1 0に処理を進めてもよい。ステップS 2 1 0では、制御部は、各誤差7 1 C～7 2 Cの合計値が小さくなるように検出器5の訓練を行ってもよい。

[0172] ステップS 1 0 3では、制御部は、保存処理部1 1 3として動作し、機械学習により構築された検出器5の構成及び演算パラメータを示す情報を学習結果データ1 2 1 Cとして保存する。これにより、本変形例に係る検出器生成装置1 Cは、製品の写る画像から欠陥（対象物）の位置を検出する能力を習得した検出器5を生成することができる。

[0173] 一方、モニタリング装置2 Cは、検出器生成装置1 Cにより構築された検出器5を利用して、生産ラインR Cを流れる製品U Cを撮影し、製品U Cに欠陥が存在する場合に、得られる画像から製品U Cの欠陥の位置を検出する。ステップS 3 0 1では、モニタリング装置2 Cの制御部は、データ取得部として動作し、生産ラインR Cを流れる製品U Cの写る画像をカメラ4 1から取得する。次のステップS 3 0 2では、制御部は、検出部として動作し、学習結果データ1 2 1 Cを参照して、学習済みの検出器5の設定を行う。続いて、制御部は、学習済みの検出器5に得られた画像を入力することで、製品U Cの欠陥を画像より検出した結果に対応する出力値を検出器から取得する。次のステップS 3 0 3では、制御部は、製品U Cの欠陥を検出した結果に関する情報を出力する。

[0174] 本変形例に係るステップS 3 0 3において、出力する情報の内容は、上記実施形態と同様に、実施の形態に応じて適宜決定されてよい。例えば、製品

UCに欠陥が存在する場合、制御部は、検出した欠陥の位置を示す情報を出力してもよい。また、例えば、制御部は、欠陥が検出された製品UCを別のラインに流す指示を生産ラインRCに出力してもよい。これにより、本変形例に係るモニタリング装置2Cは、検出器生成装置1Cにより生成された検出器5を利用して、製品UCの写る画像から欠陥を検出することで、生産ラインRCを流れる製品UCの状態を監視することができる。

[0175] <4. 7>

図9に示されるとおり、上記実施形態に係る検出器生成装置1は、検出器を生成する処理及び位置情報を更新する処理を交互に繰り返すことで、位置情報に含まれ得る人為的なノイズによる影響にロバストで、対象者の顔の器官を顔画像から検出可能な検出器を構築することができる。人為的なノイズによる影響にロバストな検出器を構築するという観点から、この検出器を生成する処理及び位置情報を更新する処理を交互に実行する部分のみを抽出して、新たな形態に係る検出器生成装置が構成されてもよい。

[0176] 図15は、本変形例に係る検出器生成装置1Dのソフトウェア構成の一例を模式的に例示する。本変形例に係る検出器生成装置1Dのハードウェア構成は、上記実施形態に係る検出器生成装置1と同様である。また、図15に示されるとおり、本変形例に係る検出器生成装置1Dのソフトウェア構成も、上記実施形態に係る検出器生成装置1と同様である。

[0177] 検出器生成装置1Dは、第2誤差72及び第3誤差73を導出する処理を省略する点を除き、上記実施形態に係る検出器生成装置1と基本的には同様に動作する。すなわち、第1回目の学習処理のステップS101では、検出器生成装置1Dの制御部は、学習データ取得部111として動作し、車両を運転する被験者の顔の写る顔画像681及び顔画像681に写る顔の器官の位置を示す第1位置情報682の組み合わせによりそれぞれ構成される複数の第1学習データセット68を取得する。第1位置情報682は、オペレータの指定により与えられてもよい。

[0178] ステップS102では、制御部は、学習処理部112として動作し、取得

した複数の第1学習データセット68を利用して、第1検出器58の機械学習を実施する。本変形例に係るステップS102では、ステップS203～S208が省略される。ステップS201では、制御部は、各顔画像681を第1検出器58に入力し、第1検出器58の演算処理を実行する。これにより、制御部は、各顔画像681に対する出力値を第1検出器58から取得するステップS202では、制御部は、各顔画像681に対して得られた出力値に基づいて特定される検出位置と対応する第1位置情報682により示される位置との誤差の合計値を算出する。

[0179] ステップS209では、制御部は、算出された誤差の合計値が閾値以下であるか否かを判定する。誤差の合計値が閾値以下であると判定した場合に、制御部は、機械学習の処理を終了し、ステップS103に処理を進める。一方、誤差の合計値が閾値を超えていると判定した場合には、制御部は、次のステップS210に処理を進める。ステップS210では、制御部は、誤差の合計値が小さくなるように第1検出器58の訓練を行う。第1検出器58のパラメータを調整する方法は、上記実施形態と同様であってよい。これにより、制御部は、各顔画像681が入力されると、対応する第1位置情報682に対応する出力値を出力するように訓練された検出器5を構築する。ステップS103では、制御部は、保存処理部113として動作し、機械学習により構築された第1検出器58の構成及び演算パラメータを示す情報を学習結果データとして保存する。なお、第1回目の学習処理では、このステップS103の処理は省略されてもよい。

[0180] 本変形例に係る検出器生成装置1Dは、このような機械学習の一連の処理と暫定的に生成された検出器を利用して位置情報を更新する処理とを交互に繰り返す。第2回目の学習処理のステップS101では、制御部は、顔画像681、及び構築された第1検出器58に顔画像681を入力することで当該第1検出器58から得られた出力値に基づいて与えられる第2位置情報692の組み合わせによりそれぞれ構成された複数の第2学習データセット69を取得する。そして、ステップS102では、制御部は、複数の第2学習

データセット 69 を利用して、第 2 検出器 59 の機械学習を実施する。この機械学習の処理は、第 1 回目の学習処理と同様である。これにより、制御部は、各顔画像 681 が入力されると、対応する第 2 位置情報 692 に対応する出力値を出力するように訓練された第 2 検出器 59 を構築する。これにより、本変形例に係る検出器生成装置 1D は、人為的なノイズによる影響にロバストな検出器（第 2 検出器 59）を生成することができる。なお、生成された第 2 検出器 59 は、上記検出器 5 と同様に利用されてよい。

[0181] なお、図 15 では、機械学習の処理を 2 回繰り返す場面を例示している。しかしながら、機械学習の処理を繰り返す回数は、このような例に限定されなくてもよく、実施の形態に応じて適宜選択されてよい。繰り返しの回数は、3 回以上であってもよい。

[0182] また、上記<4. 5>で記載のとおり、本変形例は、運転者以外の対象者の顔の器官を顔画像から検出する場面に広く適用可能である。そのため、顔画像 681 は、運転者以外の被験者の顔の写る顔画像に置き換えられてよい。これにより、運転者以外の対象者の顔の写る顔画像から器官の位置を検出する能力を習得した検出器を構築することができる。

[0183] 更に、上記<4. 6>で記載のとおり、本変形例は、何らかの対象物の位置を画像から検出するあらゆる場面に広く適用可能である。そのため、顔画像 681 は、何らかの対象物の写る画像に置き換えられてよい。各位置情報（682、692）は、対象物の位置を示す位置情報に置き換えられてよい。これにより、対象物の位置を画像から検出する能力を習得した検出器を構築することができる。

[0184] § 5 実施例

以下、本発明の実施例について説明する。ただし、本発明は、これらの実施例に限定されるものではない。

[0185] [第 1 実施例]

まず、第 1 実施例では、目尻を検出するための検出器を生成するために、各学習データセットを以下の条件で用意した。検出器には、畳み込みニュー

ラルネットワーク及び全結合ニューラルネットワークを含む学習ネットワークを用いた。学習ネットワークでは、畳み込みニューラルネットワークの出力を全結合ニューラルネットワークの入力に接続した。そして、検出器生成装置（汎用のパーソナルコンピュータ）を用意し、用意した各学習データセットを利用して、用意した検出器生成装置に上記実施形態に係る各処理を実行させた。これにより、第1実施例に係る学習済みの検出器を得た。

[0186] <学習データセットの条件>

- ・件数：3000枚
- ・顔画像：64×64ピクセルの画像（目の写る部分のみ抽出）
- ・抽出方法：オペレータの指定した目頭及び目尻の位置に基づいて目の写る部分を抽出
- ・顔画像の取得条件：近赤外線カメラにより被験者の顔を撮影
- ・位置情報：目尻の位置を示す
- ・位置情報の取得条件：オペレータの入力（オペレータにマウスを操作させて、顔画像内で目尻の位置を指定させた）

[0187] [第2実施例]

第1実施例に係る学習済みの検出器を利用して、第1実施例で利用した各学習データセットの顔画像から目尻の位置を検出し、検出結果により位置情報を更新することで、新たな各学習データセットを得た。そして、得られた各学習データセットを利用して、検出器生成装置に上記実施形態に係る機械学習の処理を実行させた。これにより、第2実施例に係る学習済みの検出器を得た。つまり、第2実施例に係る学習済みの検出器は、上記実施形態に係る機械学習の処理を2回繰り返すことで生成された学習済みの検出器に相当する。

[0188] [第3実施例]

第2実施例に係る学習済みの検出器を利用して、第2実施例で利用した各学習データセットの顔画像から目尻の位置を検出し、検出結果により位置情報を更新することで、新たな各学習データセットを得た。そして、得られた

各学習データセットを利用して、検出器生成装置に上記実施形態に係る機械学習の処理を実行させた。これにより、第3実施例に係る学習済みの検出器を得た。つまり、第3実施例に係る学習済みの検出器は、上記実施形態に係る機械学習の処理を3回繰り返すことで生成された学習済みの検出器に相当する。

[0189] [比較例]

上記第1実施例において、第2誤差72及び第3誤差73の計算を省略することで、比較例に係る検出器を得た。すなわち、比較例では、上記実施形態に係る処理手順のうち、ステップS203～S208の処理を省略した。そして、ステップS209及びS210では、各誤差71～73を第1誤差71に置き換えた。

[0190] <評価方法>

次に、第1～第3実施例及び比較例に係る検出器の検出精度を次の方法で評価した。すなわち、カメラにより対象者の顔を撮影し、学習用の顔画像と同様に、評価用の5000枚の顔画像（64×64ピクセル）を用意した。評価用の各顔画像を各検出器に入力し、各顔画像に対する目尻の位置の第1検出結果を各検出器から得た。また、評価用の各顔画像を幾何変換（平行移動+回転移動）することで変換画像を生成した。生成した各変換画像を各検出器に入力し、各変換画像に対する目尻の位置の第2検出結果を各検出器から得た。続いて、各検出器の第1検出結果の示す位置と第2検出結果の示す位置に上記幾何変換の逆変換を適用することで得られた位置との差分（ずれ）をピクセル単位で算出した。そして、得られた差分の平均及び標準偏差を算出した。差分の平均及び標準偏差の算出結果は、以下の表1のとおりである。

[0191] [表1]

	第1実施例	第2実施例	第3実施例	比較例
平均(pixel)	0.252	0.131	0.0866	0.83
標準偏差(pixel)	0.101	0.0581	0.0336	0.41

[0192] 表 1 に示されるとおり、第 2 誤差 7 2 及び第 3 誤差 7 3 の学習を実施する第 1 ～第 3 実施例に係る検出器の検出のブレは、第 2 誤差 7 2 及び第 3 誤差 7 3 の学習を実施しない比較例よりも小さかった。このことから、上記実施形態によれば、第 1 誤差 7 1 のみの学習を実施するケースよりも検出結果のブレの少ない検出器を生成可能であることが分かった。また、第 1 実施例から第 3 実施例になるにつれ、検出器による目尻の検出のブレが低減した。このことから、検出器を生成する処理と位置情報を更新する処理とを交互に繰り返すことで、検出器の検出のブレを抑えることができることが分かった。したがって、上記実施形態によれば、対象物をより高精度に検出可能な検出器を生成可能であることが分かった。

符号の説明

[0193] 1 …検出器生成装置、
1 1 …制御部、1 2 …記憶部、1 3 …通信インタフェース、
1 4 …入力装置、1 5 …出力装置、1 6 …ドライブ、
1 1 1 …学習データ取得部、1 1 2 …学習処理部、
1 1 3 …保存処理部、
1 2 1 …学習結果データ、
8 1 …検出器生成プログラム、9 1 …記憶媒体、
2 …モニタリング装置、
2 1 …制御部、2 2 …記憶部、2 3 …通信インタフェース、
2 4 …外部インタフェース、
2 5 …入力装置、2 6 …出力装置、2 7 …ドライブ、
2 1 1 …データ取得部、2 1 2 …検出部、
2 1 3 …出力部、
2 2 1 …顔画像、
8 2 …モニタリングプログラム、9 2 …記憶媒体、
3 1 …カメラ、4 1 …カメラ、
5 …検出器、

5 1 …入力層、5 2 …中間層（隠れ層）、5 3 …出力層、
6 1 …学習データセット、
6 1 1 …顔画像（第 1 顔画像）、6 1 2 …位置情報、
6 3 …顔画像（第 2 顔画像）、6 6 …顔画像（第 3 顔画像）、
7 1 …誤差（第 1 誤差）、7 2 …誤差（第 2 誤差）、
7 3 …誤差（第 3 誤差）、
A 1 …位置、
V 1 …検出位置（第 1 検出位置）、V 2 …検出位置（第 2 検出位置）、
V 3 0 ・ V 3 1 …検出位置（第 3 検出位置）、
T …被験者、D …運転者（対象者）

請求の範囲

- [請求項1] 車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、
- 、
- 前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、
- を備え、
- 前記検出器の機械学習を実施することは、
- 前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、
- 前記各学習データセットについて、前記第1出力値に基づいて特定される前記器官の第1検出位置と前記位置情報により示される位置との第1誤差の合計値を算出するステップと、
- 前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、
- 前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、
- 前記各学習データセットについて、前記第1検出位置と前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、
- 前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、
- を含む、
- 検出器生成装置。

- [請求項2] 前記検出器の機械学習を実施することは、

前記各学習データセットの前記第1顔画像に所定の幾何変換を適用することで、前記各学習データセットの前記第1顔画像に対して第3顔画像を生成するステップと、

前記各第3顔画像を前記検出器に入力することで、前記各第3顔画像に対する第3出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1検出位置と前記第3出力値に基づいて特定される前記器官の第3検出位置との第3誤差の合計値を算出するステップと、
を更に含み、

前記訓練するステップでは、前記学習処理部は、前記第1誤差の合計値、前記第2誤差の合計値及び前記第3誤差の合計値が小さくなるように前記検出器を訓練する、
請求項1に記載の検出器生成装置。

[請求項3] 前記所定の幾何変換は、平行移動及び回転移動を伴う変換である、
請求項2に記載の検出器生成装置。

[請求項4] 前記各学習データセットの前記位置情報は、当該位置情報に組み合わせられる前記第1顔画像を、機械学習を実施済みの他の検出器に入力することで、当該他の検出器から得られる出力値に基づいて与えられる、
請求項1から3のいずれか1項に記載の検出器生成装置。

[請求項5] 前記器官は、目、鼻、口、眉、顎及びこれらの組み合わせから選択される、
請求項1から4のいずれか1項に記載の検出器生成装置。

[請求項6] 前記検出器は、ニューラルネットワークにより構成される、
請求項1から5のいずれか1項に記載の検出器生成装置。

[請求項7] 車両を運転する運転者の顔の写る顔画像を取得するデータ取得部と、
請求項1から6のいずれか1項の検出器生成装置による機械学習を

実施済みの前記検出器に取得した前記顔画像を入力することで、前記運転者の前記顔の器官を検出した結果に対応する出力値を前記検出器から取得する検出部と、

前記運転者の顔の器官を検出した結果に関する情報を出力する出力部と、

を備える、

モニタリング装置。

[請求項8]

コンピュータが、

車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得するステップと、

前記複数の学習データセットを利用して、検出器の機械学習を実施するステップと、

を実行し、

前記検出器の機械学習を実施するステップは、

前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1出力値に基づいて特定される前記器官の第1検出位置と前記位置情報により示される位置との第1誤差の合計値を算出するステップと、

前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、

前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1検出位置と前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の

合計値を算出するステップと、

前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、

を含む、

検出器生成方法。

[請求項9]

コンピュータに、

車両を運転する被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得するステップと、

前記複数の学習データセットを利用して、検出器の機械学習を実施するステップと、

を実行させるための検出器生成プログラムであって、

前記検出器の機械学習を実施するステップは、

前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1出力値に基づいて特定される前記器官の第1検出位置と前記位置情報により示される位置との第1誤差の合計値を算出するステップと、

前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、

前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1検出位置と前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、

前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるよ

うに前記検出器を訓練するステップと、
を含む、
検出器生成プログラム。

[請求項10]

被験者の顔の写る第1顔画像、及び前記第1顔画像に写る前記顔の器官の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、

前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、

を備え、

前記検出器の機械学習を実施することは、

前記各学習データセットの前記第1顔画像を前記検出器に入力することで、前記各学習データセットの前記第1顔画像に対する第1出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1出力値に基づいて特定される前記器官の第1検出位置と前記位置情報により示される位置との第1誤差の合計値を算出するステップと、

前記各学習データセットの前記第1顔画像にノイズを付与することで、前記各学習データセットの前記第1顔画像に対して第2顔画像を生成するステップと、

前記各第2顔画像を前記検出器に入力することで、前記各第2顔画像に対する第2出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1検出位置と前記第2出力値に基づいて特定される前記器官の第2検出位置との第2誤差の合計値を算出するステップと、

前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、

を含む、

検出器生成装置。

[請求項11] 対象物の写る第1画像、及び前記第1画像に写る前記対象物の位置を示す位置情報の組み合わせによりそれぞれ構成される複数の学習データセットを取得する学習データ取得部と、

前記複数の学習データセットを利用して、検出器の機械学習を実施する学習処理部と、

を備え、

前記検出器の機械学習を実施することは、

前記各学習データセットの前記第1画像を前記検出器に入力することで、前記各学習データセットの前記第1画像に対する第1出力値を前記検出器から取得するステップと、

前記各学習データセットについて、前記第1出力値に基づいて特定される前記対象物の第1検出位置と前記位置情報により示される位置との第1誤差の合計値を算出するステップと、

前記各学習データセットの前記第1画像にノイズを付与することで、前記各学習データセットの前記第1画像に対して第2画像を生成するステップと、

前記各第2画像を前記検出器に入力することで、前記各第2画像に対する第2出力値を前記検出器から取得するステップと、

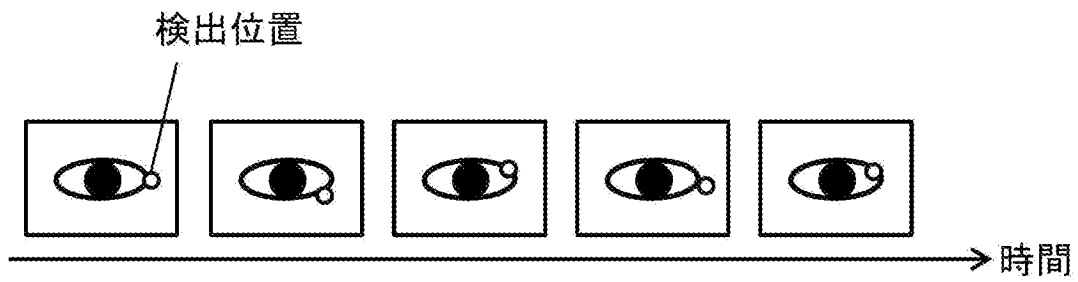
前記各学習データセットについて、前記第1検出位置と前記第2出力値に基づいて特定される前記対象物の第2検出位置との第2誤差の合計値を算出するステップと、

前記第1誤差の合計値及び前記第2誤差の合計値が小さくなるように前記検出器を訓練するステップと、

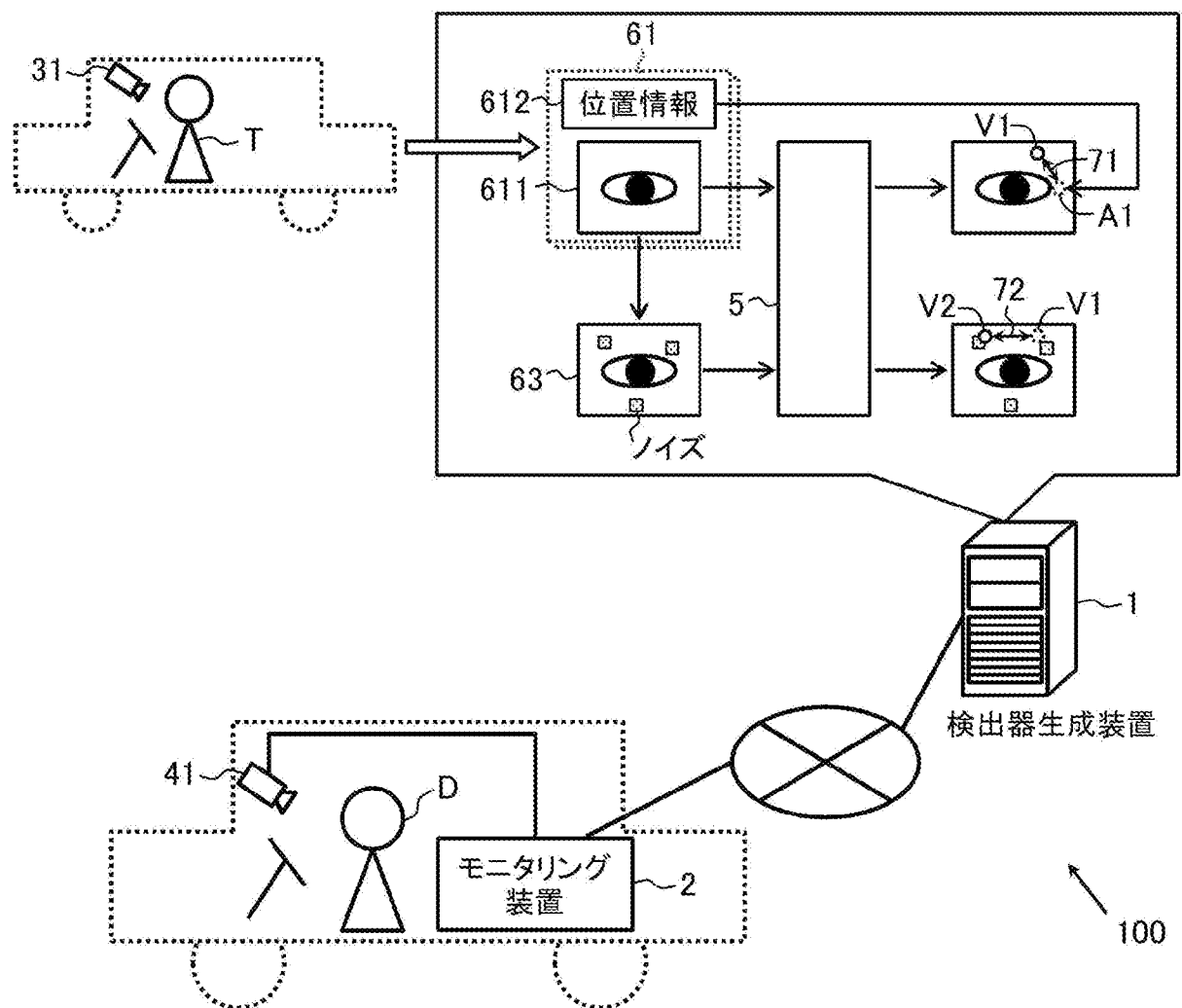
を含む、

検出器生成装置。

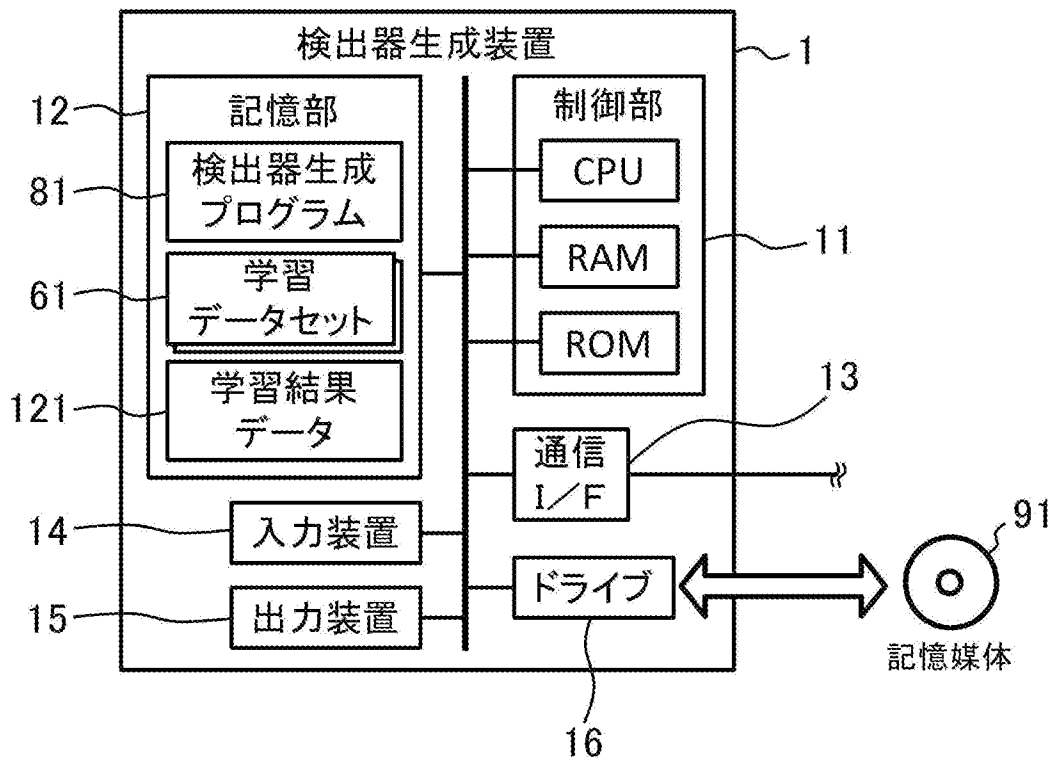
[図1]



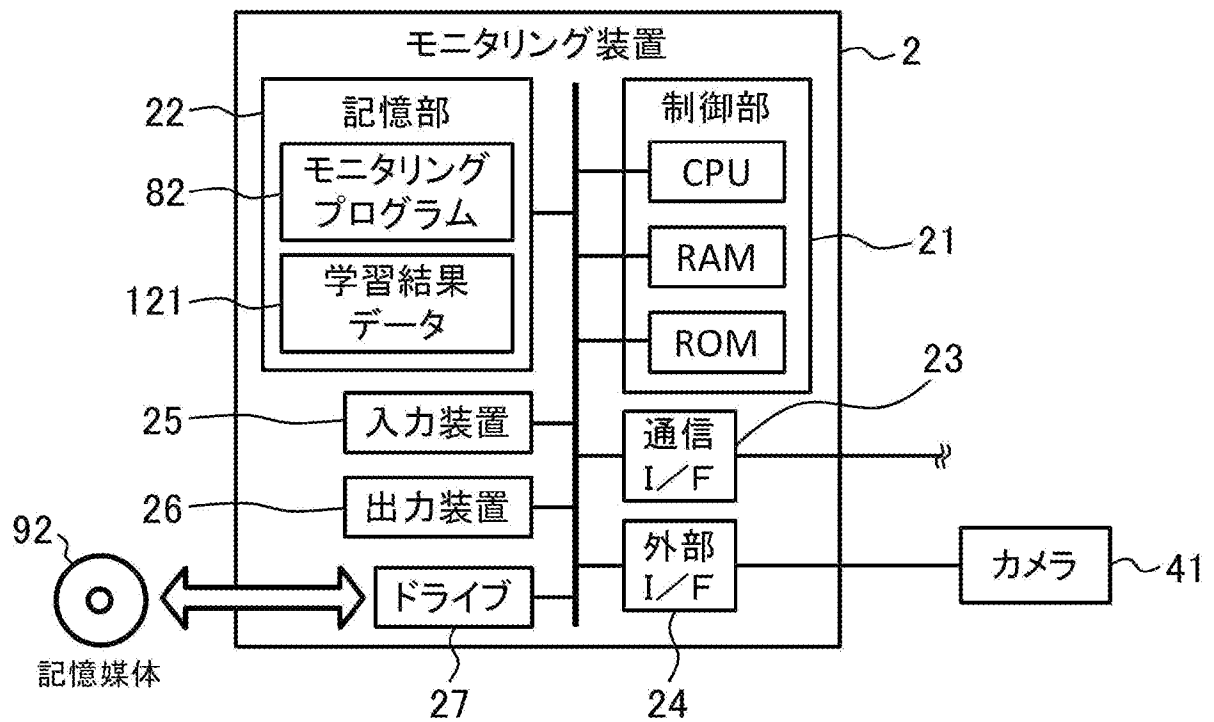
[図2]



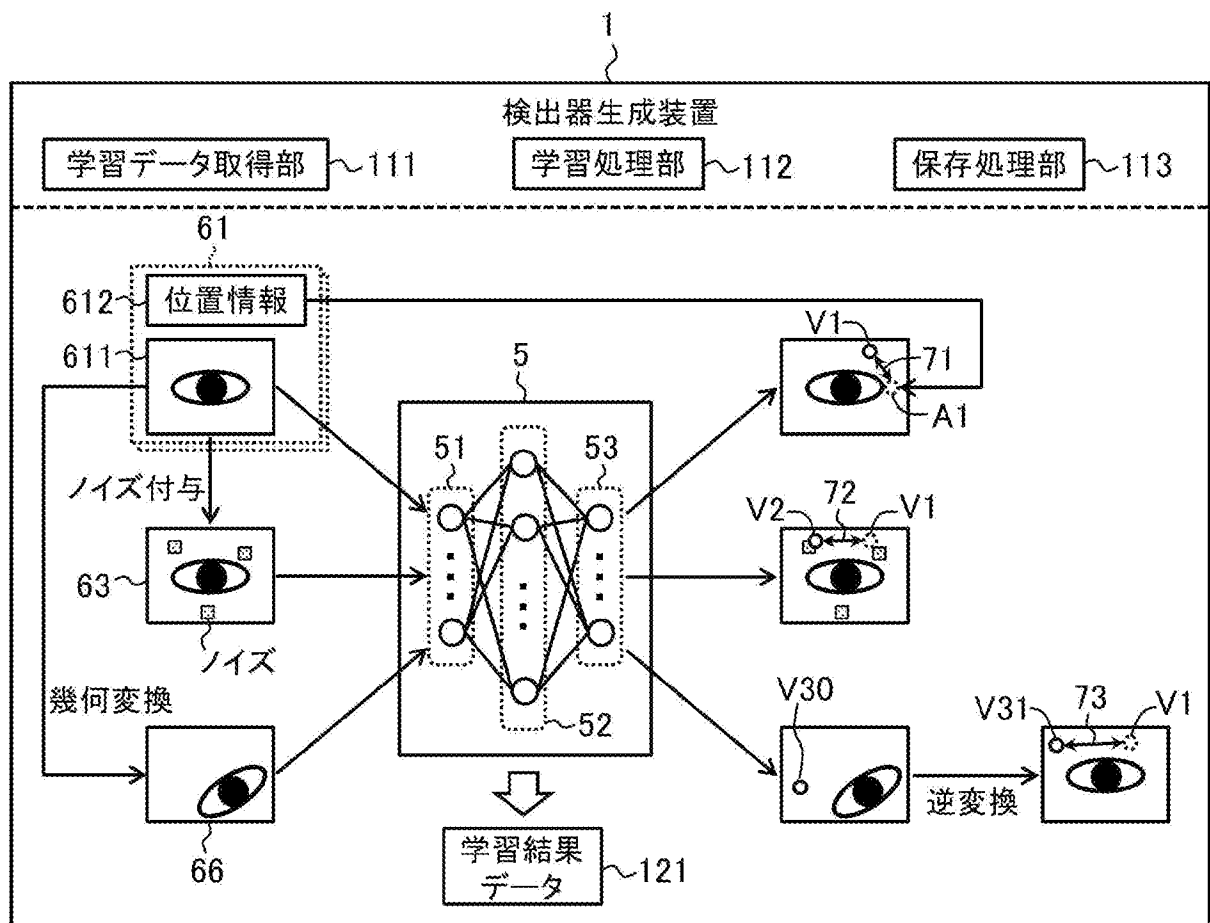
[図3]



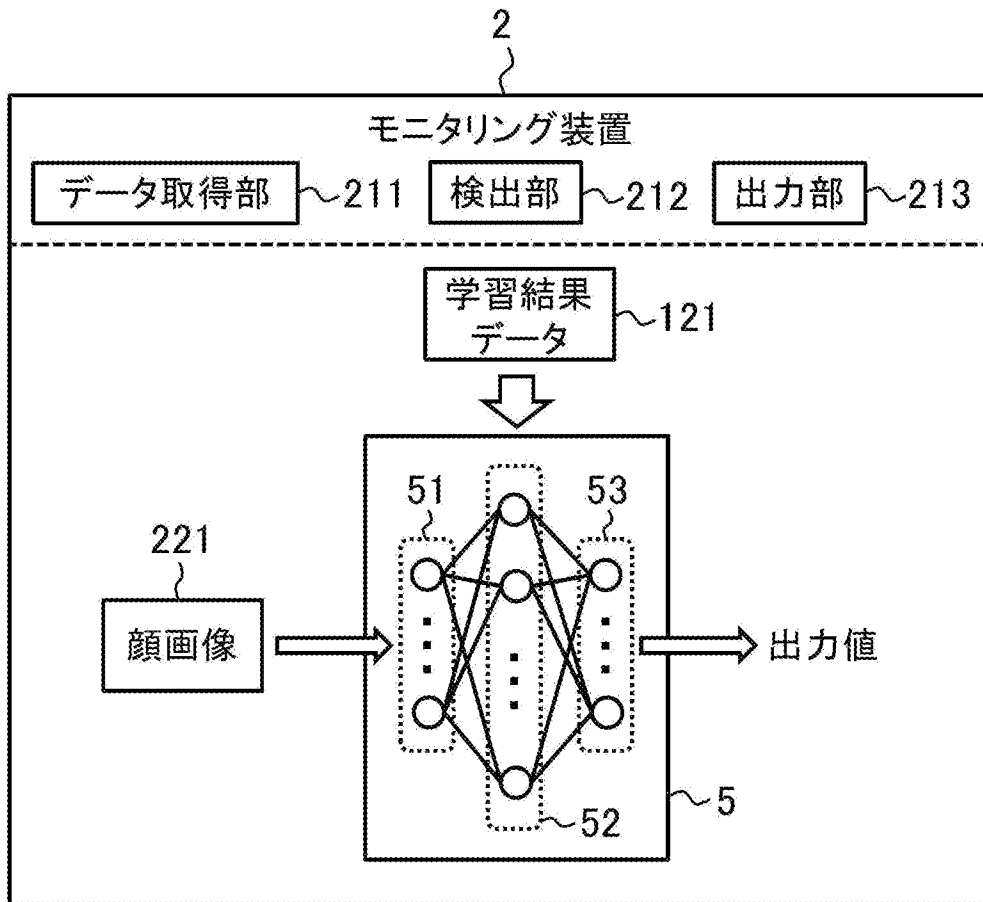
[図4]



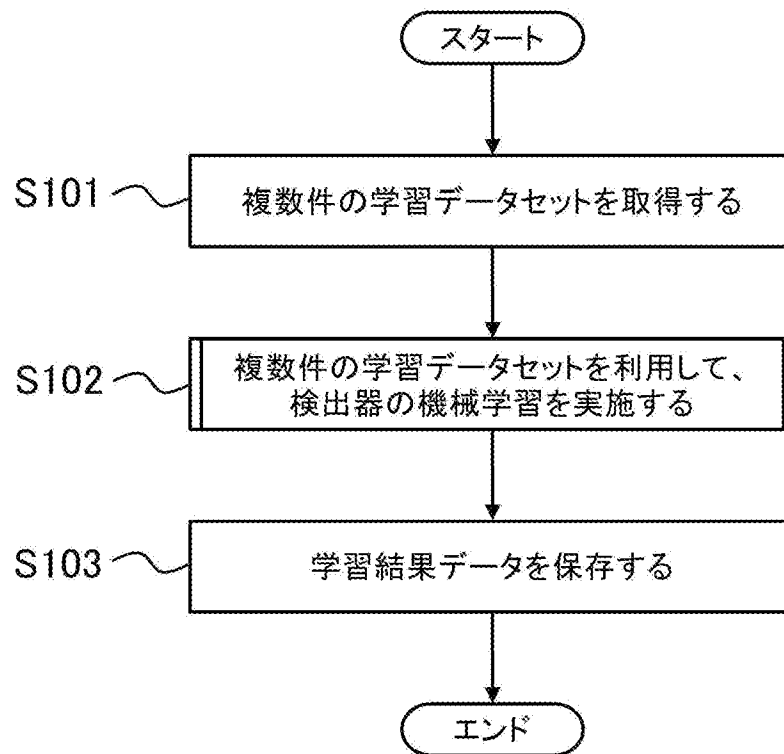
[図5]



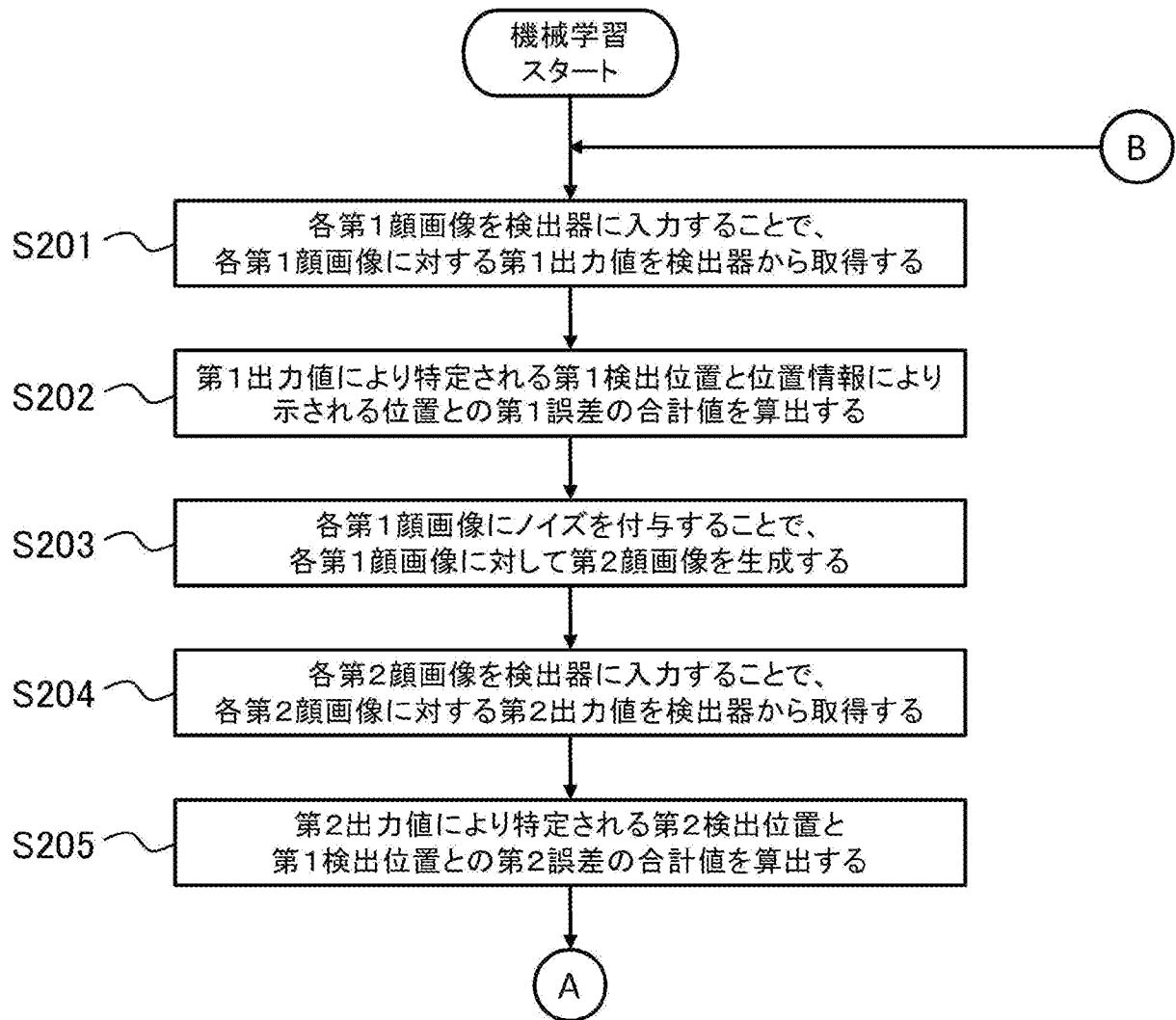
[図6]



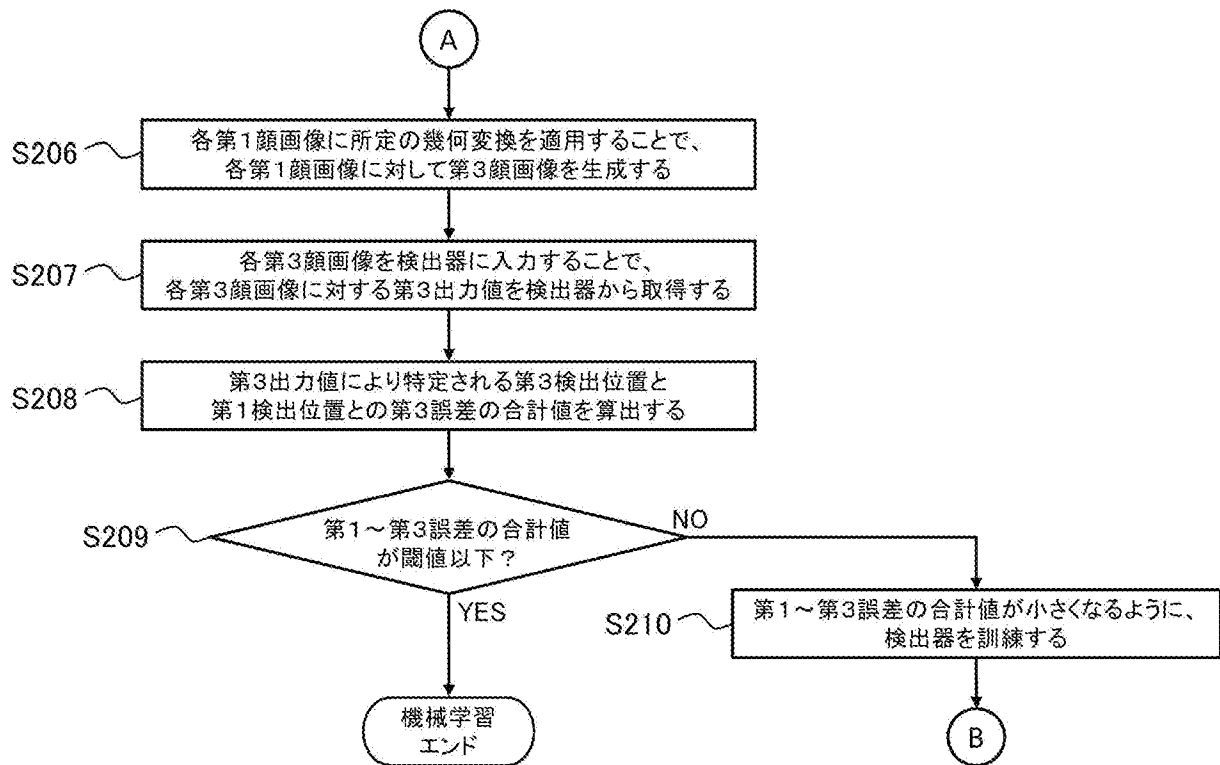
[図7]



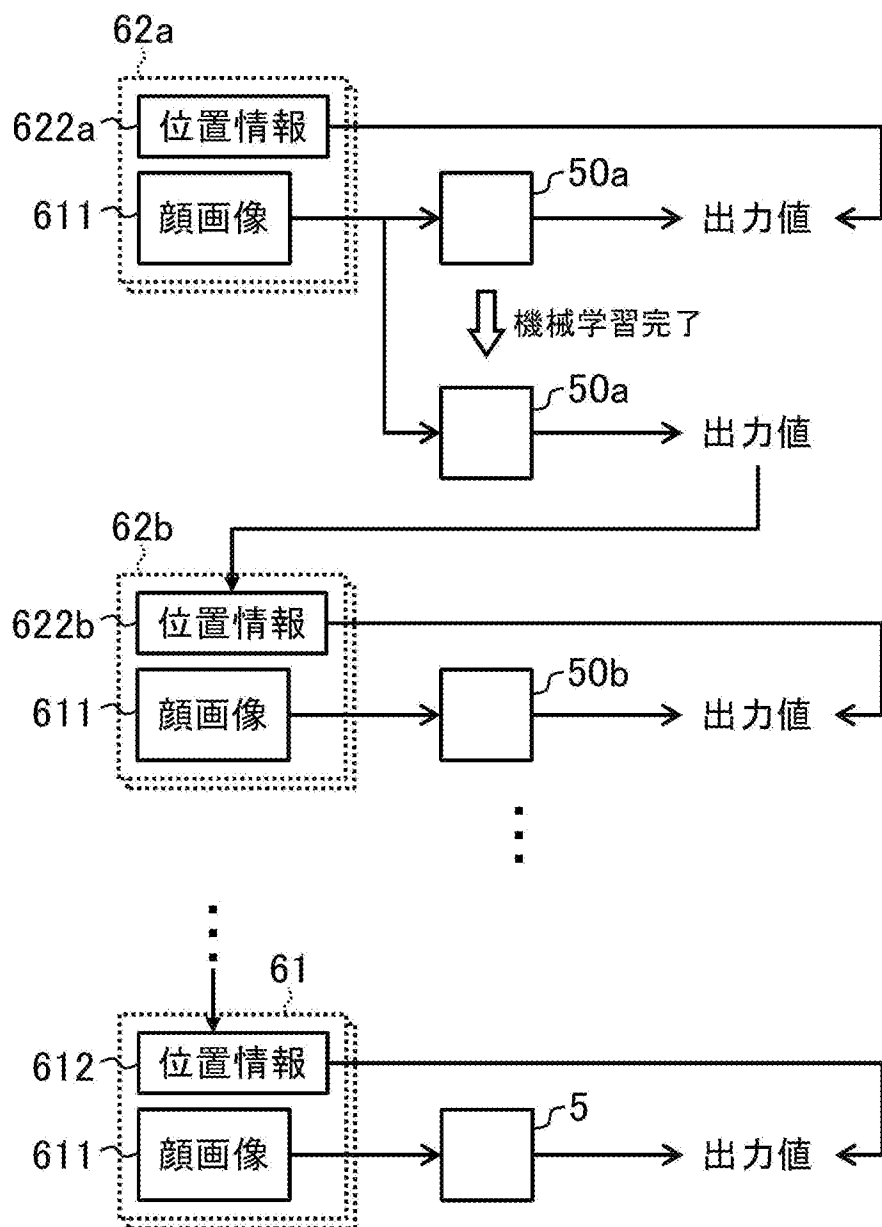
[図8A]



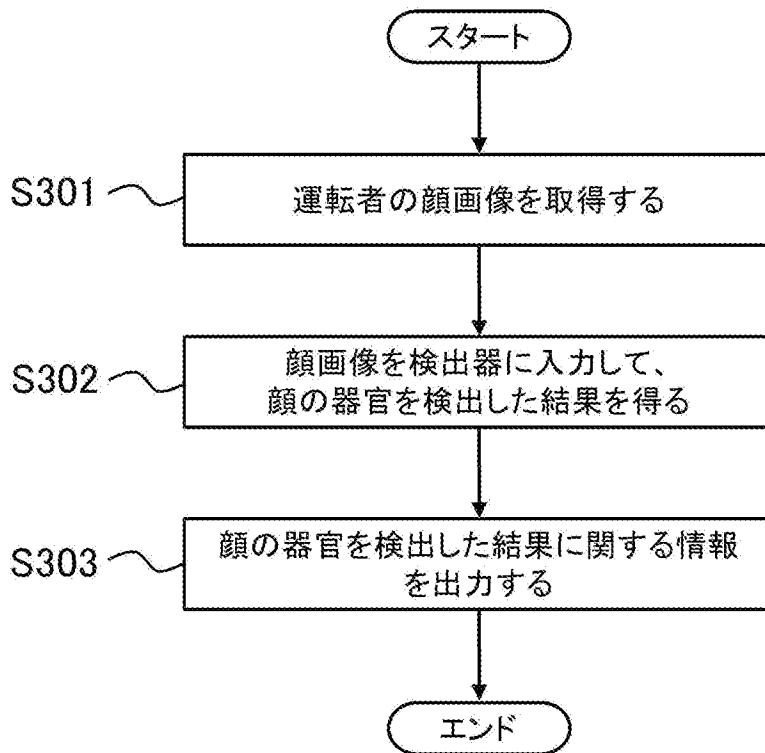
[図8B]



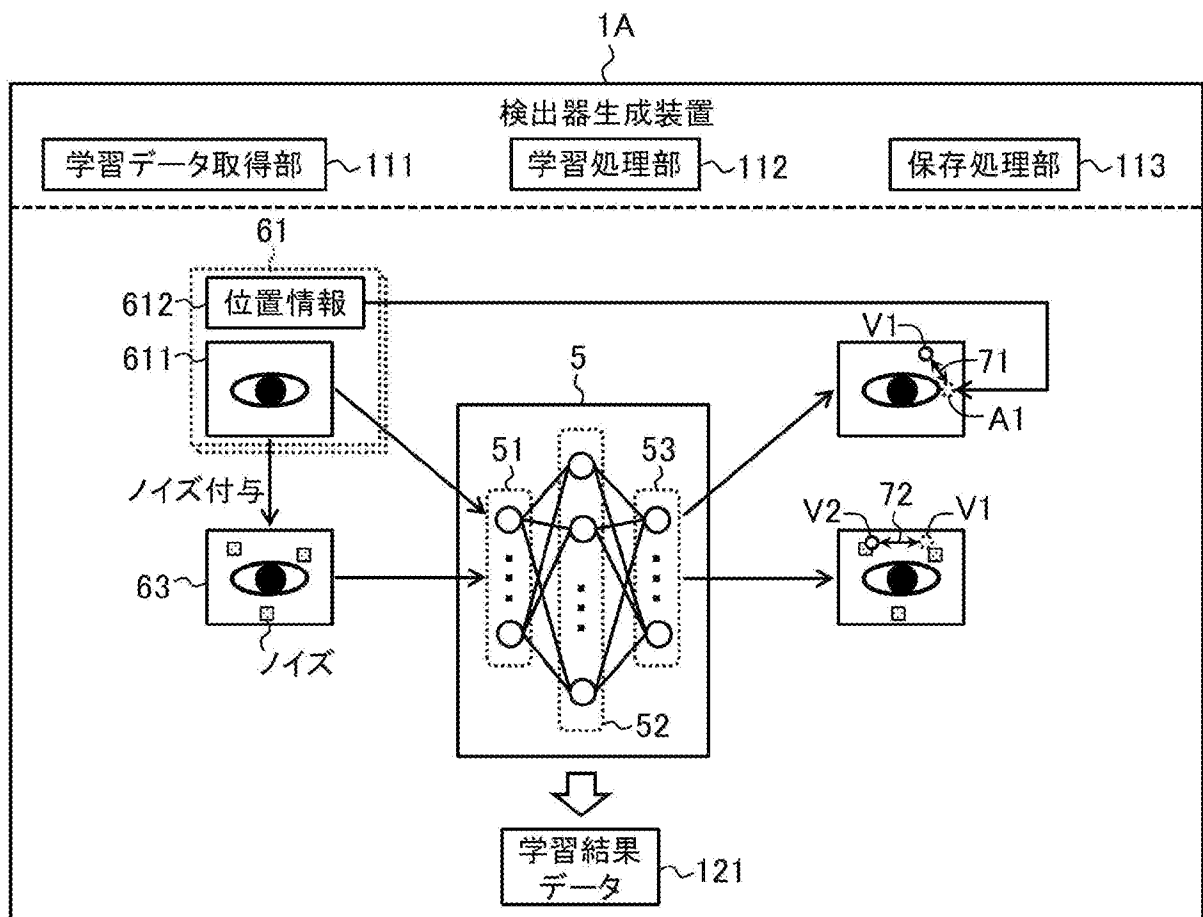
[図9]



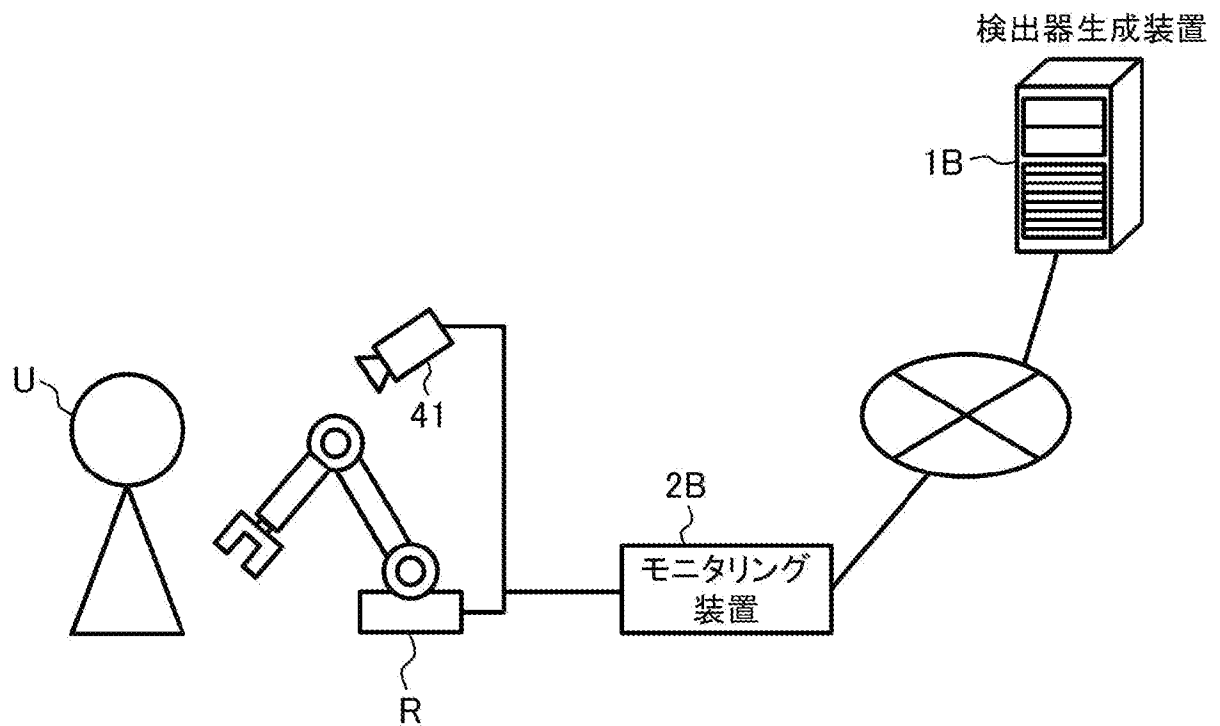
[図10]



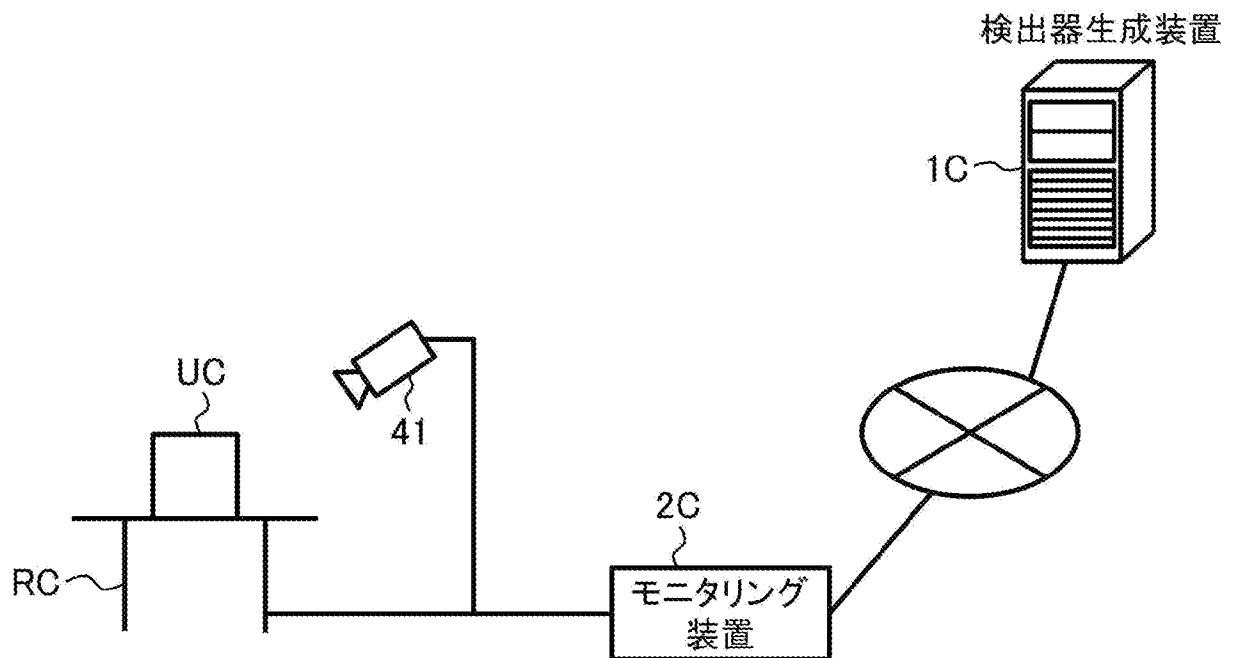
[図11]



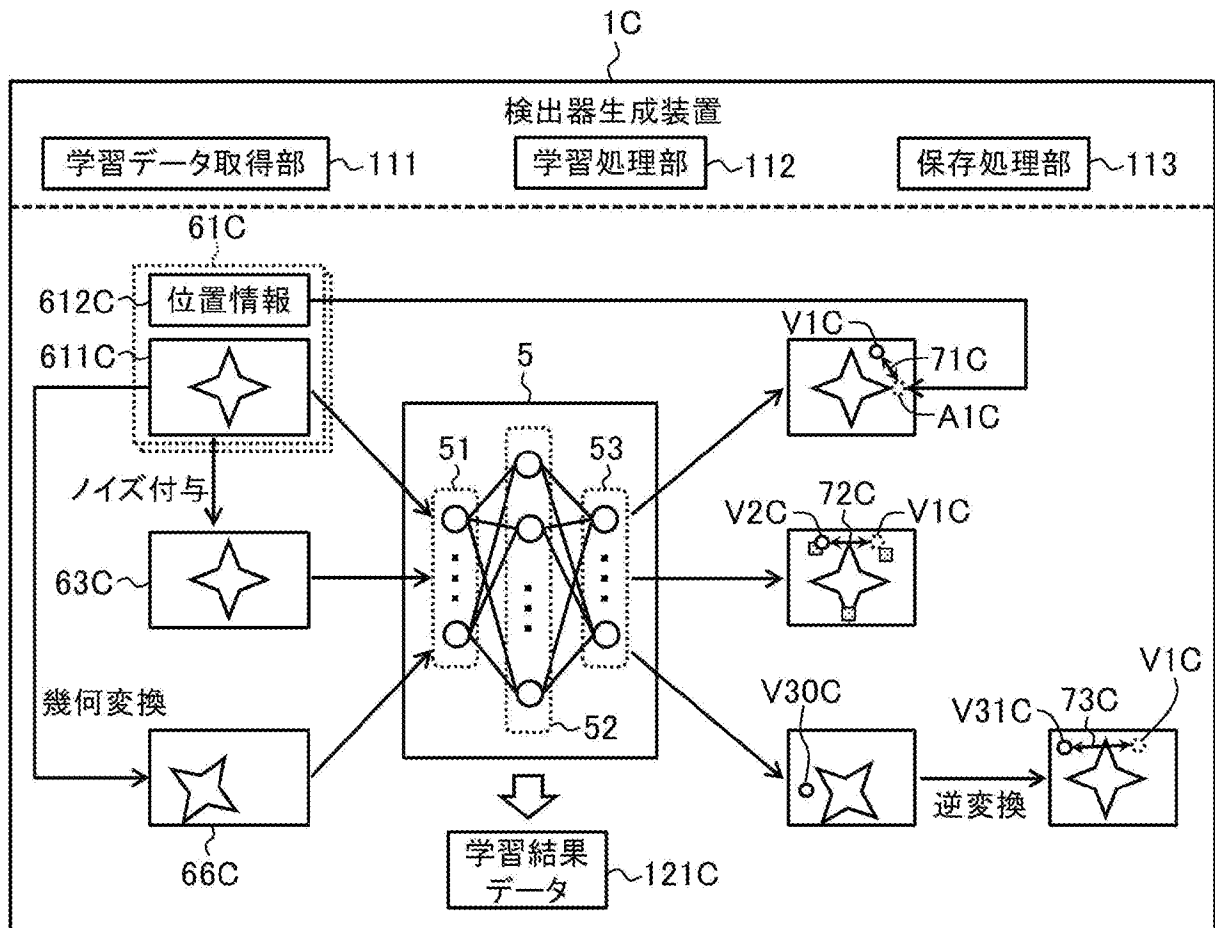
[図12]



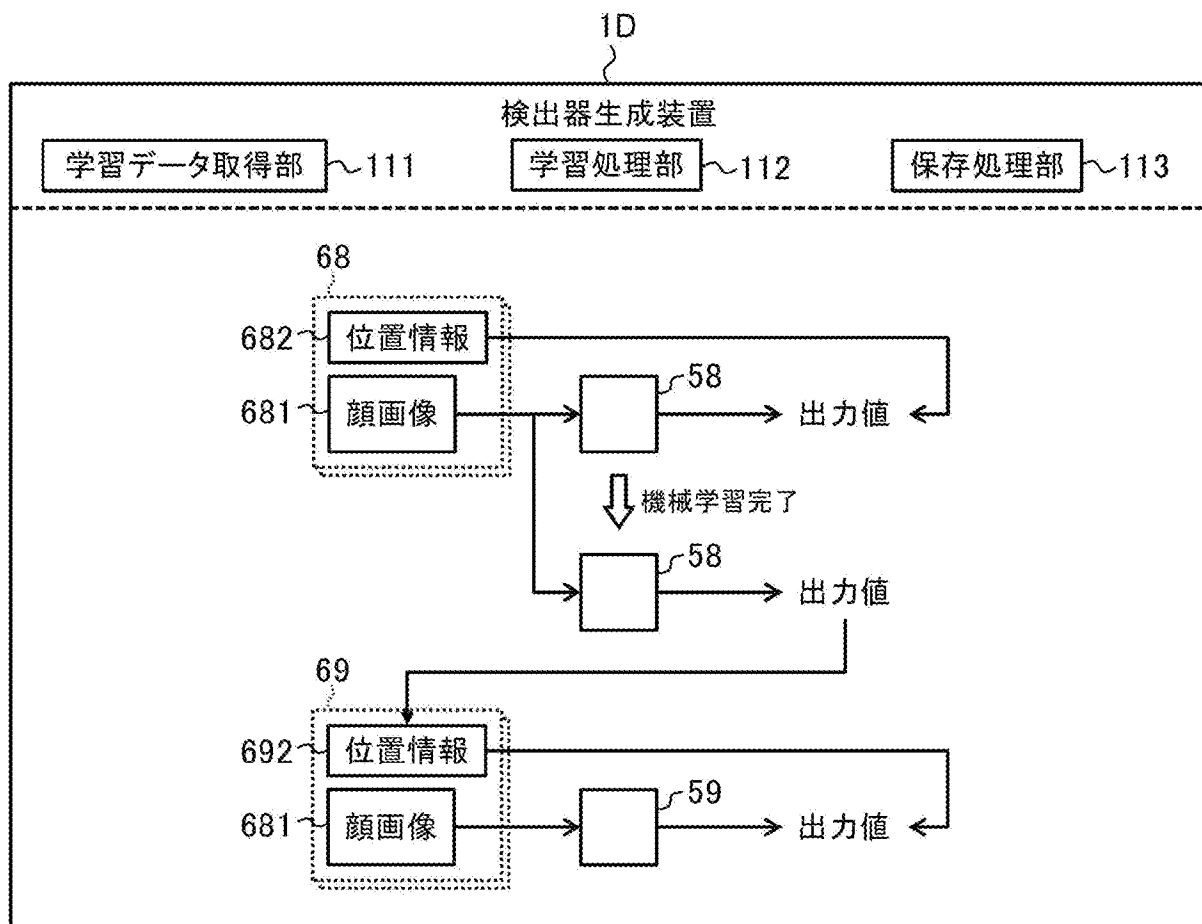
[図13]



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/010187

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. G06T7/00 (2017.01) 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. G06T7/00

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

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	山下隆義など, Deep Convolutional Neural Network による顔器官点検出における最適なミニバッチ作成方法, SSII2015 第21回画像センシングシンポジウム 講演論文集 [CD-ROM], 10 June 2015, IS2-13-1 to IS2-13-6, non-official translation (YAMASHITA, Takayoshi et al. Optimal Mini-Batch Generation Method for Facial Feature Point Detection Using Deep Convolutional Neural Network. Proceedings of the 21st Symposium on Sensing via Image Information.)	1-11
A	JP 2008-536211 A (FRANCE TELECOM) 04 September 2008, entire text, all drawings & US 2008/0201282 A1, entire document & WO 2006/103241 A2 & EP 1866834 A2 & FR 2884008 A1 & CN 101171598 A	1-11

☒ 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
07.06.2019

Date of mailing of the international search report
18.06.2019

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

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/010187

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	山下隆義, イラストで学ぶディープラーニング, 株式会社講談社 鈴木哲, 22 February 2016, pp. 70-73 (YAMASHITA, Takayoshi. An Illustrated Guide to Deep Learning. Kodansha Ltd. Tetsu Suzuki.)	1-11

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06T7/00(2017.01)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. G06T7/00

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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	山下 隆義など, Deep Convolutional Neural Networkによる顔器官点検出における最適なミニバッチ作成方法, SSI I 2015 第21回 画像センシングシンポジウム 講演論文集 [CD-ROM], 2015.06.10, IS2-13-1~IS2-13-6	1-11
A	JP 2008-536211 A (フランス テレコム) 2008.09.04, 全文、全図 & US 2008/0201282 A1, whole document & WO 2006/103241 A2, & EP 1866834 A2 & FR 2884008 A1 & CN 101171598 A	1-11

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

07.06.2019

国際調査報告の発送日

18.06.2019

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)

鍛 利孝

5 C

5 0 9 4

電話番号 03-3581-1101 内線 3541

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	山下 隆義, イラストで学ぶ ディープラーニング, 株式会社講談社 鈴木 哲, 2016.02.22, pp.70-73, (YAMASHITA, Takayoshi, An Illustrated Guide to Deep Learning)	1-11