(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9月3日(03.09.2020)

WIPO | PCT

(10) 国際公開番号
WO 2020/174623 A1

(51) 国際特許分類:
G06T 7/70 (2017.01)

(21) 国際出願番号： PCT/JP2019/007653

(22) 国際出願日： 2019年2月27日(27.02.2019)

(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: オリンパス株式会社 (**OLYMPUS CORPORATION**) [JP/JP]; 〒1928507 東京都八王子市石川町２９５１番地 Tokyo (JP). 国立大学法人東京大学 (**THE UNIVERSITY OF TOKYO**) [JP/JP]; 〒1138654 東京都文京区本郷七丁目３番１号 Tokyo (JP).

(72) 発明者: 岡澤 淳郎 (**OKAZAWA Atsuro**). 高畑 智之 (**TAKAHATA Tomoyuki**); 〒1138654 東京都文京区本郷七丁目３番１号 国立大学法人東京大学内 Tokyo (JP). 原田 達也 (**HARADA Tatsuya**); 〒1138654 東京都文京区本郷七丁目３番１号 国立大学法人東京大学内 Tokyo (JP).

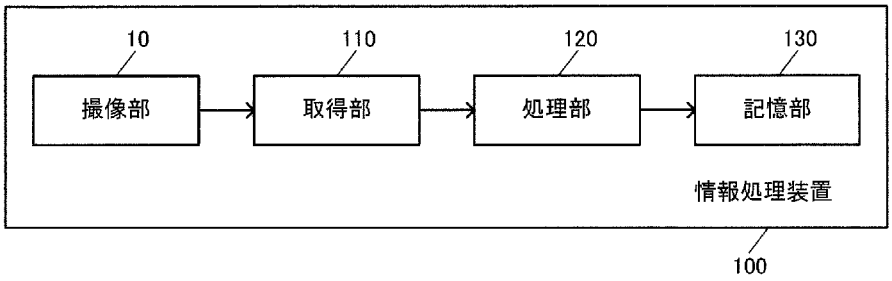
(74) 代理人: 竹 腰 昇, 外 (**TAKEKOSHI Noboru** et al.); 〒1010063 東京都千代田区神田淡路町２－８ プロステック淡路町３階 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, 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:** INFORMATION PROCESSING DEVICE, MOBILE BODY AND LEARNING DEVICE

(54) 発明の名称： 情報処理装置、移動体及び学習装置

[図1]



10 Image capture unit
100 Image processing device
110 Acquisition unit
120 Processing unit
130 Memory

(57) **Abstract:** This information processing device (100) includes an acquisition unit (110) and a processing unit (120). The acquisition unit (110) acquires: a first detection image by capturing an image of a plurality of objects including a first object and a second object which is more permeable to visible light than the first object; and a second detection image which captures an image of a plurality of objects using infrared light. The processing unit (120) determines a first feature amount on the basis of the fist detection image, determines a second feature amount on the basis of the second detection image, and calculates a third feature amount corresponding to the difference between the first feature amount and the second feature amount. The processing unit (120) detects the position of the second object in the first detection image and/or the second detection image on the basis of the third feature amount.

WO 2020/174623 A1

[続葉有]

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)）

(57) 要約：情報処理装置（１００）は、取得部（１１０）と、処理部（１２０）を含む。取得部（１１０）は、第１対象物と、第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を可視光によって撮像した第１検出用画像と、複数の対象物を赤外光によって撮像した第２検出用画像を取得する。処理部（１２０）は、第１検出用画像に基づいて第１特徴量を求め、第２検出用画像に基づいて第２特徴量を求め、第１特徴量と第２特徴量の差分に対応する第３特徴量を算出する。処理部（１２０）は、第３特徴量に基づいて、第１検出用画像及び第２検出用画像の少なくとも一方における第２対象物の位置を検出する。

明　細　書

発明の名称 ： 情報処理装置、移動体及び学習装置

技術分野

[0001] 本発明は、情報処理装置、移動体及び学習装置等に関する。

背景技術

[0002] 従来、撮像画像に基づいて、当該撮像画像に含まれる物体の認識処理を行う手法が広く知られている。例えば自律移動する車両やロボット等において、衝突回避等の移動制御を実現するために物体認識が行われる。可視光を透過するガラス等の物体を認識することも重要となるが、可視光画像にはガラスの特徴が十分に現れない。

[0003] これに対して特許文献１及び特許文献２には、赤外光を用いて撮像された画像に基づいて、ガラス等の透明物体を検出する手法が開示されている。

先行技術文献

特許文献

[0004] 特許文献1：特開２００７－７６３７８号公報

特許文献2：特開２０１０－１４６０９４号公報

発明の概要

発明が解決しようとする課題

[0005] 特許文献１においては、周囲がすべて直線エッジで構成された領域をガラス面と判断している。また特許文献２においては、赤外光画像の輝度値、領域の面積、輝度値の分散等に基づいてガラスか否かを判断する。しかし、可視光を反射する物体にも、赤外光画像における画像特徴がガラスと類似する物体がある。そのため、可視光画像の画像特徴のみ、或いは赤外光画像の画像特徴のみで、可視光を透過する物体を含む物体認識を行うことは困難であった。

[0006] 本開示のいくつかの態様によれば、可視光を透過する物体が撮像対象に含まれる場合において、精度よく物体認識を行う情報処理装置、移動体及び学

習装置等を提供できる。

課題を解決するための手段

[0007] 本開示の一態様は、第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した第１検出用画像と、前記複数の対象物を赤外光によって撮像した第２検出用画像を取得する取得部と、処理部とを含み、前記処理部は、前記第１検出用画像に基づいて第１特徴量を求め、前記第２検出用画像に基づいて第２特徴量を求め、前記第１特徴量と前記第２特徴量の差分に対応する特徴量を第３特徴量として算出し、前記第３特徴量に基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方における前記第２対象物の位置を検出する情報処理装置に関係する。

[0008] 本開示の他の態様は、第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した第１検出用画像と、前記複数の対象物を赤外光によって撮像した第２検出用画像を取得する取得部と、処理部とを含み、前記処理部は、前記第１検出用画像に基づいて第１特徴量を求め、前記第２検出用画像に基づいて第２特徴量を求め、前記第１特徴量及び前記第２特徴量に基づいて、前記第１検出用画像及び前記第２検出用画像に撮像された前記複数の対象物について前記可視光の透過度合いを表す透過スコアを算出し、前記第１検出用画像及び前記第２検出用画像に基づいて、前記第１検出用画像及び前記第２検出用画像に撮像された前記複数の対象物の形状を示す形状スコアを算出し、前記透過スコアと前記形状スコアに基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方における、前記第１対象物の位置と前記第２対象物の位置の両方を区別して検出する情報処理装置に関係する。

[0009] 本開示の他の態様は、上記のいずれかに記載の情報処理装置を含む移動体に関係する。

[0010] 本開示の他の態様は、第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した可視

光画像と、前記複数の対象物を赤外光によって撮像した赤外光画像と、前記可視光画像及び前記赤外光画像の少なくとも一方における前記第２対象物の位置情報と、を対応付けたデータセットを取得する取得部と、前記データセットに基づいて、前記可視光画像及び前記赤外光画像の少なくとも一方において、前記第２対象物の位置を検出する条件を機械学習する学習部と、を含む学習装置に関係する。

図面の簡単な説明

[0011] [図1]情報処理装置の構成例。

[図2]撮像部及び取得部の構成例。

[図3]撮像部及び取得部の構成例。

[図4]処理部の構成例。

[図5]図５（Ａ）、図５（Ｂ）は透明物体であるガラス扉の開閉を示す模式図。

[図6]可視光画像、赤外光画像、第１～第３特徴量の例。

[図7]可視光画像、赤外光画像、第１～第３特徴量の例。

[図8]第１の実施形態の処理を説明するフローチャート。

[図9]図９（Ａ）～図９（Ｃ）は情報処理装置を含む移動体の例。

[図10]処理部の構成例。

[図11]第２の実施形態の処理を説明するフローチャート。

[図12]学習装置の構成例。

[図13]ニューラルネットワークを説明する模式図。

[図14]第３の実施形態にかかる処理を説明する模式図。

[図15]学習処理を説明するフローチャート。

[図16]推論処理を説明するフローチャート。

[図17]処理部の構成例。

[図18]第４の実施形態にかかる処理を説明する模式図。

[図19]透過スコアの算出処理を説明する図。

[図20]形状スコアの算出処理を説明する図。

発明を実施するための形態

[0012] 以下、本実施形態について説明する。なお、以下に説明する本実施形態は、請求の範囲に記載された本開示の内容を不当に限定するものではない。また本実施形態で説明される構成の全てが、本開示の必須構成要件であるとは限らない。

[0013] １．第１の実施形態

上述したように、可視光を透過するガラス等の物体を検出する手法が種々開示されている。以下、可視光を透過する物体を透明物体と表記し、可視光を透過しない物体を可視物体と表記する。可視光とは、人の目で見ることが可能な光であり、例えば約３８０ｎｍ～約８００ｎｍ程度の波長帯域の光である。透明物体は可視光を透過するため、可視光画像に基づく位置検出が難しい。可視光画像とは、可視光を用いて撮像された画像である。

[0014] 特許文献１や特許文献２は、透明物体であるガラスが赤外光を吸収するという特性に着目し、赤外光画像に基づいてガラスを検出する手法を開示している。赤外光とは、可視光よりも波長が長い光であり、赤外光画像とは赤外光を用いて撮像された画像である。

[0015] 特許文献１においては、周囲がすべて直線エッジで構成された領域をガラス面と判断している。しかし、周囲がすべて直線エッジで構成された物体は、ガラスに限らず多く存在するため、それらの物体とガラスを適切に区別することが難しい。周囲がすべて直線エッジで構成された物体としては、額縁、ＰＣ（Personal Computer）等のディスプレイ、印刷物等が考えられる。例えば、画面表示がされていないディスプレイは、周囲が直線エッジで構成され、且つ、内部のコントラストが非常に低くなる。赤外光画像におけるガラスの画像特徴とディスプレイの画像特徴が類似するため、ガラスの適切な検出が困難である。

[0016] また特許文献２においては、赤外光画像の輝度値と、領域の面積、分散でガラスか否かを判断する。しかし、ガラス以外にも同様の輝度値、面積、分散等の特徴を有する物体が存在する。例えばガラスと同等の大きさであって

、画面表示がされていないディスプレイを、ガラスと区別することが難しい。以上のように、可視光画像の画像特徴のみ、或いは赤外光画像の画像特徴のみで、透明物体の位置検出を行うことは困難であった。

[0017] 図１は本実施形態の情報処理装置１００の構成例を示す図である。情報処理装置１００は、撮像部１０と、取得部１１０と、処理部１２０と、記憶部１３０を含む。撮像部１０及び取得部１１０については、図２及び図３を用いて後述する。処理部１２０については、図４を用いて後述する。記憶部１３０は、処理部１２０等のワーク領域となるもので、その機能はＲＡＭ（Random Access Memory）等のメモリーやＨＤＤ（Hard Disk Drive）などによって実現できる。なお、情報処理装置１００は図１の構成に限定されず、これらの一部の構成要素を省略したり、他の構成要素を追加するなどの種々の変形実施が可能である。例えば、情報処理装置１００から撮像部１０を省略してもよい。この場合、情報処理装置１００は、外部の撮像装置から後述する可視光画像及び赤外光画像を取得する処理を行う。

[0018] 図２は、撮像部１０及び取得部１１０の構成例を示す図である。撮像部１０は、波長分離ミラー（ダイクロイックミラー）１１と、第１光学系１２と、第１撮像素子１３と、第２光学系１４と、第２撮像素子１５を含む。波長分離ミラー１１は、所定の波長帯域の光を反射し、異なる波長帯域の光を透過する光学素子である。例えば波長分離ミラー１１は、可視光を反射し、赤外光を透過する。波長分離ミラー１１を用いることによって、光軸ＡＸに沿った対象物（被写体）からの光が２つの方向に分離される。

[0019] 波長分離ミラー１１によって反射された可視光は、第１光学系１２を経由して第１撮像素子１３に入射する。図２においては、第１光学系１２としてレンズを例示したが、第１光学系は、絞りやメカシャッター等の不図示の構成を含んでもよい。第１撮像素子１３は、ＣＣＤ（Charge Coupled Device）、ＣＭＯＳ（Complementary metal-oxide semiconductor）等の光電変換素子を含み、可視光を光電変化した結果である可視光画像信号を出力する。ここでの可視光画像信号はアナログ信号である。第１撮像素子１３は、例えば広

く知られたベイヤ配列のカラーフィルターを備えた撮像素子である。ただし、第１撮像素子１３は、補色型等の他のカラーフィルターを用いた素子であってもよいし、異なる方式の撮像素子であってもよい。

[0020] また波長分離ミラー１１を透過した赤外光は、第２光学系１４を経由して第２撮像素子１５に入射する。第２光学系１４についても、レンズに加えて、絞りやメカシャッター等の不図示の構成を含んでもよい。第２撮像素子１５は、マイクロボロメータ、ＩｎＳｂ（Indium Antimonide）等の光電変換素子を含み、赤外光を光電変化した結果である赤外光画像信号を出力する。ここでの赤外光画像信号はアナログ信号である。

[0021] 取得部１１０は、第１Ａ／Ｄ変換回路１１１と、第２Ａ／Ｄ変換回路１１２を含む。第１Ａ／Ｄ変換回路１１１は、第１撮像素子１３からの可視光画像信号に対するＡ／Ｄ変換処理を行い、デジタルデータである可視光画像データを出力する。可視光画像データは、例えばＲＧＢの３チャンネルの画像データである。第２Ａ／Ｄ変換回路１１２は、第２撮像素子１５からの赤外光画像信号に対するＡ／Ｄ変換処理を行い、デジタルデータである赤外光画像データを出力する。赤外光画像データは、例えば１チャンネルの画像データである。以下、デジタルデータである可視光画像データ及び赤外光画像データを、単に可視光画像、赤外光画像と表記する。

[0022] 図３は、撮像部１０及び取得部１１０の他の構成例を示す図である。撮像部１０は、第３光学系１６と、撮像素子１７を含む。第３光学系は、レンズに加えて、絞りやメカシャッター等の不図示の構成を含んでもよい。撮像素子１７は、可視光を受光する第１撮像素子１３－２と、赤外光を受光する第２撮像素子１５－２が光軸ＡＸに沿った方向において積層された積層型の撮像素子である。

[0023] 図３の例においては、第３光学系１６に相対的に近い第２撮像素子１５－２において、赤外光の撮像が行われる。第２撮像素子１５－２は、赤外光画像信号を取得部１１０に出力する。また第３光学系１６から相対的に遠い第１撮像素子１３－２において、可視光の撮像が行われる。第１撮像素子１３

－２は、可視光画像信号を取得部１１０に出力する。なお、撮像対象の波長帯域が異なる複数の撮像素子を、光軸方向において積層する手法については広く知られているため、詳細な説明は省略する。

[0024] 取得部１１０は、図２と同様に、第１Ａ／Ｄ変換回路１１１と、第２Ａ／Ｄ変換回路１１２を含む。第１Ａ／Ｄ変換回路１１１は、第１撮像素子１３－２からの可視光画像信号に対するＡ／Ｄ変換処理を行い、デジタルデータである可視光画像データを出力する。第２Ａ／Ｄ変換回路１１２は、第２撮像素子１５－２からの赤外光画像信号に対するＡ／Ｄ変換処理を行い、デジタルデータである赤外光画像データを出力する。

[0025] なお、取得部１１０は図２及び図３に示した構成に限定されない。例えば取得部１１０は、可視光画像信号及び赤外光画像信号に対する増幅処理を行うアナログアンプ回路を含んでもよい。取得部１１０は、増幅処理後の画像信号に対するＡ／Ｄ変換処理を行う。また、アナログアンプ回路は取得部１１０に設けられるのではなく、撮像部１０側に設けられてもよい。また、図２においては取得部１１０においてＡ／Ｄ変換を行う例を示したが、撮像部１０においてＡ／Ｄ変換が行われてもよい。この場合、撮像部１０は、デジタルデータである可視光画像及び赤外光画像を出力する。取得部１１０は、撮像部１０からのデジタルデータを取得するためのインターフェースである。

[0026] 以上のように、撮像部１０は、第１光軸を用いて対象物を可視光によって撮像し、且つ、第１光軸に対応する軸である第２光軸を用いて当該対象物を赤外光によって撮像する。ここでの対象物は、後述するように第１対象物と、第１対象物に比べて可視光を透過する第２対象物を含む複数の対象物である。具体的には、第１対象物とは可視光を反射する可視物体であり、第２対象物とは可視光を透過する透明物体である。第１光軸と第２光軸とは、狭義には図２及び図３の光軸ＡＸに示した同一の軸である。撮像部１０は情報処理装置１００に含まれてもよい。取得部１１０は、撮像部１０による撮像に基づいて、第１検出用画像及び第２検出用画像を取得する。第１検出用画像

とは可視光画像であり、第２検出用画像とは赤外光画像である。

[0027] このように、撮像部１０は、可視光と赤外光の両方において、同一の対象物を同軸で撮像可能である。そのため、可視光画像における透明物体の位置と、赤外光画像における透明物体の位置とを容易に対応付けることが可能である。例えば、可視光画像と赤外光画像とが、画角及び画素数が等しい画像である場合、所与の対象物は、可視光画像と赤外光画像の同じ位置の画素に撮像される。画素の位置とは、基準画素に対して横方向に何画素目、縦方向に何画素目かを表す情報である。よって同じ位置の画素の情報を対応付けることによって、可視光画像と赤外光画像の両方の情報を用いた処理を適切に実行できる。例えば後述するように、第１検出用画像に基づく第１特徴量と、第２検出用画像に基づく第２特徴量を用いて、透明物体である第２対象物の位置検出を適切に行うことが可能である。なお、撮像部１０は、可視光画像と赤外光画像の間で対象物の位置の対応づけが可能な構成であればよく、上記の構成には限定されない。例えば、第１光軸と第２光軸は略等しい軸であればよく、厳密に一致する必要はない。また、可視光画像の画素数と赤外光画像の画素数は同じである必要はない。

[0028] 図４は、処理部１２０の構成例を示す図である。処理部１２０は、第１特徴量抽出部１２１と、第２特徴量抽出部１２２と、第３特徴量抽出部１２３と、位置検出部１２４を含む。なお、本実施形態の処理部１２０は、下記のハードウェアによって構成される。ハードウェアは、デジタル信号を処理する回路及びアナログ信号を処理する回路の少なくとも一方を含むことができる。例えば、ハードウェアは、回路基板に実装された１又は複数の回路装置や、１又は複数の回路素子で構成することができる。１又は複数の回路装置は例えばＩＣ等である。１又は複数の回路素子は例えば抵抗、キャパシター等である。

[0029] また処理部１２０は、下記のプロセッサーによって実現されてもよい。本実施形態の情報処理装置１００は、情報を記憶するメモリーと、メモリーに記憶された情報に基づいて動作するプロセッサーと、を含む。情報は、例え

ばプログラムと各種のデータ等である。プロセッサーは、ハードウェアを含む。プロセッサーは、ＣＰＵ（Central Processing Unit）、ＧＰＵ（Graphics Processing Unit）、ＤＳＰ（Digital Signal Processor）等、各種のプロセッサーを用いることが可能である。メモリーは、ＳＲＡＭ（Static Random Access Memory）、ＤＲＡＭ（Dynamic Random Access Memory）などの半導体メモリーであってもよいし、レジスターであってもよいし、ハードディスク装置等の磁気記憶装置であってもよいし、光学ディスク装置等の光学式記憶装置であってもよい。例えば、メモリーはコンピューターによって読み取り可能な命令を格納しており、当該命令がプロセッサーによって実行されることで、情報処理装置１００の各部の機能が処理として実現されることになる。ここでの命令は、プログラムを構成する命令セットの命令でもよいし、プロセッサーのハードウェア回路に対して動作を指示する命令であってもよい。

[0030] 第１特徴量抽出部１２１は、取得部１１０の第１Ａ／Ｄ変換回路１１１から、可視光画像である第１検出用画像を取得する。第２特徴量抽出部１２２は、取得部１１０の第２Ａ／Ｄ変換回路１１２から、赤外光画像である第２検出用画像を取得する。なお、可視光画像及び赤外光画像は、取得部１１０から処理部１２０へ直接送信されるものに限定されない。例えば取得部１１０は、取得した可視光画像と赤外光画像を記憶部１３０に書き込む処理を行い、処理部１２０は、記憶部１３０に記憶された可視光画像及び赤外光画像を読み出す処理を行ってもよい。

[0031] 第１特徴量抽出部１２１は、第１検出用画像（可視光画像）の特徴量を第１特徴量として抽出する。第２特徴量抽出部１２２は、第２検出用画像（赤外光画像）の特徴量を第２特徴量として抽出する。第１特徴量及び第２特徴量は、輝度、コントラスト等の種々の特徴量を用いることが可能である。例えば第１特徴量は、可視光画像に対してエッジ抽出フィルターを適用した結果であるエッジ情報である。第２特徴量は、赤外光画像に対してエッジ抽出フィルターを適用した結果であるエッジ情報である。エッジ抽出フィルター

は、例えばラプラシアンフィルター等のハイパスフィルターである。

[0032] ここで、可視光を透過する物体である透明物体と、可視光を透過しない物体である可視物体について、可視光画像及び赤外光画像における傾向を検討する。透明物体は、可視光が透過するため、可視光画像において特徴が現れにくい。即ち、第１特徴量には透明物体の特徴が反映されにくい。また、透明物体は赤外光を吸収するため、赤外光画像には特徴が現れる。即ち、第２特徴量には透明物体の特徴が反映されやすい。これに対して、可視物体は、可視光と赤外光のいずれも透過度合いが小さい。そのため、可視物体は、可視光画像と赤外光画像の両方に特徴が現れる。即ち、第１特徴量と第２特徴量の両方に、可視物体の特徴が反映される。

[0033] 以上の点を考慮し、第３特徴量抽出部１２３は、第１特徴量と第２特徴量の差分を第３特徴量として算出する。第１特徴量と第２特徴量の差分を取ることによって、透明物体の特徴を表す情報が強調される。具体的には、赤外光画像に基づく第２特徴量が強調される。一方、第１特徴量と第２特徴量の両方に含まれる可視物体の特徴は、差分演算によってキャンセルされる。そのため、第３特徴量には透明物体の特徴量が支配的に現れる。

[0034] 位置検出部１２４は、第３特徴量に基づいて、可視光画像及び赤外光画像の少なくとも一方における透明物体の位置情報を検出した後、検出結果を出力する。例えば第３特徴量がエッジを表す情報である場合、位置検出部１２４は、透明物体のエッジの位置を表す情報、或いはエッジによって囲まれる領域の位置を表す情報を、位置情報として出力する。

[0035] なお可視光画像と赤外光画像とで、光軸、画角、画素数等の条件が等しい場合、可視光画像における透明物体の位置と、赤外光画像における透明物体の位置とは等価である。また、光軸等に差があったとしても、本実施形態では可視光画像における所与の対象物の位置と、赤外光画像における当該対象物の位置とが対応づけ可能であることを想定している。そのため、可視光画像及び赤外光画像の一方における透明物体の位置情報に基づいて、他方における透明物体の位置情報を特定することが可能である。位置検出部１２４は

、可視光画像及び赤外光画像の両方における位置情報を求めてもよいし、いずれか一方における透明物体の位置情報を求めてもよい。

[0036] 図５（Ａ）、図５（Ｂ）は、透明物体の一例であるガラス扉を示す図である。図５（Ａ）はカラス扉が閉まった状態を表し、図５（Ｂ）はガラス扉が開いた状態を表す。図５（Ａ）、図５（Ｂ）に示す例においては、Ａ１に示す矩形状の領域内に、Ａ２及びＡ３に示す２枚のガラスが配置される。２枚のガラスのうち、Ａ２に示すガラスが水平方向に移動することによって、ガラス扉が開閉される。図５（Ａ）に示す閉状態においては、Ａ１の領域のほぼ全域にＡ１とＡ２の２枚のガラスが配置される。図５（Ｂ）に示す開状態においては、Ａ１の左方領域にはガラスが存在せず、右方領域にガラスが２枚重複した状態となる。なお、Ａ１以外の領域は例えば建造物の壁面等であり、ここでは説明を簡略化するため、凹凸がなく、色味の変化も少ない一様な物体であると考える。

[0037] 図６は、ガラス扉が閉まった状態における可視光画像と赤外光画像の例、及び第１～第３特徴量の例を示す図である。図６のＢ１が可視光画像の例であり、Ｂ２が赤外光画像の例である。

[0038] 可視光はガラスを透過するため、可視光画像は、ガラスが存在する領域において、ガラスよりも奥にある対象物が撮像される。奥とは具体的には撮像部１０との距離がガラスよりも遠い側の空間を表す。図６のＢ１に示す例においては、ガラスの奥に存在する可視物体であるＢ１１～Ｂ１３が撮像される。

[0039] Ｂ３は、Ｂ１の可視光画像に対して、エッジ抽出フィルター等を適用することによって取得される第１特徴量の例である。上述したように、ガラス以外の領域には例えば建造物の壁面が撮像され、ガラスが存在する領域にはＢ１１～Ｂ１３等のガラスよりも奥にある物体が撮像される。ガラスとそれ以外の領域とで撮像される物体が異なるため、境界においてエッジが検出される。結果として、ガラス領域の境界において、第１特徴量の値が大きくなる（Ｂ３１）。また、ガラスが存在する領域の内部においては、Ｂ１１～Ｂ１

３等のガラスの奥に存在する物体に起因するエッジが検出されるため、第１特徴量の値はある程度大きくなる（Ｂ３２）。

[0040] また、赤外光はガラスによって吸収されるため、Ｂ２に示す赤外光画像においては、ガラスが存在する領域は輝度値が小さく、且つ、ローコントラストな領域として撮像される。また、ガラスよりも奥に物体が存在したとしても、当該物体は赤外光画像には撮像されない。

[0041] Ｂ４は、Ｂ２の赤外光画像に対して、エッジ抽出フィルター等を適用することによって取得される第２特徴量の例である。ガラス以外の領域と、ガラスが存在する領域とで輝度値に差が出るため、ガラス領域の境界において、第２特徴量の値が大きくなる（Ｂ４１）。また、ガラスが存在する領域は、上述したようにローコントラストであるため、第２特徴量の値は非常に小さい（Ｂ４２）。

[0042] Ｂ５は、第１特徴量と第２特徴量の差分である第３特徴量の例である。差分をとることによって、ガラスに対応する領域であるＢ５１において第３特徴量の値が大きくなる。一方、それ以外の領域においては、可視光画像と赤外光画像とで同様の特徴が検出されるため、差分によって得られる第３特徴量の値は小さくなる。例えば、ガラス領域と可視物体の境界においては、第１特徴量と第２特徴量の両方においてエッジが検出されるため、当該エッジはキャンセルされる。また、ガラス領域以外の可視物体は、第１特徴量と第２特徴量が同様の傾向を示すため、やはり値がキャンセルされる。なお、図６においては、可視物体がローコントラストである例を示したが、可視物体が何らかのエッジを有する場合であっても、差分によって特徴がキャンセルされる点は同様である。

[0043] 図６に示した例においては、位置検出部１２４は、第３特徴量の値が所与の閾値よりも大きい画素を、透明物体に対応する画素であると特定する。例えば位置検出部１２４は、第３特徴量の値が所与の閾値よりも大きい画素を連結した領域に基づいて、透明物体に対応する位置、形状を判定する。位置検出部１２４は、検出した透明物体の位置情報を、記憶部１３０に記憶する

。或いは、情報処理装置１００は不図示の表示部を含み、位置検出部１２４は、検出した透明物体の位置情報を提示するための画像データを、表示部に出力してもよい。ここでの画像データは、例えば可視光画像に対して、透明物体の位置を表す情報が付加された情報である。

[0044] また本実施形態の手法は、ガラス扉の開閉の判断に用いることが可能である。図７は、ガラス扉が開いた状態における可視光画像と赤外光画像の例、及び第１～第３特徴量の例を示す図である。図７のＣ１が可視光画像の例であり、Ｃ２が赤外光画像の例である。Ｃ３～Ｃ５が第１～第３特徴量の例である。

[0045] ガラス扉が開いた状態においては、ガラス扉の左方領域はガラスが存在しない開口となる。ガラス扉よりも奥に存在する対象物から照射される赤外光は、ガラスに吸収されることなく、撮像部１０に到達可能である。そのため、可視光画像においてＣ１１、Ｃ１２が撮像されるだけでなく、赤外光画像においても、同じ対象物（Ｃ２１、Ｃ２２）が撮像される。一方、ガラスの存在する右方領域については閉状態と同様であり、可視光画像では奥の対象物（Ｃ１３）が撮像されるのに対して、赤外光画像では当該対象物が撮像されない。

[0046] 結果として、ガラスの存在しない左方領域においては、第１特徴量と第２特徴量の両方の値が大きくなり、差分によってキャンセルされる（Ｃ３１，Ｃ４１，Ｃ５１）。一方、ガラスの存在する右方領域においては、第１特徴量は奥の物体の特徴を反映し、且つ、第２特徴量はローコントラストとなるため、差分によって第３特徴量の値が大きくなる（Ｃ３２，Ｃ４２，Ｃ５２）。

[0047] 従来手法は、可視光画像及び赤外光画像に基づいて、物体の形状やテクスチャ等の特徴を求め、当該特徴からガラスを判断する手法である。そのため、ローコントラストな四角い枠であれば、その他の対象物と分別が難しい。しかし本実施形態の手法は、図６及び図７を用いて説明したとおり、赤外光はガラスを撮像し、可視光はガラスを透過することによって奥の対象物を撮

像する、という波長帯域に応じた撮像対象の違いを利用する。透明物体が存在する領域においては、別の対象物を撮像しているため、形状やテクスチャが同じであっても、特徴の違いが大きくなる。一方、透明物体でない領域においては、同一の対象物を撮像している為、特徴の違いは大きくならない。本実施形態の手法は、第１特徴量と第２特徴量の差分に対応する第３特徴量を用いることによって、従来手法に比べて透明物体を精度よく検出することが可能になる。また図６及び図７を用いて上述したとおり、透明物体の有無だけではなく、位置や形状を検出することが可能である。また、図７を用いて上述したとおり、透明物体が移動した結果、開口となった領域を透明物体であると誤検出することを抑制できるため、可動式の透明物体の検出、具体的にはガラス扉等の開閉を判断することも可能である。

[0048] 図８は、本実施形態の処理を説明するフローチャートである。この処理が開始されると、取得部１１０は、第１検出用画像である可視光画像と、第２検出用画像である赤外光画像を取得する（Ｓ１０１、Ｓ１０２）。例えば処理部１２０は、撮像部１０及び取得部１１０の制御を行う。次に処理部１２０は、可視光画像に基づく第１特徴量の抽出、及び赤外光画像に基づく第２特徴量の抽出を行う（Ｓ１０３，Ｓ１０４）。Ｓ１０３及びＳ１０４の処理は、例えば上述したようにエッジ抽出フィルターを用いたフィルター処理である。ただし図６及び図７を用いて上述したとおり、本実施形態の手法は、撮像対象の物体が同じであるか否かに基づいて、透明物体を検出する。そのため、第１特徴量及び第２特徴量は撮像対象となる物体の特徴を反映する情報であればよく、エッジに限定されるものではない。

[0049] 次に処理部１２０は、第１特徴量と第２特徴量の差分を演算することによって、第３特徴量を抽出する（Ｓ１０５）。処理部１２０は、第３特徴量に基づいて透明物体の位置検出を行う（Ｓ１０６）。Ｓ１０６の処理は、例えば上述したように、第３特徴量の値と所与の閾値との比較処理である。

[0050] 以上のように、本実施形態の情報処理装置１００は、取得部１１０と、処理部１２０を含む。取得部１１０は、第１対象物と、第１対象物に比べて可

視光を透過する第２対象物とを含む複数の対象物を可視光によって撮像した第１検出用画像と、複数の対象物を赤外光によって撮像した第２検出用画像を取得する。処理部１２０は、第１検出用画像に基づいて第１特徴量を求め、第２検出用画像に基づいて第２特徴量を求め、第１特徴量と第２特徴量の差分に対応する特徴量を第３特徴量として算出する。処理部１２０は、第３特徴量に基づいて、第１検出用画像及び第２検出用画像の少なくとも一方における第２対象物の位置を検出する。

[0051] 以上では、第３特徴量が、第１特徴量と第２特徴量の差分そのものである例を説明した。ただし、第３特徴量は、差分に対応する演算、即ち、第１特徴量と第２特徴量の両方に含まれる特徴をキャンセル可能な演算によって求められる特徴量であればよく、具体的な演算は差分そのものに限定されない。例えば、第２特徴量の一方の符号を反転し加算する処理は、差分に対応する演算に含まれる。また第３特徴量抽出部１２３は、第１特徴量に第１係数を乗算し、第２特徴量に第２係数を乗算し、２つの乗算結果を加算することによって第３特徴量を求めてもよい。また第３特徴量抽出部１２３は、第１特徴量と第２特徴量の比率、或いはそれに準ずる情報を、差分に対応する特徴量として求めてもよい。この場合、位置検出部１２４は、比率である第３特徴量が、１から所定閾値以上乖離する画素を、透明物体であると判定する。

[0052] 本実施形態の手法によれば、可視光画像と赤外光画像からそれぞれ特徴量を求め、それらの差分に基づく特徴量を用いて透明物体を検出する。このようにすれば、可視光画像における可視物体の特徴、可視光画像における透明物体の特徴、赤外光画像における可視物体の特徴、赤外光画像における透明物体の特徴をそれぞれ考慮した、精度の高い透明物体の位置検出が可能になる。

[0053] また第１特徴量は、第１検出用画像のコントラストを表す情報であり、第２特徴量は、第２検出用画像のコントラストを表す情報である。そして処理部１２０は、第１検出用画像のコントラストと第２検出用画像のコントラス

トの差分に対応する第３特徴量に基づいて、第１検出用画像及び第２検出用画像の少なくとも一方における第２対象物の位置を検出する。

[0054] このようにすれば、コントラストを特徴量として用いることによって、透明物体の位置検出を行うことが可能になる。なおここでのコントラストは、所与の画素と、当該画素の近傍の画素との画素値の相違度合いを表す情報である。例えば上述したエッジは、画素値の変化が急峻な領域を表す情報であるため、コントラストを表す情報に含まれる。ただし、コントラストを求める画像処理は種々知られており、本実施形態においてはそれらを広く適用可能である。例えばコントラストは、所定領域における画素値の最大値と最小値の差分に基づく情報であってもよい。或いは、コントラストを表す情報は、ローコントラストである領域において値が大きくなるような情報であってもよい。

[0055] また本実施形態の手法は、上記の情報処理装置１００を含む移動体に適用できる。情報処理装置１００は、自動車、飛行機、バイク、自転車、ロボット、或いは船舶等の種々の移動体に組み込むことができる。移動体は、例えばエンジンやモーター等の駆動機構、ハンドルや舵等の操舵機構、各種の電子機器を備えて、地上や空や海上を移動する機器・装置である。移動体は、例えば情報処理装置１００と、移動体の移動制御を行う制御装置３０とを含む。図９（Ａ）～図９（Ｃ）は、本実施形態にかかる移動体の例を示す図である。なお図９（Ａ）～図９（Ｃ）においては、撮像部１０が情報処理装置１００の外部に設けられる例を示している。

[0056] 図９（Ａ）に示す例においては、移動体は、例えば自律走行を行う車椅子２０である。車椅子２０は、撮像部１０と、情報処理装置１００と、制御装置３０とを含む。なお図９（Ａ）においては情報処理装置１００と制御装置３０が一体として設けられる例を示したが、これらは別体として設けられてもよい。

[0057] 情報処理装置１００は、上述した処理を行うことによって、透明物体の位置情報を検出する。制御装置３０は、位置検出部１２４が検出した位置情報

を、情報処理装置１００から取得する。そして制御装置３０は、取得した透明物体の位置情報に基づいて、車椅子２０と透明物体との衝突を抑制するための駆動部の制御を行う。ここでの駆動部は、例えば車輪２１を回転させるためのモーターである。なお障害物との衝突を回避するための移動体制御については種々の手法が知られているため、詳細な説明は省略する。

[0058] また移動体は図９（Ｂ）に示すロボットであってもよい。ロボット４０は、頭部に設けられる撮像部１０と、本体部４１に内蔵される情報処理装置１００及び制御装置３０と、アーム４３と、ハンド４５と、車輪４７とを含む。制御装置３０は、位置検出部１２４が検出した透明物体の位置情報に基づいて、ロボット４０と透明物体との衝突を抑制するための駆動部の制御を行う。例えば、制御装置３０は、透明物体の位置情報に基づいて、当該透明物体に衝突しないようなハンド４５の移動経路を生成する処理、当該移動経路に沿ったハンド４５の移動を実現し且つアーム４３が透明物体に衝突しないアーム姿勢の生成処理、及び生成された情報に基づいて駆動部を制御する処理等を行う。ここでの駆動部は、アーム４３、ハンド４５を駆動するためのモーターである。また駆動部は車輪４７を駆動するためのモーターを含み、制御装置３０は、ロボット４０と透明物体との衝突を抑制するための、車輪駆動制御を行ってもよい。なお、図９（Ｂ）においてはアームを有するロボットを例示したが、本実施形態の手法は種々の態様のロボットに適用可能である。

[0059] また移動体は図９（Ｃ）に示す自動車６０であってもよい。自動車６０は、撮像部１０と、情報処理装置１００と、制御装置３０を含む。撮像部１０は、例えばドライレコーダー等と併用可能な車載カメラである。制御装置３０は、位置検出部１２４が検出した透明物体の位置に基づいて、自動運転のための種々の制御処理を行う。制御装置３０は、例えば個々の車輪６１のブレーキを制御する。また制御装置３０は、透明物体の検出結果を、表示部６３に表示する制御を行ってもよい。

[0060] ２．第２の実施形態

図１０は、第２の実施形態における処理部１２０の構成例を示す図である。処理部１２０は、図４に示した構成に加えて、第４特徴量を算出する第４特徴量抽出部１２５をさらに含む。

[0061] 第３特徴量抽出部１２３は、第１の実施形態と同様に、第１特徴量と第２特徴量の差分を算出することによって、透明物体に支配的な第３特徴量を算出する。第３特徴量を用いることによって、透明物体の位置を精度よく検出することが可能になる。

[0062] 第４特徴量抽出部１２５は、第１検出用画像（可視光画像）と第２検出用画像（赤外光画像）を合成した画像である第３検出用画像を用いて、可視物体の特徴量を第４特徴量として検出する。第３検出用画像とは、例えば各画素について、可視光画像の画素値と赤外光画像の画素値を合成した画像である。具体的には第４特徴量抽出部１２５は、赤色光に対応するＲ画像の画素値と、緑色光に対応するＧ画像の画素値と、青色光に対応するＢ画像の画素値と、赤外光画像の画素値と、の平均値を画素ごとに求めることによって、第３検出用画像を生成する。ここでの平均は単純平均であってもよいし加重平均であってもよい。例えば第４特徴量抽出部１２５は、ＲＧＢの３つの画像に基づいて、輝度画像信号Ｙを求め、当該輝度画像信号と赤外光画像とを合成してもよい。

[0063] 第４特徴量抽出部１２５は、例えば第３検出用画像に対してエッジ抽出フィルターを用いたフィルター処理を行うことによって第４特徴量を求める。ただし第４特徴量はエッジに限定されず、種々の変形実施が可能である。また、第４特徴量抽出部１２５は、第３検出用画像を用いて第４特徴量を算出するだけでなく、可視光画像と赤外光画像から個別に抽出した特徴量を足し合わせることによって、第４特徴量を求めてもよい。

[0064] 位置検出部１２４は、第３特徴量に基づいて透明物体の位置検出を行い、第４特徴量に基づいて可視物体の位置検出を行う。これにより、位置検出部１２４は、透明物体と可視物体の両方を区別して位置検出を行う。或いは、位置検出部１２４は、第３特徴量と第４特徴量を合わせて用いることによっ

て、可視物体と透明物体の両方を区別して位置検出を行ってもよい。

[0065] 図１１は、本実施形態の処理を説明するフローチャートである。図１１のＳ２０１～Ｓ２０５については、図８のＳ１０１～Ｓ１０５と同様であり、処理部１２０は、第１特徴量と第２特徴量に基づいて第３特徴量を求める。また処理部１２０は、可視光画像と赤外光画像に基づいて、第４特徴量を抽出する（Ｓ２０６）。例えば処理部１２０は、上述したように、可視光画像と赤外光画像を合成することによって第３検出用画像を求め、当該第３検出用画像から第４特徴量を抽出する。

[0066] 次に処理部１２０は、第３特徴量と第４特徴量に基づいて透明物体の位置検出及び可視物体の位置検出を行う（Ｓ２０７）。Ｓ２０７の処理は、例えば第３特徴量の値と所与の閾値との比較処理による透明物体の検出処理、及び、第４特徴量の値と他の閾値との比較処理による可視物体の検出処理を含む。

[0067] 以上のように、本実施形態の処理部１２０は、第１検出用画像及び第２検出用画像に基づいて、第１対象物の特徴を表す第４特徴量を求める。そして処理部１２０は、第３特徴量及び第４特徴量に基づいて、第１対象物の位置と第２対象物の位置の両方を区別して検出する。このようにすれば、画像内に可視物体と透明物体とが混在する場合にも、画像内の各物体の位置を適切に検出することが可能になる。また、暗いシーンでは可視光による特徴量が乏しいため、可視光画像のみを用いた場合、可視物体の検出精度が低下するおそれがある。その点、本実施形態の手法では、第４特徴量の抽出に可視光画像と赤外光画像の両方を用いるため、暗いシーンであっても可視物体を精度よく検出することが可能になる。

[0068] ３．第３の実施形態

第２の実施形態において、位置検出に用いる第３特徴量及び第４特徴量を求めるためには、エッジ抽出フィルター等の特性をあらかじめ設定しておく必要がある。一例としては、可視物体や透明物体の特徴が適切に抽出されるようなフィルター特性をユーザーが手動で設定する。ただし、特徴量抽出処

理を含む位置検出処理に機械学習を適用してもよい。

[0069] 本実施形態の情報処理装置１００は、学習済モデルを記憶する記憶部１３０を含む。学習済モデルは、第１学習用画像と、第２学習用画像と、第１対象物の位置情報及び第２対象物の位置情報と、を対応付けたデータセットに基づいて機械学習されている。第１学習用画像とは、第１対象物（可視物体）と第２対象物（透明物体）を含む複数の対象物を可視光によって撮像した可視光画像である。第２学習用画像とは、上記複数の対象物を赤外光によって撮像した赤外光画像である。処理部１２０は、第１検出用画像と、第２検出用画像と、学習済モデルとに基づいて、第１検出用画像及び第２検出用画像の少なくとも一方において、第１対象物の位置と第２対象物の位置の両方を区別して検出する。

[0070] このように機械学習を用いることによって、可視物体及び透明物体の位置を精度よく検出することが可能になる。以下、学習処理、及び学習済モデルを用いた推論処理について説明する。なお、以下ではニューラルネットワークを用いた機械学習について説明するが、本実施形態の手法はこれに限定されない。本実施形態においては、例えばＳＶＭ（support vector machine）等の他のモデルを用いた機械学習が行われてもよいし、ニューラルネットワークやＳＶＭ等の種々の手法を発展させた手法を用いた機械学習が行われてもよい。

[0071] ３．１ 学習処理

図１２は、本実施形態の学習装置２００の構成例を示す図である。学習装置２００は、学習に用いられる訓練データを取得する取得部２１０と、当該訓練データに基づいて機械学習を行う学習部２２０を含む。

[0072] 取得部２１０は、例えば訓練データを他の装置から取得する通信インターフェースである。或いは取得部２１０は、学習装置２００が保持する訓練データを取得してもよい。例えば、学習装置２００は不図示の記憶部を含み、取得部２１０は当該記憶部から訓練データを読み出すためのインターフェースである。本実施形態における学習は、例えば教師あり学習である。教師あ

り学習における訓練データは、入力データと正解ラベルとを対応付けたデータセットである。

[0073] 学習部２２０は、取得部２１０が取得した訓練データに基づく機械学習を行い、学習済モデルを生成する。なお、本実施形態の学習部２２０は、情報処理装置１００の処理部１２０と同様に、デジタル信号を処理する回路及びアナログ信号を処理する回路の少なくとも一方を含むハードウェアによって構成される。例えば、ハードウェアは、回路基板に実装された１又は複数の回路装置や、１又は複数の回路素子で構成することができる。また学習装置２００はプロセッサーとメモリーを含み、学習部２２０は、ＣＰＵ、ＧＰＵ、ＤＳＰ等の各種のプロセッサーによって実現されてもよい。メモリーは、半導体メモリーであってもよいし、レジスターであってもよいし、磁気記憶装置であってもよいし、光学式記憶装置であってもよい。

[0074] より具体的には、取得部２１０は、第１対象物と、第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を可視光によって撮像した可視光画像と、複数の対象物を赤外光によって撮像した赤外光画像と、可視光画像及び赤外光画像の少なくとも一方における第１対象物及び第２対象物の位置情報と、を対応付けたデータセットを取得する。学習部２２０は、当該データセットに基づいて、可視光画像及び赤外光画像の少なくとも一方において、第１対象物を検出する条件、及び第２対象物の位置を検出する条件を機械学習する。

[0075] このような機械学習を行うことによって、可視物体と透明物体の位置を精度よく検出することが可能になる。例えば第２の実施形態においては、第１特徴量、第２特徴量及び第４特徴量を抽出するためのフィルター特性をユーザーが手動で設定する必要がある。そのため、可視物体や透明物体の特徴を効率的に抽出可能なフィルターを多数設定することが難しい。その点、機械学習を用いることによって、多数のフィルター特性を自動的に設定することが可能である。そのため、第２の実施形態に比べて、可視物体及び透明物体の位置を精度よく検出することが可能になる。

[0076] 図１３は、ニューラルネットワークを説明する模式図である。ニューラルネットワークは、データが入力される入力層と、入力層からの出力に基づいて演算を行う中間層と、中間層からの出力に基づいてデータを出力する出力層を有する。図１３においては、中間層が２層であるネットワークを例示するが、中間層は１層であってもよいし、３層以上であってもよい。また各層に含まれるノード（ニューロン）の数は図１３の例に限定されず、種々の変形実施が可能である。なお精度を考慮すれば、本実施形態の学習は多層のニューラルネットワークを用いた深層学習（ディープラーニング）を用いることが望ましい。ここでの多層とは、狭義には４層以上である。

[0077] 図１３に示すとおり、所与の層に含まれるノードは、隣接する層のノードと結合される。各結合には重みが設定されている。例えば、所与の層に含まれる各ノードが、次の層の全てのノードと接続される全結合のニューラルネットワークを用いる場合、当該２つの層の間の重みは、所与の層に含まれるノード数と、次の層に含まれるノード数とを乗算した値の集合となる。各ノードは、前段のノードの出力と重みを乗算し、乗算結果の合計値を求める。さらに各ノードは、合計値に対してバイアスを加算し、加算結果に活性化関数を適用することによって当該ノードの出力を求める。活性化関数としては、ＲｅＬＵ関数が知られている。ただし、活性化関数は種々の関数を利用可能であることが知られており、シグモイド関数を用いてもよいし、ＲｅＬＵ関数を改良した関数を用いてもよいし、他の関数を用いてもよい。

[0078] 以上の処理を、入力層から出力層へ向けて順次実行することによって、ニューラルネットワークの出力が求められる。ニューラルネットにおける学習は、適切な重み（バイアスを含む）を決定する処理である。具体的な学習手法として、誤差逆伝播法等の種々の手法が知られており、本実施形態においてはそれらを広く適用可能である。なお誤差逆伝播法については公知であるため、詳細な説明は省略する。

[0079] ただし、ニューラルネットワークは図１３に示した構成に限定されない。例えば、学習処理及び推論処理において、畳み込みニューラルネットワーク

（ＣＮＮ：Convolutional Neural Network）が用いられてもよい。ＣＮＮは、例えば畳み込み演算を行う畳み込み層とプーリング層を含む。畳み込み層は、フィルター処理を行う層である。プーリング層は、縦方向、横方向のサイズを縮小するプーリング演算を行う層である。ＣＮＮの畳み込み層における重みは、フィルターのパラメータである。即ち、ＣＮＮにおける学習とは、畳み込み演算に用いるフィルター特性の学習を含む。

[0080] 図１４は、本実施形態におけるニューラルネットワークの構成を示す模式図である。図１４のＤ１は、３チャンネルの可視光画像を入力として受け付け、畳み込み演算を含む処理を行うことによって、第１特徴量を求めるブロックである。第１特徴量は、例えば可視光画像に対して、２５６通りのフィルター処理を行うことによって求められた、２５６チャンネルの第１特徴マップである。なお特徴マップのチャンネル数は２５６に限定されず、種々の変形実施が可能である。

[0081] Ｄ２は、１チャンネルの赤外光画像を入力として受け付け、畳み込み演算を含む処理を行うことによって、第２特徴量を求めるブロックである。第２特徴量は、例えば２５６チャンネルの第２特徴マップである。

[0082] Ｄ３は、第１特徴マップと第２特徴マップの差分を求める処理を行うことによって、第３特徴量を求めるブロックである。第３特徴量は、例えば第１特徴マップのｉ（ｉは１以上２５６以下の整数）チャンネル目の特徴マップの各画素値から、第２特徴マップのｉチャンネル目の特徴マップの各画素値を減算する処理を、チャンネルごと行うことによって取得される２５６チャンネルの第３特徴マップである。

[0083] Ｄ４は、３チャンネルの可視光画像及び１チャンネルの赤外光画像を合わせた４チャンネルの画像を入力として受け取り、畳み込み演算を含む処理を行うことによって、第４特徴量を求めるブロックである。第４特徴量は、例えば２５６チャンネルの第４特徴マップである。

[0084] なお図１４においては、Ｄ１，Ｄ２，Ｄ４の各ブロックが、１つの畳み込み層と１つのプーリング層を含む例を示した。しかし畳み込み層及びプーリ

ング層の少なくとも一方を２層以上に増やしてもよい。また図１４においては省略されているが、Ｄ１，Ｄ２，Ｄ４の各ブロックにおいて、例えば畳み込み演算の結果に対して活性化関数を適用する演算処理が行われる。

[0085] Ｄ５は、第３特徴マップと第４特徴マップを合わせた５１２チャンネルの特徴マップに基づいて、可視物体と透明物体の位置検出を行うブロックである。図１４においては、５１２チャンネルの特徴マップに対して、畳み込み層、プーリング層、アップサンプリング層、畳み込み層、ソフトマックス層による演算を行う例を示したが、具体的な構成については種々の変形実施が可能である。アップサンプリング層は、縦方向及び横方向のサイズを拡大する層であり、逆プーリング層と言い換えてもよい。ソフトマックス層とは、公知のソフトマックス関数による演算を行う層である。

[0086] 例えば可視物体と、透明物体と、その他の物体とを分類する場合、ソフトマックス層の出力は３チャンネルの画像データである。各チャンネルの画像データは、例えば入力である可視光画像及び赤外光画像と同じ画素数の画像である。第１チャンネルの各画素は、当該画素が可視物体である確率を表す０以上１以下の数値データである。第２チャンネルの各画素は、当該画素が透明物体である確率を表す０以上１以下の数値データである。第３チャンネルの各画素は、当該画素がその他の物体である確率を表す０以上１以下の数値データである。本実施形態におけるニューラルネットワークの出力は、上記３チャンネルの画像データである。或いはニューラルネットワークの出力は、各画素について、最も確率が高い物体を表すラベルと、その確率が対応付けられた画像データであってもよい。例えばラベルは（０，１，２）の３通りであり、０が可視物体、１が透明物体、２がその他の物体である。例えばある画素において、可視物体である確率が０．３、透明物体である確率が０．５、その他の物体である確率が０．２となった場合、出力データにおける当該画素には、透明物体を表す“１”というラベルと、０．５という確率が割り当てられる。なお、ここでは３つの物体を分類する例を示したが、分類数はこれに限定されない。例えば処理部１２０は、可視物体をさらに人と

道路等に分類する等、４種類以上の物体を分類してもよい。

[0087] 本実施形態における訓練データは、同軸で撮像された可視光画像及び赤外光画像と、当該画像に対応付けられた位置情報である。位置情報は、例えば各画素について、（０，１，２）のいずれかのラベルが付与された情報である。上述したように、ここでのラベルは、０が可視物体、１が透明物体、２がその他の物体を表す。

[0088] 学習処理においては、まずニューラルネットワークに入力データを入力し、そのときの重みを用いて順方向の演算を行うことによって、出力データを取得する。本実施形態においては、入力データは、３チャンネルの可視光画像、１チャンネルの赤外光画像、３チャンネルの可視光画像及び１チャンネルの赤外光画像を合わせた４チャンネルの画像、の３つである。順方向の演算によって求められる出力データは、例えば上述したソフトマックス層の出力であり、１つの画素について、可視物体である確率ｐ０、透明物体である確率ｐ１、その他の物体である確率ｐ２（ｐ０～ｐ２はそれぞれ０以上１以下、且つｐ０＋ｐ１＋ｐ２＝１を満たす数）の３つが対応付けられた３チャンネルのデータである。

[0089] 学習部２２０は、求められた出力データと、正解ラベルとに基づいて誤差関数（損失関数）を演算する。正解ラベルが０である場合、当該画素は可視物体であるため、可視物体である確率ｐ０は１となり、透明物体である確率ｐ１及びその他の物体である確率ｐ２は０となるべきである。よって学習部２２０は、１とｐ０の相違度を誤差関数として算出し、誤差が小さくなる方向に重みを更新する。なお誤差関数は種々の形式が知られており、本実施形態においてはそれらを広く適用可能である。また重みの更新は例えば誤差逆伝播法を用いて行われるが、他の手法を用いてもよい。また学習部２２０は、０とｐ１の相違度、及び０とｐ２の相違度に基づいて誤差関数を演算し、重みを更新してもよい。

[0090] 以上が１組のデータセットに基づく学習処理の概要である。学習処理においては、多数のデータセットを用意しておき、上記処理を繰り返すことによ

って、適切な重みを学習する。例えば、学習段階において図９（Ａ）～図９（Ｃ）に示した移動体を移動させることによって、可視光画像と赤外光画像が取得されてもよい。可視光画像及び赤外光画像に対して、ユーザーが正解ラベルである位置情報を付加することによって、訓練データが取得される。この場合、図１２に示した学習装置２００は、情報処理装置１００と一体として構成されてもよい。或いは、学習装置２００は移動体とは別体として設けられ、移動体から可視光画像及び赤外光画像を取得することによって学習処理を行ってもよい。或いは、学習段階においては、移動体自体を用いずに、撮像部１０と同様の構成の撮像装置を用いて可視光画像と赤外光画像が取得されてもよい。

[0091] 図１５は、学習装置２００における処理を説明するフローチャートである。この処理が開始されると、学習装置２００の取得部２１０は、可視光画像である第１学習用画像と、赤外光である第２学習用画像を取得する（Ｓ３０１，Ｓ３０２）。また取得部２１０は、第１学習用画像と第２学習用画像に対応する位置情報を取得する（Ｓ３０３）。位置情報は、例えば上述したように、ユーザーによって付与された情報である。

[0092] 次に学習部２２０は、取得した訓練データに基づいて学習処理を行う（Ｓ３０４）。Ｓ３０４の処理は、例えば１組のデータセットに基づいて、順方向の演算、誤差関数の算出、誤差関数に基づく重みの更新、の各処理を１回行う処理である。次に学習部２２０は、機械学習を終了するか否かを判定する（Ｓ３０５）。例えば学習部２２０は、取得した多数のデータセットを、訓練データと検証データに分けておく。そして学習部２２０は、訓練データに基づいて学習処理を行うことによって取得された学習済モデルに対して、検証データを用いた処理を行うことによって精度を判定する。検証データは、正解ラベルである位置情報が対応付けられているため、学習部２２０は、学習済モデルに基づいて検出された位置情報が正解であるか否かを判定可能である。学習部２２０は、検証データに対する正解率が所定閾値以上である場合に、学習を終了すると判定し（Ｓ３０５でＹｅｓ）、処理を終了する。

或いは、学習部２２０は、Ｓ３０４に示す処理を所定回数実行した場合に、学習を終了すると判定してもよい。

[0093] 以上のように、本実施形態における第１特徴量は、第１検出用画像に対して、第１フィルターを用いた畳み込み演算を行うことによって求められる第１特徴マップである。第２特徴量は、第２検出用画像に対して、第２フィルターを用いた畳み込み演算を行うことによって求められる第２特徴マップである。第１フィルターとは、図１４のＤ１１に示す畳み込み層における演算に用いられるフィルター群であり、第２フィルターとは、図１４のＤ２１に示す畳み込み層における演算に用いられるフィルター群である。このように、可視光画像と赤外光画像に対して、それぞれ異なる空間フィルターを用いて畳み込み演算を行うことによって第１特徴量と第２特徴量が求められる。そのため、可視光画像に含まれる特徴と、赤外光画像に含まれる特徴を適切に抽出することが可能になる。

[0094] また、第１フィルター及び第２フィルターは、機械学習によってフィルター特性が設定されている。このように、フィルター特性を機械学習を用いて設定することによって、可視光画像と赤外光画像に含まれる各物体の特徴を適切に抽出することが可能になる。例えば、図１４に示したように２５６チャンネル等の多様な特徴を抽出することも可能であるため、特徴量に基づく位置検出処理の精度が向上する。

[0095] また、第４特徴量は、第１検出用画像と第２検出用画像に対して、第４フィルターを用いた畳み込み演算を行うことによって求められる第４特徴マップである。このように、可視光画像と赤外光画像の両方を入力とした畳み込み演算を行うことによって、第４特徴量が求めることが可能になる。また、第４フィルターは、機械学習によってフィルター特性が設定されている。

[0096] また以上では、可視物体と透明物体の両方を区別して検出する場合において、機械学習を適用する手法を説明した。ただし第１の実施形態と同様に、透明物体の位置検出を行う手法に機械学習を用いることも妨げられない。

[0097] この場合、学習装置２００の取得部２１０は、第１対象物と第２対象物と

を含む複数の対象物を可視光によって撮像した可視光画像と、複数の対象物を赤外光によって撮像した赤外光画像と、可視光画像及び赤外光画像の少なくとも一方における第２対象物の位置情報と、を対応付けたデータセットを取得する。学習部２２０は、データセットに基づいて、可視光画像及び赤外光画像の少なくとも一方において、第２対象物の位置を検出する条件を機械学習する。このようにすれば、透明物体の位置検出を精度よく行うことが可能になる。

[0098] ３．２　推論処理

本実施形態における情報処理装置１００の構成例は、図１と同様である。ただし、記憶部１３０は、学習部２２０における学習処理の結果である学習済モデルを記憶する。

[0099] 図１６は、情報処理装置１００における推論処理を説明するフローチャートである。この処理が開始されると、取得部１１０は、可視光画像である第１検出用画像と、赤外光画像である第２検出用画像を取得する（Ｓ４０１，Ｓ４０２）。そして処理部１２０は、記憶部１３０に記憶された学習済モデルからの指令に従って動作することによって、可視光画像及び赤外光画像における、可視物体及び透明物体の位置を検出する処理を行う（Ｓ４０３）。具体的には、処理部１２０は、可視光画像単体、赤外光画像単体、可視光画像と赤外光画像の両方、の３通りのデータを入力データとするニューラルネットワーク演算を行う。

[0100] このようにすれば、可視物体及び透明物体の位置情報を、学習済モデルに基づいて推定することが可能になる。多数の訓練データを用いて機械学習を行うことによって、学習済モデルを用いた処理を高い精度で実行することが可能になる。

[0101] なお学習済モデルは、人工知能ソフトウェアの一部であるプログラムモジュールとして利用される。処理部１２０は、記憶部１３０に記憶された学習済モデルからの指令に従って、入力である可視光画像と赤外光画像における可視物体の位置情報及び透明物体の位置情報を表すデータを出力する。

[0102] なお、学習済モデルに従った処理部１２０おける演算、即ち、入力データに基づいて出力データを出力するための演算は、ソフトウェアによって実行されてもよいし、ハードウェアによって実行されてもよい。換言すれば、ＣＮＮにおける畳み込み演算等は、ソフトウェア的に実行されてもよい。或いは上記演算は、ＦＰＧＡ（field-programmable gate array）等の回路装置によって実行されてもよい。また、上記演算は、ソフトウェアとハードウェアの組み合わせによって実行されてもよい。このように、記憶部１３０に記憶された学習済モデルからの指令に従った処理部１２０の動作は、種々の態様によって実現可能である。

[0103] ４．第４の実施形態

図１７は、第４の実施形態における処理部１２０の構成例を示す図である。情報処理装置１００の処理部１２０は、第２の実施形態における第３特徴量抽出部１２３、第４特徴量抽出部１２５に代えて、それぞれ透過スコア算出部１２６、形状スコア算出部１２７を含む。

[0104] 透過スコア算出部１２６は、第１特徴量と第２特徴量に基づいて、可視光画像及び赤外光画像における各対象物について、対象物の可視光が透過する度合いを示す透過スコアを算出する。例えば、ガラス等の透明物体は、可視光は透過し、赤外光は吸収するため、第１特徴量には特徴量が現れにくく、主に第２特徴量に特徴量が現れる。したがって、透過スコアを第１特徴量と第２特徴量の差分で算出する場合、透明物体の透過スコアは可視物体に対して高くなる。ただし、本実施形態における透過スコアは、可視光の透過度合いを表す情報であればよく、第１特徴量と第２特徴量の差分に対応する情報に限定されない。

[0105] 形状スコア算出部１２７は、第１検出用画像と第２検出用画像を合成した第３学習用画像に基づいて、第１検出用画像および第２検出用画像における各対象物について、対象物の形状を示す形状スコアを算出する。第３検出用画像は、例えば第１検出用画像と第２検出用画像の輝度を各画素について足し合わせて生成される。第３検出用画像は、撮影シーンの明暗に対するロバ

スト性が高く、形状に関する情報を安定して取得することができる。その反面、可視光画像と赤外光画像の輝度を合成しているため、可視光の透過度合いに関する情報は失われている。したがって、形状スコア算出部は、可視光の透過度合いに依存しない対象物の形状のみを示す形状スコアを算出する。

[0106] 位置検出部１２４は、透過スコアと形状スコアに基づいて、透明物体と可視物体の両方を区別して位置検出を行う。例えば位置検出部１２４は、透過スコアが比較的高い値であり、且つ、形状スコアが透明物体に対応する所定形状を示す値である場合は、当該対象物を透明物体と判定する。

[0107] このように本実施形態の情報処理装置１００の処理部１２０は、第１特徴量及び第２特徴量に基づいて、第１検出用画像及び第２検出用画像に撮像された複数の対象物について可視光の透過度合いを表す透過スコアを算出する。また処理部１２０は、第１検出用画像及び第２検出用画像に基づいて、第１検出用画像及び第２検出用画像に撮像された複数の対象物の形状を示す形状スコアを算出する。そして処理部１２０は、透過スコアと形状スコアに基づいて、第１検出用画像及び第２検出用画像の少なくとも一方における、第１対象物の位置と第２対象物の位置の両方を区別して検出する。このように、第１特徴量と第２特徴量と個別に求めることによって透過スコアを算出し、且つ、可視光画像と赤外光画像の両方を用いて形状スコアを算出する。適切な入力に基づいて各スコアを算出できるため、可視物体及び透過物体を精度よく検出することが可能になる。

[0108] また、透過スコアと形状スコアを算出する手法に、機械学習を適用してもよい。この場合、情報処理装置１００の記憶部１３０は、学習済モデルを記憶する。学習済モデルは、複数の対象物を可視光によって撮像した第１学習用画像と、複数の対象物を赤外光によって撮像した第２学習用画像と、第１学習用画像及び第２学習用画像の少なくとも一方における第１対象物の位置情報及び第２対象物の位置情報と、を対応付けたデータセットに基づいて機械学習されている。処理部１２０は、第１検出用画像と、第２検出用画像と、学習済モデルとに基づいて、形状スコア及び透過スコアを算出した後、透

過スコアと形状スコアに基づいて、第１対象物の位置と第２対象物の位置の両方を区別して検出する。

[0109] 図１８は、本実施形態におけるニューラルネットワークの構成を示す模式図である。図１８のＥ１及びＥ２は、図１４のＤ１及びＤ２と同様である。Ｅ３は、第１特徴マップと第２特徴マップに基づいて、透過スコアを求めるブロックである。本実施形態においては、第１特徴量と第２特徴量を対象とした演算は、差分に基づく演算に限定されない。例えば、それぞれ２５６チャンネルの特徴マップである第１特徴マップと第２特徴マップを結合した５１２チャンネルの特徴マップに対して、畳み込み演算を行うことによって透過スコアが算出される。またここでの演算は畳み込み層を用いた演算に限定されず、例えば全結合層による演算等が用いられてもよいし、他の演算が用いられてもよい。このようにすれば、第１特徴量と第２特徴量に基づく透過スコアの演算についても、学習処理の対象とすることが可能になる。換言すれば、透過スコアを求める演算の内容が機械学習によって最適化されるため、第３特徴量とは異なり、透過スコアは差分に対応する特徴量に限定されない。

[0110] Ｅ４は、３チャンネルの可視光画像及び１チャンネルの赤外光画像を合わせた４チャンネルの画像を入力として受け取り、畳み込み演算を含む処理を行うことによって、形状スコアを求めるブロックである。Ｅ４の構成は、図１４のＤ４と同様である。

[0111] Ｅ５は、形状スコアと透過スコアに基づいて、可視物体と透明物体の位置検出を行うブロックである。図１８においては、図１４のＤ５と同様に、畳み込み層、プーリング層、アップサンプリング層、畳み込み層、ソフトマックス層による演算を行う例を示したが、具体的な構成については種々の変形実施が可能である。

[0112] 具体的な学習処理については第３の実施形態と同様である。即ち、学習部２２０は、可視光画像と、赤外光画像と、位置情報とを対応付けたデータセットに基づいて、フィルター特性等の重みを更新する処理を行う。なお、機

械学習を行う場合、Ｅ３の出力が透過度合いを表す情報であり、Ｅ４の出力が形状を表す情報であることをユーザーが明示的に指定するわけではない。ただし、Ｅ４においては可視光画像と赤外光画像の両方を合わせた処理が行われるため、ロバスト性が高い形状認識が可能な反面、透過度合いに関する情報が失われる。一方、Ｅ３においては、第１特徴量と第２特徴量を個別に処理することが可能であり、透過度合いに関する情報が残っている。即ち、透明物体の位置検出精度を向上させようとする機械学習を行った場合、Ｅ１～Ｅ３における重みは、適切な透過スコアを出力するための値となり、Ｅ４における重みは、適切な形状スコアを出力するための値となることが期待される。換言すれば、３通りの入力を行い、各入力に対して独立に処理を行った後に、処理結果を合成する図１８の構成を用いることによって、形状スコア及び透過スコアに基づいて対象物の位置検出を行う学習済モデルを構築することが可能である。

[0113] 図１９は、透過スコア算出処理を説明する模式図である。図１９のＦ１が可視光画像であり、Ｆ１１が透明物体が存在する領域を表し、Ｆ１２が透明物体の奥に存在する可視物体を表す。Ｆ２は赤外光画像であり、Ｆ２１に示す透明物体が撮像され、Ｆ１２に対応する可視物体は撮像されない。

[0114] Ｆ３は、可視光画像のうちのＦ１３に対応する領域の画素値を表す。可視光画像において、Ｆ１３は、可視物体であるＦ１２と背景の境界である。ここでは背景が明るいため、左及び中央の列において画素値が小さくなり、右の列で画素値が大きくなる。なお図１９及び後述する図２０における画素値は、－１から＋１の範囲となるように正規化された値を示している。Ｆ３の領域に対して、Ｆ５に示す特性を有するフィルターを適用した演算を行うことによって、ある程度大きいスコア値Ｆ７が出力される。Ｆ５は、学習の結果として特性が設定されたフィルターのうちの１つであり、例えば縦エッジを抽出するフィルターである。

[0115] Ｆ４は、赤外光画像のうちのＦ２３に対応する領域の画素値を表す。赤外光画像において、Ｆ２３は透明物体に対応するため、ローコントラストとな

る。具体的には、Ｆ４の全域において画素値が同程度となる。そのため、Ｆ６に示す特性を有するフィルターを適用した演算を行うことによって、ある程度絶対値が大きく、負の値となるスコア値Ｆ８が出力される。Ｆ６は、学習の結果として特性が設定されたフィルターのうちの１つであり、例えば平坦領域を抽出するフィルターである。

[0116] 図１９の例においては、処理部１２０は、Ｆ７からＦ８を減算することによって、透過スコアを求めることが可能である。ただし、本実施形態の手法では第１特徴量と第２特徴量とをどのように用いて透過スコアを求めるかについても、機械学習の対象となる。そのため、設定されたフィルター特性に合わせて柔軟な処理によって透過スコアを算出することが可能である。

[0117] 図２０は、形状スコア算出処理を説明する模式図である。図２０のＧ１が可視光画像であり、Ｇ１１が可視物体を表す。Ｇ２は赤外光画像であり、Ｇ１１と同様の可視物体Ｇ２１が撮像される。

[0118] Ｇ３は、可視光画像のうちのＧ１２に対応する領域の画素値を表す。可視光画像において、Ｇ１２は、可視物体であるＧ１１と背景の境界である。ここでは背景が明るいため、左及び中央の列において画素値が小さくなり、右の列で画素値が大きくなる。そのため、Ｇ５に示す特性を有するフィルターを適用した演算を行うことによって、ある程度大きいスコア値Ｇ７が出力される。Ｇ５は、学習の結果として特性が設定されたフィルターのうちの１つであり、例えば縦エッジを抽出するフィルターである。

[0119] Ｇ４は、赤外光画像のうちのＧ２２に対応する領域の画素値を表す。赤外光画像において、Ｇ２２は、可視物体であるＧ２１と、背景との境界である。赤外光画像においては、人等の可視物体は熱源となるため、背景領域に比べて明るく撮像される。そのため、左及び中央の列において画素値が大きくなり、右の列で画素値が小さくなる。そのため、Ｇ６に示す特性を有するフィルターを適用した演算を行うことによって、ある程度大きいスコア値Ｇ８が出力される。Ｇ６は、学習の結果として特性が設定されたフィルターのうちの１つであり、例えば縦エッジを抽出するフィルターである。なおＧ５と

Ｇ６は、勾配方向が異なる。

[0120] 形状スコアは４チャンネルの画像に対する畳み込み演算によって求められる。例えば、形状スコアは、Ｇ７とＧ８を加算した結果を含む特徴マップである。図２０の例であれば、物体のエッジに対応する領域において値が大きくなる情報が形状スコアとして算出される。

[0121] なお、上記のように本実施形態について詳細に説明したが、本実施形態の新規事項および効果から実体的に逸脱しない多くの変形が可能であることは当業者には容易に理解できるであろう。従って、このような変形例はすべて本開示の範囲に含まれるものとする。例えば、明細書又は図面において、少なくとも一度、より広義または同義な異なる用語と共に記載された用語は、明細書又は図面のいかなる箇所においても、その異なる用語に置き換えることができる。また本実施形態及び変形例の全ての組み合わせも、本開示の範囲に含まれる。また情報処理装置、学習装置、移動体等の構成及び動作等も、本実施形態で説明したものに限定されず、種々の変形実施が可能である。

符号の説明

[0122] ＡＸ…光軸、１０…撮像部、１１…波長分離ミラー、１２…第１光学系、１３，１３－２…第１撮像素子、１４…第２光学系、１５，１５－２…第２撮像素子、１６…第３光学系、１７…撮像素子、２０…車椅子、２１…車輪、３０…制御装置、４０…ロボット、４１…本体部、４３…アーム、４５…ハンド、４７…車輪、６０…自動車、６１…車輪、６３…表示部、１００…情報処理装置、１１０…取得部、１１１…第１Ａ／Ｄ変換回路、１１２…第２Ａ／Ｄ変換回路、１２０…処理部、１２１…第１特徴量抽出部、１２２…第２特徴量抽出部、１２３…第３特徴量抽出部、１２４…位置検出部、１２５…第４特徴量抽出部、１２６…透過スコア算出部、１２７…形状スコア算出部、１３０…記憶部、２００…学習装置、２１０…取得部、２２０…学習部

請求の範囲

[請求項1] 第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した第１検出用画像と、前記複数の対象物を赤外光によって撮像した第２検出用画像を取得する取得部と、

処理部とを含み、

前記処理部は、

前記第１検出用画像に基づいて第１特徴量を求め、

前記第２検出用画像に基づいて第２特徴量を求め、

前記第１特徴量と前記第２特徴量の差分に対応する特徴量を第３特徴量として算出し、

前記第３特徴量に基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方における前記第２対象物の位置を検出する、

ことを特徴とする情報処理装置。

[請求項2] 請求項１に記載の情報処理装置において、

前記第１特徴量は、前記第１検出用画像のコントラストを表す情報であり、

前記第２特徴量は、前記第２検出用画像のコントラストを表す情報であり、

前記処理部は、

前記第１検出用画像のコントラストと前記第２検出用画像のコントラストの差分に対応する前記第３特徴量に基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方における前記第２対象物の位置を検出することを特徴とする情報処理装置。

[請求項3] 請求項１又は２に記載の情報処理装置において、

前記処理部は、

前記第１検出用画像及び前記第２検出用画像に基づいて、前記第１対象物の特徴を表す第４特徴量を求め、

前記第３特徴量及び前記第４特徴量に基づいて、前記第１対象物の位置と前記第２対象物の位置の両方を区別して検出することを特徴とする情報処理装置。

[請求項4] 請求項３に記載の情報処理装置において、

学習済モデルを記憶する記憶部を含み、

前記学習済モデルは、

前記複数の対象物を前記可視光によって撮像した第１学習用画像と、前記複数の対象物を前記赤外光によって撮像した第２学習用画像と、前記第１学習用画像及び前記第２学習用画像の少なくとも一方における前記第１対象物の位置情報及び前記第２対象物の位置情報と、を対応付けたデータセットに基づいて機械学習されており、

前記処理部は、

前記第１検出用画像と、前記第２検出用画像と、前記学習済モデルとに基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方において、前記第１対象物の位置と前記第２対象物の位置の両方を区別して検出することを特徴とする情報処理装置。

[請求項5] 請求項４に記載の情報処理装置において、

前記第１特徴量は、前記第１検出用画像に対して、第１フィルターを用いた畳み込み演算を行うことによって求められる第１特徴マップであり、

前記第２特徴量は、前記第２検出用画像に対して、第２フィルターを用いた畳み込み演算を行うことによって求められる第２特徴マップであることを特徴とする情報処理装置。

[請求項6] 請求項５に記載の情報処理装置において、

前記第１フィルター及び前記第２フィルターは、前記機械学習によってフィルター特性が設定されていることを特徴とする情報処理装置。

[請求項7] 請求項４乃至６のいずれか一項に記載の情報処理装置において、

前記第４特徴量は、前記第１検出用画像と前記第２検出用画像に対して、第４フィルターを用いた畳み込み演算を行うことによって求められる第４特徴マップであることを特徴とする情報処理装置。

[請求項8] 請求項１に記載の情報処理装置において、

学習済モデルを記憶する記憶部を含み、

前記学習済モデルは、

前記複数の対象物を前記可視光によって撮像した第１学習用画像と、前記複数の対象物を前記赤外光によって撮像した第２学習用画像と、前記第１学習用画像及び前記第２学習用画像の少なくとも一方における前記第２対象物の位置情報と、を対応付けたデータセットに基づいて機械学習されており、

前記処理部は、

前記第１検出用画像と、前記第２検出用画像と、前記学習済モデルとに基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方において、前記第２対象物の位置を検出することを特徴とする情報処理装置。

[請求項9] 請求項８に記載の情報処理装置において、

前記第１特徴量は、前記第１検出用画像に対して、第１フィルターを用いた畳み込み演算を行うことによって求められる第１特徴マップであり、

前記第２特徴量は、前記第２検出用画像に対して、第２フィルターを用いた畳み込み演算を行うことによって求められる第２特徴マップであり、

前記第１フィルター及び前記第２フィルターは、前記機械学習によってフィルター特性が設定されていることを特徴とする情報処理装置。

[請求項10] 第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した第１検出用画

像と、前記複数の対象物を赤外光によって撮像した第２検出用画像を取得する取得部と、

処理部とを含み、

前記処理部は、

前記第１検出用画像に基づいて第１特徴量を求め、

前記第２検出用画像に基づいて第２特徴量を求め、

前記第１特徴量及び前記第２特徴量に基づいて、前記第１検出用画像及び前記第２検出用画像に撮像された前記複数の対象物について前記可視光の透過度合いを表す透過スコアを算出し、

前記第１検出用画像及び前記第２検出用画像に基づいて、前記第１検出用画像及び前記第２検出用画像に撮像された前記複数の対象物の形状を示す形状スコアを算出し、

前記透過スコアと前記形状スコアに基づいて、前記第１検出用画像及び前記第２検出用画像の少なくとも一方における、前記第１対象物の位置と前記第２対象物の位置の両方を区別して検出する、

ことを特徴とする情報処理装置。

[請求項11] 請求項１０に記載の情報処理装置において、

学習済モデルを記憶する記憶部を含み、

前記学習済モデルは、

前記複数の対象物を前記可視光によって撮像した第１学習用画像と、前記複数の対象物を前記赤外光によって撮像した第２学習用画像と、前記第１学習用画像及び前記第２学習用画像の少なくとも一方における前記第１対象物の位置情報及び前記第２対象物の位置情報と、を対応付けたデータセットに基づいて機械学習されており、

前記処理部は、

前記第１検出用画像と、前記第２検出用画像と、前記学習済モデルとに基づいて、前記形状スコア及び前記透過スコアを算出し、前記透過スコアと前記形状スコアに基づいて、前記第１対象物の位置と前記

第２対象物の位置の両方を区別して検出する、

ことを特徴とする情報処理装置。

[請求項12] 請求項１乃至１１のいずれか一項に記載の情報処理装置において、

第１光軸を用いて前記複数の対象物を前記可視光によって撮像し、且つ、前記第１光軸に対応する軸である第２光軸を用いて前記複数の対象物を前記赤外光によって撮像する撮像部をさらに含み、

前記取得部は、

前記撮像部による撮像に基づいて、前記第１検出用画像及び前記第２検出用画像を取得することを特徴とする情報処理装置。

[請求項13] 請求項１乃至１２のいずれか一項に記載の情報処理装置を含むことを特徴とする移動体。

[請求項14] 第１対象物と、前記第１対象物に比べて可視光を透過する第２対象物とを含む複数の対象物を前記可視光によって撮像した可視光画像と、前記複数の対象物を赤外光によって撮像した赤外光画像と、前記可視光画像及び前記赤外光画像の少なくとも一方における前記第２対象物の位置情報と、を対応付けたデータセットを取得する取得部と、

前記データセットに基づいて、前記可視光画像及び前記赤外光画像の少なくとも一方において、前記第２対象物の位置を検出する条件を機械学習する学習部と、

を含むことを特徴とする学習装置。

[請求項15] 請求項１４において、

前記データセットは、

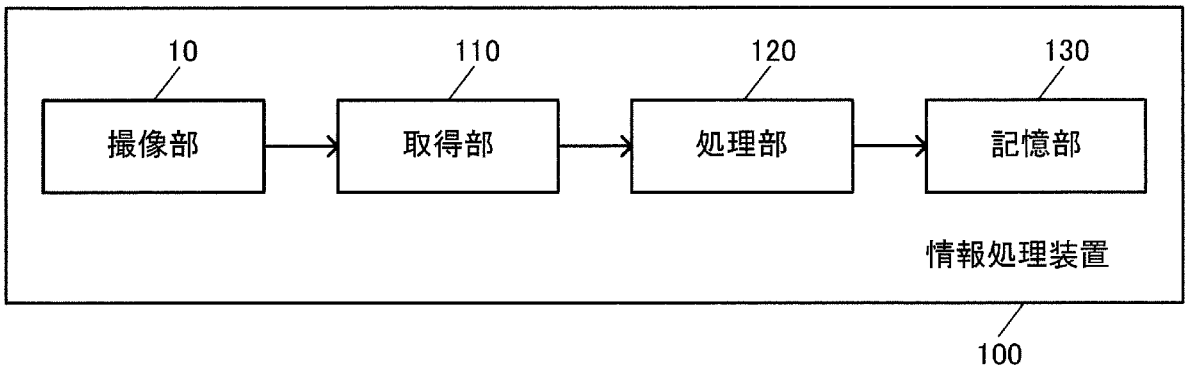
前記可視光画像と、前記赤外光画像と、前記第２対象物の前記位置情報と、前記可視光画像及び前記赤外光画像の少なくとも一方における前記第１対象物の位置情報と、を対応付けたデータセットであり、

前記学習部は、

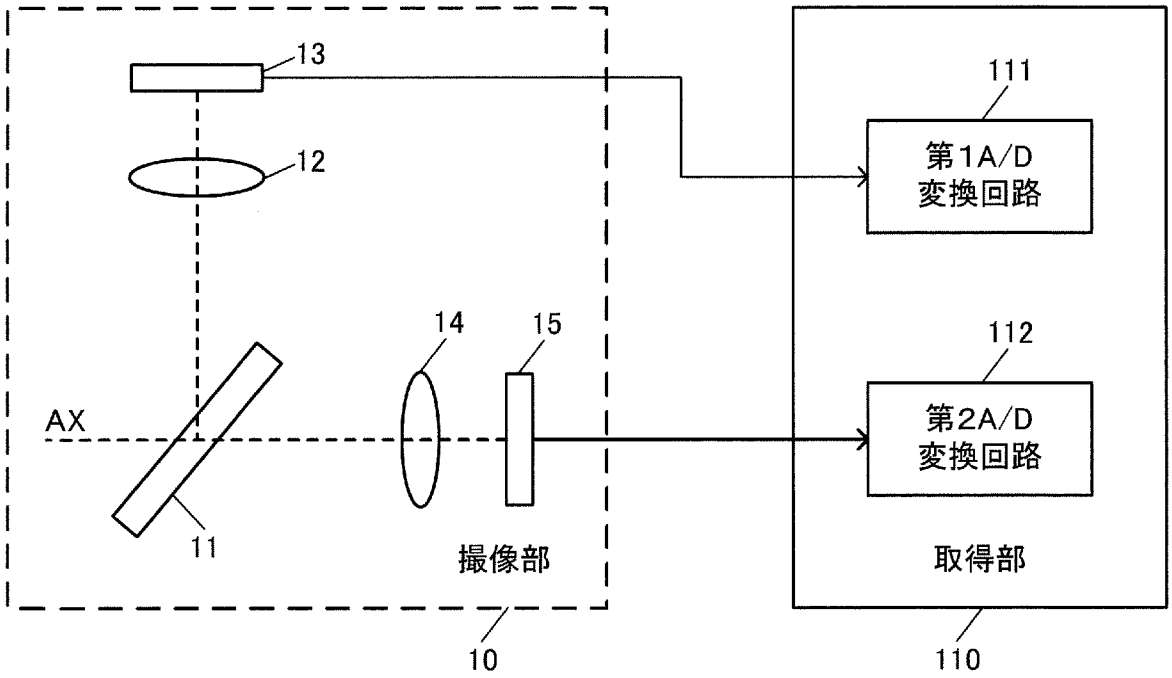
前記データセットに基づいて、前記可視光画像及び前記赤外光画像の少なくとも一方において、前記第１対象物の位置と前記第２対象物

の位置の両方を区別して検出する条件を機械学習することを特徴とする学習装置。

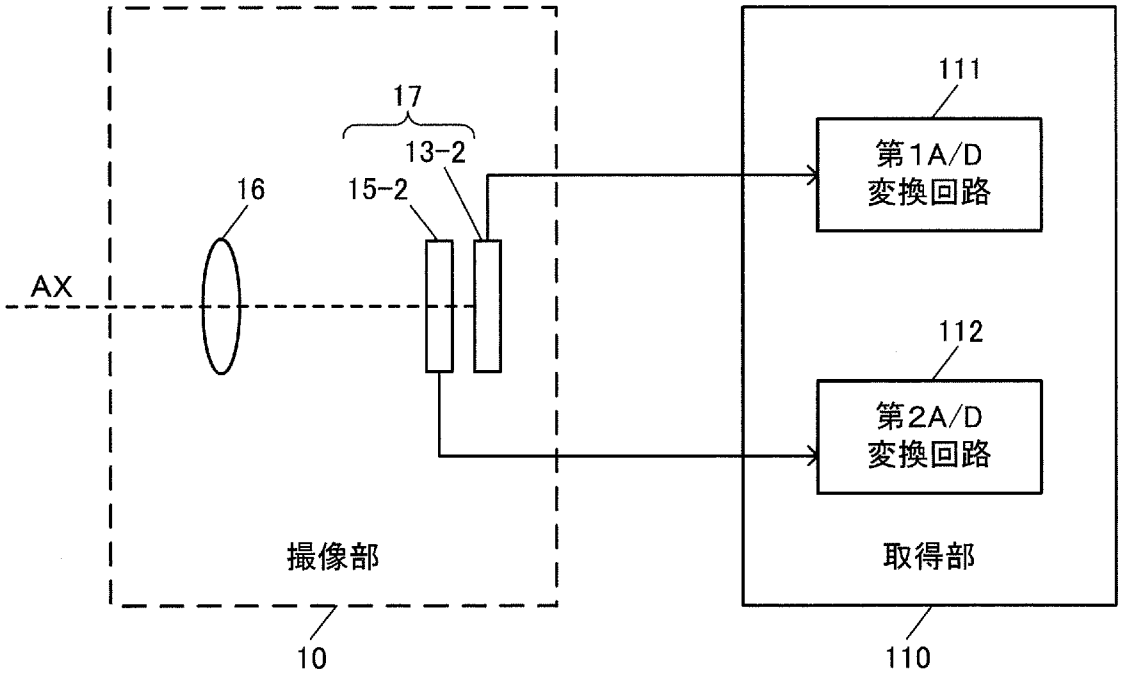
[図1]



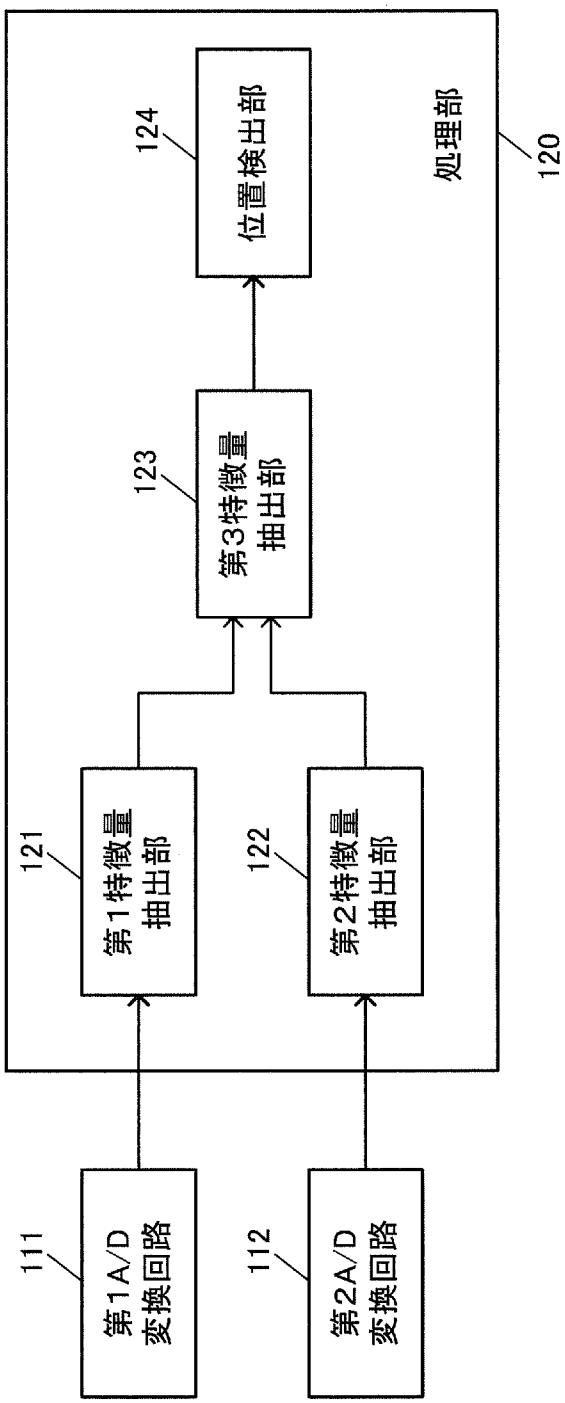
[図2]


13
12
14
15
AX
11
撮像部
10
111
第1A/D
変換回路
112
第2A/D
変換回路
取得部
110

[図3]

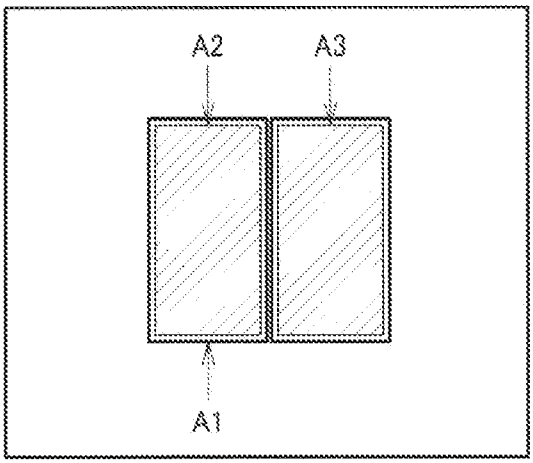

17
13-2
16
15-2
AX
撮像部
10
111
第1A/D
変換回路
112
第2A/D
変換回路
取得部
110

[図4]

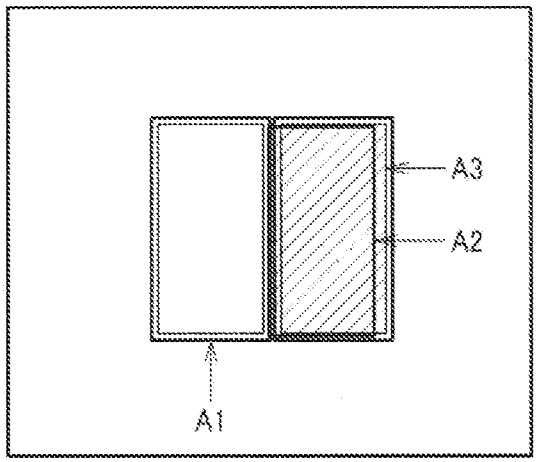

処理部
120
位置検出部
124
第3特徴量
抽出部
123
第1特徴量
抽出部
121
第2特徴量
抽出部
122
第1A/D
変換回路
111
第2A/D
変換回路
112

[図5]

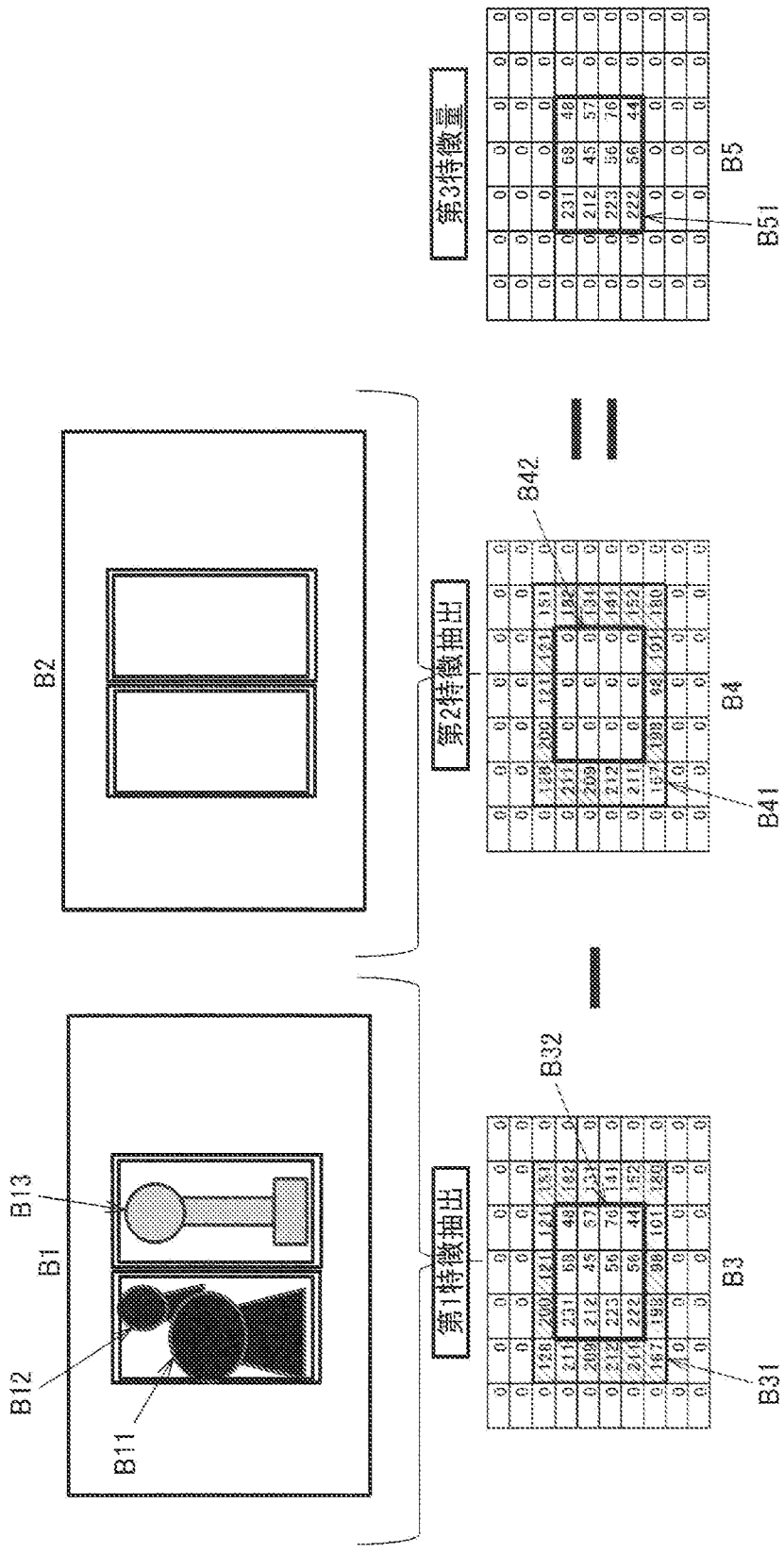
(A)



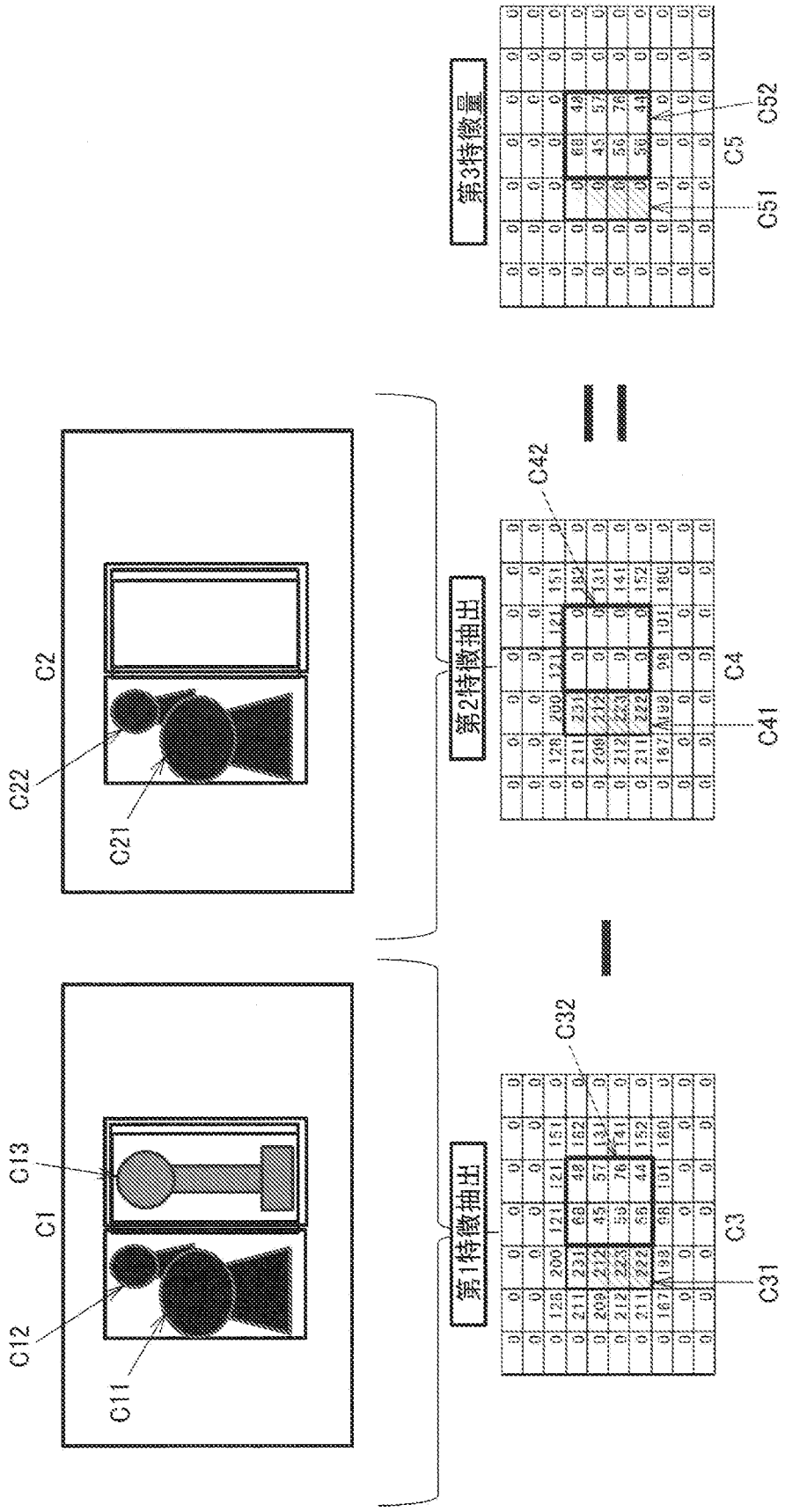
(B)



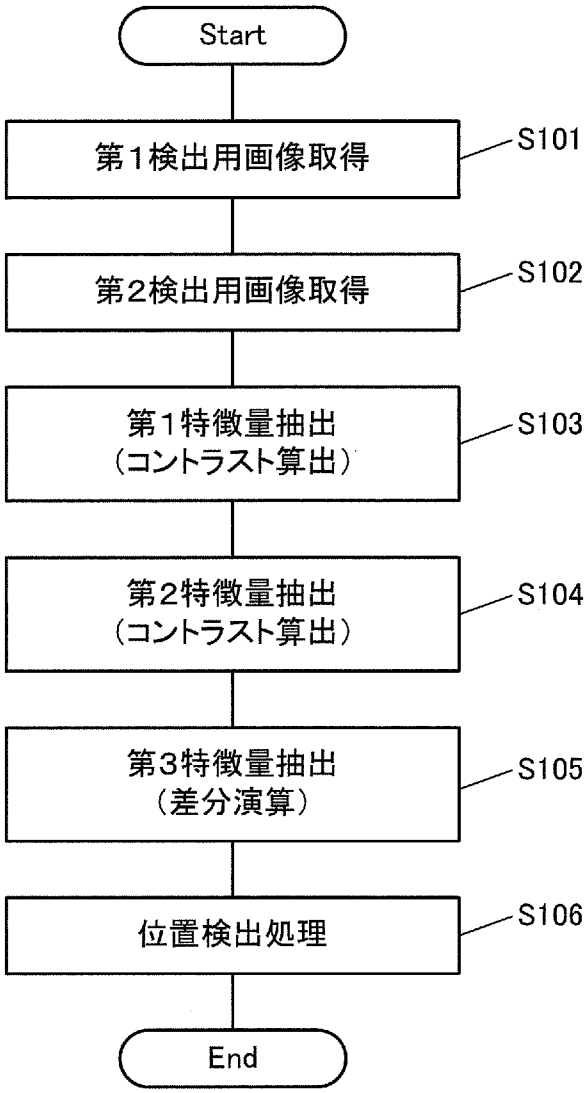
[図6]


B1
B11
B12
B13
B2
第1特徴抽出
第2特徴抽出
第3特徴量
B3
B31
B32
B4
B41
B42
B5
B51

[図7]

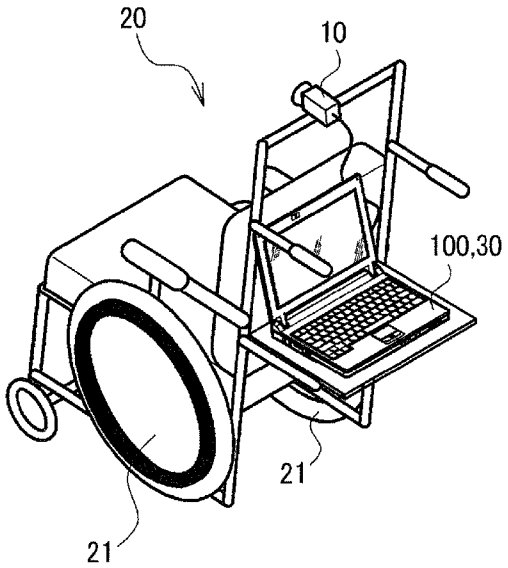

C1
C11
C12
C13
C2
C21
C22
第1特徴抽出
第2特徴抽出
第3特徴量
C3
C31
C32
C4
C41
C42
C5
C51
C52

[図8]

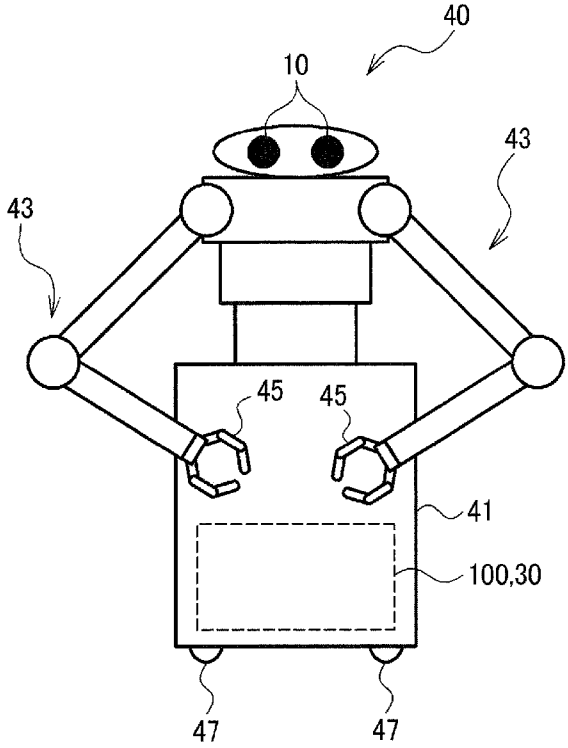

Start
第1検出用画像取得
S101
第2検出用画像取得
S102
第1特徴量抽出
（コントラスト算出）
S103
第2特徴量抽出
（コントラスト算出）
S104
第3特徴量抽出
（差分演算）
S105
位置検出処理
S106
End

[図9]

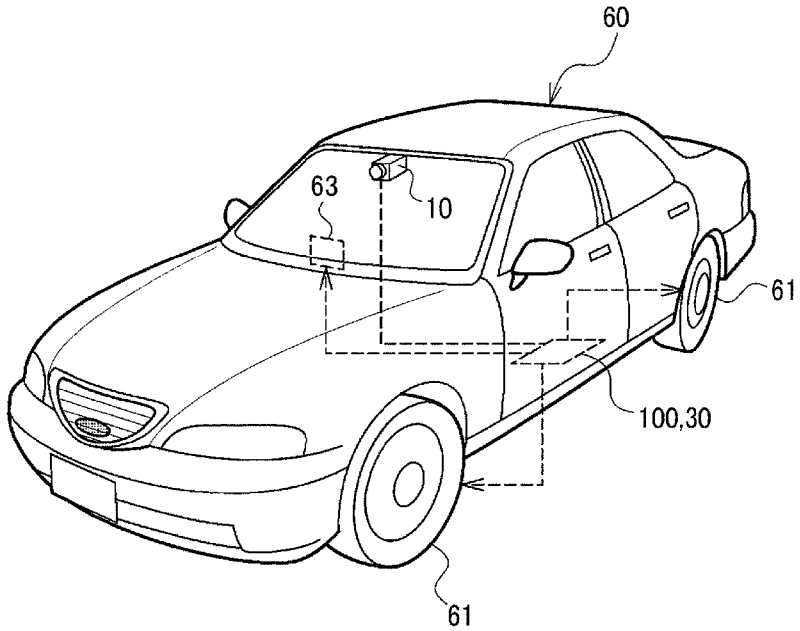
(A)


20
10
100,30
21
21

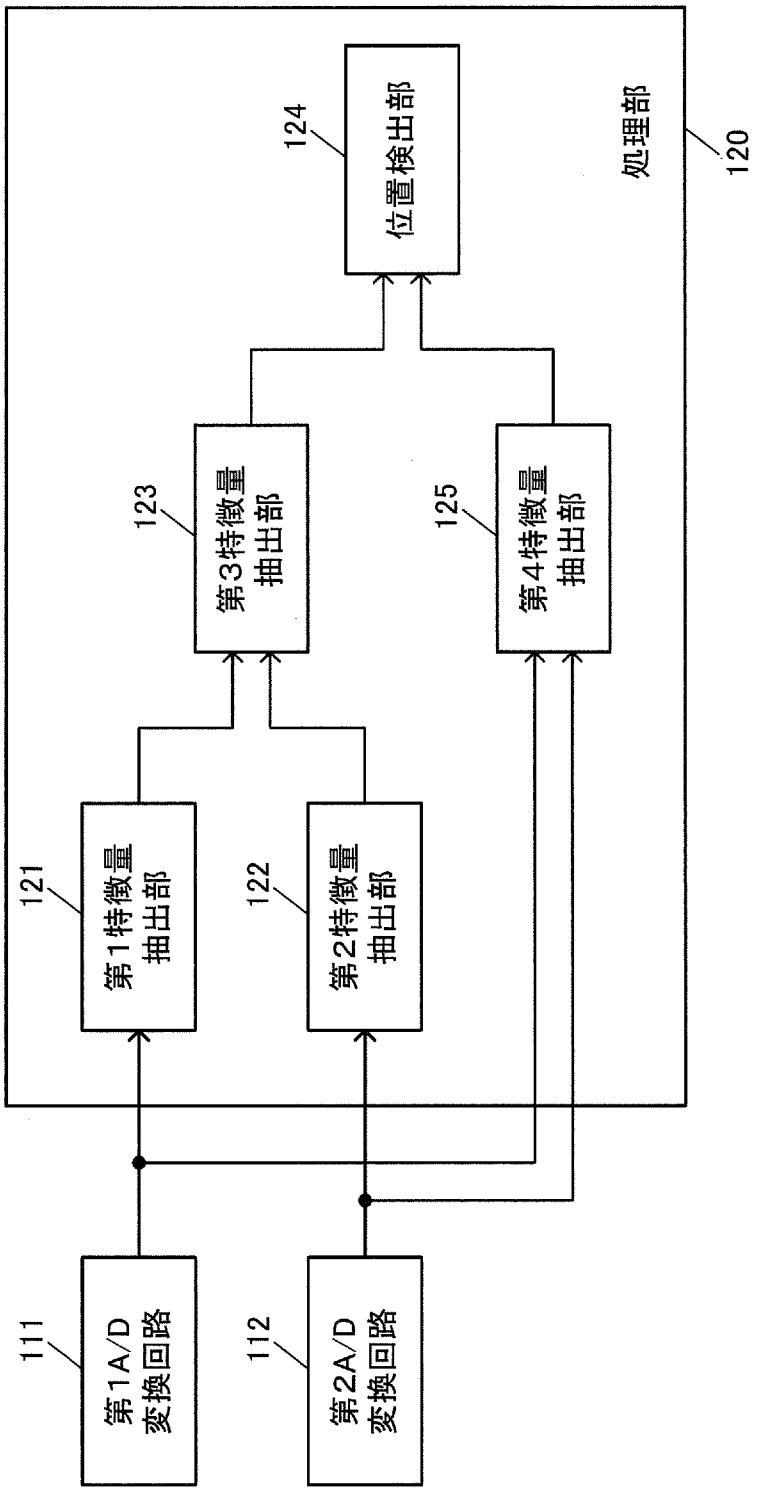
(B)


40
10
43
43
45
45
41
100,30
47
47

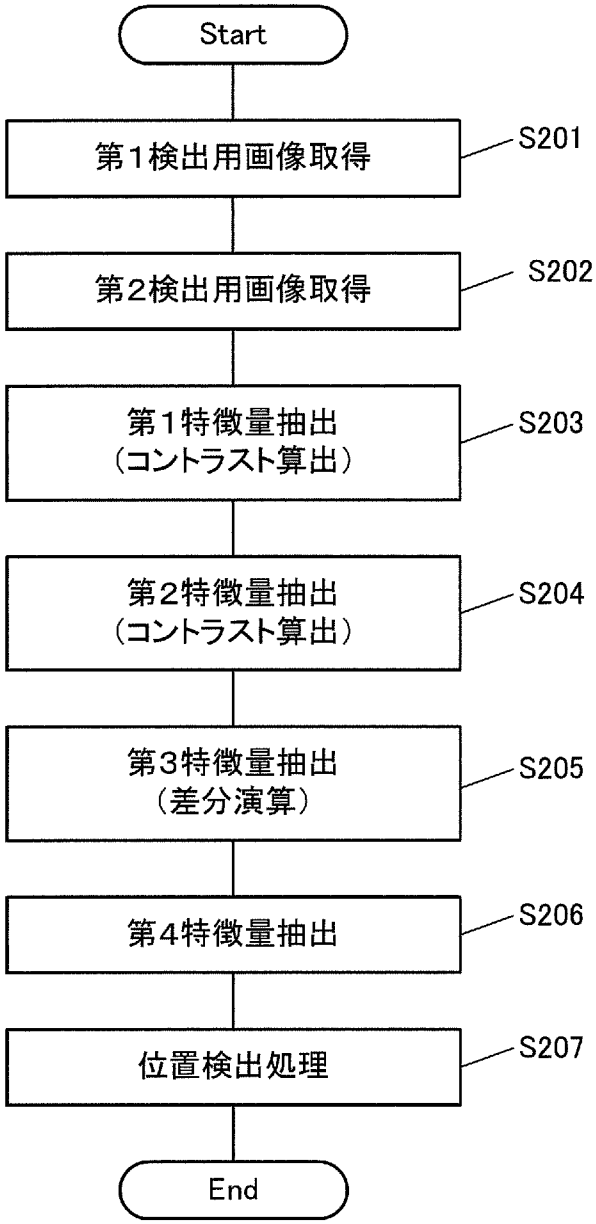
(C)


60
63
10
61
100,30
61

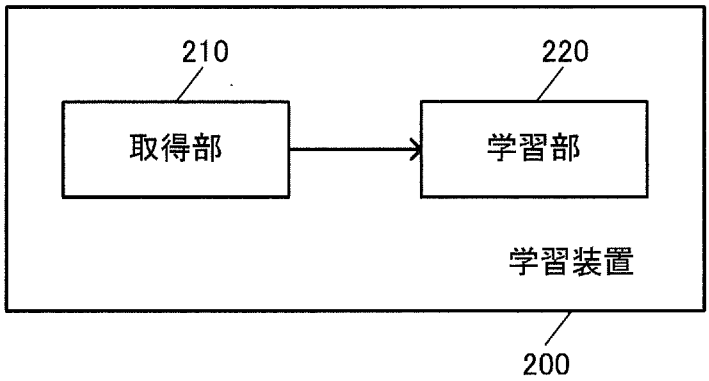
[図10]


処理部
120
位置検出部
124
第3特徴量
抽出部
123
第4特徴量
抽出部
125
第1特徴量
抽出部
121
第2特徴量
抽出部
122
第1A/D
変換回路
111
第2A/D
変換回路
112

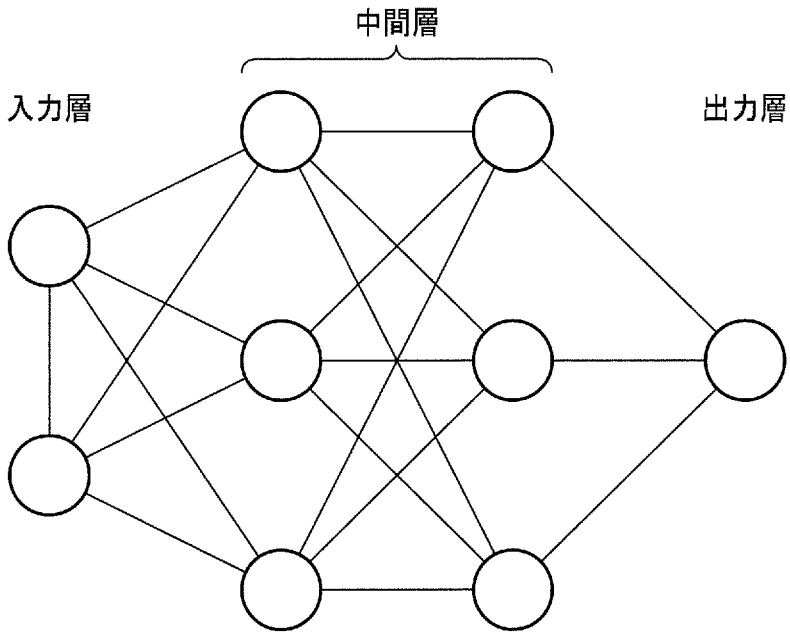
[図11]


Start
第1検出用画像取得
S201
第2検出用画像取得
S202
第1特徴量抽出
（コントラスト算出）
S203
第2特徴量抽出
（コントラスト算出）
S204
第3特徴量抽出
（差分演算）
S205
第4特徴量抽出
S206
位置検出処理
S207
End

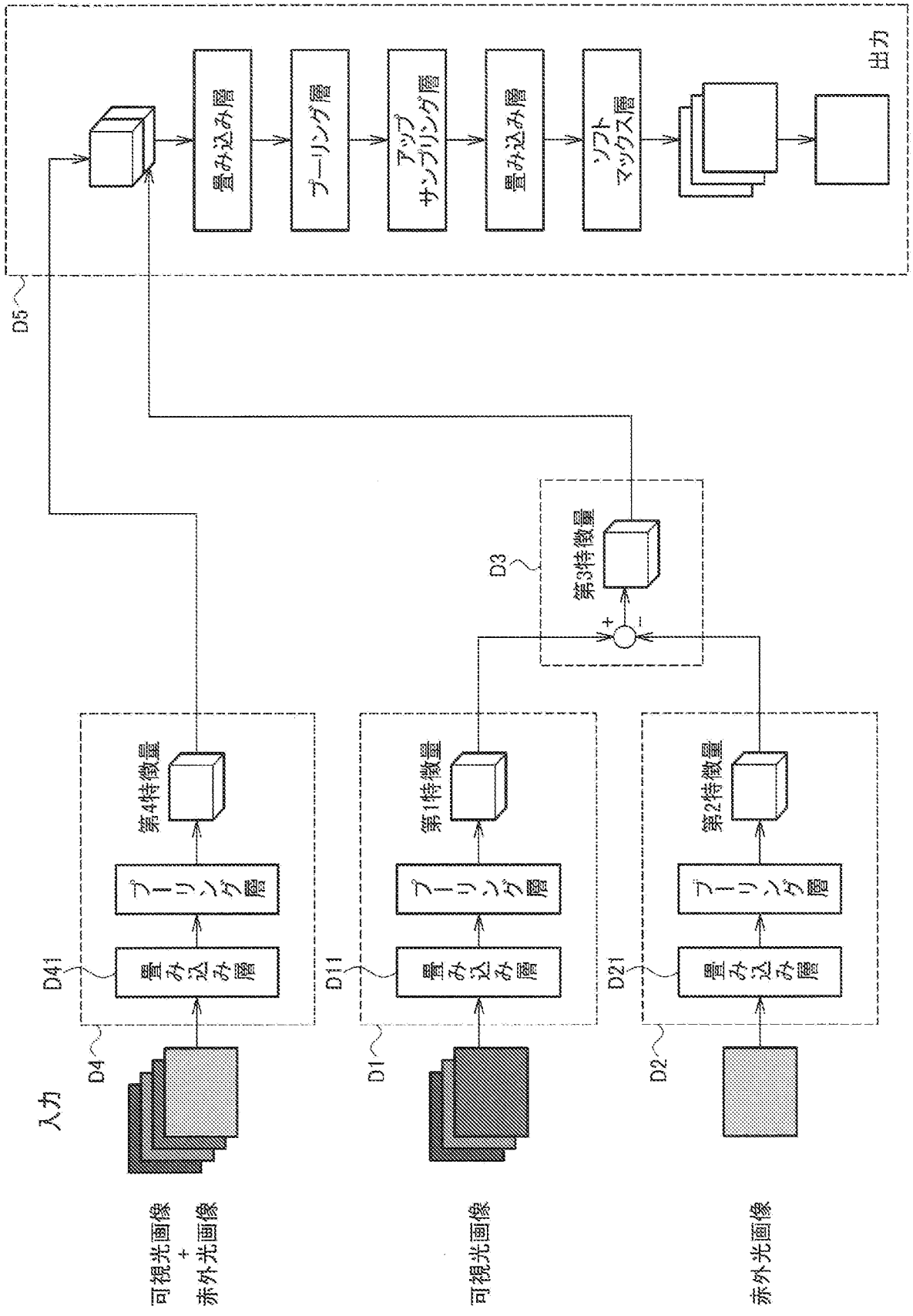
[図12]



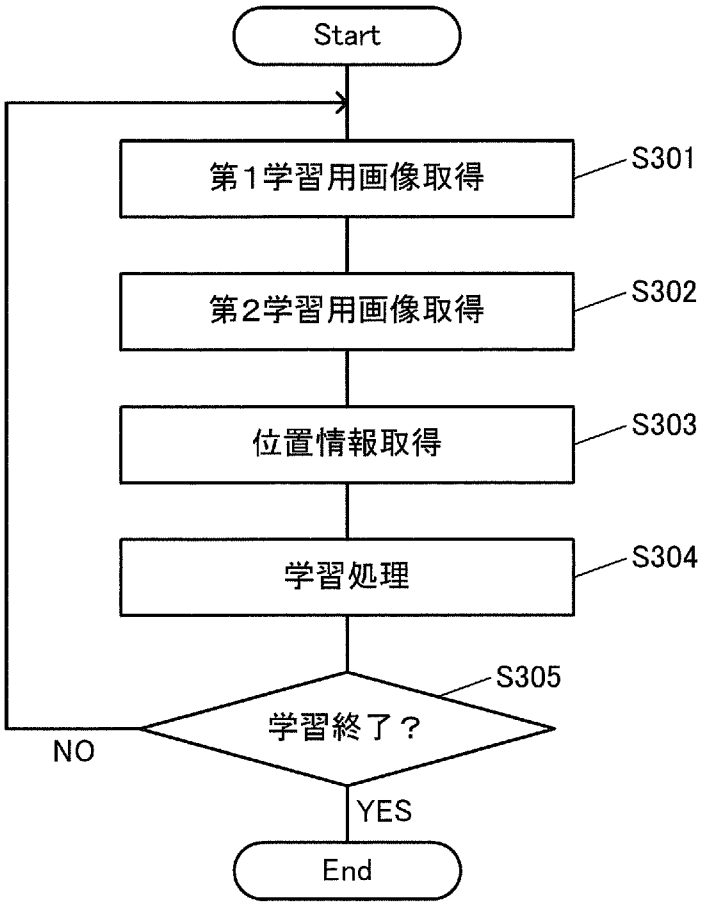
[図13]



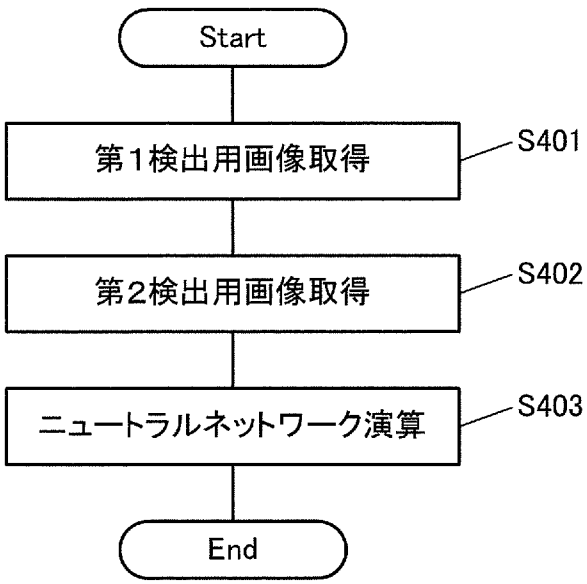
[図14]


入力
可視光画像
+
赤外光画像
可視光画像
赤外光画像
D4
D1
D2
D41
D11
D21
畳み込み層
プーリング層
第4特徴量
第1特徴量
第2特徴量
D3
第3特徴量
+
−
D5
畳み込み層
プーリング層
アップサンプリング層
畳み込み層
ソフトマックス層
出力

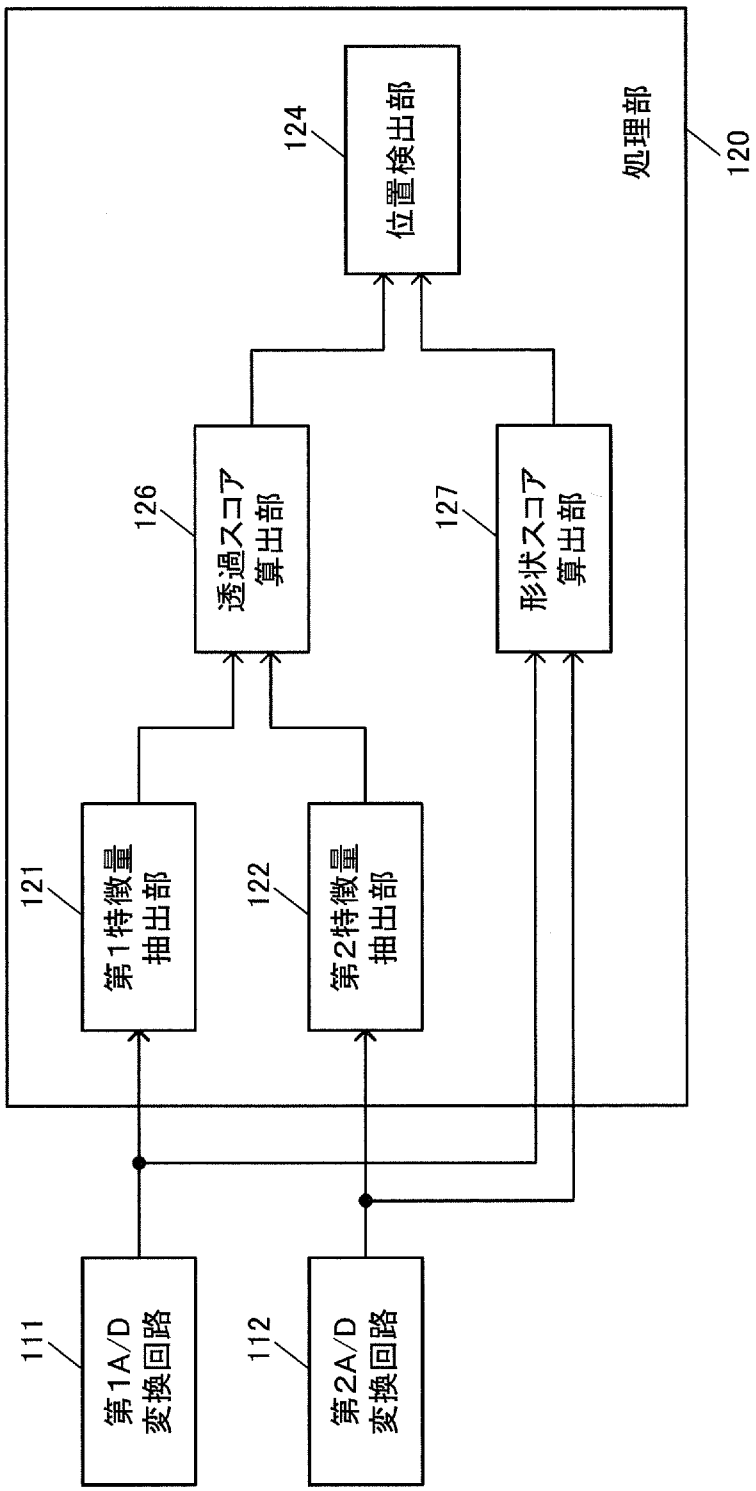
[図15]


Start
第1学習用画像取得
S301
第2学習用画像取得
S302
位置情報取得
S303
学習処理
S304
学習終了？
S305
NO
YES
End

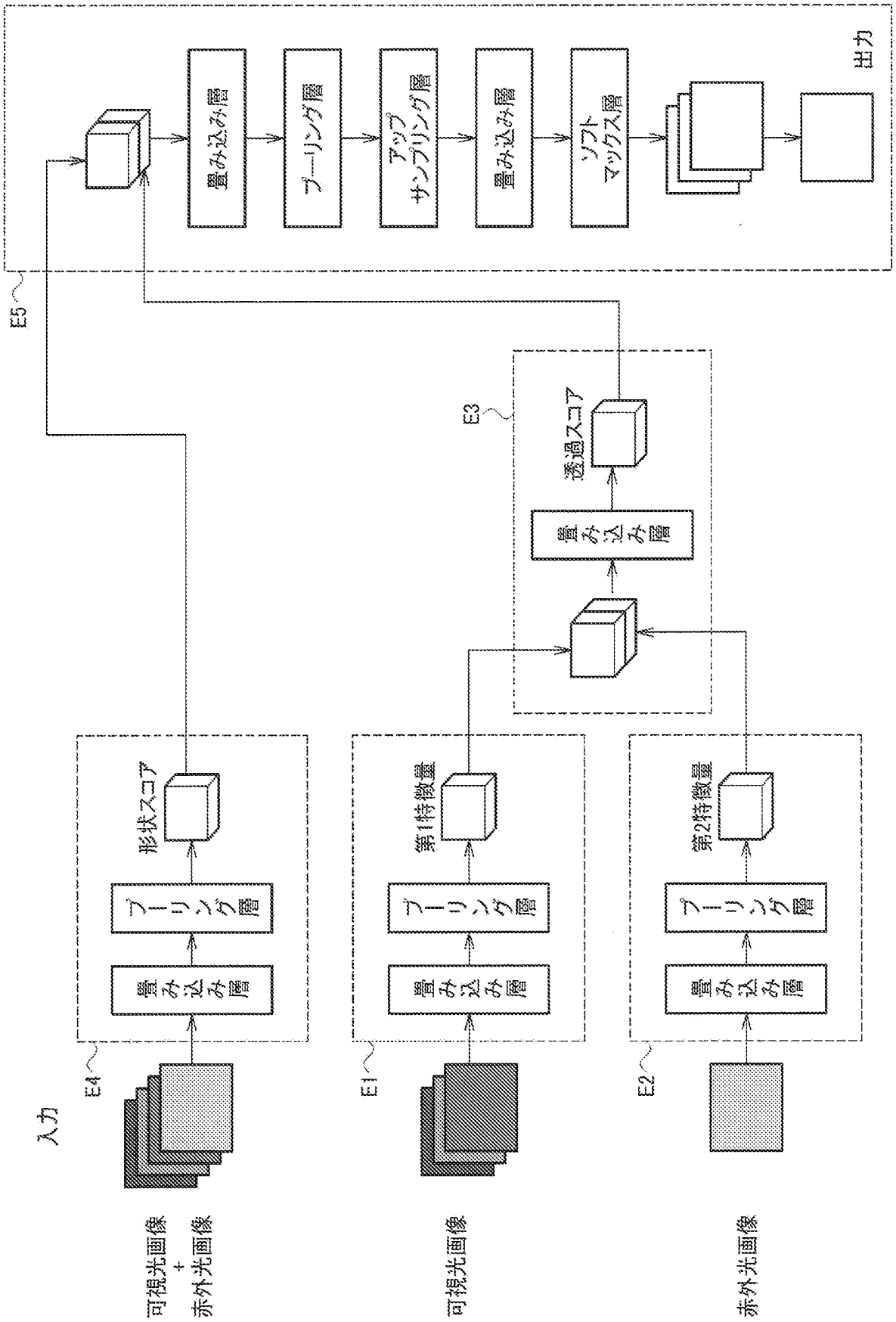
[図16]


Start
第1検出用画像取得
S401
第2検出用画像取得
S402
ニュートラルネットワーク演算
S403
End

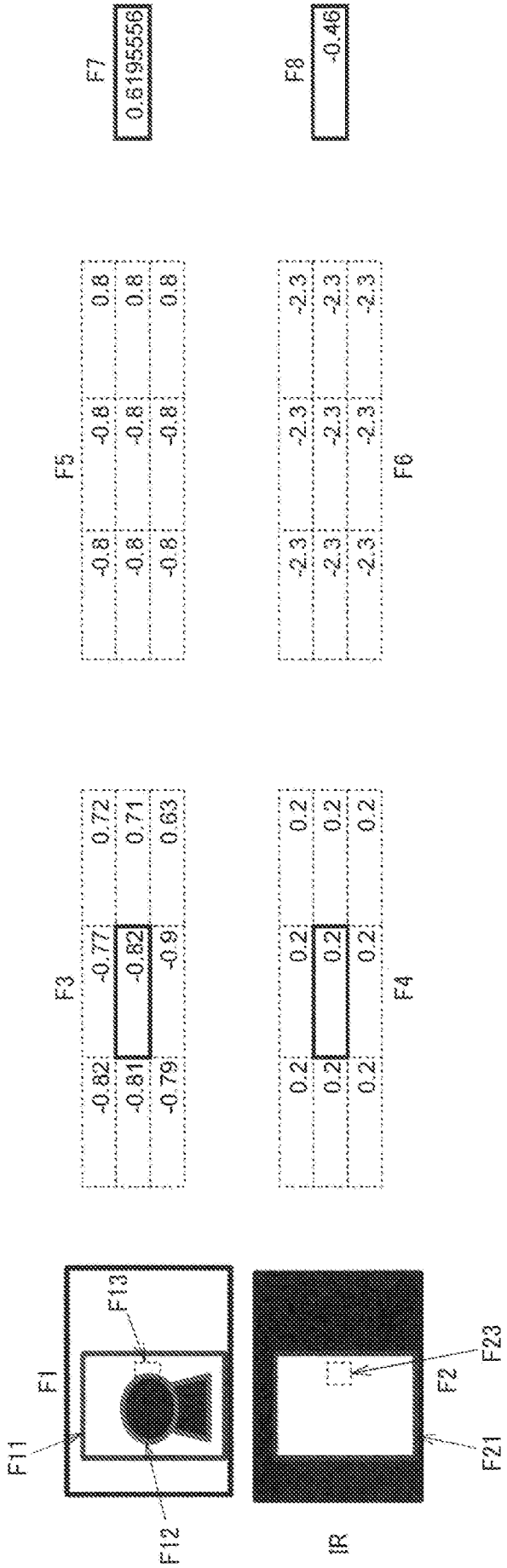
[図17]


処理部
120
位置検出部
124
透過スコア算出部
126
形状スコア算出部
127
第1特徴量抽出部
121
第2特徴量抽出部
122
第1A/D変換回路
111
第2A/D変換回路
112

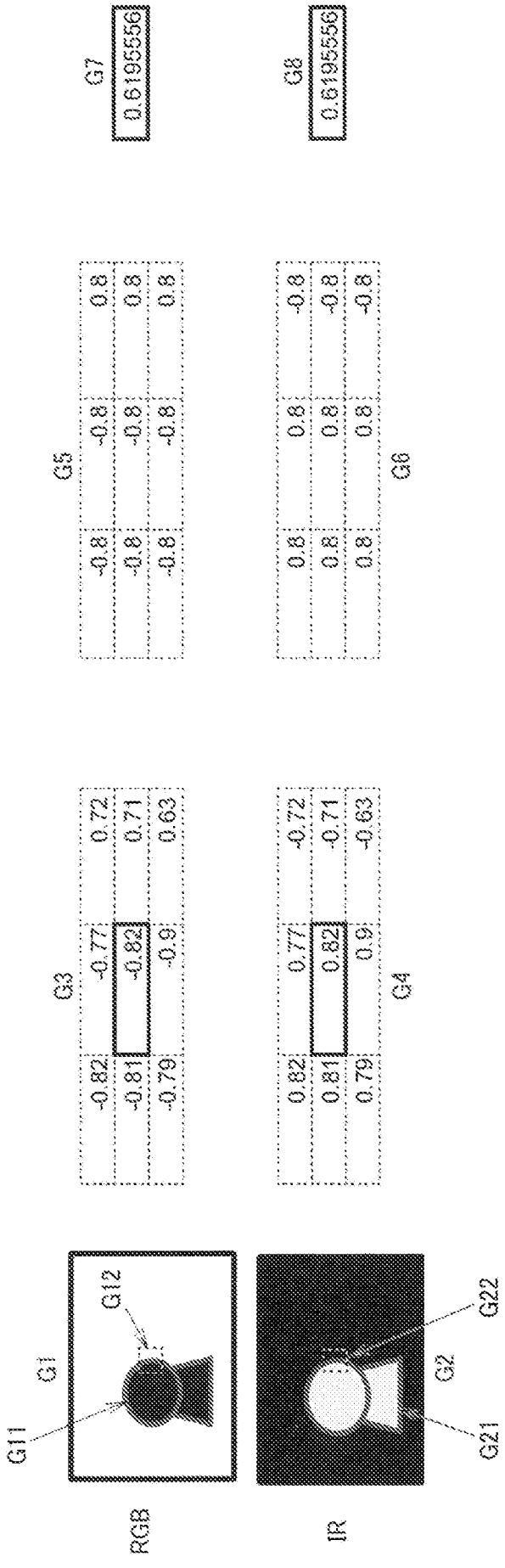
[図18]


出力
畳み込み層
プーリング層
アップサンプリング層
畳み込み層
ソフトマックス層
E5
E3
透過スコア
畳み込み層
形状スコア
プーリング層
畳み込み層
第1特徴量
プーリング層
畳み込み層
第2特徴量
プーリング層
畳み込み層
E4
E1
E2
入力
可視光画像
+
赤外光画像
可視光画像
赤外光画像

[図19]


F7
0.8195556
F8
-0.46
F5
-0.8 -0.8 0.8
-0.8 -0.8 0.8
-0.8 -0.8 0.8
F6
-2.3 -2.3 -2.3
-2.3 -2.3 -2.3
-2.3 -2.3 -2.3
F3
-0.82 -0.77 0.72
-0.81 -0.82 0.71
-0.79 -0.9 0.63
F4
0.2 0.2 0.2
0.2 0.2 0.2
0.2 0.2 0.2
F1
F13
F11
F12
IR
F2
F23
F21

[図20]


G7
0.6195556
G8
0.6195556
G5
-0.8 -0.8 0.8
-0.8 -0.8 0.8
-0.8 -0.8 0.8
G6
0.8 0.8 -0.8
0.8 0.8 -0.8
0.8 0.8 -0.8
G3
-0.82 -0.77 0.72
-0.81 -0.82 0.71
-0.79 -0.9 0.63
G4
0.82 0.77 -0.72
0.81 0.82 -0.71
0.79 0.9 -0.63
G11
G1
G12
G22
G2
G21
RGB
IR

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2019/007653

A. CLASSIFICATION OF SUBJECT MATTER
Int. Cl. G06T7/70(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. G06T7/70

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.
X Y A	JP 2017-220923 A (PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) 14 December 2017, paragraphs [0073]-[0159], fig. 5A-9 & US 2017/0352275 A1, paragraphs [0108]-[0185], fig. 5A-9	1, 3, 10 4-9, 12-15 11
X	JP 2017-191470 A (CANON KABUSHIKI KAISHA) 19 October 2017, paragraphs [0087]-[0089], fig. 9, claims 1, 2 (Family: none)	1, 2
Y	WO 2018/235777 A1 (NATIONAL UNIVERSITY CORPORATION SHIZUOKA UNIVERSITY) 27 December 2018, paragraph [0025] (Family: none)	4-9, 14, 15

☒ 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	Date of mailing of the international search report
17.04.2019	07.05.2019
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/JP2019/007653

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	高畑智之,下山勲, 防滴性のある可視光・遠赤外光同軸撮影システム, 第36回日本ロボット学会学術講演会, 04 September 2018, non-official translation (TAKAHATA, Tomoyuki, SHIMOYAMA, Isao. Drip-Proof Visible Light/Far Infrared Coaxial Imaging System. The 36th Annual Conference of the Robotics Society of Japan.)	12, 13
Y	JP 2001-178001 A (TOYODA AUTOMATIC LOOM WORKS, LTD.) 29 June 2001, paragraphs [0028], [0030], [0031], [0038], fig. 1, 2, 4 & US 2001/0004201 A1, paragraphs [0040], [0042], [0043], [0051], fig. 1, 2, 4	1-8

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１９／００７６５３

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06T7/70(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.　G06T7/70

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
X Y A	JP 2017-220923 A（パナソニックＩＰマネジメント株式会社） 2017.12.14, 段落［００７３］－［０１５９］[図５Ａ]－[図９] & US 2017/0352275 A1, [0108]-[0185], FIG.5A-FIG.9	1,3,10 4-9,12-15 11
X	JP 2017-191470 A（キヤノン株式会社）2017.10.19, 段落［００８７］－［００８９］[図９][請求項１][請求項２]（ファミリーなし）	1,2

☑　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	国際調査報告の発送日
１７．０４．２０１９	０７．０５．２０１９

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｃ　４８８１
片岡　利延
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
Y	WO 2018/235777 A1 (国立大学法人静岡大学) 2018.12.27, 段落［００２５］（ファミリーなし）	4-9,14,15
Y	高畑智之,下山勲, 防滴性のある可視光・遠赤外光同軸撮影システム, 第３６回日本ロボット学会学術講演会, 2018.09.04	12,13