

(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26日(26.03.2020)



(10) 国際公開番号

WO 2020/059564 A1

(51) 国際特許分類:

G06T 7/00 (2017.01)

(21) 国際出願番号:

PCT/JP2019/035395

(22) 国際出願日:

2019年9月9日(09.09.2019)

(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 2018-174000 2018年9月18日(18.09.2018) JP

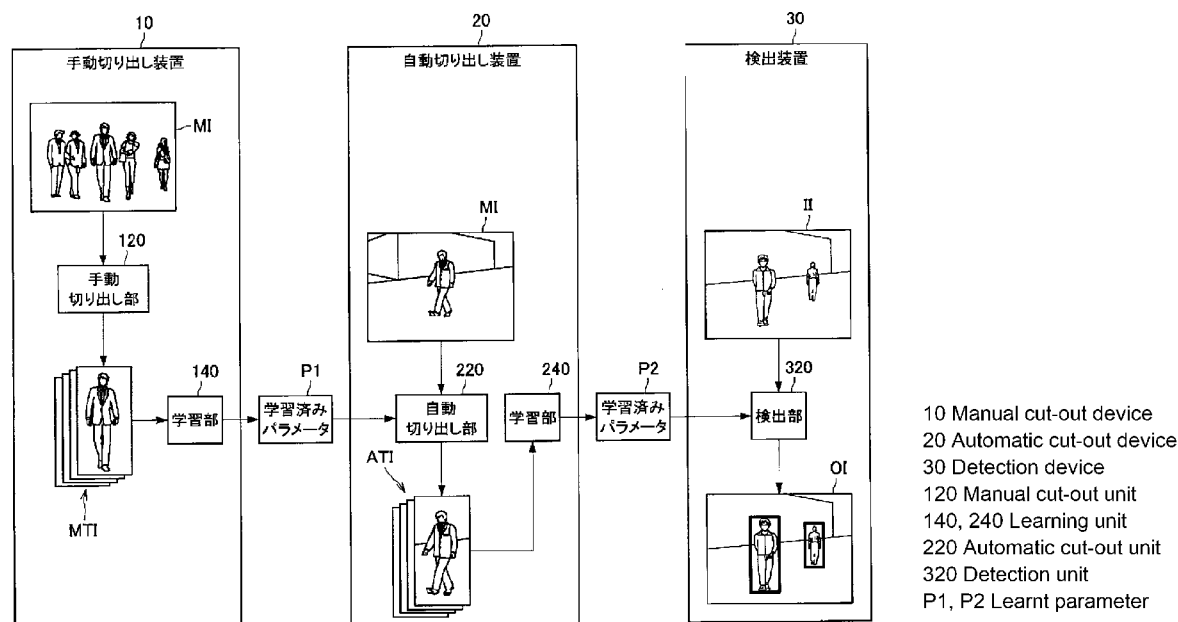
(71) 出願人: ソニーセミコンダクタソリューションズ株式会社(SONY SEMICONDUCTOR SOLUTIONS CORPORATION) [JP/JP]; 〒2430014 神奈川県厚木市旭町四丁目1番1号 Kanagawa (JP).

(72) 発明者: 石川 博隆 (ISHIKAWA, Hirotaka);

〒2430014 神奈川県厚木市旭町四丁目1番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 小川 勘治(OGAWA, Kanji); 〒2430014 神奈川県厚木市旭町四丁目1番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 山内 敏之(YAMAUCHI, Toshiyuki); 〒2430014 神奈川県厚木市旭町四丁目1番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 武藤 明宏(MUTO, Akihiro); 〒2430014 神奈川県厚木市旭町四丁目1番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 大岡 知樹(OOOKA, Tomoki); 〒2430014 神奈川県厚木市旭町四丁

(54) Title: INFORMATION PROCESSING METHOD, PROGRAM, AND INFORMATION PROCESSING SYSTEM

(54) 発明の名称: 情報処理方法、プログラム、および情報処理システム



(57) Abstract: According to the present invention, learning images, which reduce cost and have high accuracy, are generated evenly on a large scale. A processor includes: automatically cutting out a region including an object from a material image and generating an automatic cut out image; and performing learning pertaining to detection of the object on the basis of the automatic cut out image, wherein the generating of the automatic cut out image further includes generating the automatic cut out image by using an automatic cut-out apparatus that is generated by learning on the basis of manual cut out images obtained by manually cutting out a region including the object from the material image.

目 1 4 番 1 号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP).

- (74) 代理人: 特許業務法人 酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関 3 丁目 8 番 1 号 虎の門三井ビルディング 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, 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:

- 一 国際調査報告 (条約第21条(3))

(57) 要約: コストを低減すると共に精度の高い学習用画像を偏りなく大量に生成する。プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行うことと、を含み、前記自動切り出し画像を生成することは、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成すること、をさらに含む、情報処理方法が提供される。

明 細 書

発明の名称：

情報処理方法、プログラム、および情報処理システム

技術分野

[0001] 本開示は、情報処理方法、プログラム、および情報処理システムに関する。

背景技術

[0002] 近年、画像中における対象物を検出し、当該検出の結果に基づいて動作を行う装置が広く普及している。また、上記のような検出機能の構築に用いる学習用画像を自動で生成する手法も提案されている。例えば、特許文献1には、知識不要の比較的単純な非パターン認識系やパターン認識系アルゴリズムを用いて、素材画像から学習用画像の切り出しを自動で行う技術が開示されている。また、例えば、特許文献2には、対象物に係る画像をCG（Computer Graphics）により生成し、検出器の学習に用いる技術が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2014-59729号公報

特許文献2：特開2014-123366号公報

発明の概要

発明が解決しようとする課題

[0004] しかし、特許文献に開示される技術では、学習用画像の切り出しに係る精度が十分とは言えず、汎化性能が高い検出器を生成することが困難である。また、特許文献2に開示されるようなCG画像は、実空間において撮影された画像と乖離があることから、対象物の検出性能を向上させるには限界がある。

課題を解決するための手段

- [0005] 本開示によれば、プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行うことと、を含み、前記自動切り出し画像を生成することは、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成すること、をさらに含む、情報処理方法が提供される。
- [0006] また、本開示によれば、コンピュータを、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する自動切り出し部と、前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習部と、を備え、前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成する、情報処理装置、として機能させるためのプログラムが提供される。
- [0007] また、本開示によれば、対象物を含む素材画像を撮影する撮影部と、前記素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する自動切り出し部と、前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習部と、前記学習部による学習の結果に基づいて、前記対象物の検出を行う検出部と、を備え、前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成する、情報処理システムが提供される。

図面の簡単な説明

- [0008] [図1]本開示の第1の実施形態に係る情報処理システムの概要を示す図である。
- [図2]同実施形態に係る手動切り出し装置の機能構成例を示すブロック図である。

[図3]同実施形態に係る自動切り出し装置の機能構成例を示すブロック図である。

[図4]同実施形態に係る検出装置の機能構成例を示すブロック図である。

[図5]同実施形態に係る手動切り出し装置による処理の流れを示すフローチャートである。

[図6]同実施形態に係る自動切り出し装置による処理の流れを示すフローチャートである。

[図7]同実施形態に係る自動切り出し画像の生成の流れを示す図である。

[図8]同実施形態に係る自動切り出し器のネットワーク構造例を示す図である。

[図9]同実施形態に係る検出装置による処理の流れを示すフローチャートである。

[図10]本開示の第2の実施形態に係る自動切り出し装置による処理の流れを示す図である。

[図11]本開示の第3の実施形態に係る検出装置の機能構成を示すブロック図である。

[図12]同実施形態に係る検出装置による検出器のパラメータ更新処理の流れを示すフローチャートである。

[図13]本開示の一実施形態に係るハードウェア構成例を示す図である。

発明を実施するための形態

[0009] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0010] なお、説明は以下の順序で行うものとする。

1. 第1の実施形態

1. 1. 概要

1. 2. 手動切り出し装置10の機能構成例

1. 3. 自動切り出し装置 20 の機能構成例

1. 4. 検出装置 30 の機能構成例

1. 5. 処理の流れ

2. 第 2 の実施形態

3. 第 3 の実施形態

4. ハードウェア構成例

5. まとめ

[0011] <1. 第 1 の実施形態>

<<1. 1. 概要>>

まず、本開示の第 1 の実施形態の概要について説明する。上述したように、近年においては、対象物の検出結果に基づいて動作を行う装置が広く普及している。上記のような装置の一例としては、例えば、撮影された画像から歩行者や走行車両などの対象物を検出し、アラートの出力や運転補助を行う車載装置などが挙げられる。

[0012] ここで、上記のような車載装置に高性能な対象物検出機能を実現させるためには、検出器の生成段階において、学習用画像として用いる対象物の切り出し画像を大量に確保することが求められる。しかし、人手により上記のような切り出し画像を作成する場合、人件費や作業時間などのコストが増大することとなる。

[0013] このために、上述した特許文献 1 や特許文献 2 のように、学習用画像を自動で生成する手法も提案されている。しかし、例えば、特許文献 2 に記載されるように、CG により生成した合成画像を学習用画像とする場合、現実との乖離から検出器の性能を十分に向上させることが困難な場合がある。

[0014] また、例えば、特許文献 1 に開示される比較的単純な非パターン認識系やパターン認識系アルゴリズムを用いた切り出し画像の生成では、十分な切り出し精度の確保が困難であることから、例えば、歩行者の検出を行う車載装置などの高い精度が求められる装置への適用は現実的とはいえない。

[0015] さらに、特許文献 1 には、上記のように生成した切り出し画像を用いて

単一の機械学習手法を繰り返す技術が開示されているが、このような方法では学習に偏りが生じてしまい、汎化性能が高い検出器を生成することが困難である。

[0016] 本開示に係る技術思想は、上記の点に着目して発想されたものであり、コストを低減すると共に精度の高い学習用画像を偏りなく大量に生成することを可能とする。このために、本開示の一実施形態に係る情報処理方法は、プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、当該自動切り出し画像に基づいて対象物の検出に係る学習を行うことと、を含む。また、上記自動切り出し画像を生成することは、素材画像中から対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、自動切り出し画像を生成すること、をさらに含む。

[0017] 図1は、本開示の第1の実施形態に係る情報処理システムの概要を示す図である。図1に示すように、本実施形態に係る情報処理システムは、手動切り出し装置10、自動切り出し装置20、および検出装置30を含み得る。また、図1においては、検出装置30が対象物として歩行者を検出する場合の一例が示されている。

[0018] (手動切り出し装置10)

本実施形態に係る手動切り出し装置10は、自動切り出し装置20に備えられる自動切り出し機能を実現するための機械学習を行う情報処理装置である。本実施形態に係る手動切り出し装置10は、例えば、PC(Personal Computer)などで有り得る。

[0019] 本実施形態に係る手動切り出し装置10の手動切り出し部120は、作業者による操作に基づいて、対象物である歩行者を含む素材画像M1から歩行者を含む領域を切り出した手動切り出し画像MT1を生成する。すなわち、作業者は、手動切り出し装置10を操作し、素材画像M1から歩行者が映っている領域をトリミングし、当該歩行者に係る切り出し画像を生成することができる。手動切り出し画像MT1は、例えば、1万6千枚程度生成されて

もよい。

[0020] また、本実施形態に係る手動切り出し装置10の学習部140は、上記のように人手により切り出された手動切り出し画像MT1に基づいて、歩行者の特徴に係る機械学習を行う。本実施形態に係る学習部140は、DNN（Deep Neural Network）を用いた高性能の機械学習を行ってよい。本実施形態に係る学習部140によれば、作業員、すなわち人が実際に歩行者であると認識した被写体に基づき、歩行者の特徴を精度高く学習することが可能である。学習部140による学習の結果として得られる学習済みパラメータP1は、自動切り出し装置20の自動切り出し部220に備えられる自動切り出し器に適用される。上記のパラメータ適用によれば、自動切り出し装置20の自動切り出し部220が、高性能の機械学習の結果に基づいて精度高く対象物の切り出しを行うことが可能となる。なお、本実施形態に係るパラメータには、例えば、重みやバイアスなど、機械学習において一般的に用いられる各種のパラメータが広く含まれる。

[0021] （自動切り出し装置20）

本実施形態に係る自動切り出し装置20は、手動切り出し装置10による機械学習の結果として得られた学習済みパラメータP1が適用されたDNNを用いて、素材画像M1から歩行者を含む領域を自動で切り出す機能を有する情報処理装置である。本実施形態に自動切り出し装置20の自動切り出し部220は、学習済みパラメータP1が適用されたDNNを自動切り出し器として用いることで、自動切り出し画像AT1の生成することを特徴の一つとする。本実施形態に係る自動切り出し部220によれば、人手により生成された手動切り出し画像MT1に基づいて学習された学習済みパラメータP1を用いることで、精度の高い自動切り出し画像AT1を大量かつ高速で生成することが可能となる。なお、自動切り出し部220は、自動切り出し画像AT1を、例えば、100万枚程度生成してもよい。

[0022] また、本実施形態に係る自動切り出し装置20の学習部240は、自動切り出し部220により自動生成された自動切り出し画像AT1を入力として

歩行者の検出に係る中程度の性能の機械学習を行う。学習部240による学習の結果として得られる学習済みパラメータP2は、検出装置30の検出部320に備えられる検出器に適用される。本実施形態に係る自動切り出し装置20によれば、切り出し精度の高い大量の自動切り出し画像ATIを用いた機械学習を行うことで、汎化性能が高い検出器を低コストで生成することが可能となる。なお、本実施形態に係る自動切り出し装置20は、例えば、PCで有り得る。

[0023] (検出装置30)

本実施形態に係る検出装置30は、自動切り出し装置20による機械学習の結果として得られたパラメータP2を適用した中程度の性能の検出器を用いて、歩行者の検出を行う情報処理装置である。本実施形態に係る検出装置30は、例えば、車両に搭載されるカメラモジュールなどであってもよい。なお、本実施形態に係る検出器は、HOG (Histograms of Oriented Gradients) 特徴量とSVM (Support Vector Machine) を用いて対象物の検出を行う。このように、本実施形態では、DNNを用いた高性能の機械学習により、SVMを用いた中程度の性能の機械学習の学習データを自動生成することで、偏りのない学習を実現することができる。

[0024] 本実施形態に係る検出装置30の検出部320は、上記の検出器を用いて、後述する撮影部310が撮影した入力画像I1から歩行者をリアルタイムに検出することができる。また、本実施形態に係る検出部320は、検出結果を出力画像O1として出力してもよい。出力画像OMは、例えば、車両に搭載されるディスプレイ装置などに表示されてもよい。

[0025] 以上、本実施形態に係る情報処理システムの概要について説明した。以上説明したように、本実施形態に係る情報処理方法によれば、精度の高い切り出し画像を低コストで大量に自動生成することができ、結果として汎化性能が高い検出器を生成することが可能となる。

[0026] なお、本開示では、検出装置30が車両に搭載されるカメラモジュールで

あり、また検出に係る対象物が歩行者である場合を主な例として説明するが、本開示の一実施形態に係る対象物や検出装置 30 は係る例に限定されない。本開示の一実施形態に係る対象物は、車両を含む各種の移動体の移動時における障害物を広く含んでよい。障害物の一例としては、例えば、歩行者を含む人のほか、別の移動体、動物、移動ルート上における設置物などが挙げられる。また、上記の移動体の一例としては、例えば、船舶、ドローンなどを含む航空機、各種の自律移動多型ロボットなどが挙げられる。さらには、本開示の一実施形態に係る検出装置 30 は、例えば、防犯カメラなどであってもよい。本開示の技術思想は、対象物の検出を行う各種の検出器の生成に広く適用され得る。

[0027] <<1. 2. 手動切り出し装置 10 の機能構成例>>

次に、本実施形態に係る手動切り出し装置 10 の機能構成例について詳細に説明する。図 2 は、本実施形態に係る手動切り出し装置 10 の機能構成例を示すブロック図である。図 2 に示すように、本実施形態に係る手動切り出し装置 10 は、画像供給部 110、手動切り出し部 120、画像蓄積部 130、および学習部 140 を備える。

[0028] (画像供給部 110)

本実施形態に係る画像供給部 110 は、手動切り出し部 120 に対象物を被写体とした素材画像を供給する機能を有する。画像供給部 110 は、例えば、予め人手により登録された素材画像を手動切り出し部 120 により供給してもよい。また、画像供給部 110 は、例えば、インターネット上から自動で供給した素材画像を手動切り出し部 120 に供給することも可能である。

[0029] (手動切り出し部 120)

本実施形態に係る手動切り出し部 120 は、作業者による操作に基づいて、素材画像から対象物を含む領域を切り出し、手動切り出し画像を生成する機能を有する。このために、本実施形態に係る手動切り出し部 120 は、例えば、トリミング操作などが可能な画像編集インタフェースを作業者に提供

してもよい。

[0030] (画像蓄積部 130)

本実施形態に係る画像蓄積部 130 は、手動切り出し部 120 により生成された手動切り出し画像を蓄積する。

[0031] (学習部 140)

本実施形態に係る学習部 140 は、手動切り出し部 120 により生成され、画像蓄積部 130 により蓄積される手動切り出し画像を入力として対象物の特徴に係る機械学習を行う。上述したように、本実施形態に係る学習部 140 は、DNN などを用いた機械学習を行ってよい。学習部 140 による学習の結果として得られた学習済みパラメータは、自動切り出し装置 20 に備えられる自動切り出し器に適用される。

[0032] 以上、本実施形態に係る手動切り出し装置 10 の機能構成例について説明した。なお、図 2 を用いて説明した上記の構成はあくまで一例であり、本実施形態に係る手動切り出し装置 10 の機能構成は係る例に限定されない。本実施形態に係る手動切り出し装置 10 の機能構成は、仕様や運用に応じて柔軟に変形され得る。

[0033] <<1. 3. 自動切り出し装置 20 の機能構成例>>

次に、本実施形態に係る自動切り出し装置 20 の機能構成例について詳細に説明する。図 3 は、本実施形態に係る自動切り出し装置 20 の機能構成例を示すブロック図である。図 3 に示すように、本実施形態に係る自動切り出し装置 20 は、画像供給部 210、自動切り出し部 220、画像蓄積部 230、および学習部 240 を備える。

[0034] (画像供給部 210)

本実施形態に係る画像供給部 210 は、自動切り出し部 220 に対象物を被写体とした素材画像を供給する機能を有する。画像供給部 210 は、予め人手により登録された素材画像や、インターネット上から自動で供給した素材画像を自動切り出し部 220 に供給してよい。

[0035] (自動切り出し部 220)

本実施形態に係る自動切り出し部２２０は、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する機能を有する。上述したように、本実施形態に係る自動切り出し部２２０は、手動切り出し装置１０による機械学習の結果として得られた学習済みパラメータを適用した自動切り出し器を用いて、自動切り出し画像を生成することを特徴の一つとする。本実施形態に係る自動切り出し部２２０が有する機能の詳細については、別途後述する。

[0036] （画像蓄積部２３０）

本実施形態に係る画像蓄積部２３０は、自動切り出し部２２０により生成された自動切り出し画像を蓄積する。

[0037] （学習部２４０）

本実施形態に係る学習部２４０は、自動切り出し部２２０により生成され、画像蓄積部２３０により蓄積される自動切り出し画像を入力として対象物の検出に係る機械学習を行う。本実施形態に係る学習部２４０は、ＨＯＧ特徴量とＳＶＭなどによる機械学習を行ってよい。学習部２４０による学習の結果として得られた学習済みパラメータは、検出装置３０に備えられる検出器に適用される。

[0038] 以上、本実施形態に係る自動切り出し装置２０の機能構成例について説明した。なお、図３を用いて説明した上記の構成はあくまで一例であり、本実施形態に係る自動切り出し装置２０の機能構成は係る例に限定されない。本実施形態に係る自動切り出し装置２０の機能構成は、仕様や運用に応じて柔軟に変形され得る。

[0039] ＜＜１．４．検出装置３０の機能構成例＞＞

次に、本実施形態に係る検出装置３０の機能構成例について詳細に説明する。図４は、本実施形態に係る検出装置３０の機能構成例を示すブロック図である。図４に示すように、本実施形態に係る検出装置３０は、撮影部３１０、検出部３２０、動作制御部３３０、および動作部３４０を備える。

[0040] （撮影部３１０）

本実施形態に係る撮影部３１０は、車両の周囲における画像（ＲＧＢ画像）を撮影する機能を有する。なお、上記の画像には、動画像および静止画像が含まれる。

[0041] （検出部３２０）

本実施形態に係る検出部３２０は、自動切り出し装置２０による学習の結果として得られた学習済みパラメータが適用された検出器を用いて、撮影部３１０が撮影した画像から対象物をリアルタイムに検出する機能を有する。本実施形態に係る検出部３２０は、対象物の検出結果を動作制御部３３０に出力する。本実施形態に係る検出部３２０は、例えば、マイクロコントローラにより実現される。

[0042] （動作制御部３３０）

本実施形態に係る動作制御部３３０は、検出部３２０による対象物の検出結果に基づいて、動作部３４０の動作を制御する機能を有する。本実施形態に係る動作制御部３３０は、例えば、検出部３２０が車両の前方に対象物を検出したことに基づいて、動作部３４０にアラートを出力させたり、動作部３４０にブレーキを作動させてもよい。また、検出部３２０が対象物の検出結果を図１に例示したような出力画像０１として出力する場合、動作制御部３３０は、動作部３４０が備えるディスプレイ装置に出力画像０１を表示させてもよい。

[0043] （動作部３４０）

本実施形態に係る動作部３４０は、動作制御部３３０による制御に基づいて各種の動作を実行する。本実施形態に係る動作部３４０は、例えば、アクセル、ブレーキ、ステアリング、ディスプレイ装置、スピーカなどを備えてもよい。

[0044] 以上、本実施形態に係る検出装置３０の機能構成例について説明した。なお、図４を用いて説明した上記の構成はあくまで一例であり、本実施形態に係る検出装置３０の機能構成は係る例に限定されない。本実施形態に係る検出装置３０の機能構成は、仕様や運用に応じて柔軟に変形され得る。

[0045] <<1. 5. 処理の流れ>>

次に、本実施形態に係る手動切り出し装置10、自動切り出し装置20、検出装置30による処理の流れについて詳細に説明する。

[0046] まず、本実施形態に係る手動切り出し装置10による処理の流れについて説明する。図5は、本実施形態に係る手動切り出し装置10による処理の流れを示すフローチャートである。

[0047] 図5を参照すると、まず、画像供給部110が、手動切り出し部120に素材画像を供給する(S1101)。

[0048] 次に、手動切り出し部120が、ユーザ操作に基づいて、ステップS1101において供給された素材画像から対象物を含む領域を切り出し、手動切り出し画像を生成する(S1102)。

[0049] 次に、画像蓄積部130が、ステップS1102において生成された手動切り出し画像を蓄積する(S1103)。

[0050] 次に、学習部140が、ステップS1103において蓄積された手動切り出し画像を入力として対象物の特徴に係る学習を実行し、パラメータを生成する(S1104)。

[0051] 続いて、学習部140は、自動切り出し装置20に備えられる自動切り出し器にステップS1104において得られた学習済みパラメータを適用する(S1105)。

[0052] 以上、本実施形態に係る手動切り出し装置10による処理の流れについて詳細に説明した。次に、本実施形態に係る自動切り出し装置20による処理の流れについて詳細に説明する。図6は、本実施形態に係る自動切り出し装置20による処理の流れを示すフローチャートである。

[0053] 図6を参照すると、まず、画像供給部210が、自動切り出し部220に素材画像を供給する(S1201)。

[0054] 次に、自動切り出し部220が、図5のステップS1105において学習済みパラメータが適用された自動切り出し器を用いて、ステップS1201において供給された素材画像から対象物を含む領域を自動で切り出し、自動

切り出し画像を生成する（S 1 2 0 2）。

[0055] 次に、画像蓄積部 2 3 0 が、ステップ S 1 2 0 2 において生成された自動切り出し画像を蓄積する（S 1 2 0 3）。

[0056] 次に、学習部 2 4 0 が、ステップ S 1 2 0 3 において蓄積された自動切り出し画像を入力として対象物の検出に係る学習を実行し、パラメータを生成する（S 1 2 0 4）。

[0057] 続いて、学習部 2 4 0 は、検出装置 3 0 に備えられる検出器にステップ S 1 2 0 4 において得られた学習済みパラメータを適用する（S 1 2 0 5）。

[0058] 以上、本実施形態に係る自動切り出し装置 2 0 による処理の流れについて説明した。続いて、本実施形態に係る自動切り出し部 2 2 0 による自動切り出し画像の生成について、より詳細に説明する。

[0059] 図 7 は、本実施形態に係る自動切り出し画像の生成の流れを示す図である。本実施形態に係る自動切り出し部 2 2 0 による自動切り出し画像の生成は、大きく 2 つの処理により実現される。

[0060] 具体的には、本実施形態に係る自動切り出し部 2 2 0 は、まず、ディープラーニングを用いた高速物体検出手法により、画像供給部 2 1 0 により供給された素材画像 M 1 から、対象物を含むおおよその領域である暫定領域 R i を特定する。上記の高速物体検出手法には、例えば、SSD (Single Shot multibox Detector) が挙げられる。本実施形態に係る自動切り出し部 2 2 0 は、SSD のような高速物体検出手法を用いることで、素材画像 M 1 において対象物を含むおおよその暫定領域 R i を比較的高速に特定することが可能である（例えば、自動切り出し装置 2 0 が、出願時における最新の GPU (Graphics Processing Unit) を搭載している場合、100 msec 以内）。なお、自動切り出し部 2 2 0 は、SSD 等により導出される領域よりも少し広めの領域を暫定領域 R i として特定してよい。

[0061] また、上記のような SSD を利用した暫定領域 R i の特定には、例えば、インターネット上から取得可能な、コンテスト画像などで学習済みのパラメ

ータを転用することが可能である。上記のような学習済みパラメータを転用することにより、学習に要する時間や、学習用画像を生成するコストを大幅に削減することができる。

[0062] なお、本実施形態に係るディープラーニングを用いた高速物体検出手法は、SSDに限定されない。本実施形態に係る自動切り出し部220は、例えば、Faster RCNN (Regions with Convolutional Neural Networks) やYOLO (You Only Look Once) などを用いて暫定領域 R_i の特定を行ってもよい。

[0063] 次に、本実施形態に係る自動切り出し部220は、上述した自動切り出し器を用いて、SSDにより特定された暫定領域 R_i から対象物を含む詳細領域 R_d を自動で切り出し、自動切り出し画像ATIを生成する。ここで、本実施形態に係る詳細領域 R_d とは、暫定領域 R_i よりも小さな領域であり、かつ対象物を含まない領域を極力排除した矩形領域であってもよい。例えば、対象物が歩行者である場合、詳細領域 R_d の上端は歩行者の頭部上端とおおよそ一致し、詳細領域 R_d の下端は、歩行者の足の下端とおおよそ一致してよい。

[0064] このように、本実施形態に係る自動切り出し部220は、2つの異なるニューラルネットワークを用いて、暫定領域 R_i の特定処理と詳細領域 R_d の自動切り出し処理を実行することを特徴の一つとする。

[0065] 本実施形態に係る自動切り出し部220が有する上記の特徴によれば、自動切り出しの前段階として、対象物を含むおおよその領域である暫定領域 R_i を特定することで、素材画像 M_i を全走査することなく詳細領域 R_d を切り出し可能となり、自動切り出し画像ATIの生成に要する時間を大幅に削減することができる。

[0066] 以上、本実施形態に係る自動切り出し画像ATIの生成の流れについて詳細に説明した。なお、自動切り出し部220により上記のように生成された自動切り出し画像ATIは、図7の下段に示すように、正しい領域が切り出

されているかの作業者確認を以って画像蓄積部230に蓄積されてもよい。

[0067] また、上記では、本実施形態に係る自動切り出し部220が、SSDなどの高速物体検出手法を用いて素材画像から暫定領域を特定する場合を述べたが、素材画像が動画像の連続したフレームである場合、自動切り出し部220は、当該連続したフレームから動きベクトルなどの情報を取得することで、対象物が映ったおおよその領域を特定することも可能である。

[0068] 続いて、本実施形態に係る自動切り出し器のネットワーク構造について詳細に説明する。図8は、本実施形態に係る自動切り出し器のネットワーク構造例を示す図である。なお、図8に示すネットワーク構造はあくまで一例であり、本実施形態に係る自動切り出し器のネットワーク構造は柔軟に変形可能である。

[0069] 図8における「Conv」は「Convolution」層を、「Lrn」は「局所応答正規化（Local response normalization）」を、「Pool」は「Pooling」を、「FC」は「Fully Connected」層をそれぞれ示している。また、各層における「S」はストライドを、「D」はフィルタ数をそれぞれ示している。

[0070] なお、活性化関数には例えばReLUを、Poolingには、Max Poolingを採用してもよい。また、出力層は恒等関数とし、損失関数には2乗和誤差を用いてもよい。

[0071] 以上、本実施形態に係る自動切り出し器のネットワーク構成について一例を示した。本実施形態では、上記のようなネットワークに、頭部の上端と足の下端とが画像の上下端にそれぞれ一致する切り出し画像をポジティブ画像として、頭部の上端と足の下端とが画像の上下端とずれている画像をネガティブ画像として与え学習を行わせる。この際、例えば、教示データとして、ポジティブ画像には、256.0を、ネガティブ画像には0.0をそれぞれ与え、出力値の128.0を閾値として画像が正確に切り出されているか否かを判断してもよい。

[0072] また、学習データとしては、画像内正規化を施したポジティブ画像を1万

6千枚程度、同様に画像内正規化を施したネガティブ画像を20万4千枚程度与えることで、自動切り出しの精度を確保することができる。また、色や明るさ、コントラストの変化に対し寛容な自動切り出し器を生成するため、色、明るさ、コントラストをランダムに変化させるデータ拡張を行った画像を学習に用いてもよい。さらには、画像を左右方向にミラーリングするデータ拡張を行った画像を用いてもよい。上記のようなデータ拡張は、頭部や足の位置に影響を与えないことから、自動切り出しの精度を低下させることなく、学習データ数を確保することが可能である。

[0073] なお、その他のデータ拡張の一例としては、例えば、FCN (Fully Convolutional Networks) によるセグメンテーションを利用したバリエーションの増加が想定される。FCNによれば、例えば、画像中における歩行者の服の領域を特定し服の色を変えた画像を生成したり、肌の領域を特定し肌の色を変えることで人種の違いを表現した画像を生成することも可能である。また、FCNによれば、画像における背景のビルや道路などを特定した、異なる建造物等に変更することなども可能である。さらには、画像中における人の手の位置を物体検出手法を用いて特定し、鞆などを上書きすることで、素材画像のバリエーションを多様化させることも可能である。

[0074] 以上、本実施形態に係る自動切り出し装置20による処理の流れについて詳細に説明した。次に、本実施形態に係る検出装置30による処理の流れについて説明する。図9は、本実施形態に係る検出装置30による処理の流れを示すフローチャートである。

[0075] 図9を参照すると、まず、撮影部310が車両周囲におけるRGB画像を撮影する(S1301)。

[0076] 次に、検出部320が、自動切り出し装置20による学習の結果として得られた学習済みパラメータが適用された検出器を用いて、ステップS1301において撮影されたRGB画像から対象物を検出する(S1302)。

[0077] 次に、動作制御部330が、ステップS1302における検出結果に基づ

いて、動作部340に各種の動作を実行させる（S1303）。動作制御部330は、例えば、検出された対象物の領域を示す画像を動作部340に表示させてもよいし、対象物が検出されたことに基づいて、動作部340にブレーキなどを差動させてもよい。

[0078] <2. 第2の実施形態>

次に、本開示の第2の実施形態について説明する。上述した第1の実施形態では、自動切り出し装置20の自動切り出し部220にRGB画像のみが供給される場合について述べた。一方、本開示の第2の実施形態では、RGB画像に加え、同時に撮影された距離画像が併せて自動切り出し部220に供給されてよい。上記の距離画像には、例えば、ToF（Time of Flight）画像が挙げられる。上記のRGB画像および距離画像は、例えば、RGB-Dカメラなどの撮影装置により同時に撮影されてもよいし、並列に設置された異なる2つの撮影装置によりそれぞれ撮影されてもよい。

[0079] この際、本実施形態に係る自動切り出し部220は、RGB画像と同時に撮影された距離画像に基づいて、自動切り出し画像の採択に係る判定を行ってよい。すなわち、本実施形態に係る学習部240は、素材画像の撮影時における対象物と撮影装置との距離に基づいて採択された自動切り出し画像に基づいて学習を行うことができる。

[0080] 以下、本実施形態に係る距離画像を用いた自動切り出し画像の採択について詳細に説明する。なお、以下においては、第1の実施形態との差異について中心に述べ、共通する機能や効果については重複する説明を省略する。

[0081] 図10は、本開示の第2の実施形態に係る自動切り出し装置20による処理の流れを示す図である。

[0082] 図10を参照すると、まず、画像供給部210が、対象物を含むRGB画像と当該RGB画像と同時に撮影された距離画像を自動切り出し部220に供給する（S2101）。

[0083] 次に、自動切り出し部220は、SSDなどの高速物体検出手法を用いてRGB画像における暫定領域を特定し、距離画像に基づいて対象物と撮影装

置との距離が所定の範囲内であるか否かを判定する（S 2 1 0 2）。

[0084] ここで、対象物と撮影装置との距離が所定の範囲内でない場合（S 2 1 0 2 : N o）、自動切り出し部 2 2 0 は、供給された R G B 画像に対する自動切り出しを行わずに処理を終了してもよい。これによれば、後述するように学習データとして適さない自動切り出し画像を不要に生成することなく処理を効率化することができる。

[0085] 一方、対象物と撮影装置との距離が所定の範囲内である場合（S 2 1 0 2 : Y e s）、自動切り出し部 2 2 0 が自動切り出し画像を生成し、画像蓄積部 2 3 0 が当該自動切り出し画像を蓄積する（S 2 1 0 3）。

[0086] 次に、学習部 2 4 0 は、ステップ S 2 1 0 3 において蓄積された自動切り出し画像を入力とした画像を実行し、パラメータを生成する（S 2 1 0 4）。

[0087] 続いて、学習部 2 4 0 は、検出装置 3 0 に備えられる検出器にステップ S 2 1 0 4 において得られた学習済みパラメータを適用する（S 2 0 1 5）。

[0088] 以上、本実施形態に係る自動切り出し装置 2 0 の処理の流れについて一例を示した。このように、本実施形態では、撮影時における対象物との距離が所定の範囲内にある自動切り出し画像のみが学習部 2 4 0 に入力される。

[0089] なお、上記の所定の範囲には、対象物と撮影装置との距離が近すぎず、かつ遠すぎない値が設定される。対象物と撮影装置との距離が近すぎる画像では、対象物が歪んで映る場合も多く想定されるが、歪みが生じている画像を学習に用いることで、検出器の性能が低下することが予測される。このため、距離画像に基づいて、対象物と撮影装置との距離が近すぎる画像を排除することで、生成される検出器の性能を確保することが可能となる。

[0090] また、対象物と撮影装置との距離が遠すぎる画像では、対象物を含む領域にノイズが多く加わることとなる。ノイズが多い画像は、検出器の性能低下を招く要因となり得る。このため、距離画像に基づいて、対象物と撮影装置との距離が遠すぎる画像を排除することで、生成される検出器の性能を確保することが可能となる。

[0091] 以上、本開示の第2の実施形態について説明した。本実施形態に係る自動切り出し装置20によれば、より汎化性能の高い検出器を生成することが可能となる。なお、上記では、本実施形態に係る自動切り出し部220が、対象物と撮影装置との距離が所定の範囲内である場合にのみ、素材画像から自動切り出しを行う場合を例に説明したが、採択の判定タイミングは、係る例に限定されない。

[0092] 本実施形態に係る自動切り出し部220は、例えば、自動切り出し画像を生成した後に、距離画像に基づいて、当該自動切り出し画像の採択を判定してもよい。自動切り出し部220は、例えば、まず距離に依らず自動切り出し画像の生成を行い、対象物と撮影装置との距離が所定の範囲内である場合にのみ、自動切り出し画像を画像蓄積部230に蓄積させてもよい。また、素材画像を自動切り出し部220に供給する前の段階において、例えば、作業者が目視により大まかに画像の選別を行うことも可能である。

[0093] <3. 第3の実施形態>

次に、本開示の第3の実施形態について説明する。上記の第1の実施形態および第2の実施形態では、検出器を用いた対象物の検出と、検出器を生成するための学習とが、それぞれ別途の装置（すなわち、自動切り出し装置20および検出装置30）により行われる場合を述べた。一方、本開示の第3の実施形態では、単一の装置により上記の両機能を実現する構成について言及する。すなわち、本実施形態に係る検出装置30は、撮影した画像に基づいて自己学習を行い、自動的に検出器のパラメータを更新することができてよい。

[0094] 以下、本開示の第3の実施形態に係る検出装置30が有する機能について詳細に説明する。なお、以下においては、第1の実施形態および第2の実施形態との差異について中心に述べ、第1の実施形態および第2の実施形態と共通する機能や効果については、重複する説明を省略する。

[0095] 図11は、本開示の第3の実施形態に係る検出装置40の機能構成を示すブロック図である。図11に示すように、本実施形態に係る検出装置40は

、撮影部410、画像供給部420、自動切り出し部430、画像蓄積部440、学習部450、検出部460、動作制御部470、および動作部480を備える。

[0096] すなわち、本実施形態に係る検出装置40は、第1の実施形態および第2の実施形態に係る検出装置30の構成に加え、自動切り出し機能および学習機能をさらに有するといえる。上記のような構成によれば、車両の走行に伴い撮影される画像を素材画像として利用することで、検出に係る学習を継続的に行いパラメータを更新し続けることができる。

[0097] なお、本実施形態に係る学習部450は、評価用画像セットを用いることで、新たに生成したパラメータの性能が現在のパラメータの性能を上回っているか否かを判定することができる。この際、本実施形態に係る学習部450は、新たに生成したパラメータの性能が現在のパラメータの性能を上回っている場合にのみ、検出器のパラメータを自動で更新してよい。本実施形態に係る学習部450が有する上記の機能によれば、走行すればするほど検出器の汎化性能を向上させることが可能となる。

[0098] 図12は、本実施形態に係る検出装置40による検出器のパラメータ更新処理の流れを示すフローチャートである。図12を参照すると、まず撮影部410が、撮影した画像を画像供給部420に入力する(S3101)。

[0099] 次に、画像供給部420が、ステップS3101において入力された画像を素材画像として自動切り出し部430に供給する(S3102)。

[0100] 次に、自動切り出し部430が、ステップS3102において供給された素材画像から対象物を含む自動切り出し画像を生成する(S3103)。この際、自動切り出し部430は、第2の実施形態で説明したように、対象物までの距離に基づいて自動切り出し画像の採択に係る判定を行ってもよい。

[0101] 次に、画像蓄積部440が、ステップS3103において生成された自動切り出し画像を蓄積する(S3104)。

[0102] 次に、学習部450が、ステップS3104において蓄積された自動切り出し画像を入力として対象物の検出に係る学習を実行し、パラメータを生成

する（Ｓ３１０５）。

[0103] 続いて、学習部４５０は、評価用画像セットに基づいて、ステップＳ３１０５において新たに生成したパラメータにより検出器の性能が向上するか否かを判定する（Ｓ３１０６）。

[0104] ここで、新たに生成したパラメータにより検出器の性能が向上しないと判定した場合（Ｓ３１０６：Ｎｏ）、学習部４５０は検出器のパラメータを更新を行わず、検出装置４０はステップＳ３１０１に復帰する。

[0105] 一方、新たに生成したパラメータにより検出器の性能が向上すると判定した場合（Ｓ３１０６：Ｙｅｓ）、学習部４５０は検出器のパラメータを更新し、検出装置４０はステップＳ３１０１に復帰する。

[0106] 以上、本実施形態に係る検出器のパラメータ更新処理の流れについて説明した。本実施形態に係る検出装置４０による上記の処理によれば、検出に係る学習を継続的に行いパラメータを更新することで、対象物の検出性能を効率的に向上させることが可能となる。

[0107] なお、本実施形態に係る検出装置４０は、学習したパラメータを、ネットワークを介して外部のサーバなどに送信してもよい。この場合、上記サーバは、複数の検出装置４０から収集したパラメータのうち最も高い検出性能を実現するパラメータを特定し、当該パラメータを、各検出装置４０に配布することができる。このような構成によれば、複数の検出装置４０の間で、性能が高いパラメータを広く共有することができ、検出性能の高い検出装置４０を効率的に量産することが可能となる。

[0108] 一方、上記のようなパラメータの共有は、搭載される車両や撮影部４１０の仕様が共通しており、また個体差が少ない検出装置４０の間に限定されてよい。搭載される車両や撮影部４１０の個体差が大きい場合には、各検出装置４０で自己学習したパラメータのみを用いることで、対象物の検出性能を効果的に向上させることができる。

[0109] なお、検出装置４０が、ネットワークを介して外部のサーバなどと通信を行う場合、自動切り出し部４３０が有する自動切り出し画像の生成機能は、

必ずしも検出装置４０の機能として実現されなくてもよい。例えば、検出装置４０は、ネットワークを介して上記サーバに素材画像を送信し、上記サーバにより生成された自動切り出し画像を受信することで、当該自動切り出し画像を用いた学習を行ってもよい。この場合、検出装置４０の電力消費を低減し、また筐体の小型化を実現することができる。

[0110] <４．ハードウェア構成例>

次に、本開示の一実施形態に係る自動切り出し装置２０のハードウェア構成例について説明する。図１３は、本開示の一実施形態に係る自動切り出し装置２０のハードウェア構成例を示すブロック図である。図１３を参照すると、自動切り出し装置２０は、例えば、プロセッサ８７１と、ＲＯＭ８７２と、ＲＡＭ８７３と、ホストバス８７４と、ブリッジ８７５と、外部バス８７６と、インタフェース８７７と、入力装置８７８と、出力装置８７９と、ストレージ８８０と、ドライブ８８１と、接続ポート８８２と、通信装置８８３と、を有する。なお、ここで示すハードウェア構成は一例であり、構成要素の一部が省略されてもよい。また、ここで示される構成要素以外の構成要素をさらに含んでもよい。

[0111] （プロセッサ８７１）

プロセッサ８７１は、例えば、演算処理装置又は制御装置として機能し、ＲＯＭ８７２、ＲＡＭ８７３、ストレージ８８０、又はリムーバブル記録媒体９０１に記録された各種プログラムに基づいて各構成要素の動作全般又はその一部を制御する。プロセッサ８７１は、例えば、ＧＰＵやＣＰＵを含む。

[0112] （ＲＯＭ８７２、ＲＡＭ８７３）

ＲＯＭ８７２は、プロセッサ８７１に読み込まれるプログラムや演算に用いるデータ等を格納する手段である。ＲＡＭ８７３には、例えば、プロセッサ８７１に読み込まれるプログラムや、そのプログラムを実行する際に適宜変化する各種パラメータ等が一時的又は永続的に格納される。

[0113] （ホストバス８７４、ブリッジ８７５、外部バス８７６、インタフェース

８７７）

プロセッサ８７１、ＲＯＭ８７２、ＲＡＭ８７３は、例えば、高速なデータ伝送が可能なホストバス８７４を介して相互に接続される。一方、ホストバス８７４は、例えば、ブリッジ８７５を介して比較的データ伝送速度が低速な外部バス８７６に接続される。また、外部バス８７６は、インタフェース８７７を介して種々の構成要素と接続される。

[0114] （入力装置８７８）

入力装置８７８には、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチ、及びレバー等が用いられる。さらに、入力装置８７８としては、赤外線やその他の電波を利用して制御信号を送信することが可能なリモートコントローラ（以下、リモコン）が用いられることもある。また、入力装置８７８には、マイクロフォンなどの音声入力装置が含まれる。

[0115] （出力装置８７９）

出力装置８７９は、例えば、ＣＲＴ（Ｃａｔｈｏｄｅ Ｒａｙ Ｔｕｂｅ）、ＬＣＤ、又は有機ＥＬ等のディスプレイ装置、スピーカ、ヘッドホン等のオーディオ出力装置、プリンタ、携帯電話、又はファクシミリ等、取得した情報を利用者に対して視覚的又は聴覚的に通知することが可能な装置である。また、本開示に係る出力装置８７９は、触覚刺激を出力することが可能な種々の振動デバイスを含む。

[0116] （ストレージ８８０）

ストレージ８８０は、各種のデータを格納するための装置である。ストレージ８８０としては、例えば、ハードディスクドライブ（ＨＤＤ）等の磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、又は光磁気記憶デバイス等が用いられる。

[0117] （ドライブ８８１）

ドライブ８８１は、例えば、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリ等のリムーバブル記録媒体９０１に記録された情報を読み出し、又はリムーバブル記録媒体９０１に情報を書き込む装置である。

[0118] (リムーバブル記録媒体 901) リムーバブル記録媒体 901 は、例えば、DVD メディア、Blu-ray (登録商標) メディア、HD DVD メディア、各種の半導体記憶メディア等である。もちろん、リムーバブル記録媒体 901 は、例えば、非接触型 IC チップを搭載した IC カード、又は電子機器等であってもよい。

[0119] (接続ポート 882)

接続ポート 882 は、例えば、USB (Universal Serial Bus) ポート、IEEE1394 ポート、SCSI (Small Computer System Interface)、RS-232C ポート、又は光オーディオ端子等のような外部接続機器 902 を接続するためのポートである。

[0120] (外部接続機器 902)

外部接続機器 902 は、例えば、プリンタ、携帯音楽プレーヤ、デジタルカメラ、デジタルビデオカメラ、又は IC レコーダ等である。

[0121] (通信装置 883)

通信装置 883 は、ネットワークに接続するための通信デバイスであり、例えば、有線又は無線 LAN、Bluetooth (登録商標)、又は WUSB (Wireless USB) 用の通信カード、光通信用のルータ、ADSL (Asymmetric Digital Subscriber Line) 用のルータ、又は各種通信用のモデム等である。

[0122] <5. まとめ>

以上説明したように、本開示の一実施形態に係る自動切り出し装置 20 によれば、プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、自動切り出し画像に基づいて対象物の検出に係る学習を行うことと、を含み、自動切り出し画像を生成することは、素材画像中から対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、上記自動切り出し画像を生成すること、をさらに含む、情報処理方法が実現される。本開示の一

実施形態に係る情報処理方法によれば、コストを低減すると共に精度の高い学習用画像を偏りなく大量に生成することが可能となる。

[0123] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0124] また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

[0125] また、コンピュータに内蔵されるプロセッサ、ROMおよびRAMなどのハードウェアに、自動切り出し装置20や検出装置30、検出装置40が有する構成と同等の機能を発揮させるためのプログラムも作成可能であり、当該プログラムを記録した、コンピュータに読み取り可能な非一過性の記録媒体も提供され得る。

[0126] また、本明細書の各装置の処理に係る各ステップは、必ずしもフローチャートに記載された順序に沿って時系列に処理される必要はない。例えば、自動切り出し装置20や検出装置30、検出装置40の処理に係る各ステップは、フローチャートに記載された順序と異なる順序で処理されても、並列的に処理されてもよい。

[0127] なお、以下のような構成も本開示の技術的範囲に属する。

(1)

プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、

前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行うことと、

を含み、

前記自動切り出し画像を生成することは、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成すること、
をさらに含む、情報処理方法。

(2)

前記自動切り出し器は、前記手動切り出し画像を入力としたニューラルネットワークによる学習の結果として得られた学習済みパラメータに基づいて生成される、前記(1)に記載の情報処理方法。

(3)

前記自動切り出し画像を生成することは、前記素材画像中における前記対象物を含む暫定領域を特定することと、

前記暫定領域から前記対象物を含む詳細領域を自動で切り出すことと、
をさらに含む、前記(2)に記載の情報処理方法。

(4)

前記暫定領域を特定することと、前記詳細領域を自動で切り出すことは、それぞれ異なるニューラルネットワークを用いて処理を実行する、前記(3)に記載の情報処理方法。

(5)

前記暫定領域を特定することは、ディープラーニングを用いた高速物体検出手法により前記暫定領域を特定する、前記(3)または(4)に記載の情報処理方法。

(6)

前記詳細領域を自動で切り出すことは、前記自動切り出し器を用いて前記詳細領域の自動切り出しを実行し、前記自動切り出し画像を生成する、前記(4)または(5)に記載の情報処理方法。

(7)

前記学習することは、前記素材画像の撮影時における前記対象物と撮影装

置との距離に基づいて採択された前記自動切り出し画像に基づいて学習を行う、前記（１）～（６）のいずれかに記載の情報処理方法。

（８）

前記自動切り出し画像を生成することは、前記対象物と前記撮影装置との距離が所定の範囲内である場合にのみ、前記素材画像に係る前記自動切り出し画像を学習データとして採択する、前記（７）に記載の情報処理方法。

（９）

前記自動切り出し画像を生成することは、前記対象物と前記撮影装置との距離が所定の範囲内である場合にのみ、前記素材画像中から前記対象物を含む領域を自動で切り出し、前記自動切り出し画像を生成する、前記（７）に記載の情報処理方法。

（１０）

前記素材画像は、ＲＧＢ画像を含み、

前記自動切り出し画像を生成することは、前記ＲＧＢ画像と同時に撮影された距離画像に基づいて、前記自動切り出し画像の採択に係る判定を行う、前記（７）～（９）のいずれかに記載の情報処理方法。

（１１）

前記対象物は、移動体の移動時における障害物を含む、前記（１）～（１０）のいずれかに記載の情報処理方法。

（１２）

コンピュータを、

素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する自動切り出し部と、

前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習部と、

を備え、

前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し

器を用いて、前記自動切り出し画像を生成する、
情報処理装置、として機能させるためのプログラム。

(13)

対象物を含む素材画像を撮影する撮影部と
、 前記素材画像中から対象物を含む領域を自動で切り出し、自動切り出し
画像を生成する自動切り出し部と、

前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習
部と、

前記学習部による学習の結果に基づいて、前記対象物の検出を行う検出部
と、

を備え、

前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動
で切り出した手動切り出し画像に基づく学習により生成された自動切り出し
器を用いて、前記自動切り出し画像を生成する、
情報処理システム。

符号の説明

[0128]	10	手動切り出し装置
	20	自動切り出し装置
	210	画像供給部
	220	自動切り出し部
	230	画像蓄積部
	240	学習部
	30	検出装置
	310	撮影部
	320	検出部
	330	動作制御部
	340	動作部

請求の範囲

- [請求項1] プロセッサが、素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成することと、
- 前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行うことと、
- を含み、
- 前記自動切り出し画像を生成することは、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成すること、
- をさらに含む、情報処理方法。
- [請求項2] 前記自動切り出し器は、前記手動切り出し画像を入力としたニューラルネットワークによる学習の結果として得られた学習済みパラメータに基づいて生成される、請求項1に記載の情報処理方法。
- [請求項3] 前記自動切り出し画像を生成することは、前記素材画像中における前記対象物を含む暫定領域を特定することと、
- 前記暫定領域から前記対象物を含む詳細領域を自動で切り出すことと、
- をさらに含む、請求項2に記載の情報処理方法。
- [請求項4] 前記暫定領域を特定することと、前記詳細領域を自動で切り出すことは、それぞれ異なるニューラルネットワークを用いて処理を実行する、請求項3に記載の情報処理方法。
- [請求項5] 前記暫定領域を特定することは、ディープラーニングを用いた高速物体検出手法により前記暫定領域を特定する、請求項3に記載の情報処理方法。
- [請求項6] 前記詳細領域を自動で切り出すことは、前記自動切り出し器を用いて前記詳細領域の自動切り出しを実行し、前記自動切り出し画像を生成する、請求項4に記載の情報処理方法。

- [請求項7] 前記学習することは、前記素材画像の撮影時における前記対象物と撮影装置との距離に基づいて採択された前記自動切り出し画像に基づいて学習を行う、請求項1に記載の情報処理方法。
- [請求項8] 前記自動切り出し画像を生成することは、前記対象物と前記撮影装置との距離が所定の範囲内である場合にのみ、前記素材画像に係る前記自動切り出し画像を学習データとして採択する、請求項7に記載の情報処理方法。
- [請求項9] 前記自動切り出し画像を生成することは、前記対象物と前記撮影装置との距離が所定の範囲内である場合にのみ、前記素材画像中から前記対象物を含む領域を自動で切り出し、前記自動切り出し画像を生成する、請求項7に記載の情報処理方法。
- [請求項10] 前記素材画像は、RGB画像を含み、
前記自動切り出し画像を生成することは、前記RGB画像と同時に撮影された距離画像に基づいて、前記自動切り出し画像の採択に係る判定を行う、
請求項7に記載の情報処理方法。
- [請求項11] 前記対象物は、移動体の移動時における障害物を含む、請求項1に記載の情報処理方法。
- [請求項12] コンピュータを、
素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する自動切り出し部と、
前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習部と、
を備え、
前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成する、
情報処理装置、として機能させるためのプログラム。

[請求項13]

対象物を含む素材画像を撮影する撮影部と、

前記素材画像中から対象物を含む領域を自動で切り出し、自動切り出し画像を生成する自動切り出し部と、

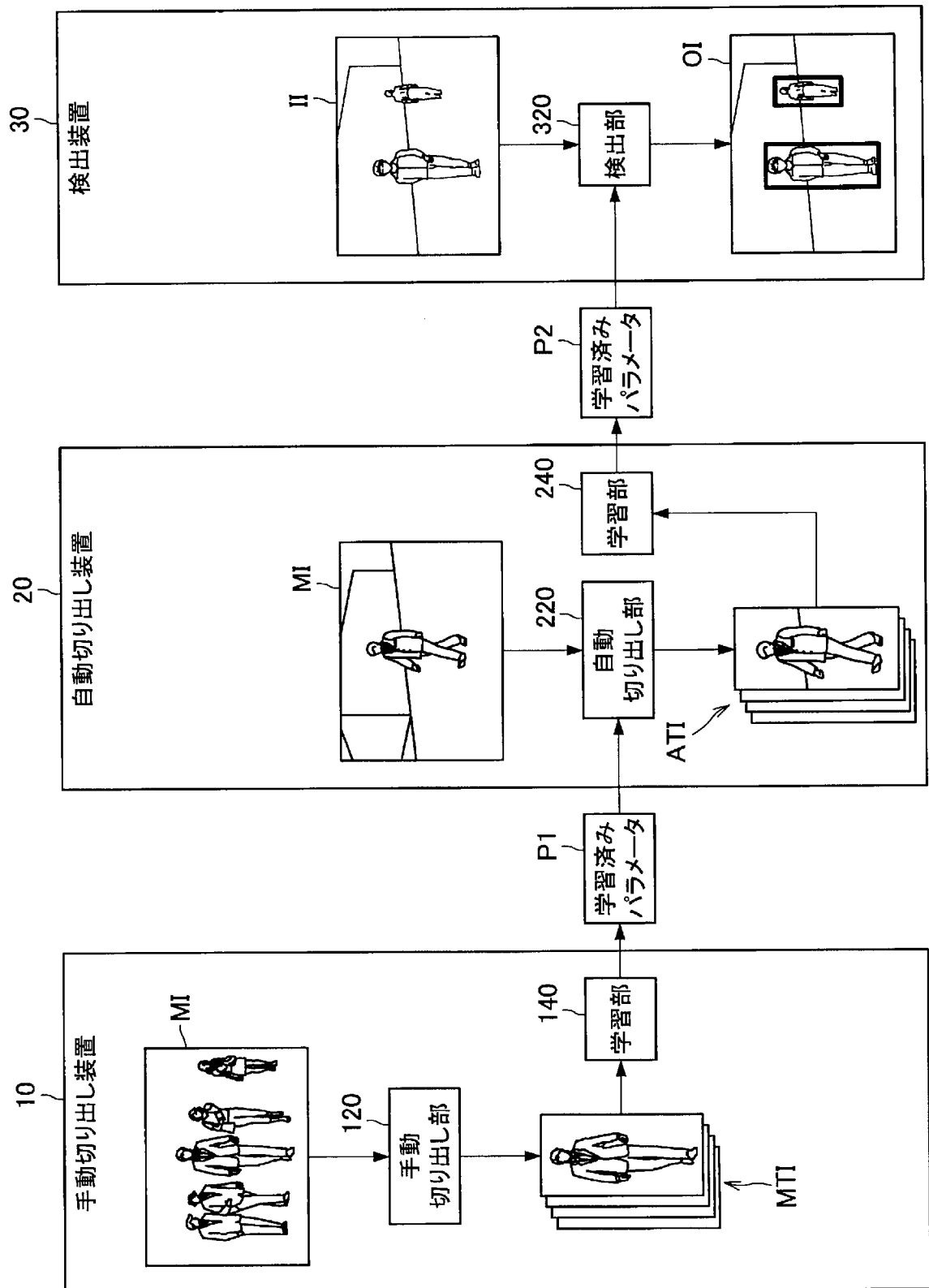
前記自動切り出し画像に基づいて前記対象物の検出に係る学習を行う学習部と、

前記学習部による学習の結果に基づいて、前記対象物の検出を行う検出部と、

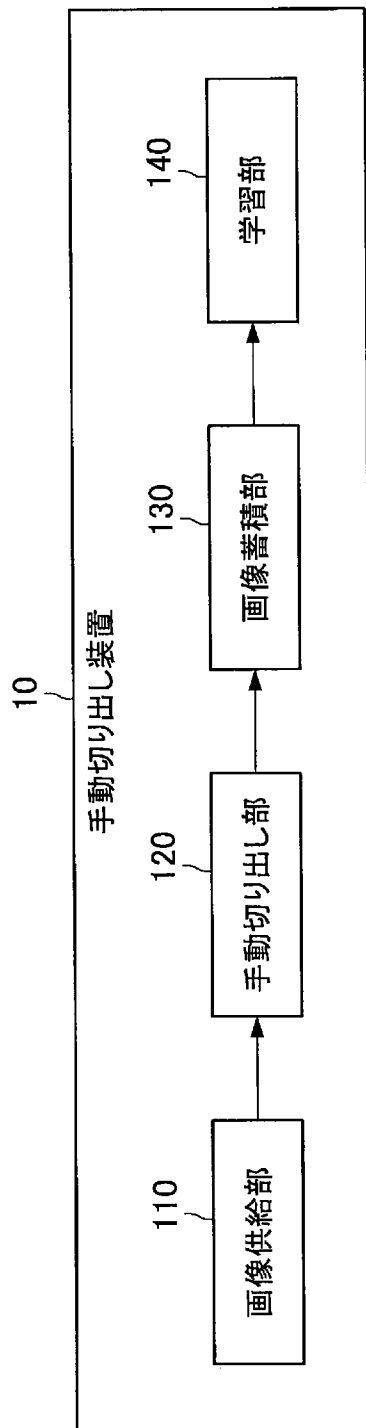
を備え、

前記自動切り出し部は、前記素材画像中から前記対象物を含む領域を手動で切り出した手動切り出し画像に基づく学習により生成された自動切り出し器を用いて、前記自動切り出し画像を生成する、情報処理システム。

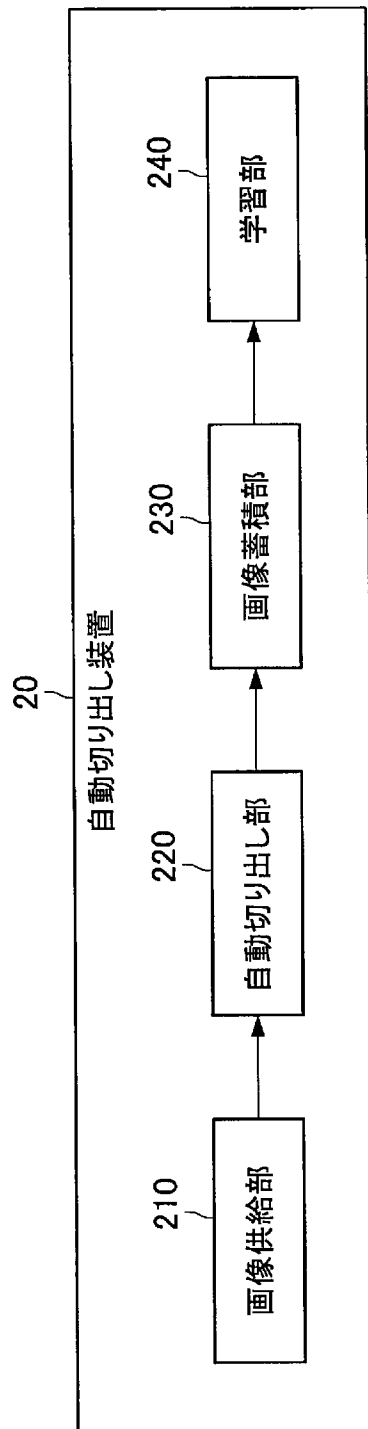
[図1]



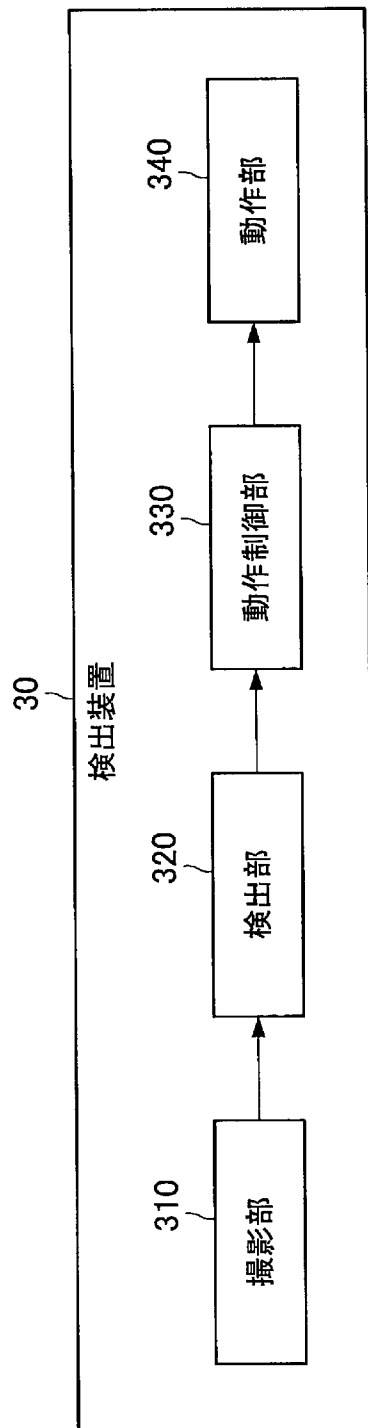
[図2]



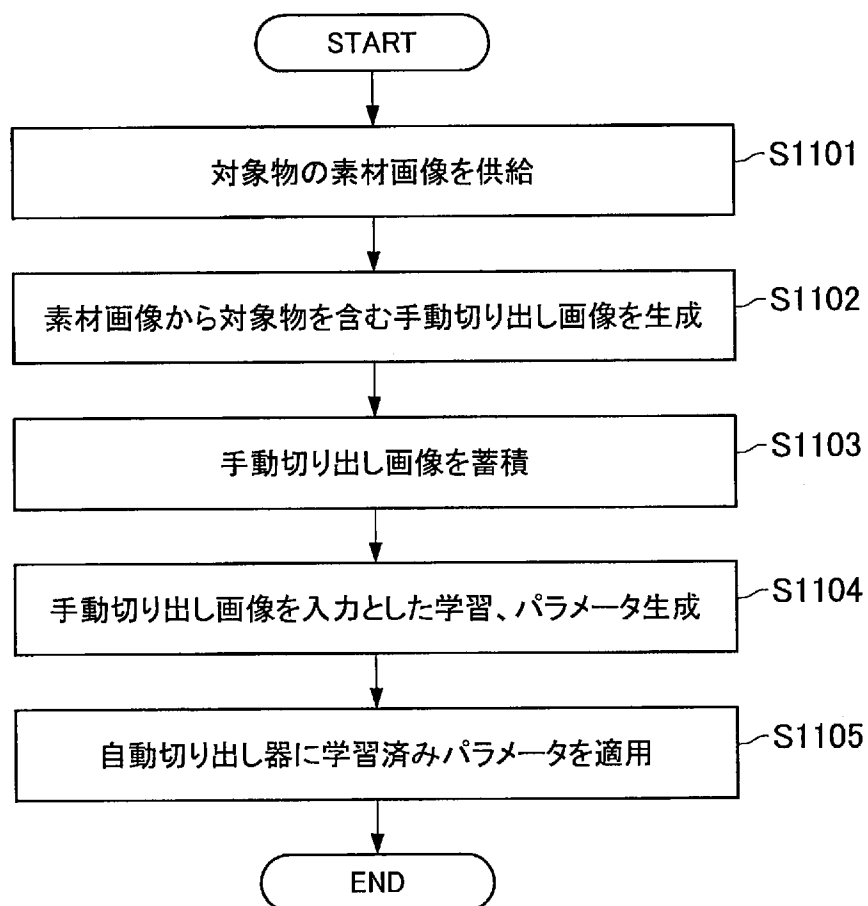
[図3]



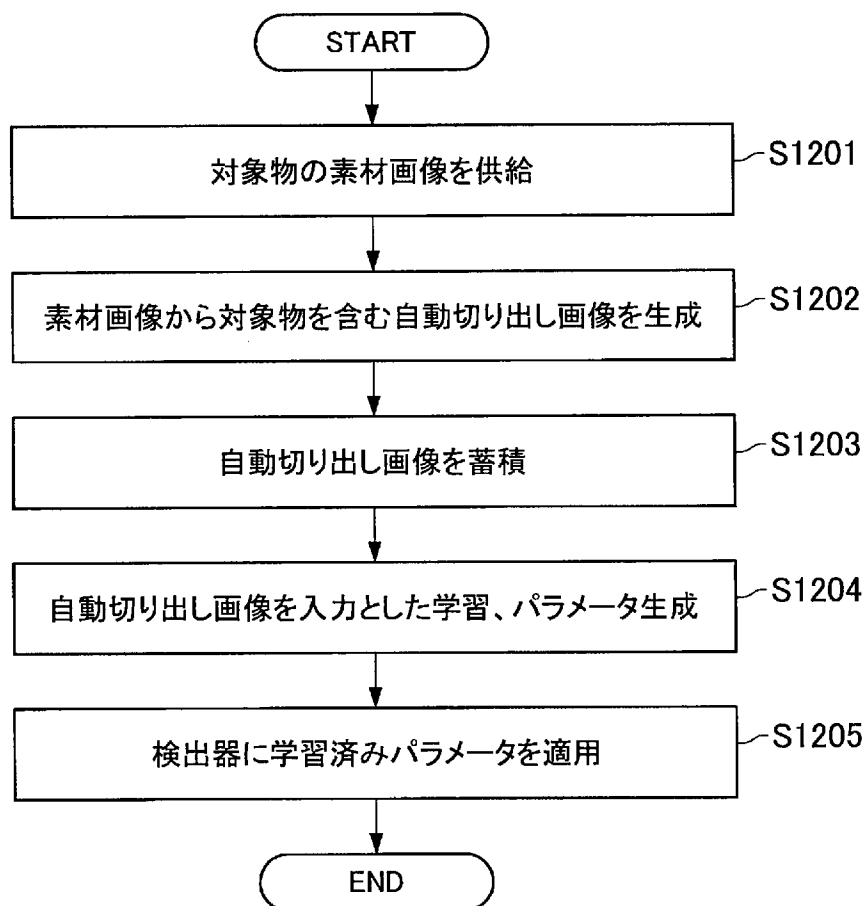
[図4]



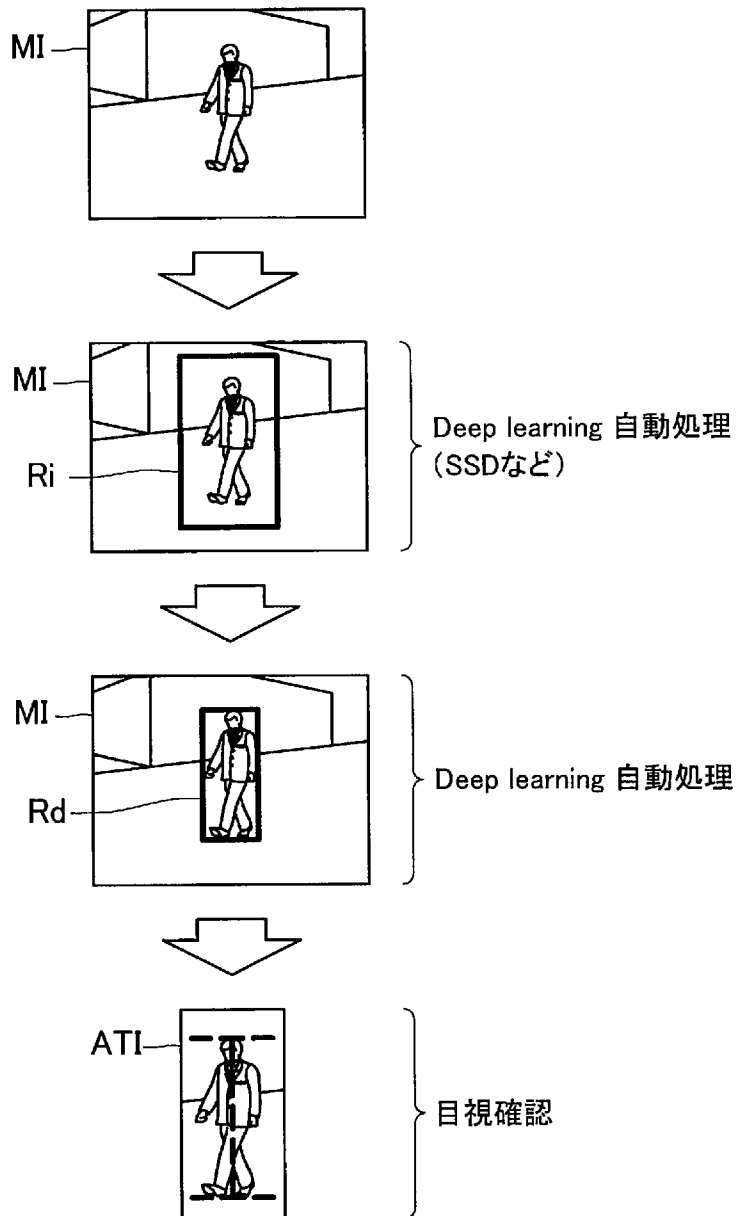
[図5]



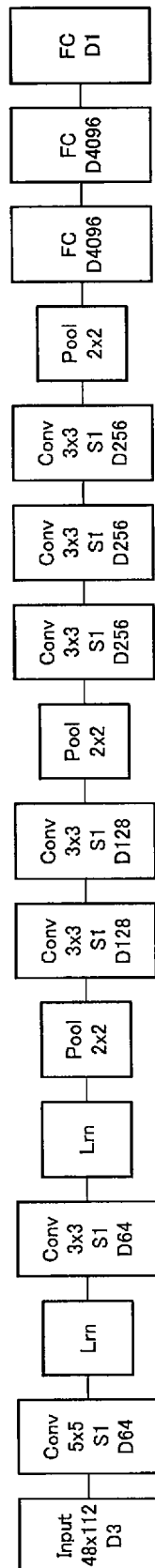
[図6]



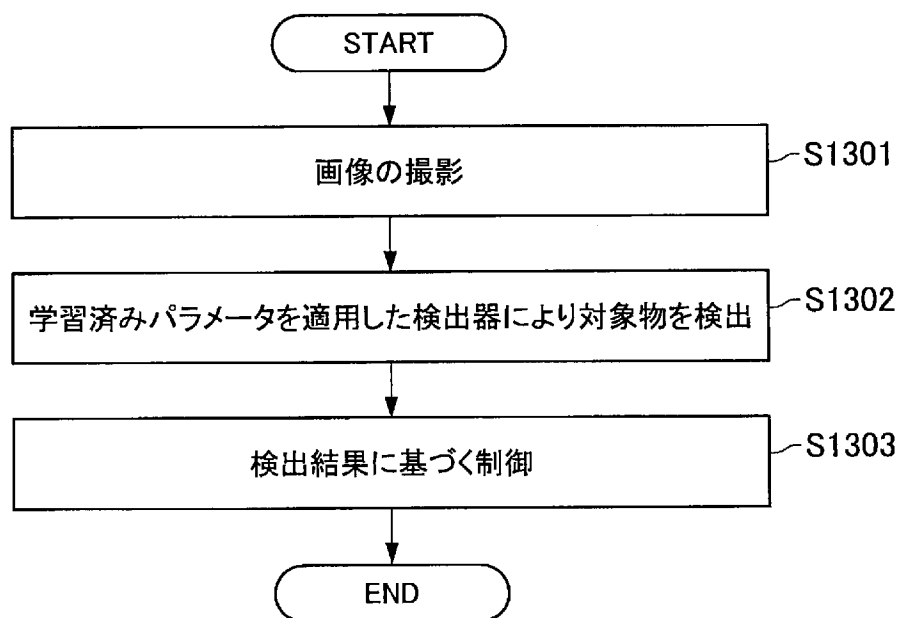
[図7]



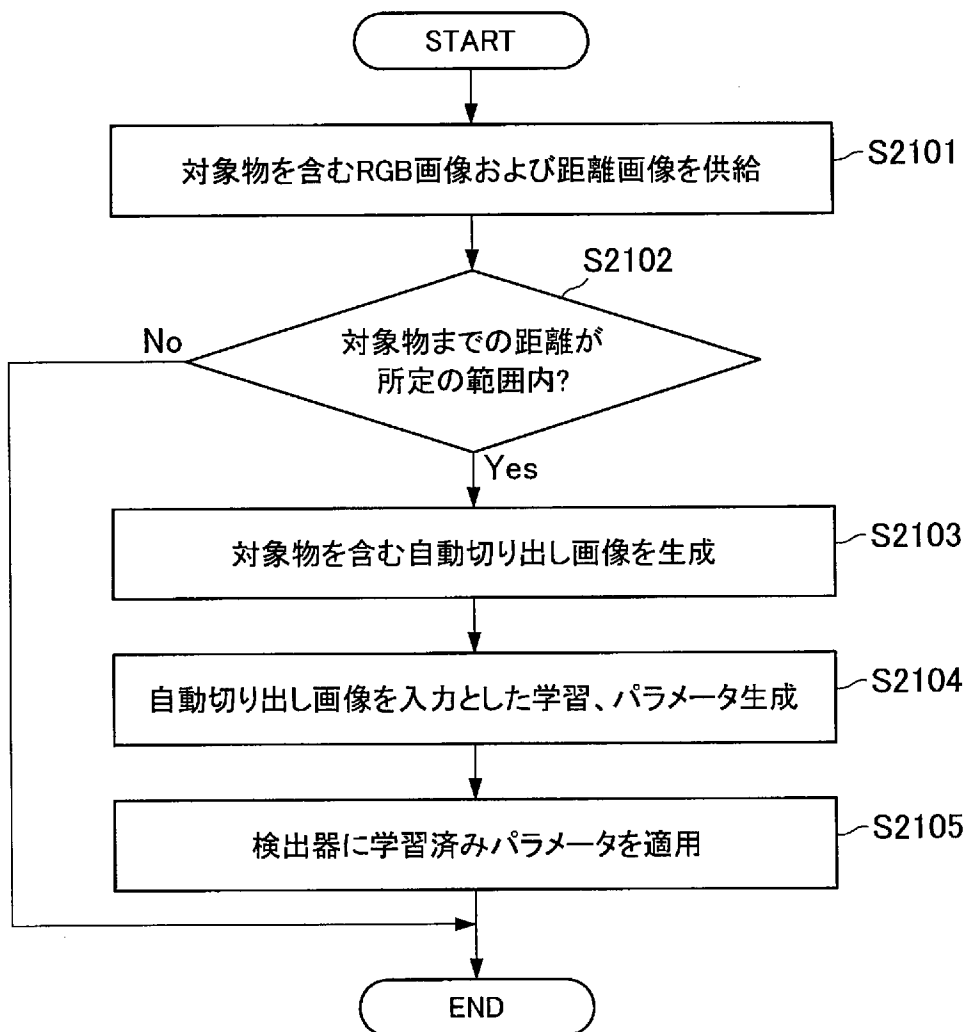
[図8]



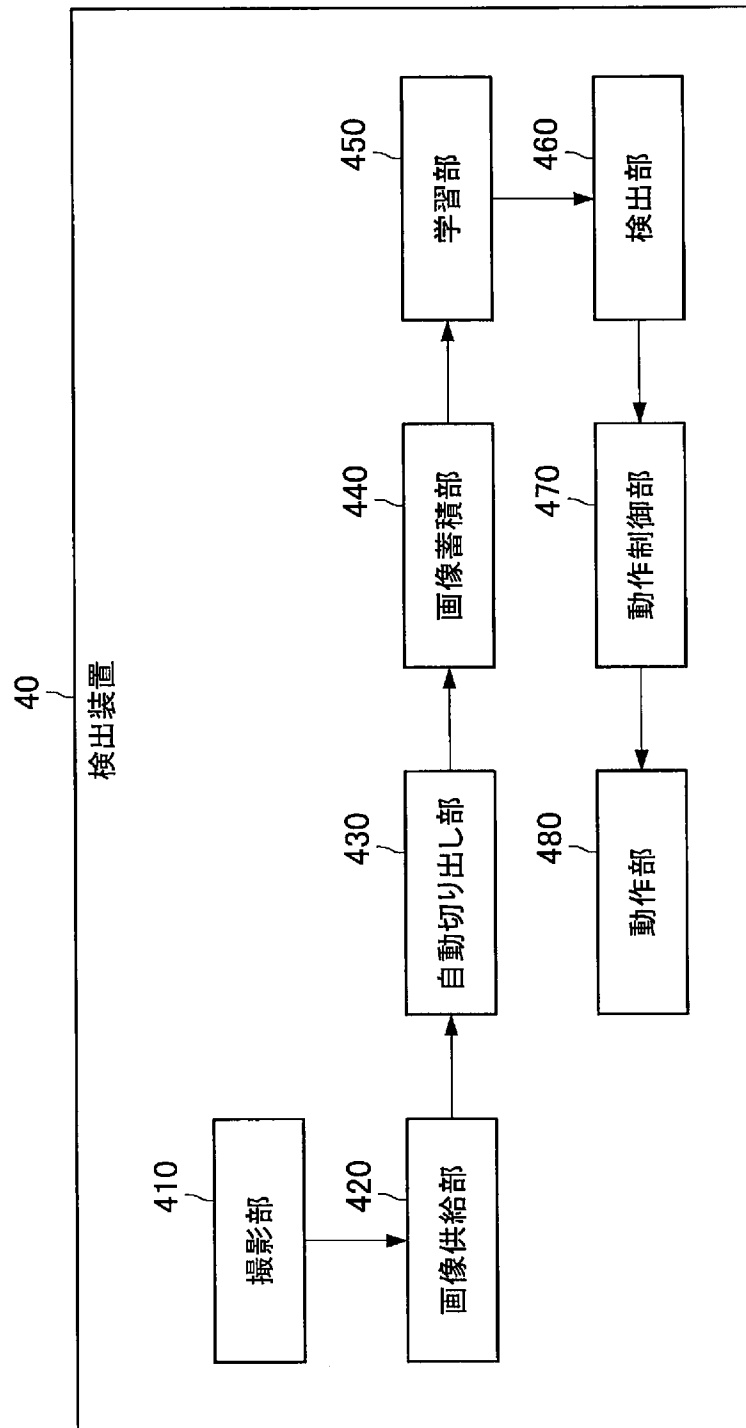
[図9]



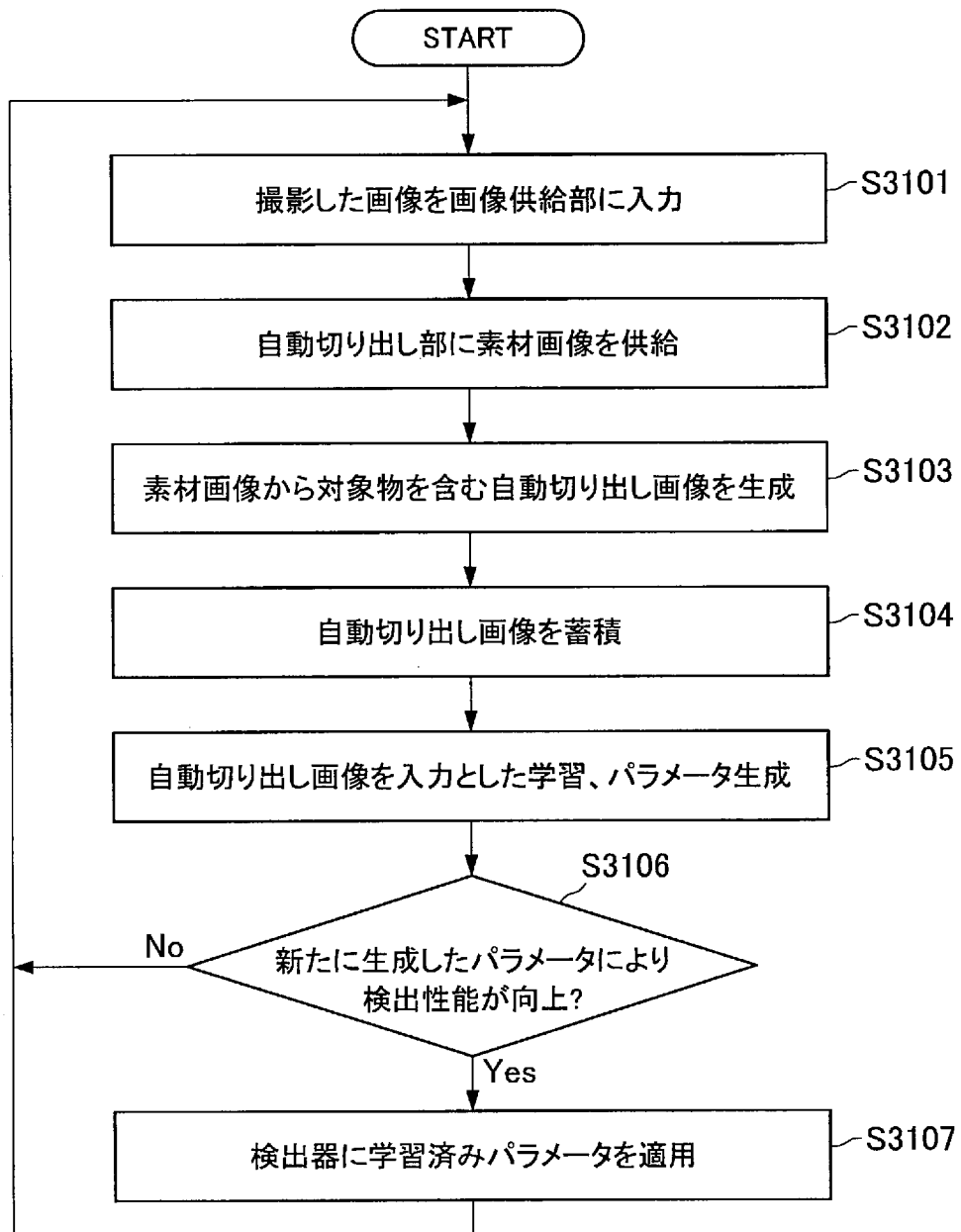
[図10]



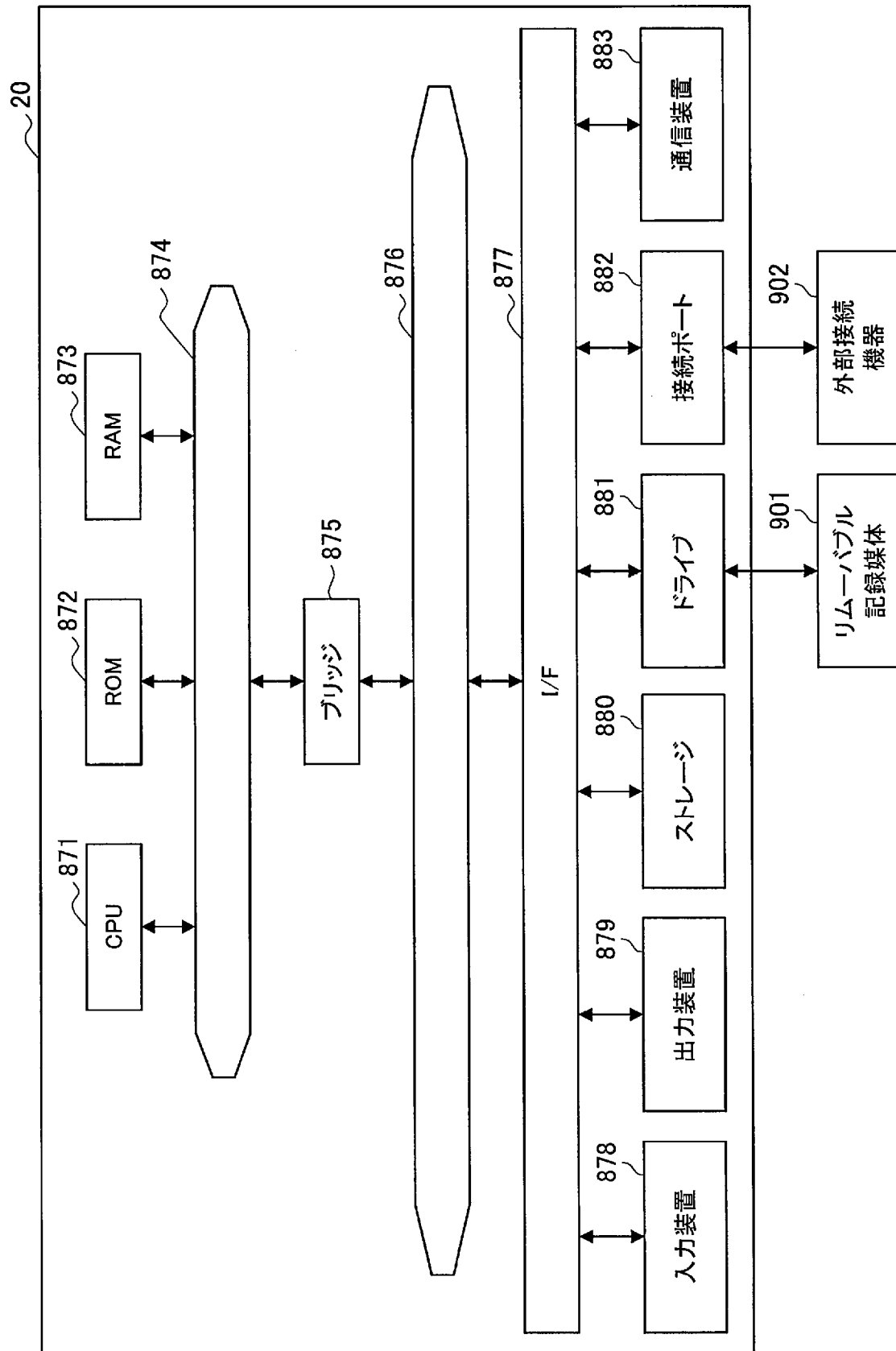
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/035395

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)

IEEE Xplore

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	FUJISANX, “画像データの切り抜きを CNN と openCV で自動化する (“Automated image data cropping using CNN and openCV”)”, In: ゼロから始める機械学習 (Starting Machine Learning from scratch), [online], Hatena Blog, 03 July 2017, pp. 1-11, [In Japanese], [retrieved on 12 November 2019], Retrieved from the Internet: <URL:http://zero-ai.hatenablog.com/entry/2017/07/03/211918>,	1, 2, 12, 13 11 3-10



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
13 November 2019 (13.11.2019)Date of mailing of the international search report
26 November 2019 (26.11.2019)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/JP2019/035395

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	渡部直樹 (外 1 名) (WATANABE, N. et al.), 「層状ニューラルネットワークによる自動車画像認識 ("Recognition of Car Images Using A Multi-Layered Neural Network")」, 電子情報通信学会技術研究報告 (IEICE Technical Report), 25 March 1994, vol. 93, no. 537, 東京: 社団法人電子情報通信学会 (Tokyo: The Institute of Electronics, Information and Communication Engineers), pp. 191-198, [In Japanese], especially, section 1.	11
A	KENNEDY, L. et al., "Learning Crop Regions for Content-Aware Generation of Thumbnail Images", Proceedings of the 1st ACM International Conference on Multimedia Retrieval (ICMR'11), 20 April 2011, article no. 30, pp. 1-8, ISBN: 978-1-4503-0336-1, <DOI:10.1145/1991996.1992026>, especially, abstract; fig. 2, [NOTE] Due to a reason of copyright (the reason for a contract with database provider), a copy of this document will not be sent	1-13

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)

IEEE Xplore

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	FUJISANX, "画像データの切り抜きを CNN と openCV で自動化する (Automated image data cropping using CNN and openCV)", In: ゼロから始める機械学習 (Starting Machine Learning from scratch), [online], Hatena Blog, 2017.07.03, Pages 1-11,	1, 2, 12, 13
Y	[In Japanese], [retrieved on 2019.11.12], Retrieved from the Internet: <URL: http://zero-ai.hatenablog.com/entry/2017/07/03/211918>.	11
A		3-10

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

13.11.2019

国際調査報告の発送日

26.11.2019

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)

久保 光宏

5 C

9189

電話番号 03-3581-1101 内線 3539

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	渡部直樹 (外 1 名) (WATANABE, N., et al.), 「層状ニューラルネットによる自動車画像認識 (Recognition of Car Images Using A Multi-Layered Neural Network)」, 電子情報通信学会技術研究報告 (IEICE Technical Report), 1994.03.25, Vol.93, No.537, 東京: 社団法人電子情報通信学会 (Tokyo: The Institute of Electronics, Information and Communication Engineers), Pages 191-198, [In Japanese]. Especially, Section 1.	1 1
A	KENNEDY, L., et al., "Learning Crop Regions for Content-Aware Generation of Thumbnail Images", Proceedings of the 1st ACM International Conference on Multimedia Retrieval (ICMR'11), 2011.04.20, Article No.30, Pages 1-8, ISBN: 978-1-4503-0336-1, <DOI: 10.1145/1991996.1992026>. Especially, Abstract; Figure 2. [NOTE] 著作権上の理由 (データベース提供者との契約上の理由) により、この文献のコピーは送付されない。	1 - 1 3