

(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
2015年12月10日(10.12.2015)



(10) 国際公開番号
WO 2015/186341 A1

- (51) 国際特許分類:
G06T 7/20 (2006.01) H04N 7/18 (2006.01)
G06T 7/00 (2006.01)
- (21) 国際出願番号: PCT/JP2015/002768
- (22) 国際出願日: 2015年6月2日(02.06.2015)
- (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 2014-115205 2014年6月3日(03.06.2014) JP
- (71) 出願人: 日本電気株式会社(NEC CORPORATION)
[JP/JP]; 〒1088001 東京都港区芝五丁目7番1号
Tokyo (JP).
- (72) 発明者: 宮野 博義 (MIYANO, Hiroyoshi); 〒
1088001 東京都港区芝五丁目7番1号日本電気
株式会社内 Tokyo (JP).
- (74) 代理人: 下坂 直樹 (SHIMOSAKA, Naoki); 〒
1088001 東京都港区芝五丁目7番1号日本電気
株式会社内 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, DK, DM, DO, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN,
IR, IS, JP, KE, KG, KN, KP, KR, 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).

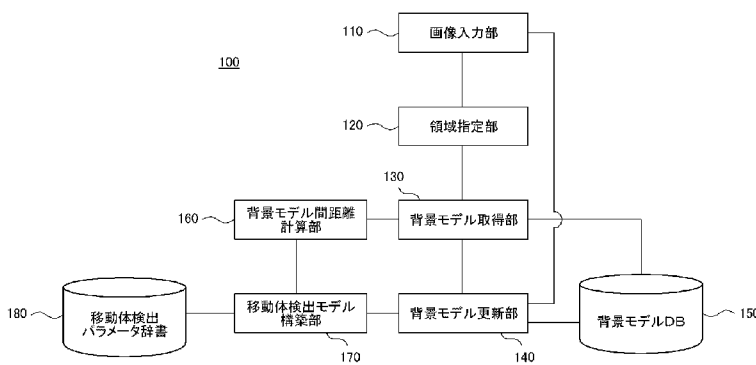
添付公開書類:

- 国際調査報告 (条約第 21 条(3))

(54) Title: IMAGE PROCESSING SYSTEM, IMAGE PROCESSING METHOD, AND PROGRAM STORAGE MEDIUM

(54) 発明の名称: 画像処理システム、画像処理方法及びプログラム記憶媒体

【図4】



- 110 Image input unit
- 120 Region designation unit
- 130 Background model acquisition unit
- 140 Background model update unit
- 150 Background model database
- 160 Background model-to-model distance computing unit
- 170 Mobile object detection model constructing unit
- 180 Mobile object detection parameter dictionary

(57) Abstract: An image processing system, an image processing method, and a program for preferably detecting a mobile object are provided. The image processing system is provided with: an image input unit (110) that receives an input with respect to several image frames of different times among a plurality of image frames configuring a picture, the input indicating whether, in the image frame of a processing time, one or more arbitrary pixels that have been selected are pixels showing a mobile object or pixels not showing the mobile object; and a mobile object detection model constructing unit (170) that learns a parameter for detecting the mobile object on the basis of the input.

(57) 要約: 好適に移動体を検出するための画像処理システム、画像処理方法及びプログラムを提供する。画像処理システムは、映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける画像入力部110と、当該入力に基

づき、移動体を検出するためのパラメータを学習する移動体検出モデル構築部170とを備える。



WO 2015/186341 A1

明 細 書

発明の名称：

画像処理システム、画像処理方法及びプログラム記憶媒体

技術分野

[0001] 本発明に係るいくつかの態様は、画像処理システム、画像処理方法及びプログラム記憶媒体に関する。

背景技術

[0002] 近年、ビデオ監視などの用途において、人や車等の移動体を検出し追跡することに対するニーズが高まっている。このようなニーズの高まりにあわせて、移動体を検出し、検出した移動体を追跡する手法が数多く提案されている。ここで移動体は、画像に映る物体のうち、常に移動をつづける物体のみに限られるものではなく、一時的に停止（静止／滞留ともいう。）する場合をも含む。換言すれば、移動体は、画像中の背景とみなされる部分以外に映る物体全般を指す。例えば、ビデオ監視での監視対象として一般的である人や車は、常に動いているわけではなく、一時的に停止したり駐車したりといった静止状態が発生する。そのため、一時的に停止する場合でも検出できるようにすることが、ビデオ監視などの用途において重要である。

[0003] 移動体を検出する手法の1つとして、背景差分法が知られている（例えば、非特許文献1及び非特許文献2参照）。背景差分法は、背景として記憶している画像と、カメラで撮影された画像とを比較して、違いのある領域を移動体として抽出する手法である。ここで、背景差分を用いて移動体を検出する場合、解析を行う時点における正確な背景抽出が必要となる。単純に観測開始時のデータを固定的に背景として利用する場合には、照明の変化など環境が変化していくことに伴う背景変化の影響を受けて誤検出が多く発生してしまうからである。そこで、このような課題を回避すべく、最新の一定時間内の観測画像から例えば各画素について平均値を算出する等の手法により、解析を行う時点での背景を推定することが多い。例えば非特許文献1には、

背景の更新を逐次行いながら背景差分法を適用する方法が開示されている。

[0004] 一方、置き去りにされた物体や一定時間滞留する人物のような一時的な静止する物体に限定して抽出する技術も存在する（例えば、特許文献1参照）。特許文献1には、シーン中の動きを複数の異なる時間幅の背景モデルで解析する手法が開示されている。当該手法では、長時間のレンジで解析された長期背景モデルと、短時間のレンジで解析された短期背景モデルを作成する。そして、短期背景モデルに基づく背景差分で移動体が検知されず、長期背景モデルに基づく背景差分で移動体が検知される状態が所定回数継続すれば、一時的な静止物が存在するものとして、当該静止物が移動体として検出される。

先行技術文献

特許文献

[0005] 特許文献1：特許第5058010号公報

非特許文献

[0006] 非特許文献1：川端敦、谷藤真也、諸岡泰男。「移動体像の抽出技術」、情報処理学会、vol.28、no.4、pp.395-402、1987

非特許文献2：C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking”, Proceedings of CVPR, vol. 2, pp. 246-252, 1999

発明の概要

発明が解決しようとする課題

[0007] ここで、非特許文献1のように、逐次背景画像を更新していきながら解析対象の画像と背景画像との差分（「背景差分」ともいう。）を行う手法において、人や車等の移動体が背景画像を解析する時間幅よりも長時間滞在する場合を考える。この場合、当該移動体は背景画像の一部と判定されてしまうため、検知できないという課題がある。一方で、一時的な静止物を検出するために解析する時間幅を長くすると、照明変動のような外部ノイズによる背

景変化の影響を受けやすくなるため、静止物以外の背景画像の一時的な変化を多く誤検出する課題が生じる。

[0008] また、特許文献1は一時的な静止物を検出することを意図したものであるが、長期背景モデルに基づく背景差分が観測画像取得時の真の背景を表現できることを前提としている。そのため、照明変動のような時々刻々と背景が変化する環境では、長期背景モデルが観測画像取得時の真の背景との違いが大きくなるため、誤検出を十分に抑止するのは困難であった。

[0009] 本発明のいくつかの態様は前述の課題に鑑みてなされたものであり、好適に移動体を検出するための画像処理システム、画像処理方法及びプログラム記憶媒体を提供することを目的の1つとする。

課題を解決するための手段

[0010] 本発明に係る1の画像処理システムは、映像を構成する複数の画像フレームのうちいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受け取る入力手段と、前記入力に基づき、移動体を検出するためのパラメータを学習する学習手段とを備える。

[0011] 本発明に係る1の画像処理システムは、撮影時刻の異なる複数の画像フレームの入力を受け取る入力手段と、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する検出手段とを備える。

[0012] 本発明に係る1の画像処理方法は、映像を構成する複数の画像フレームのうちいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受け取るステップと

、前記入力に基づき、移動体を検出するためのパラメータを学習するステップとをコンピュータが行う。

[0013] 本発明に係る1の画像処理方法は、撮影時刻の異なる複数の画像フレームの入力を受けるステップと、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出するステップとをコンピュータが行う。

[0014] 本発明に係る1のプログラムは、映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける処理と、前記入力に基づき、移動体を検出するためのパラメータを学習する処理とをコンピュータに実行させる。

[0015] 本発明に係る1のプログラムは、撮影時刻の異なる複数の画像フレームの入力を受ける処理と、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する処理とをコンピュータに実行させる。

[0016] なお、本発明において、「部」や「手段」、「装置」、「システム」とは、単に物理的手段を意味するものではなく、その「部」や「手段」、「装置」、「システム」が有する機能をソフトウェアによって実現する場合も含む。また、1つの「部」や「手段」、「装置」、「システム」が有する機能が2つ以上の物理的手段や装置により実現されても、2つ以上の「部」や「手段」、「装置」、「システム」の機能が1つの物理的手段や装置により実現

されても良い。

発明の効果

[0017] 本発明によれば、好適に移動体を検出するための画像処理システム、画像処理方法及びプログラム記憶媒体を提供することができる。

図面の簡単な説明

[0018] [図1]背景モデルと入力画像フレームとの関係を説明するための図である。

[図2]表示画面の具体例を説明するための図である。

[図3]本実施形態に係る移動体検出手法を説明するための図である。

[図4]第1実施形態に係る画像処理システムの概略構成を示す機能ブロック図である。

[図5]第1実施形態に係る画像処理システムの概略構成を示す機能ブロック図である。

[図6]図4に示す画像処理システムの処理の流れを示すフローチャートである。

[図7]図5に示す画像処理システムの処理の流れを示すフローチャートである。

[図8]図4及び5に示す画像処理システムを実装可能なハードウェアの構成を示すブロック図である。

[図9]第2実施形態に係る画像処理システムの概略構成を示す機能ブロック図である。

[図10]第2実施形態に係る画像処理システムの概略構成を示す機能ブロック図である。

発明を実施するための形態

[0019] 以下に本発明の実施形態を説明する。以下の説明及び参照する図面の記載において、同一又は類似の構成には、それぞれ同一又は類似の符号が付されている。

[0020] (1 第1実施形態)

(1.1 概要)

図1乃至図7は、実施形態を説明するための図である。以下、これらの図を参照しながら説明する。

- [0021] 本実施形態は、カメラ等の撮影装置により撮影された映像中から、人物や車などの、移動や一時的な滞留を繰り返す移動体を検出するための画像処理システムに関する。特に、本実施形態に係る画像処理システムでは、照明変動のような時々刻々と環境が変化する場合であっても、人や車などの移動体を好適に検出する。
- [0022] そのために本実施形態に係る画像処理システムは、図1に示すように、映像から切りだされた各時刻の画像フレームを元にそれぞれ作成される3つの背景モデルを生成し、これらの背景モデルを用いて移動体の検出を行う。これらの3つの背景モデルは、それぞれ背景モデルの元となる複数の画像フレームが撮影される時刻の時間幅（解析対象の時間幅）が異なる。以下、これらの3つの背景モデルを、長期背景モデル、中期背景モデル、及び短期背景モデルと呼ぶ。
- [0023] ここで、本実施形態に係る画像処理システムでは、短期背景モデル、中期背景モデル、及び長期背景モデルに対して非線形関数を適用することにより、移動体の写る領域（移動体領域）と移動体の存在しない背景領域とを判定する。より具体的には、本実施形態に係る画像処理システムは、非線形関数として、CNN（Convolution Neural Network。畳み込みニューラルネットワークとも呼ぶ。）を用いて、移動体領域の判別を行う。当該手法は、（1）移動体を判別するための移動体検出モデル（パラメータ）を学習（教師付き学習）するフェーズと、（2）生成した移動体検出モデルを用いて移動体を検出するフェーズ、の大きく2つに分かれる。
- [0024] まず、移動体検出モデルを生成するための正解データの生成手法について説明する。画像処理システムは、正解データとして、入力された各撮影時刻の画像フレームに対して、移動体領域の画素と背景領域の画素との指定をユーザから受ける。図2に、移動体領域の画素と背景領域の画素との指定を受

けるGUI (Graphical User Interface) 画面の画面例を示す。

[0025] 図2の例では、画像フレーム20上に、カーソル21が表示されている。ユーザはマウス等のポインティングデバイスによりカーソル21を操作して、背景領域にはアイコン23を配置し、人物が写る移動体領域にはアイコン25を配置する。なお、ユーザは、画像フレーム20の全画素について移動体領域であるか背景領域であるかを指定する必要はない。画像処理システムは、このように移動体領域か背景領域かの指定を受けた画素を用いてCNNの移動体検出モデルを生成する。

[0026] 図3に、本実施形態に係る画像処理システムが利用可能なCNNの具体例を示す。図3に示した例では、まず、移動体か否かを判別したい画素位置に対し、当該画素位置を中心とする5画素×5画素の画像を、短期背景モデルと中期背景モデルとの差分画像、短期背景モデルと長期背景モデルとの差分画像、及び中期背景モデルと長期背景モデルとの差分画像のそれぞれから抽出する。

[0027] その上で、8種類の3×3×3フィルタを用いて、8種類の畳み込み処理を行うことにより、3画素×3画素の画像を8枚生成する。さらに、それぞれの画像内の画素値xに対して、以下の数式f(x)を適用することにより、非線形変換を行う。

[0028] [数1]

$$f(x) = \frac{1}{\exp(x-a) + 1}$$

[0029] ここで、aは8種類のフィルタで得られる画像の画素毎に定められるパラメータであり、教師付き学習により決定される。生成された3画素×3画素×8枚の画像は、ニューラルネットワークのノードに相当する。

[0030] 同様に、これらの3画素×3画素×8枚の画像に対して、15種類の3×3×8フィルタを用いて、15種類の畳み込み処理を行うことにより、1画

素×1画素の画像を15枚生成する。その上で、それぞれの画像内の画素値 x に対して、上記 $f(x)$ を適用する。

$f(x)$ に含まれるパラメータ a は、上記の場合と同様に、15種類のフィルタで得られる画像の画素毎に定められるパラメータであり、上記教師付き学習により決定される。生成された1画素×1画素×15枚の画像は、ニューラルネットワークのノードに相当する。

[0031] 最後に、これらの1画素×1画素×15枚の画像に対して、1種類の1×1×15フィルタを用いて畳み込み処理を行い、1つの値を計算する。その上で、当該1つの値に対して上記 $f(x)$ を適用する。 $f(x)$ に含まれるパラメータ a は、上記の場合と同様に、当該1つのフィルタで得られる値に対して定められるパラメータであり、上記教師付き学習により決定される。

[0032] このような処理により得られる値を、ここでは移動体らしさ v とする。移動体であるか否かの判定は、 v に対して予め定めた閾値 T との比較によりおこなわれる。 $v \geq T$ であれば、処理対象の画素位置は移動体であると判別され、 $v < T$ であれば背景領域と判別される。閾値 T の値は予め定められるパラメータである。

[0033] 上述の通り、パラメータ a や T 等のCNNに利用するパラメータは、教師付き学習によって推定され、後述の移動体検出パラメータ辞書に格納される。パラメータの学習を行うことで、例えば風による木の揺れ等、動きを検出しやすい時間幅の背景モデルをあまり利用せずに、人や車などの特定の移動体の動きを正しく検出しやすい時間幅の背景モデルを利用するといった、特定の移動体を検出しやすい移動体検出モデルを構築することが可能である。

[0034] よって、本実施形態に係る画像処理システムは、照明変動や風等の外部ノイズによる背景変化の影響をうける環境下であっても、人や車などの移動体を、安定的に移動物体を検出できる。

[0035] なお、図3の移動体検出モデルの例では、3枚の背景モデルの差分画像から、8枚及び15枚の中間画像を生成し、最終的な移動体らしさ v を算出しているが、これに限られるものではない。例えば、入力となる背景モデルの

差分画像は4以上であっても良いし、中間画像の枚数の数ももっと多く、若しくは少なくすることも可能である。

[0036] (1. 2 システム構成)

以下、図4及び図5を用いて、本実施形態に係る画像処理システムのシステム構成を説明する。図4は移動体検出のための移動体検出モデル（パラメータ）の生成に係る学習を行う画像処理システム100のシステム構成を示す。また図5は生成された移動体検出モデルを用いて移動体を検出する画像処理システム200のシステム構成を示す。なお、図4及び図5に係る画像処理システム100及び200は、同一の装置上で実現してもよいし、異なる装置上で実現してもよい。

[0037] (1. 2. 1 学習に用いる画像処理システム100のシステム構成)

まず、図4を参照しながら、移動体検出のためのCNNの移動体検出モデルを生成するための画像処理システム100のシステム構成について説明する。画像処理システム100は、画像入力部110、領域指定部120、背景モデル取得部130、背景モデル更新部140、背景モデルデータベース(DB)150、背景モデル間距離計算部160、移動体検出モデル構築部170、及び移動体検出パラメータ辞書180を含む。

[0038] (1. 2. 1. 1 画像入力部110)

画像入力部110は、図示しないカメラなどの撮影装置から映像を構成する画像フレーム、すなわちそれぞれ撮影時刻の異なる画像フレームの入力を受ける。ここで、画像フレームはモノクロ画像であっても良いし、カラー画像であっても良い。モノクロ画像であれば、画像フレームには各画素に1つの値が含まれる。カラー画像であれば、画像フレームには各画素に3つの値（例えばRGB、YCbCr等の色表現）を有する。或いは画像フレームには、TOF (Time of Flight) カメラなどにより得られる距離情報など、画素毎に4以上の値を有してもよい。

[0039] (1. 2. 1. 2 領域指定部120)

領域指定部120は、画像フレームに対して、正解データを入力するため

のGUIをユーザに提供し、ユーザからの入力に応じて、画像フレーム内に含まれる画素に対して、移動体領域と背景領域との指定を行う。領域指定部120が表示装置に表示する表示画面の具体例は、図2に示した通りである。

[0040] これにより、領域指定部120は、当該画像フレーム内のユーザの選択した画素についての、正解データ（移動体領域若しくは背景領域の区別）を用意することができる。

[0041] なお、移動体領域と背景領域との指定は、画素に対して行うことからわかるように、画面上の様々な位置に対して点として入力する。様々な位置に対して点として入力することにより、少ない入力数でも多様性のある学習データが生成され、学習効率が良くなる。また、移動体領域と背景領域との指定は、さまざまな時刻の画像に対して実施する。これにより、より多様性のある学習データ（正解データ）が生成され、学習効率が良くなる。

[0042] (1. 2. 1. 3 背景モデル取得部130)

背景モデル取得部130は、画像入力部110から入力された画像フレーム、及び、背景モデルDB150に格納されている短期背景モデル、中期背景モデル、及び長期背景モデルの3つの背景モデルを読み込む。

[0043] (1. 2. 1. 4 背景モデルDB150)

背景モデルDB150は、解析元となる画像フレームの撮影時刻の時刻幅の異なる短期背景モデル、中期背景モデル、及び長期背景モデルを含む複数の背景モデルを格納する。

ここで、各背景モデルの形式は種々考えられるが、例えば画像入力部110から入力される画像フレームと同様の画像形式とすることができる。例えば背景モデルをモノクロ画像とするのであれば、各画素毎に1つの値が、カラー画像とするのであれば各画素毎に3つの値が含まれる。

[0044] 或いは、背景モデルは、各画素毎に、その元となった各画像フレームの画素値の尤度を示す、各画素毎の分布関数とすることもできる。ここで分布関数は、例えばヒストグラムとすることも考えられるし、或いは、複数のガウ

シアンの和による分布関数としてもよい。

[0045] 前述のとおり、短期背景モデル、中期背景モデル、及び長期背景モデルは、元となる画像フレームの撮影時刻の時間幅がそれぞれ異なり、短期背景モデル、中期背景モデル、長期背景モデルの順に、時間幅が長くなる。特に短期背景モデルに関しては、画像入力部 110 から入力された画像フレームを、そのまま短期背景モデルとして採用することも考えられる。その場合には、背景モデル DB 150 では短期背景モデルを管理しないことも考えられる。

[0046] (1. 2. 1. 5 背景モデル更新部 140)

背景モデル更新部 140 は、背景モデル取得部 130 が取得した処理時刻の画像フレーム、及び背景モデル DB 150 に記憶された背景モデルから、処理時刻の画像フレーム（最も新しい時刻の画像フレーム）を考慮した短期背景モデル、中期背景モデル、及び長期背景モデルを生成する。生成された背景モデルは、背景モデル DB 150 に格納される。

[0047] 本実施形態において、短期背景モデル、中期背景モデル、及び長期背景モデルは、それぞれ元となる画像フレームの撮影時刻の時間幅が異なる。図 1 に示すように、短期背景モデルは処理時刻から最も短い時間幅に撮影された画像フレームから、中期背景モデルはそれよりも長い時間幅に撮影された画像フレームから、長期背景モデルは最も長い時間幅に撮影された画像フレームから、それぞれ生成される。

[0048] 背景モデルの生成方法としては、例えば、各背景モデルごとに定められた時間幅分の画像フレームについて、画素値の平均値や最頻値を取ることが考えられる。或いは、前述のとおり背景モデルを画素毎の分布関数とするのであれば、時間幅内に含まれる各画像フレームの画素値の分布関数を生成してもよい。

[0049] なお、本実施形態では、短期背景モデル、中期背景モデル、及び長期背景モデルは、それぞれ元となる画像フレームの撮影時刻の時間幅が異なるものとして説明しているが、これに限られるものではない。短期背景モデル、中

期背景モデル、及び長期背景モデルは、処理時刻（最も新しい時刻）の画像フレームが与える影響の大きさの異なる背景モデルであると理解することもできる。すなわち、短期背景モデルは、処理時刻の画像フレームの与える影響が最も大きく、長期背景モデルは、処理時刻の画像フレームの与える影響は最も小さい。よって、時間幅という概念を用いる代わりに更新係数という概念を導入し、画像入力部 110 から入力された画像フレームを用いて背景モデルを更新する際の更新係数を短期背景モデル、中期背景モデル及び長期背景モデルで変えるようにしても良い。

[0050] この場合、例えば、背景モデルが I_{bg} であり、画像入力部 110 から入力された画像フレームを I とすると、

[0051] [数2]

$$I_{bg} \leftarrow (1 - a) \times I_{bg} + a \times I$$

[0052] として、背景モデルを更新することができる。この式において、 a は 0 以上 1 以下の定数であり、短期背景モデルと中期背景モデルと長期背景モデルとで異なる値を取る。短期背景モデル、中期背景モデル、及び長期背景モデルの定数を a_1 、 a_2 、 a_3 とすると、

[0053] [数3]

$$0 \leq a_3 < a_2 < a_1 \leq 1$$

[0054] との関係が成立する。 $a_1 = 1$ の場合には、短期背景モデルは新しい画像フレームで常に置き換えられる。また、 $a_3 = 0$ の場合には、長期背景モデルは固定の背景モデルを利用することを意味する。背景モデルとして固定の背景モデルを用いる場合であっても、同様の方式により更新することができる。

[0055] (1. 2. 1. 6 背景モデル間距離計算部 160)

背景モデル間距離計算部 160 は、背景モデル取得部 130 が取得した 3 つの背景モデル間の差異を数値で示す距離値を、各画素毎に計算する。具体

的には、各画素毎に、短期背景モデルと中期背景モデルとの距離、短期背景モデルと長期背景モデルとの距離、中期背景モデルと長期背景モデルとの距離を、背景モデル間距離計算部160はそれぞれ算出する。

[0056] 例えば、背景モデルを画像形式とする場合には、背景モデル間距離計算部160は、各画素の画素値の差分値もしくは差分ベクトルを算出した上で、その絶対値若しくは大きさを距離として計算することが考えられる。背景モデルが画素毎に複数の値を持っている場合、例えばRGBやYCbCr、HSV等のカラー画像形式である場合には、それぞれの値について差分値を算出した上で、それらの差分値の絶対値総和を、各画素の距離とすることも考えられる。或いは、処理対象の画素位置を中心とした近傍3画素×3画素画像や5画素×5画素画像等の近傍部分画像を抽出した上で、抽出された2つの近傍部分画像の画素値をそれぞれ2つのベクトルと捉え、当該2つのベクトルのベクトル間距離や正規化相関 r を計算してもよい。この場合、例えばモノクロ画像形式の背景モデルで近傍3×3画像により距離を算出する場合には、9次元ベクトル同士の距離を算出することになる。

また、RGBカラー画像で近傍5×5画像により距離を算出する場合には、75次元(5×5×3)ベクトル同士の距離を算出することになる。

[0057] なお、正規化相関 r を距離に用いる場合には、相関 r は1が最大値であり、 r が1に近いほど同一に近いことを示す。よって、距離の尺度に変換するため、 $1-r$ を距離を示す値として用いることができる。或いは、エッジ強調フィルタなどで上記近傍部分画像に対して前処理を行った上で、距離を計算してもよい。

[0058] また、背景モデルにヒストグラムなどの分布関数を用いている場合には、背景モデル間距離計算部160は、2つのヒストグラムの共通部分の面積や、バタチャリヤ距離などのヒストグラム距離計算手法を用いて、背景モデル間の距離を計算することができる。

[0059] なお、上述の手法では、背景モデル間距離計算部160は画素毎に距離を計算するものとして説明したが、これに限られるものではない。例えば、画

像をいくつかの領域単位であるメッシュ状に区切った上で、当該メッシュ単位毎に距離を算出する等の手法を用いることもできる。なお、距離はマイナス値を取るようにしても良い。

[0060] また、短期背景モデル、中期背景モデル、及び長期背景モデルは、それぞれ異なる形式であることも考えられる。例えば、短期背景モデルは画像形式とし、中期背景モデルは画素毎の分布関数としてもよい。この場合、距離の計算方法としては、例えば、短期背景モデルで保持されている画素値を中心として、予め定めた標準偏差の正規分布のヒストグラムを生成する。そして、当該ヒストグラムを短期背景モデルにおける分布関数とみなして、当該ヒストグラムと中期背景モデルのヒストグラムとを比較することにより距離を算出する手法を採用することが考えられる。或いは、中期背景モデルの各画素の分布関数から画素毎に平均値を算出した上で、当該平均値の集合として生成される画像形式の中期背景モデルと、短期背景モデルとを比較することにより距離を算出することも考えられる。

[0061] (1. 2. 1. 7 移動体検出モデル構築部170)

移動体検出モデル構築部170は、領域指定部120により与えられた正解データを用いて、CNNを用いて移動体を検出するための移動体検出モデル(CNNに対応するパラメータ)を生成する。より具体的には、移動体検出モデル構築部170は、与えられた正解データに最も近くなるように設定される。具体的には、N個の画素 x_i (i は $1 \leq i \leq N$ の整数)に対してそれぞれ正解データ y_i が与えられるとする。ここで、例えば画素 x_i が背景領域であれば $y_i = 0$ 、移動体領域であれば $y_i = 1$ とする。

[0062] CNNの各パラメータは、初期値としてはランダムな値を設定すれば良い。その後、上記N個の画素においてCNNを用いて移動体らしさを計算する。ここで、画素 x_i に対する移動体領域若しくは背景領域のいずれであることを示す推定結果(移動体らしさ)を v_i とする。このとき、以下の評価値Sを考える。

[0063]

[数4]

$$S = \sum_i^N (v_i - y_i)^2$$

[0064] 当該評価値Sは、CNNによる推定結果が正解データに近いほど小さい値となる。よって、Sがなるべく小さくなるように、例えば確率的降下法などの勾配法を用いてCNNのパラメータを求めれば良い。

[0065] なお、評価値Sの算出方法は異なる方式を用いても良い。例えば、交差エントロピーに相当する値である

[0066] [数5]

$$S = \sum_i^N \left\{ -y_i \times \log_e v_i - (1 - y_i) \times \log_e (1 - v_i) \right\}$$

[0067] とすることも考えられる。

[0068] なお、上記の領域指定部120による正解データの生成方法は、ユーザがランダムな位置に対して行うものである。このような手法においては、学習データが十分ではなく、特定箇所の推定精度が悪くなる場合もある。そこで、領域指定部120により与えられた正解データにより学習を行うことにより移動体検出モデルを生成した上で、推定結果を見て推定精度の低いところに関して、追加で正解データを与え、当該正解データを用いて再度学習を行う、ということを繰り返すようにしても良い。

[0069] このような処理により移動体検出モデル構築部170が生成した移動体検出パラメータは、移動体検出パラメータ辞書180に格納される。

[0070] (1. 2. 2 移動体検出に用いる画像処理システム200のシステム構成)

次に、上記画像処理システム100が生成した移動体検出モデルを用いて移動体を検出する画像処理システム200のシステム構成について説明する。画像処理システム200は、画像入力部210、背景モデル取得部220

、背景モデル更新部230、背景モデルDB240、背景モデル間距離計算部250、移動体検出部260、移動体検出パラメータ辞書270、及び結果出力部280を含む。

[0071] ここで、画像入力部210、背景モデル取得部220、背景モデル更新部230、背景モデルDB240、背景モデル間距離計算部250、及び移動体検出パラメータ辞書270の機能は、それぞれ画像入力部110、背景モデル取得部130、背景モデル更新部140、背景モデルDB150、背景モデル間距離計算部160、及び移動体検出パラメータ辞書180と同様であるため、説明を省略する。

[0072] 移動体検出部260は、移動体検出パラメータ辞書270に格納されたパラメータを用いた移動体検出モデルであるCNNにより移動体か否かを判定する。移動体の具体的な検出方法については、図3を参照しながら前述したため、ここでは説明を省略する。なお、移動体検出部260は、静止中の移動体を、移動体検出パラメータ辞書270に格納されたパラメータを用いた移動体検出モデルにより検出し、移動中の移動体を、中期背景モデルと、短期背景モデルとの差分にもとづき検出するようにしても良い。

[0073] 結果出力部280は、移動体検出部260で得られた移動体の情報を出力する。出力方法は種々考えられるが、例えば、移動体領域を1とし、それ以外の領域を0とした2値画像によって出力することができる。或いは、当該2値画像に対してラベリング処理を施すことによって連結部分を生成し、連結部分毎に外接矩形を出力することも考えられる。

[0074] 若しくは、移動体検出部260で移動中の移動体と一時的に静止している移動体とを、移動体検出部260が検出できる場合を考える。この場合、例えば、移動中の移動体として検出された画素の画素値を1、一時的に静止している移動体として検出された画素の画素値を2、それ以外の画素値を0として、3値で出力することも考えられる。場合によっては、移動中の物体であるか静止中の物体であるかを判別しづらい場合も考えられるが、その場合には移動中の物体として1を出力してもよい。或いは、全体として4値をと

ることを可能とした上で、どちらかわからない画素に画素値 3 を出力するようによい。

[0075] (1. 3 処理の流れ)

以下、画像処理システム 100 及び画像処理システム 200 の処理の流れを、図 6 及び図 7 を参照しながら説明する。図 6 は画像処理システム 100 の処理の流れを示すフローチャート、図 7 は画像処理システム 200 の処理の流れを示すフローチャートである。

[0076] なお、後述の各処理ステップは、処理内容に矛盾を生じない範囲で、任意に順番を変更して若しくは並列に実行することができ、また、各処理ステップ間に他のステップを追加してもよい。更に、便宜上 1 つのステップとして記載されているステップは複数のステップに分けて実行することもでき、便宜上複数に分けて記載されているステップを 1 ステップとして実行することもできる。

[0077] (1. 3. 1 画像処理システム 100 の処理の流れ)

まず、移動体検出のためのパラメータを学習する画像処理システム 100 の処理の流れを、図 6 を参照しながら説明する。

[0078] 画像入力部 110 は、新しい画像フレーム（処理時刻の画像フレーム）の入力を受ける（S601）。また、領域指定部 120 は、入力を受けた画像フレームに対する正解データの入力を受けるべく、図 2 に具体例を示した表示画面を表示し、当該画像フレーム内のランダムな位置の 1 以上の画素に対して、ユーザから移動体領域であるか背景領域であるかの指定を受ける（S603）。当該指定により、領域指定部 120 は移動体領域であるか否かを示す正解データを生成する。ここで、画像入力部 110 による入力の受付は、所定枚数の複数の異なる時刻の画像フレームに対して行われる。ここで、所定枚数は任意であり、予め定められていても良く、また、ユーザによる指定を受け付けてもよい。

[0079] また、背景モデル取得部 130 は、背景モデル DB 150 に格納された短期背景モデル、中期背景モデル、及び長期背景モデルを読み込む（S605

）。背景モデル間距離計算部160は、各画素に対して、短期背景モデルと中期背景モデルとの間の距離、中期背景モデルと長期背景モデルとの間の距離、及び短期背景モデルと長期背景モデルとの間の距離を算出する（S607）。

[0080] 移動体検出モデル構築部170は、背景モデル間距離計算部160が算出した各背景モデルの、正解データが用意された画素に対してCNNを適用し、評価値Sが小さくなるようなCNNのパラメータを求める（S609）。すなわち、学習により移動体検出モデルを構築する。

[0081] 画像処理システム100は、学習した移動体検出モデルによりサンプルデータの移動体検出結果を表示画面に表示する。画像処理システム100は、表示の際、移動体検出結果に対する信頼度として、前述した推定結果（移動体らしさ） v_i を検出した移動体領域の近辺に表示しても良い。または、 v_i の値に応じて、ヒートマップのように表示しても良い。具体的には、画像処理システム100は、 v_i の値が高ければ高いほど赤く、低ければ低いほど青く表示するなどしても良い。このようにして、学習した移動体検出モデルによる移動体検出結果を、人が目視で判断する。移動体検出モデルによる移動体検出精度が十分であれば（S611のYes）、移動体検出モデル構築部170は算出したパラメータを移動体検出パラメータ辞書180に出力する（S613）。

[0082] もし移動体検出モデルによる移動体検出精度が不十分であれば（S611のNo）、領域指定部120は図2のような表示画面により、特に検出精度の低い領域のランダムな画素に対して、移動体領域であるか否かの入力をユーザから受ける（S615）。その後、S605に戻って新たな正解データを用いた移動体検出モデルの構築を行う。

[0083] (1. 3. 2 画像処理システム200の処理の流れ)

続いて、画像処理システム100により生成されたCNNのパラメータを用いた移動体検出に係る処理の流れを、図7を参照しながら説明する。

[0084] 画像入力部210は、新しい画像フレーム（処理時刻の画像フレーム）の

入力を受ける（S701）。また、背景モデル取得部220は、背景モデルDB240に格納された短期背景モデル、中期背景モデル及び長期背景モデルを読み込む（S703）。

[0085] 背景モデル間距離計算部250は、各画素に対して、短期背景モデルと中期背景モデルとの間の距離、中期背景モデルと長期背景モデルとの間の距離、及び短期背景モデルと長期背景モデルとの間の距離を算出する（S705）。移動体検出部260は、移動体検出パラメータ辞書270に格納された、画像処理システム100が生成したパラメータを用いて、背景モデル間距離計算部250が算出した各背景モデル間の距離を入力として、各画素が移動体の映る領域であるか否かを判別する（S707）。結果出力部280は、検出結果を出力する（S709）。

[0086] また、背景モデル更新部230は、画像入力部210から入力された画像フレームを用いて各背景モデルを更新し、更新した背景モデルを背景モデルDB240に記憶させる（S711）。

[0087] (1.4 ハードウェア構成の具体例)

以下、図8を参照しながら、上述してきた画像処理システム100及び200をコンピュータ800により実現する場合のハードウェア構成の一例を説明する。なお、前述のとおり、画像処理システム100及び200は同一のコンピュータ上に実装してもよいし、画像処理システム100及び200の機能を、それぞれ複数台のコンピュータ上に実装してもよい。

[0088] 図8に示すように、コンピュータ800は、プロセッサ801、メモリ803、記憶装置805、入力インタフェース（I/F）部807、データI/F部809、通信I/F部811、及び表示装置813を含む。

[0089] プロセッサ801は、メモリ803に記憶されているプログラムを実行することにより、画像処理システム100及び200の様々な処理を制御する。例えば、図4に示した画像入力部110、領域指定部120、背景モデル取得部130、背景モデル更新部140、背景モデル間距離計算部160、及び移動体検出モデル構築部170にかかる処理は、メモリ803に一時記

憶された上で主にプロセッサ801上で動作するプログラムとして実現可能である。また、図5に示した画像入力部210、背景モデル取得部220、背景モデル更新部230、背景モデル間距離計算部250、移動体検出部260、及び結果出力部280にかかる処理も、同様に、メモリ803に一時記憶された上で主にプロセッサ801上で動作するプログラムとして実現可能である。

[0090] メモリ803は、例えばRAM (Random Access Memory) 等の記憶媒体である。メモリ803は、プロセッサ801によって実行されるプログラムのプログラムコードや、プログラムの実行時に必要となるデータを一時的に記憶する。

[0091] 記憶装置805は、例えばハードディスクやフラッシュメモリ等の不揮発性の記憶媒体である。記憶装置805は、オペレーティングシステムや、画像処理システム100及び200の各機能を実現するための各種プログラムや、背景モデルDB150及び240、並びに移動体検出パラメータ辞書180及び270を含む各種データ等を記憶できる。記憶装置805に記憶されているプログラムやデータは、必要に応じてメモリ803にロードされることにより、プロセッサ801から参照される。

[0092] 入力I/F部807は、ユーザからの入力を受け付けるためのデバイスである。例えば、領域指定部120が提供する図2のような表示画面において、背景領域であるか移動体領域であるかの指定を行うためのユーザ操作は、入力I/F部807により入力される。

入力I/F部807の具体例としては、キーボードやマウス、タッチパネル等が挙げられる。入力I/F部807は、例えばUSB (Universal Serial Bus) 等のインタフェースを介してコンピュータ800に接続されても良い。

[0093] データI/F部809は、コンピュータ800の外部からデータを入力するためのデバイスである。データI/F部809の具体例としては、各種記憶装置に記憶されているデータを読み取るためのドライブ装置等がある。デ

ータ I / F 部 8 0 9 は、コンピュータ 8 0 0 の外部に設けられても良い。その場合、データ I / F 部 8 0 9 は、例えば USB 等のインタフェースを介してコンピュータ 8 0 0 へと接続される。

[0094] 通信 I / F 部 8 1 1 は、コンピュータ 8 0 0 の外部の装置、例えば撮影装置（ビデオカメラや監視カメラ、デジタルカメラ）などとの間で有線又は無線によりデータ通信するためのデバイスである。通信 I / F 部 8 1 1 は画像処理システム 1 0 0 の外部に設けられても良い。その場合、通信 I / F 部 8 1 1 は、例えば USB 等のインタフェースを介してコンピュータ 8 0 0 に接続される。

[0095] 表示装置 8 1 3 は、例えば、図 2 に例示した背景領域 / 移動体領域を指定するための表示画面や、結果出力部 2 8 0 が出力する移動体の検出結果等を表示するためのデバイスである。表示装置 8 1 3 の具体例としては、例えば液晶ディスプレイや有機 EL (Electro-Luminescence) ディスプレイ等が挙げられる。表示装置 8 1 3 は、コンピュータ 8 0 0 の外部に設けられても良い。その場合、表示装置 8 1 3 は、例えばディスプレイケーブルなどを介してコンピュータ 8 0 0 に接続される。

[0096] (1. 5 本実施形態に係る効果)

以上説明したように、本実施形態に係る画像処理システム 1 0 0 及び 2 0 0 は、短期背景モデル、中期背景モデル、及び長期背景モデル間のそれぞれの差異を用いて CNN により移動体を検出する。特に、学習を用いることにより、例えば風による木の揺れなどで動きを検出しやすい時間幅の背景モデルをあまり利用せずに、人や車などの移動体を検出しやすい時間幅の背景モデルの寄与度を大きくすることにより、特定の移動体を誤検出を抑制しつつ検知することが可能である。また、照明変動のような外部ノイズによる背景変化の影響を受ける環境下においても、安定的に移動体を検出することができる。

[0097] (2 第 2 実施形態)

以下、第 2 実施形態を、図 9 及び図 1 0 を参照しながら説明する。図 9 は

、本実施形態に係る画像処理システム900の機能構成を示すブロック図である。画像処理システム900は、入力部910、及び学習部920を含む。

[0098] 入力部910は、映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける。

[0099] 学習部920は、入力部910から入力された、移動体の映る画素であるか移動体の映らない画素であるかの入力に基づき、移動体を検出するためのパラメータを学習する。

[0100] また、画像処理システム1000は、入力部1010及び検出部1020を含む。

[0101] 入力部1010は、撮影時刻の異なる複数の画像フレームの入力を受ける。

[0102] 検出部1020は、処理時刻の画像フレームを元に生成された第1の背景モデルと、第1の背景モデルよりも処理時刻の画像フレームの影響が小さい第2の背景モデルと、第2の背景モデルよりも処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する。

[0103] このように実装することで、本実施形態に係る画像処理システム900及び画像処理システム1000は、好適に移動体を検出することができる。

[0104] (3 付記事項)

なお、前述の実施形態の構成は、組み合わせたり或いは一部の構成部分を入れ替えたりしてもよい。また、本発明の構成は前述の実施形態のみに限定されるものではなく、本発明の要旨を逸脱しない範囲内において種々変更を加えてもよい。

[0105] なお、前述の各実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。また、本発明のプログラムは、上記の各実

施形態で説明した各動作を、コンピュータに実行させるプログラムであれば良い。

[0106] (付記1)

映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける入力手段と、前記入力に基づき、移動体を検出するためのパラメータを学習する学習手段とを備える画像処理システム。

[0107] (付記2)

前記処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとの間で、それぞれ差異を算出する算出手段、を更に備え、前記学習手段は、前記入力に基づき、前記第1の背景モデル、前記第2の背景モデル、及び前記第3の背景モデルを用いて移動体を検出するためのパラメータを学習する付記1に記載の画像処理システム。

[0108] (付記3)

前記学習手段は、前記1以上の画素の近傍領域の背景モデルの値を用いて畳み込み計算を1以上行うことにより移動体を検出するための検出モデルに用いられる前記パラメータを学習する、付記1又は付記2に記載の画像処理システム。

[0109] (付記4)

前記学習手段は、前記畳み込み計算、及び、前記畳み込み計算の結果得られる値と比較される閾値とを前記パラメータとして学習する付記1乃至付記3のいずれか1項記載の画像処理システム。

[0110] (付記5)

撮影時刻の異なる複数の画像フレームの入力を受ける入力手段と、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モ

デルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する検出手段とを備える画像処理システム。

[0111] (付記6)

前記第1の背景モデルと、前記第2の背景モデルと、前記第3の背景モデルとは、考慮される画像フレームの撮影時刻の時間幅が異なる付記5記載の画像処理システム。

[0112] (付記7)

前記第1の背景モデルとして、前記処理時刻の画像フレームを用いる、付記5又は付記6記載の画像処理システム。

[0113] (付記8)

映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受け取るステップと、前記入力に基づき、移動体を検出するためのパラメータを学習するステップとをコンピュータが行う画像処理方法。

[0114] (付記9)

前記処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとの間で、それぞれ差異を算出する算出手段、を更に備え、前記学習手段は、前記入力に基づき、前記第1の背景モデル、前記第2の背景モデル、及び前記第3の背景モデルを用いて移動体を検出するためのパラメータを学習する付記8に記載の画像処理方法。

[0115] (付記10)

前記 1 以上の画素の近傍領域の背景モデルの値を用いて畳み込み計算を 1 以上行うことにより移動体を検出するための検出モデルに用いられる前記パラメータを学習する、付記 8 又は付記 9 記載の画像処理方法。

[0116] (付記 1 1)

前記畳み込み計算、及び、前記畳み込み計算の結果得られる値と比較される閾値とを前記パラメータとして学習する付記 8 乃至付記 10 のいずれか 1 項記載の画像処理方法。

[0117] (付記 1 2)

撮影時刻の異なる複数の画像フレームの入力を受けるステップと、処理時刻の画像フレームを元に生成された第 1 の背景モデルと、前記第 1 の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第 2 の背景モデルと、前記第 2 の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第 3 の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を 1 回以上行うことにより移動体を検出するステップとをコンピュータが行う画像処理方法。

[0118] (付記 1 3)

前記第 1 の背景モデルと、前記第 2 の背景モデルと、前記第 3 の背景モデルとは、考慮される画像フレームの撮影時刻の時間幅が異なる付記 12 記載の画像処理方法。

[0119] (付記 1 4)

前記第 1 の背景モデルとして、前記処理時刻の画像フレームを用いる、付記 12 又は付記 13 記載の画像処理方法。

[0120] (付記 1 5)

映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の 1 以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける処理と、

前記入力に基づき、移動体を検出するためのパラメータを学習する処理と

をコンピュータに実行させるプログラム。

[0121] (付記 16)

前記処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとの間で、それぞれ差異を算出する算出手段、を更に備え、前記学習手段は、前記入力に基づき、前記第1の背景モデル、前記第2の背景モデル、及び前記第3の背景モデルを用いて移動体を検出するためのパラメータを学習する付記15に記載のプログラム。

[0122] (付記 17)

前記1以上の画素の近傍領域の背景モデルの値を用いて畳み込み計算を1以上行うことにより移動体を検出するための検出モデルに用いられる前記パラメータを学習する、付記15又は付記16に記載のプログラム。

[0123] (付記 18)

前記畳み込み計算、及び、前記畳み込み計算の結果得られる値と比較される閾値とを前記パラメータとして学習する付記15乃至付記17のいずれか1項記載のプログラム。

[0124] (付記 19)

撮影時刻の異なる複数の画像フレームの入力を受ける処理と、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する処理とをコンピュータに実行させるプログラム。

[0125] (付記 20)

前記第1の背景モデルと、前記第2の背景モデルと、前記第3の背景モデルとは、考慮される画像フレームの撮影時刻の時間幅が異なる付記19記載

のプログラム。

[0126] (付記 21)

前記第1の背景モデルとして、前記処理時刻の画像フレームを用いる、付記19又は付記20記載のプログラム。

[0127] 以上、上述した実施形態を模範的な例として本発明を説明した。しかしながら、本発明は、上述した実施形態には限定されない。即ち、本発明は、本発明の範囲内において、当業者が理解し得る様々な態様を適用することができる。

[0128] この出願は、2014年6月30日に提出された日本出願特願2014-115205を基礎とする優先権を主張し、その開示の全てをここに取り込む。

符号の説明

- [0129] 20 : 画像フレーム
21 : カーソル
23、25 : アイコン
100 : 画像処理システム
110 : 画像入力部
120 : 領域指定部
130 : 背景モデル取得部
140 : 背景モデル更新部
150 : 背景モデルデータベース
160 : 背景モデル間距離計算部
170 : 移動体検出モデル構築部
180 : 移動体検出パラメータ辞書
200 : 画像処理システム
210 : 画像入力部
220 : 背景モデル取得部
230 : 背景モデル更新部

- 240 : 背景モデルデータベース
- 250 : 背景モデル間距離計算部
- 260 : 移動体検出部
- 270 : 移動体検出パラメータ辞書
- 280 : 結果出力部
- 800 : コンピュータ
- 801 : プロセッサ
- 803 : メモリ
- 805 : 記憶装置
- 807 : 入力インタフェース部
- 809 : データインタフェース部
- 811 : 通信インタフェース部
- 813 : 表示装置
- 900 : 画像処理システム
- 910 : 入力部
- 920 : 学習部
- 1000 : 画像処理システム
- 1010 : 入力部
- 1020 : 検出部

請求の範囲

- [請求項1] 映像を構成する複数の画像フレームのうちの一つの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける入力手段と、
前記入力に基づき、移動体を検出するためのパラメータを学習する学習手段と
を備える画像処理システム。
- [請求項2] 前記処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとの間で、それぞれ差異を算出する算出手段、
を更に備え、
前記学習手段は、前記入力に基づき、前記第1の背景モデル、前記第2の背景モデル、及び前記第3の背景モデルを用いて移動体を検出するためのパラメータを学習する
請求項1に記載の画像処理システム。
- [請求項3] 前記学習手段は、前記1以上の画素の近傍領域の背景モデルの値を用いて畳み込み計算を1以上行うことにより移動体を検出するための検出モデルに用いられる前記パラメータを学習する、
請求項1又は請求項2に記載の画像処理システム。
- [請求項4] 前記学習手段は、前記畳み込み計算、及び、前記畳み込み計算の結果得られる値と比較される閾値とを前記パラメータとして学習する
請求項1乃至請求項3のいずれか1項に記載の画像処理システム。
- [請求項5] 撮影時刻の異なる複数の画像フレームの入力を受ける入力手段と、
処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さ

い第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する検出手段とを備える画像処理システム。

[請求項6] 前記第1の背景モデルと、前記第2の背景モデルと、前記第3の背景モデルとは、考慮される画像フレームの撮影時刻の時間幅が異なる請求項5記載の画像処理システム。

[請求項7] 前記第1の背景モデルとして、前記処理時刻の画像フレームを用いる、請求項5又は請求項6記載の画像処理システム。

[請求項8] コンピュータが映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのうち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受信し、前記入力に基づき、移動体を検出するためのパラメータを学習する画像処理方法。

[請求項9] コンピュータが撮影時刻の異なる複数の画像フレームの入力を受信し、処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する画像処理方法。

[請求項10] 映像を構成する複数の画像フレームのうちのいくつかの異なる時刻の画像フレームに対する入力であって、処理時刻の画像フレームのう

ち、選択された任意の1以上の画素に対し、移動体の映る画素であるか移動体の映らない画素であるかの入力を受ける処理と、

前記入力に基づき、移動体を検出するためのパラメータを学習する処理と

をコンピュータに実行させるプログラムを記憶するプログラム記憶媒体。

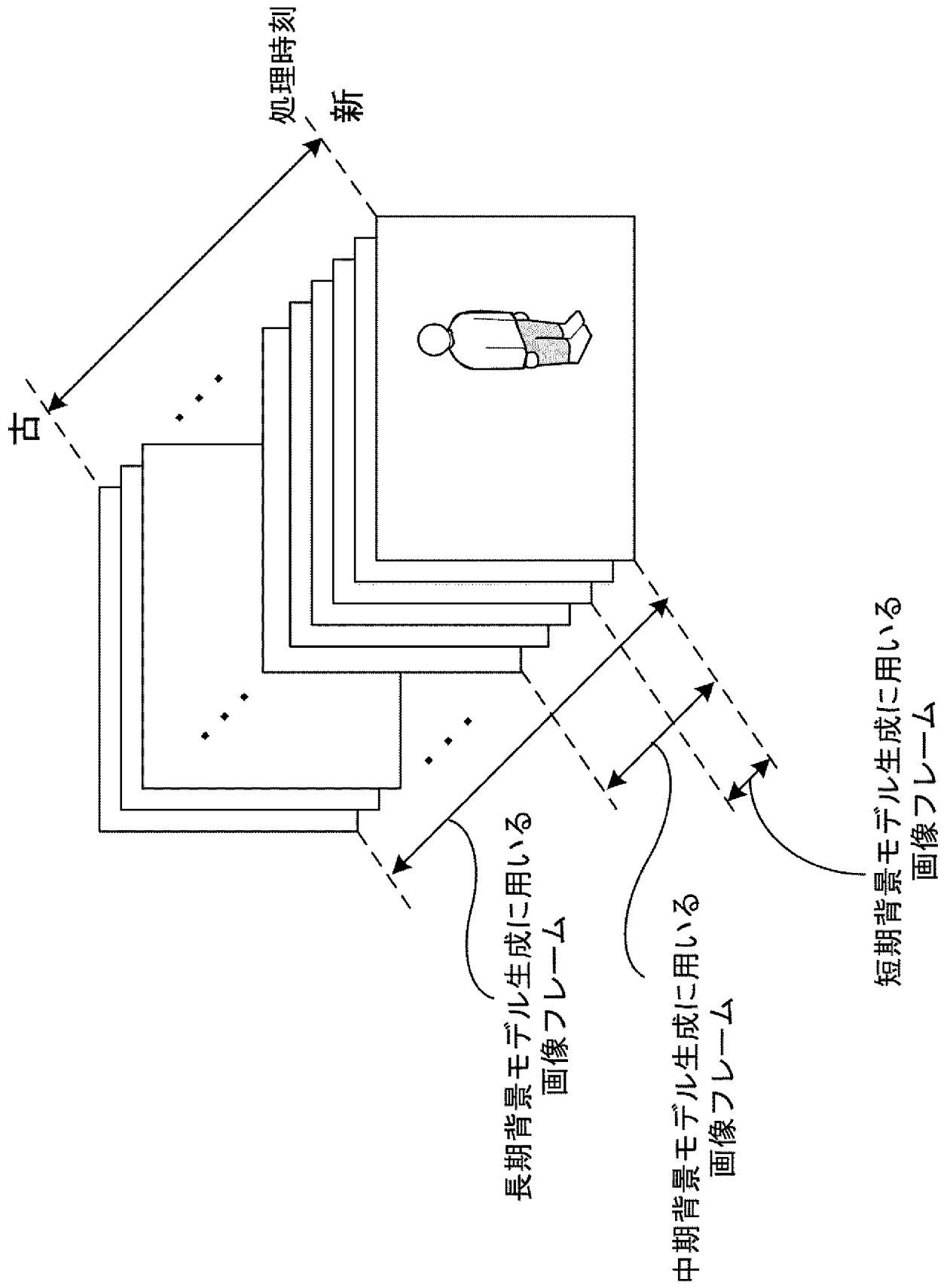
[請求項11]

撮影時刻の異なる複数の画像フレームの入力を受ける処理と、

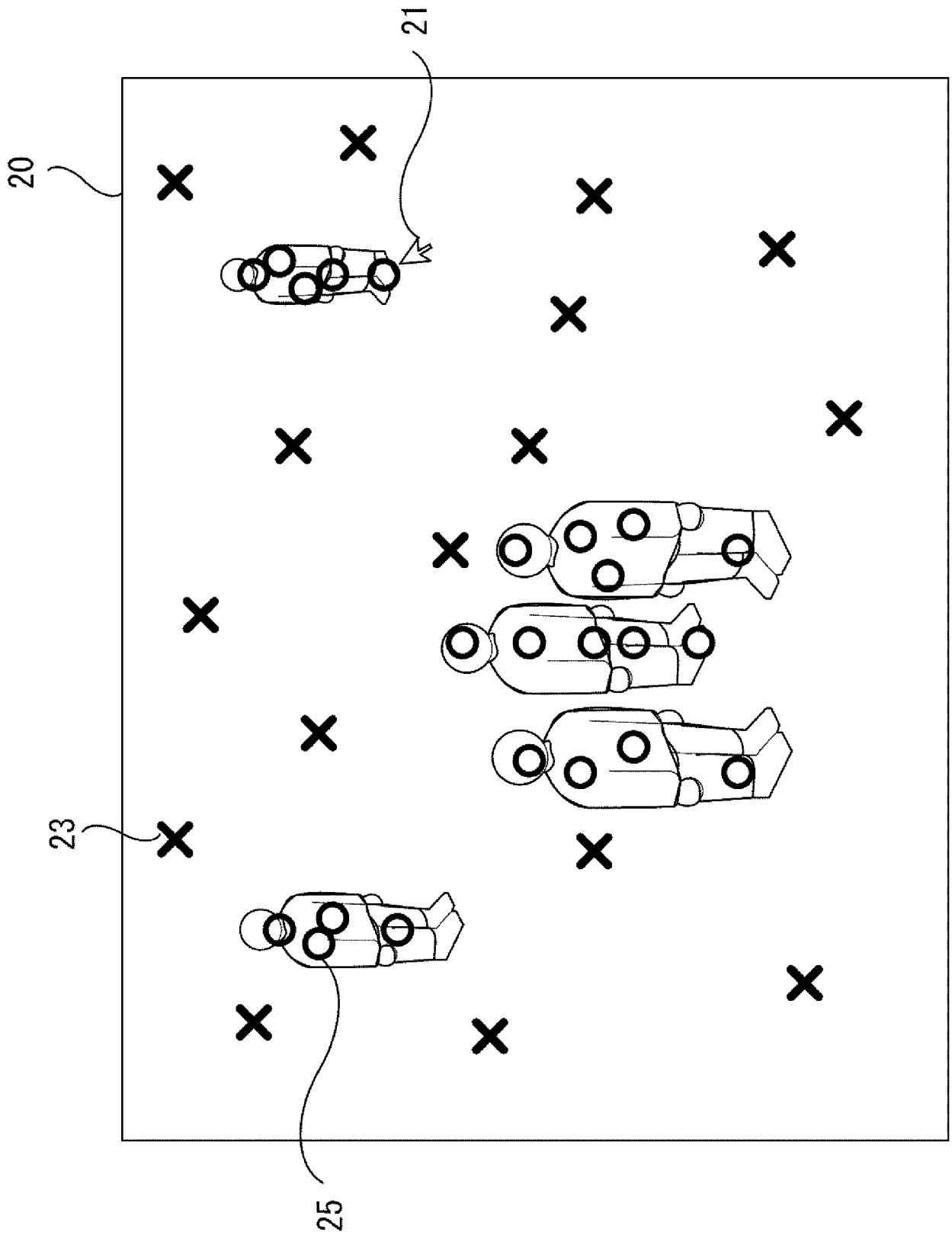
処理時刻の画像フレームを元に生成された第1の背景モデルと、前記第1の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第2の背景モデルと、前記第2の背景モデルよりも前記処理時刻の画像フレームの影響が小さい第3の背景モデルとを用いて、対象画素の近傍領域の背景モデルの値を用いて畳み込み計算を1回以上行うことにより移動体を検出する処理と

をコンピュータに実行させるプログラムを記憶するプログラム記憶媒体。

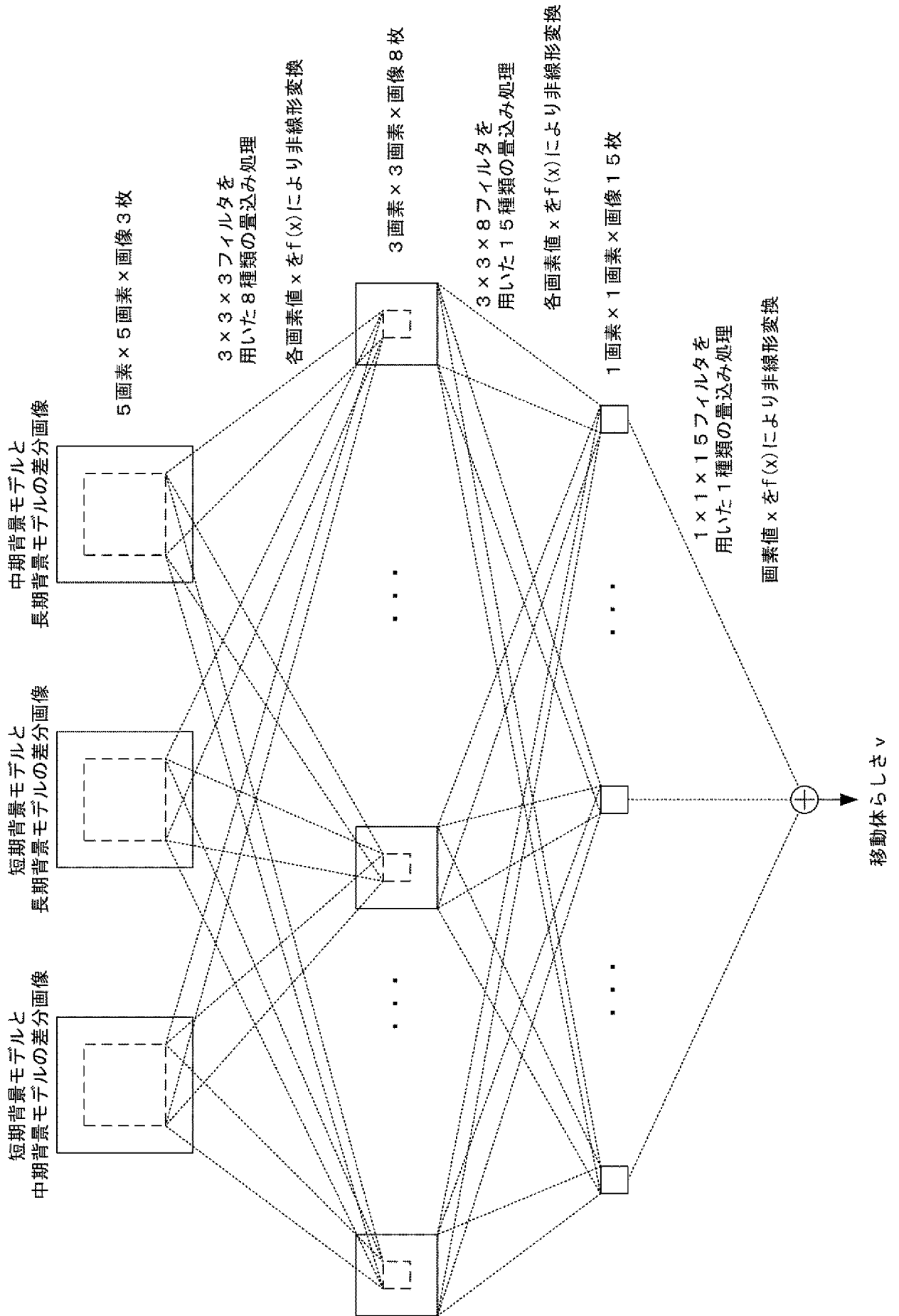
[図1]



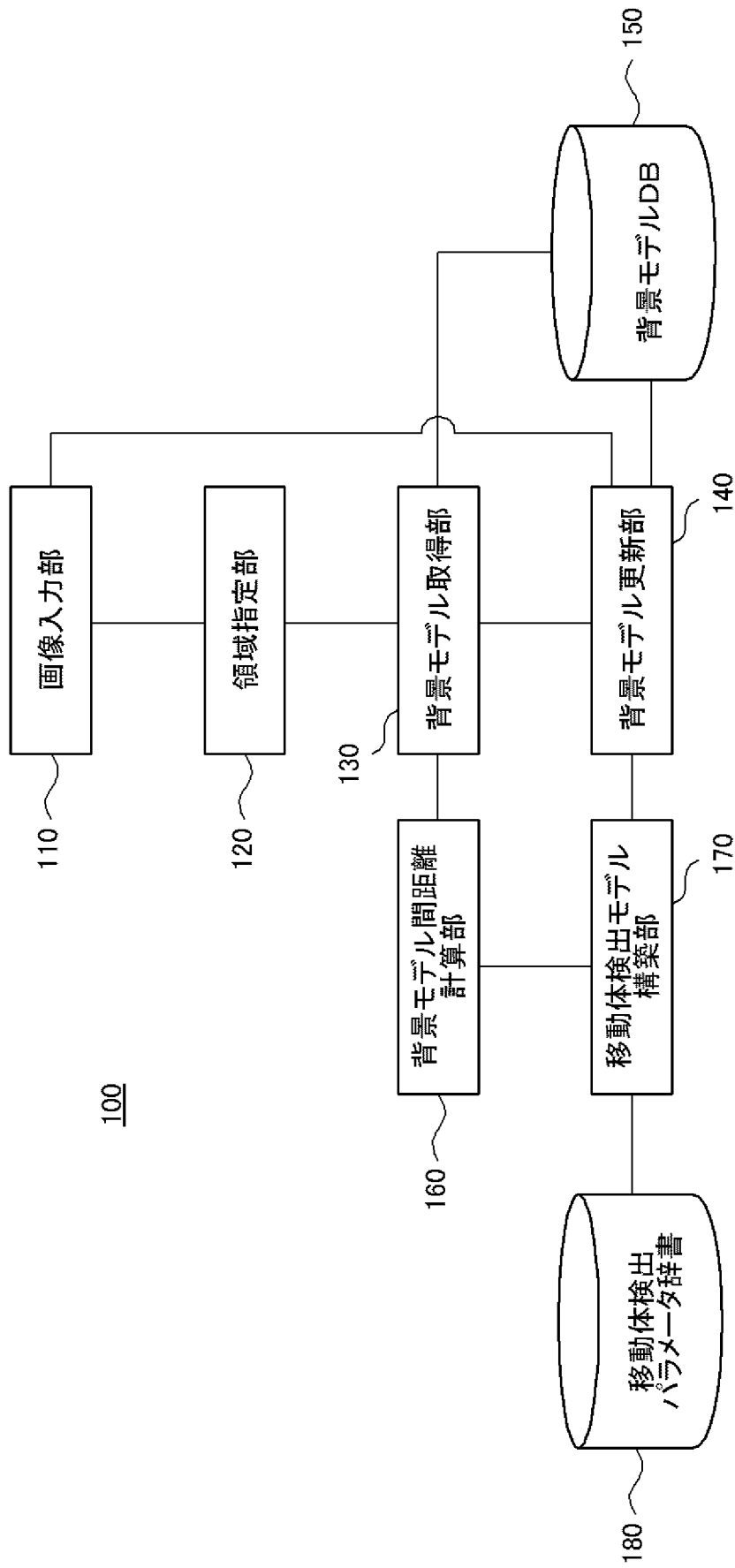
[図2]



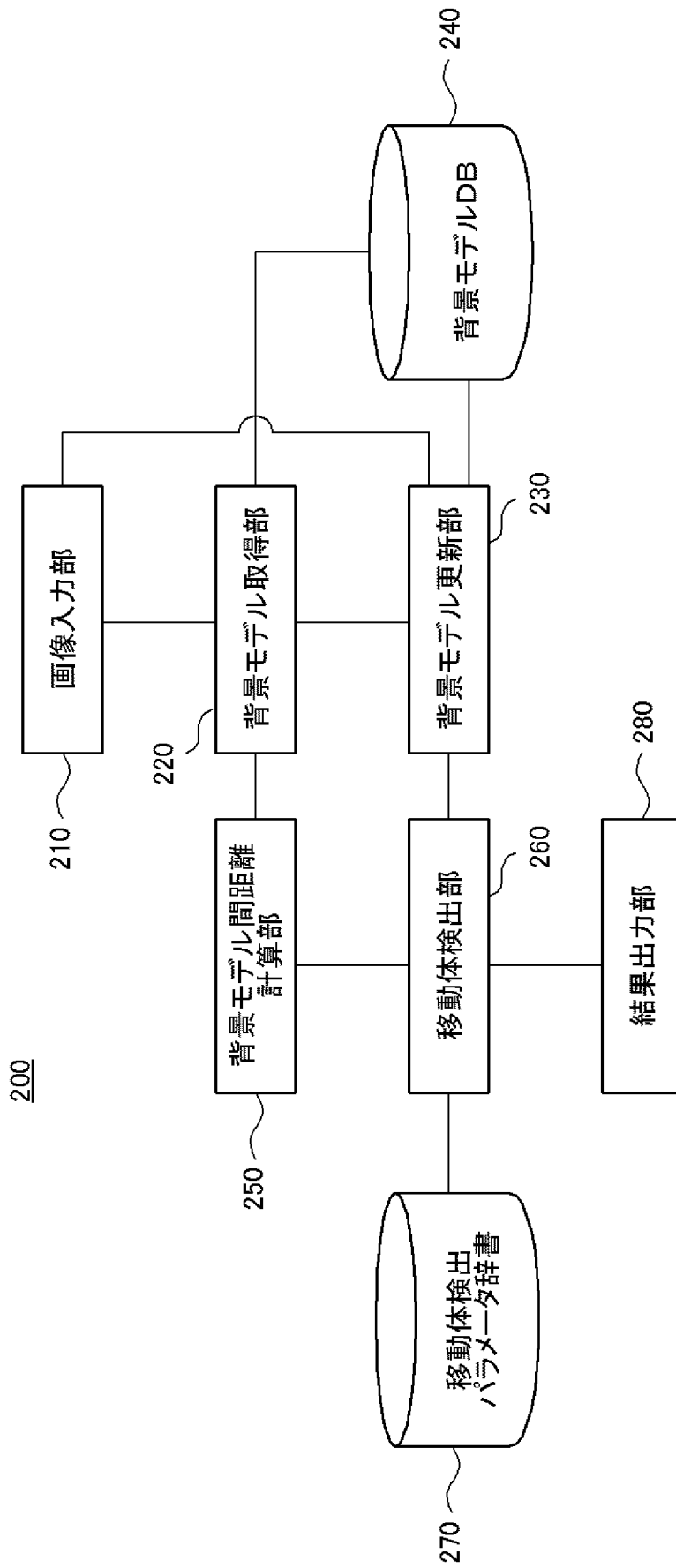
[図3]



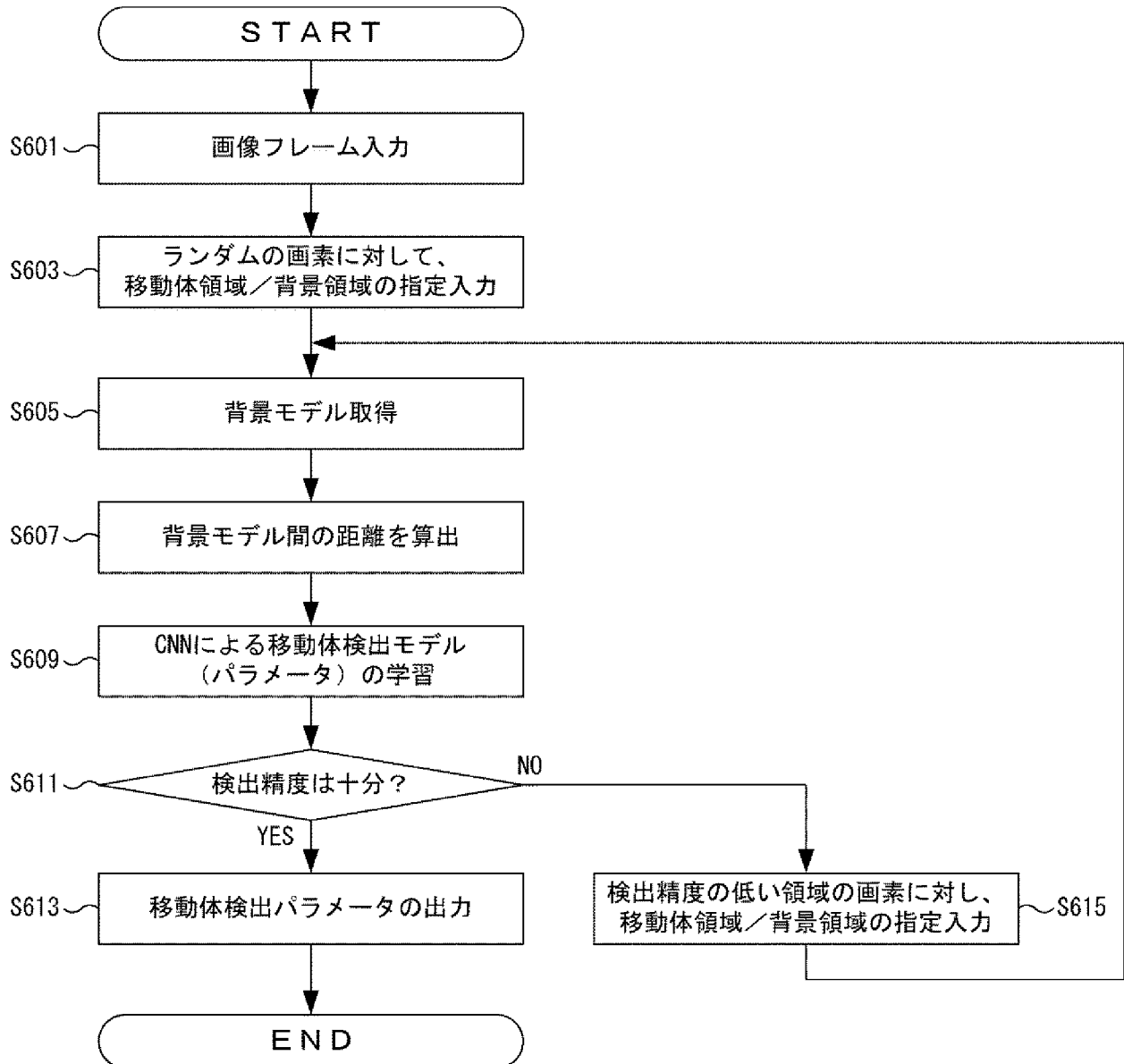
[図4]



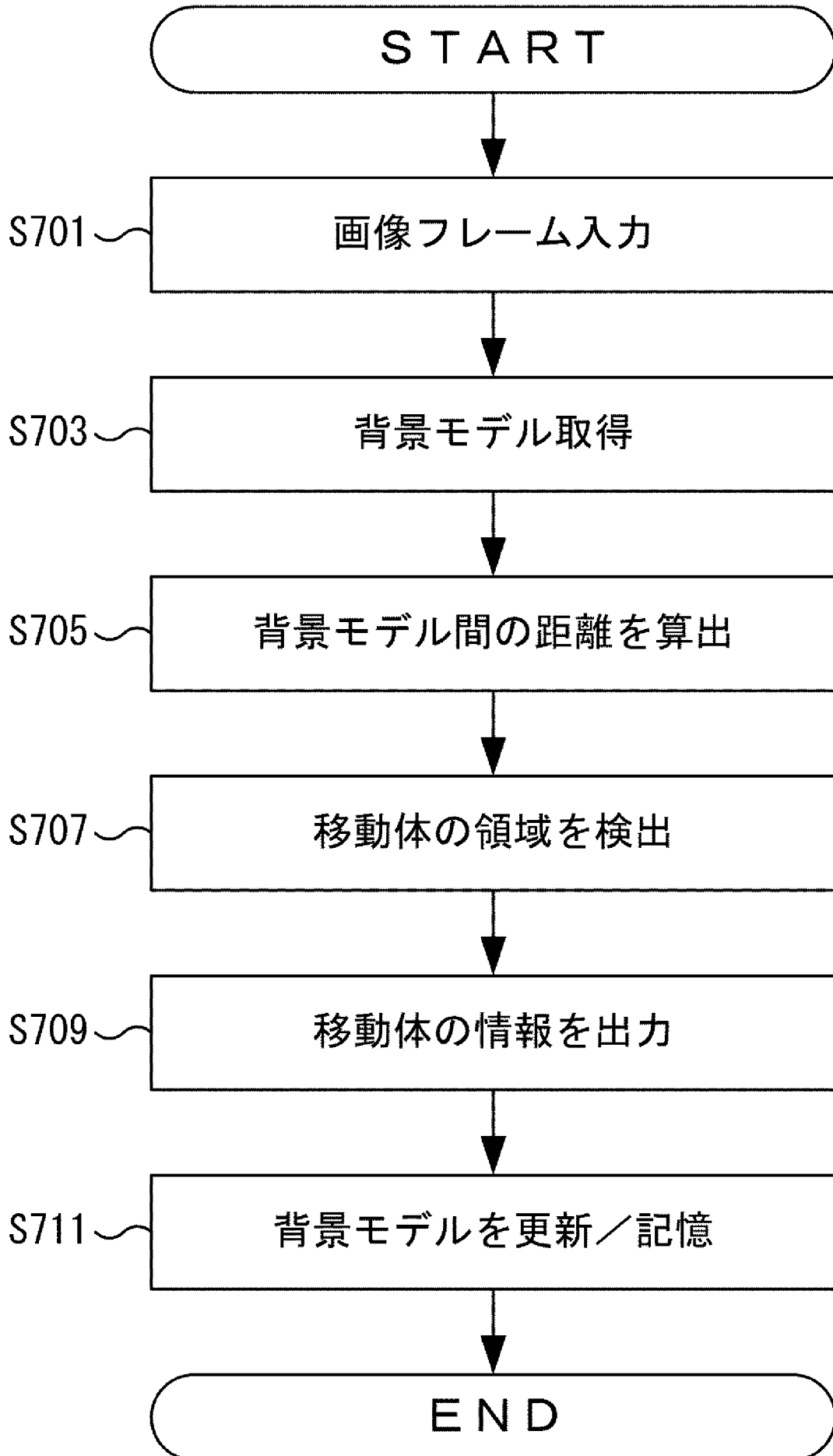
[図5]



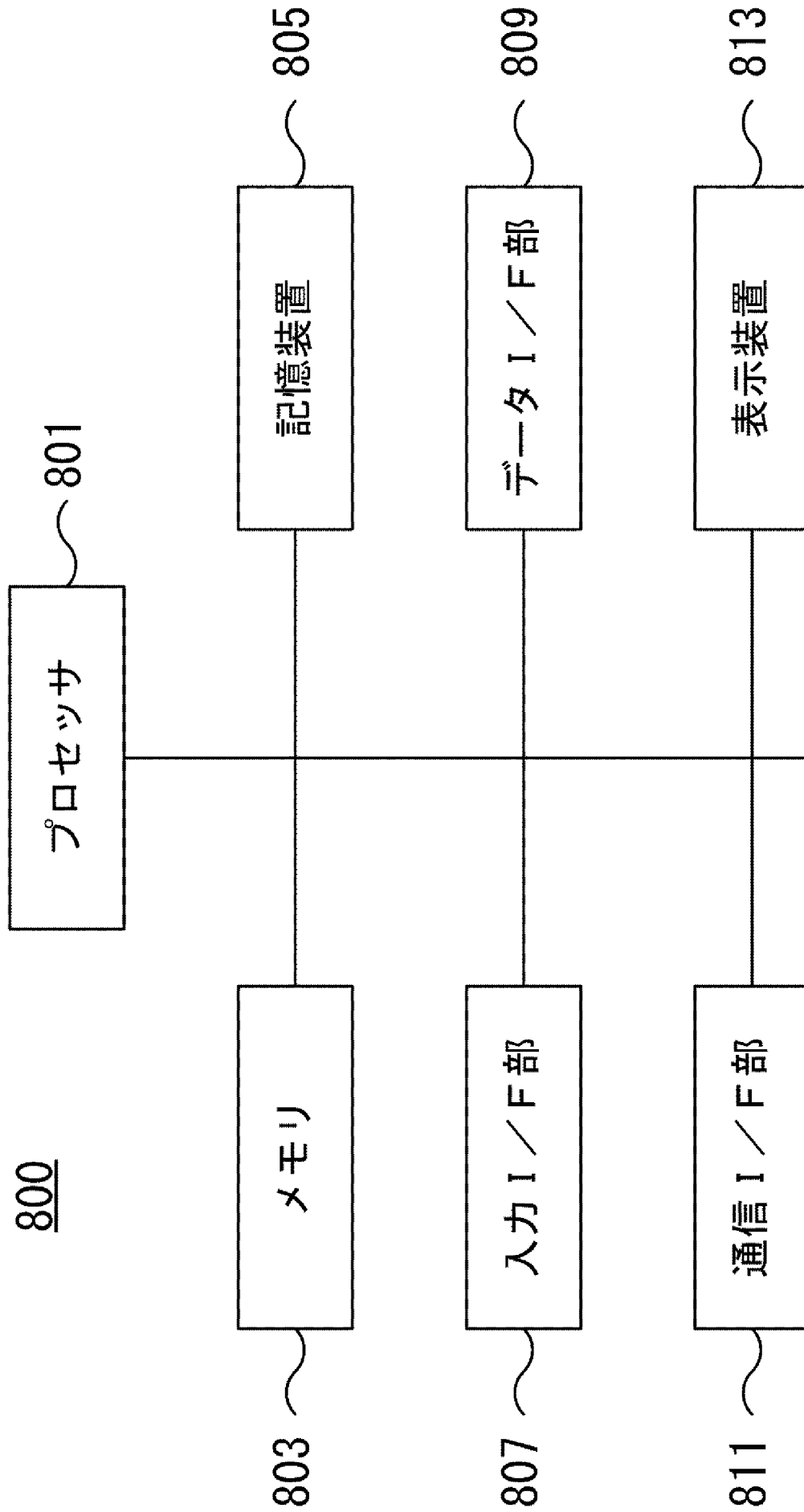
[図6]



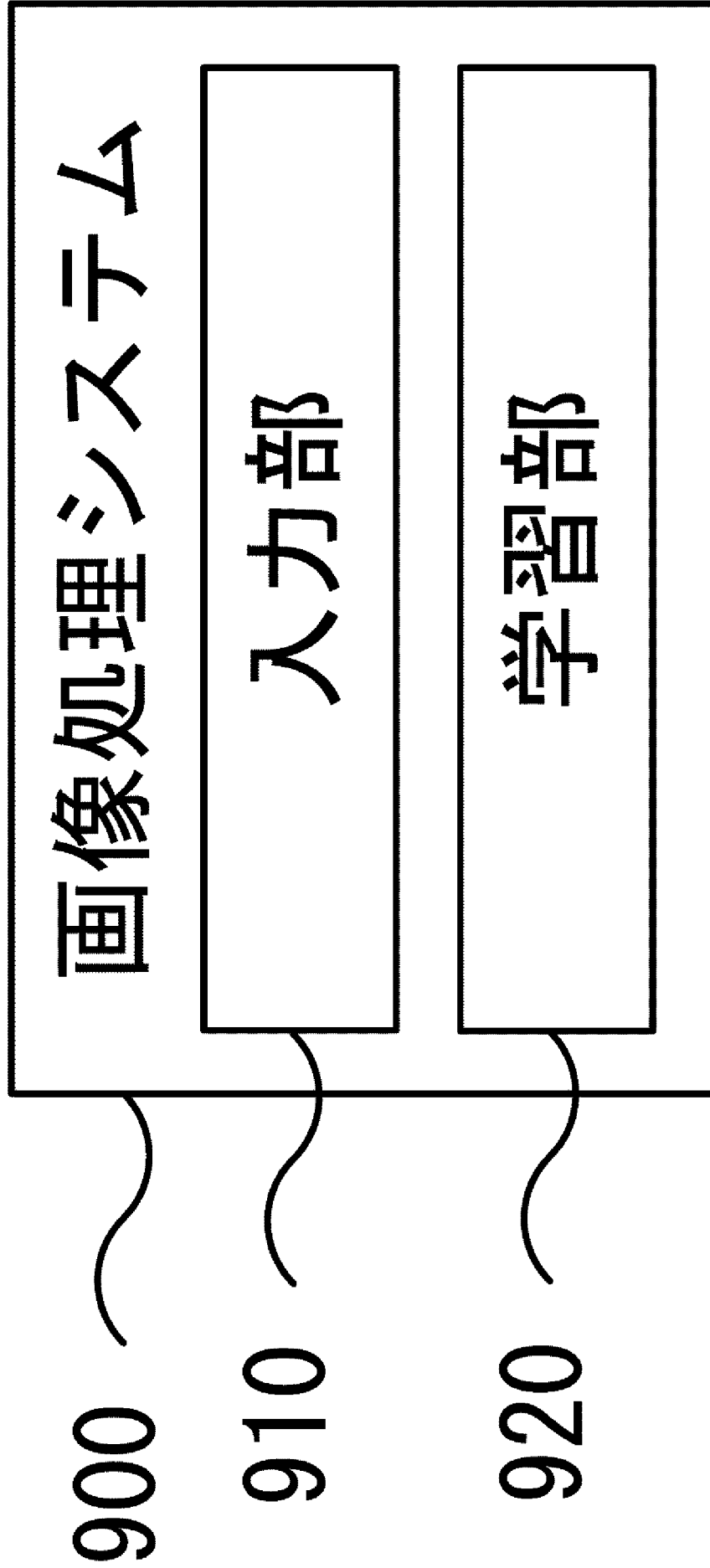
[図7]



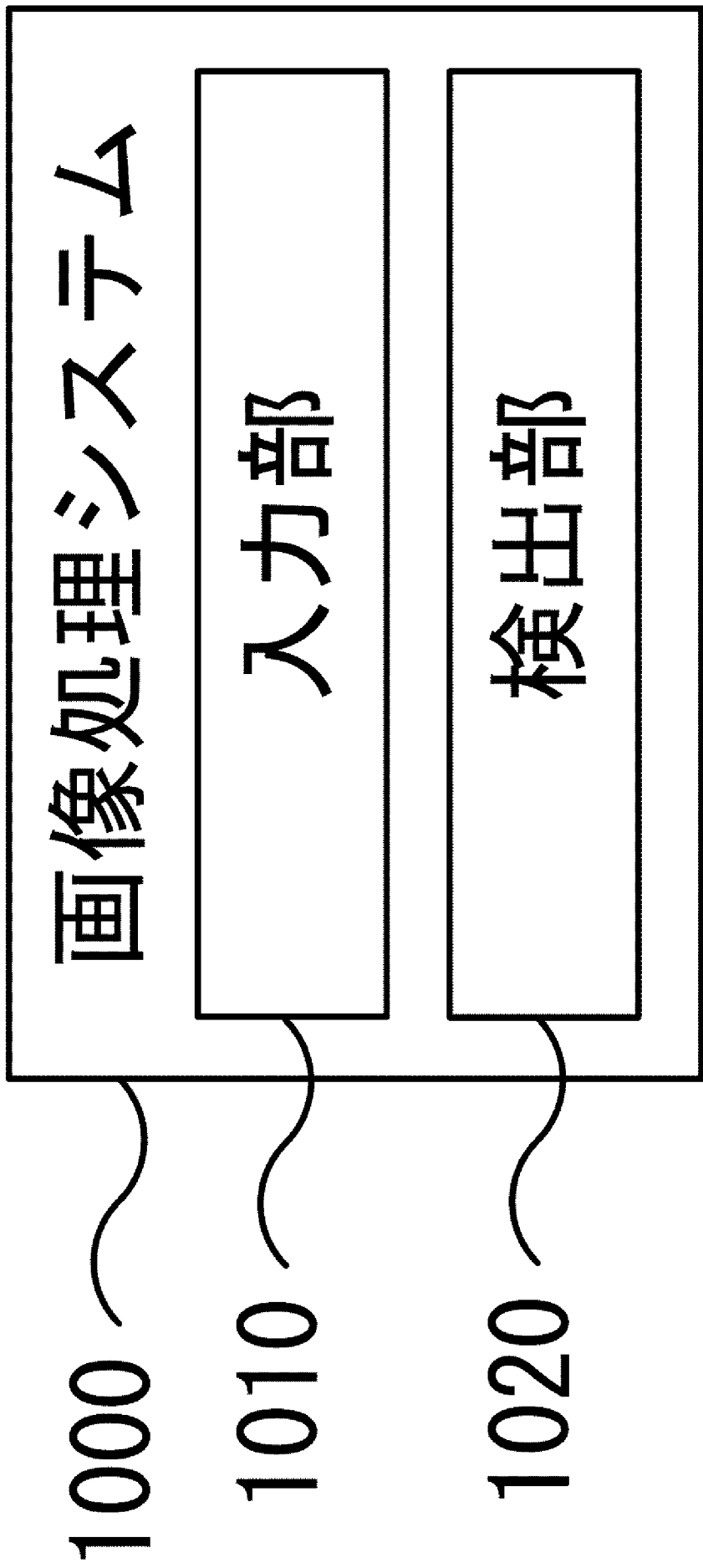
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/002768

A. CLASSIFICATION OF SUBJECT MATTER

G06T7/20(2006.01)i, G06T7/00(2006.01)i, H04N7/18(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06T7/20, G06T7/00, H04N7/18, G08B13/194, G08B25/00

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

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2015
Kokai Jitsuyo Shinan Koho 1971-2015 Toroku Jitsuyo Shinan Koho 1994-2015

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	JP 2012-059224 A (Toshiba Corp.), 22 March 2012 (22.03.2012), entire text; all drawings & US 2013/0050502 A1 & WO 2011/102416 A1 & KR 10-2012-0120499 A & MX 2012009579 A	1-11
A	JP 10-285581 A (Fujitsu Ltd.), 23 October 1998 (23.10.1998), entire text; all drawings & US 6128396 A	1-11
A	JP 2013-125529 A (Panasonic Corp.), 24 June 2013 (24.06.2013), entire text; all drawings & US 2014/0355882 A1 & WO 2013/088665 A1	1-11

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search
03 August 2015 (03.08.15)

Date of mailing of the international search report
11 August 2015 (11.08.15)

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/JP2015/002768

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2008-257693 A (Mitsubishi Electric Research Laboratories, Inc.), 23 October 2008 (23.10.2008), entire text; all drawings & US 2008/0247599 A1 & EP 1978470 A1 & CN 101281596 A	1-11
A	JP 2012-257173 A (Canon Inc.), 27 December 2012 (27.12.2012), entire text; all drawings (Family: none)	1-11
A	JP 2013-065151 A (Sony Corp.), 11 April 2013 (11.04.2013), entire text; all drawings & US 2013/0071032 A1 & EP 2570992 A1 & CN 103198490 A	1-11

A. 発明の属する分野の分類（国際特許分類（IPC））
 Int.Cl. G06T7/20(2006.01)i, G06T7/00(2006.01)i, H04N7/18(2006.01)i

B. 調査を行った分野
 調査を行った最小限資料（国際特許分類（IPC））
 Int.Cl. G06T7/20, G06T7/00, H04N7/18, G08B13/194, G08B25/00

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 1971-2015年
 日本国実用新案登録公報 1996-2015年
 日本国登録実用新案公報 1994-2015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	JP 2012-059224 A（株式会社東芝）2012.03.22, 全文、全図 & US 2013/0050502 A1 & WO 2011/102416 A1 & KR 10-2012-0120499 A & MX 2012009579 A	1-11
A	JP 10-285581 A（富士通株式会社）1998.10.23, 全文、全図 & US 6128396 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	の日の後に公表された文献
「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	「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
「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	「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
「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）	「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
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日 03.08.2015	国際調査報告の発送日 11.08.2015
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） 佐田 宏史 電話番号 03-3581-1101 内線 3531

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	JP 2013-125529 A (パナソニック株式会社) 2013.06.24, 全文、全図 & US 2014/0355882 A1 & WO 2013/088665 A1	1-11
A	JP 2008-257693 A (三菱・エレクトリック・リサーチ・ラボラ トリーズ・インコーポレイテッド) 2008.10.23, 全文、全図 & US 2008/0247599 A1 & EP 1978470 A1 & CN 101281596 A	1-11
A	JP 2012-257173 A (キヤノン株式会社) 2012.12.27, 全文、全図 (フ ァミリーなし)	1-11
A	JP 2013-065151 A (ソニー株式会社) 2013.04.11, 全文、全図 & US 2013/0071032 A1 & EP 2570992 A1 & CN 103198490 A	1-11