

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7327083号

(P7327083)

(45)発行日 令和5年8月16日(2023.8.16)

(24)登録日 令和5年8月7日(2023.8.7)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00

C

G 0 6 T 7/593(2017.01)

G 0 6 T 7/593

G 0 8 G 1/16 (2006.01)

G 0 8 G 1/16

C

請求項の数 10 (全26頁)

(21)出願番号 特願2019-197743(P2019-197743)

(22)出願日 令和1年10月30日(2019.10.30)

(65)公開番号 特開2021-71885(P2021-71885A)

(43)公開日 令和3年5月6日(2021.5.6)

審査請求日 令和4年7月8日(2022.7.8)

(73)特許権者 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1

番1号

(74)代理人 100104190

弁理士 酒井 昭徳

(72)発明者 草薨 一平

神奈川県川崎市中原区上小田中4丁目1

番1号 富士通株式会社内

(72)発明者 北浦 麻子

神奈川県川崎市中原区上小田中4丁目1

番1号 富士通株式会社内

(72)発明者 藤田 卓志

神奈川県川崎市中原区上小田中4丁目1

番1号 富士通株式会社内

最終頁に続く

(54)【発明の名称】 領域切り出し方法および領域切り出しプログラム

(57)【特許請求の範囲】

【請求項1】

画像から対象物の領域を切り出す領域切り出し方法であって、
情報処理装置が、
前記対象物の3次元位置情報を記憶した記憶部を参照し、
前記画像の撮影位置および撮影姿勢に関する情報を用いて、前記3次元位置情報から、
前記画像における前記対象物の領域を算出する、
ことを特徴とする領域切り出し方法。

【請求項2】

前記画像は、撮像装置によって撮影された画像であることを特徴とする請求項1に記載
の領域切り出し方法。

10

【請求項3】

前記画像は、前記撮像装置によって移動しながら撮影された画像であることを特徴とす
る請求項2に記載の領域切り出し方法。

【請求項4】

前記画像の撮影位置および撮影姿勢に関する情報および前記撮像装置のパラメータに関
する情報を用いて、前記画像における前記対象物の領域を算出することを特徴とする請求
項2または3に記載の領域切り出し方法。

【請求項5】

透視変換処理により、前記画像における前記対象物の領域を算出することを特徴とする

20

請求項 1 ~ 4 のいずれか一つに記載の領域切り出し方法。

【請求項 6】

前記対象物が前記画像内に存在するかを判断し、

前記判断の結果、前記対象物が前記画像内に存在する場合に、当該画像における当該対象物の領域を算出することを特徴とする請求項 1 ~ 5 のいずれか一つに記載の領域切り出し方法。

【請求項 7】

前記画像における前記対象物の領域の算出結果に基づいて、当該対象物の画像を切り出すことを特徴とする請求項 1 ~ 6 のいずれか一つに記載の領域切り出し方法。

【請求項 8】

前記画像の撮影位置および撮影姿勢に関する情報は、当該画像を用いた V - S L A M によって取得することを特徴とする請求項 1 ~ 7 のいずれか一つに記載の領域切り出し方法。

【請求項 9】

前記 3 次元位置情報は、三角測量を用いて決定することを特徴とする請求項 1 ~ 8 のいずれか一つに記載の領域切り出し方法。

【請求項 10】

画像から対象物の領域を切り出す領域切り出しプログラムであって、
情報処理装置に、

前記対象物の 3 次元位置情報を記憶した記憶部を参照し、

前記画像の撮影位置および撮影姿勢に関する情報を用いて、前記 3 次元位置情報から、
前記画像における前記対象物の領域を算出する、

処理を実行させることを特徴とする領域切り出しプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像の領域を切り出す領域切り出し方法および領域切り出しプログラムに関する。

【背景技術】

【0002】

従来、たとえば地物などの対象物の画像を、機械学習用の教師データとして作成する場合には、以下の 2 つの作業をおこなう。まず、大量の動画データの中から教師データにしたい対象物が写る画像を、目視により探すという画像探索作業をおこなう。つぎに、画像探索作業において探した画像に対して、対象物が画像内に写り込む画像領域を手作業にて指定するという領域指定作業をおこなう。

【0003】

関連する先行技術としては、対象地物の位置情報を含む地物情報を用いて、画像情報に対する前記対象地物の画像認識処理をおこなう技術がある。また、関連する技術としては、仮想 3 次元空間内に設定された注目点と同一の位置を注視点とする視点に基づいて、仮想 3 次元空間を 2 次元座標系に透視変換する技術がある。

【先行技術文献】

【特許文献】

【0004】

【文献】特開 2008 - 298698 号公報
特開 2009 - 053922 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、従来技術では、対象物の画像探索および領域指定の作業は、上述のように人手によりおこなわなければならない、多くの時間と工数がかかる。そのため、対象物の画像を機械学習用の教師データとして用いるにあたり、十分な量のデータを効率よく収集

10

20

30

40

50

することができないという問題点がある。

【 0 0 0 6 】

一つの側面では、本発明は、画像内に写る対象物の画像領域を自動で特定することを目的とする。

【課題を解決するための手段】

【 0 0 0 7 】

一つの実施態様では、画像から対象物の領域を切り出す領域切り出し方法であって、情報処理装置が、前記対象物の３次元位置情報を記憶した記憶部を参照し、前記画像の撮影位置および姿勢に関する情報を用いて、前記３次元位置情報から、前記画像における前記対象物の領域を算出する、領域切り出し方法が提供される。

10

【発明の効果】

【 0 0 0 8 】

本発明の一側面によれば、画像内に写る対象物の画像領域を自動で特定することができる。

【図面の簡単な説明】

【 0 0 0 9 】

【図 1】図 1 は、領域切り出し方法および領域切り出しプログラムの概要の一例を示す説明図である。

【図 2】図 2 は、領域切り出し方法を実現する情報処理装置の機能的構成の一例を示す説明図である。

20

【図 3】図 3 は、情報処理装置のハードウェア構成の一例を示すブロック図である。

【図 4】図 4 は、領域切り出し方法の一連の処理の手順の一例を示すフローチャートである。

【図 5】図 5 は、各構成部および各情報の内容の一例を示す説明図である。

【図 6】図 6 は、撮影位置・姿勢 A a のデータ構成の一例を示す説明図である。

【図 7】図 7 は、カメラパラメータ A p のデータ構成の一例を示す説明図である。

【図 8】図 8 は、対象物の３次元位置 A o のデータ構成の一例を示す説明図である。

【図 9】図 9 は、三角測量の処理内容を示す説明図である。

【図 10】図 10 は、画像内位置結果 A r のデータ構成の一例を示す説明図である。

【図 11】図 11 は、画像内位置算出部の一連の処理の手順の一例を示すフローチャートである。

30

【図 12】図 12 は、透視変換の処理内容を示す説明図である。

【図 13】図 13 は、位置情報 A i のデータ構成の一例を示す説明図である。

【図 14】図 14 は、領域特定部の一連の処理の手順の一例を示すフローチャートである。

【図 15】図 15 は、領域画像 A b のデータ構成の一例を示す説明図である。

【図 16】図 16 は、切り出された領域画像の一例を示す説明図である。

【図 17】図 17 は、領域出力部の一連の処理の手順の一例を示すフローチャートである。

【発明を実施するための形態】

【 0 0 1 0 】

以下に図面を参照して、本発明にかかる領域切り出し方法および領域切り出しプログラムの実施の形態を詳細に説明する。

40

【 0 0 1 1 】

(実施の形態)

まず、図 1 を用いて、本実施の形態にかかる領域切り出し方法および領域切り出しプログラムの概要について説明する。図 1 は、領域切り出し方法および領域切り出しプログラムの概要の一例を示す説明図である。

【 0 0 1 2 】

図 1 において、たとえば移動体に搭載された撮像装置であるカメラ 100 は、画像を撮影する。撮影した画像には、対象となる地物（標識）が写っている。符号 101 a は、カメラ撮像面（側面）であり、符号 101 b は、カメラ撮像面（正面）である。カメラ撮像

50

面（正面）101bにおいては、移動体が走行中に撮影した画像に、道路の様子が写っていることを示している。カメラ撮像面101bの道路の様子には、移動体の前を走る自動車、対向車のほか、ガードレール、信号機、道路標識、建物、電柱などの地物が含まれている。

【0013】

ここで、切り出しの対象となる地物は、制限速度50km/hの道路標識であり、符号102は、この地物（標識）の3次元位置、すなわち、この地物が実際に存在する座標位置を示している。また、符号103は、対象物存在直線を示している。そして、図1に示すように、対象物存在直線103と、カメラ撮像面（側面）101aとが交わった交点の位置が、地物（対象物）の画像内位置104となる。

10

【0014】

この画像内位置104を算出するために、画像の撮影位置・姿勢とカメラパラメータとを用いて、地物（標識）の3次元位置102を透視変換する。そうすることによって、地物（標識）の3次元位置102から、当該地物が、カメラ撮像面101内のどの位置に写っているかを推定（算出）することができる。

【0015】

このように、地物（対象物）の画像内位置104を推定（算出）することができれば、推定（算出）した画像内位置に基づいて、地物の画像領域を自動切り出しすることができる。符号105は、自動切り出しをした地物画像（制限速度50km/hの道路標識）であり、この自動切り出しをした地物画像は、機械学習用の教師データとして用いることができる。

20

【0016】

ここで、画像は、撮像装置であるカメラ100によって撮影された画像であってもよく、また、他の方法によって取得した画像であってもよい。また、撮像装置によって移動しながら撮影された画像（動画）であってもよい。地物を切り出すためには、地物の3次元位置102に関する情報と、画像の撮影位置・姿勢に関する情報があればよい。さらに、カメラパラメータに関する情報があればよい。

【0017】

ところで、移動体の一例である一般車の車載機や、ドライブレコーダのデータ（映像）を収集（プローブ）し、収集した一般車映像から、道路上や道路周辺の地物や移動物体の情報を取得し、交通情報や路上障害物情報を更新提供したり、自動運转向けの高精度地図を更新したり、自動運转向けに撮影時の周辺状況を解析したりするサービスを実施している。

30

【0018】

移動体が移動中に取得した周辺状況に関するデータを入力とし、移動体の走行経路と周辺環境地図を同時に作成するSLAM（Simultaneous Localization and Mapping）という技術がある。SLAMは、移動中に取得した周辺状況に関する車載データ、たとえばLIDAR（Laser Imaging Detection and Ranging）データなどを入力とし、自車走行経路（自車の位置と姿勢）と周辺環境地図（周辺物の3次元位置マップなど）を同時に作成する技術の総称である。

40

【0019】

また、SLAM技術の中でも、移動体が撮影した映像を入力として、移動体の走行時のカメラ位置・姿勢を推定する技術は、Visual-SLAM（以後、「V-SLAM」という）と呼ばれる。V-SLAMは、車載カメラで撮影した映像を入力とし、撮影した映像に写る被写体の変化を用いて、自車走行経路（自車位置・姿勢）と周辺環境地図（周辺の被写体の画像特徴点群の3次元位置マップ）を推定作成することができる技術であり、一般車の映像から自車位置と姿勢を推定することができる。

【0020】

V-SLAMは、移動するカメラで撮影した、時系列に連続する連続画像（動画）から

50

、それぞれ空間上の点に対応付け可能な画像特徴を多数抽出する。そして、抽出した各画像特徴について、類似性を用いて連続画像間での対応付けをおこない、それらすべてが連続画像間で満たすべき幾何学条件について解を求める。これによって、各画像特徴の空間内での位置と、各画像の撮影位置および姿勢を求めることができる。

【 0 0 2 1 】

そこで、カメラ 1 0 0 によって撮影される画像の撮影位置および姿勢に関する情報は、当該画像を用いた V - S L A M によって取得するようにしてもよい。なお、これら V - S L A M などによるカメラ位置・姿勢の推定は、後述する車載機、情報処理装置 2 0 0 のいずれで実施してもよい。

【 0 0 2 2 】

また、一般的には、車載機は、カメラ 1 0 0 とともに移動体に搭載され、衛星からの G N S S 情報および車載カメラの映像を収集する。移動体は、具体的には、たとえば、一般乗用車やタクシーなどの営業車、二輪車（自動二輪や自転車）、大型車（バスやトラック）などである。また、情報を収集するコネクテッド・カーであってもよい。また、移動体には、水上を移動する船舶や上空を移動する航空機、無人航空機（ドローン）、自動走行ロボットなどであってもよい。

【 0 0 2 3 】

車載機は、カメラ 1 0 0 の撮影映像に関する情報を収集する。たとえば、撮影映像時の移動体およびカメラの位置情報、カメラパラメータを特定するためのカメラ 1 0 0 の情報のいずれかを収集する。この位置情報は、測位情報の一例である G N S S 情報などから得てよい。また、V - S L A M などを用いて撮影位置・姿勢を算出する代わりに、車載機はさらに移動体の情報として、後述するように慣性計測装置 I M U (i n e r t i a l m e a s u r e m e n t u n i t) などの姿勢センサから収集する移動体の姿勢情報などを取得してもよい。また、車載機は、撮影時刻などに関する情報も収集するようにしてもよい。

【 0 0 2 4 】

車載機は、移動体に搭載された専用の装置であってもよく、取り外し可能な機器であってもよい。また、スマートフォンやタブレット端末装置などの通信機能を備えた携帯端末装置を移動体において利用するものであってもよい。また、車載機の各種機能の全部または一部を、移動体が備えている機能を用いて実現するようにしてもよい。

【 0 0 2 5 】

したがって、車載機の『車載』という表現は、移動体に搭載された専用装置という意味には限定されない。車載機は、移動体における情報を収集する機能を持った装置であれば、どのような形態の装置であってもよい。

【 0 0 2 6 】

車載機は、撮影映像に関する情報および G N S S 情報を含む移動体の情報（車載データ）を取得し、取得した車載データを保存する。そして、保存した車載データを、無線通信によって、ネットワーク（後述する、図 3 に示すネットワーク 3 1 0 ）を介して、サーバ（後述する、図 2 に示す情報処理装置 2 0 0 ）へ送信するようにしてもよい。

【 0 0 2 7 】

また、車載機は、通信手段を備えていなくてもよい。すなわち、車載機は、サーバとは、ネットワーク 3 1 0 を介して接続されていなくてもよい。その場合は、車載機に蓄積されたデータは、オフラインで（たとえば、記録メディアを介して人手などにより）、サーバに入力することができる。

【 0 0 2 8 】

このように、車載機によって、カメラ 1 0 0 が撮影した画像からの地物画像の領域を切り出す処理の全部または一部をおこなうようにしてもよい。

【 0 0 2 9 】

（情報処理装置の機能的構成例）

図 2 は、領域切り出し方法を実現する情報処理装置の機能的構成の一例を示す説明図で

10

20

30

40

50

ある。図 2 において、情報処理装置 200 は、具体的には、たとえば、1 台ないし複数台のサーバなどのコンピュータによって、その機能を実現する。複数台のサーバは、ネットワーク（後述する図 3 におけるネットワーク 310）によって接続される。また、情報処理装置 200 は、図示は省略するが、クラウドコンピューティングシステムによって、その機能を実現するようにしてもよい。

【0030】

情報処理装置（サーバ）200 は、画像入力部 201、画像内位置算出部 202、領域特定部 203、領域出力部 204 を有する。これらの構成部 201 ~ 204 によって、情報処理装置 200 の制御部を構成することができる。また、情報処理装置 200 は、入力される、画像（画像 A）211、撮影位置・姿勢（撮影位置・姿勢 Aa）212、カメラパラメータ（カメラパラメータ Ap）213、対象物の 3 次元位置（対象物の 3 次元位置 Ao）214 の各情報を持つ。

10

【0031】

また、情報処理装置 200 は、記憶部 205 を有していてもよい。記憶部 205 は、情報処理装置 200 が有する代わりに、図示を省略する他の情報処理装置が有していてもよい。そして、ネットワーク（ネットワーク 310）を介して当該他の情報処理装置から取得するようにしてもよい。

【0032】

画像入力部 201 は、画像 211 から対象物（地物画像 105）の領域を切り出すにあたり、画像 211 の入力を受け付ける。画像 211 の入力は、すでに撮影され所定の記憶領域に記憶された画像データを入力するようにしてもよく、また、撮影装置（図 1 に示したカメラ 100）によって撮影された映像データの中から所定の画像を抽出するようにしてもよい。

20

【0033】

画像入力部 201 は、入力の対象となる画像 211 に対応する、当該画像 211 の撮影位置および姿勢に関する情報（撮影位置・姿勢 212）、当該画像 211 を撮影した撮像装置のパラメータに関する情報（カメラパラメータ 213）を取得する。撮影位置・姿勢 212 は、前述した車載機で収集したセンサデータ、または、車載機または情報処理装置 200 で V-SLAM などの任意の既知の処理で求めたデータである。なお、カメラパラメータ 213 の情報は、処理対象のカメラがあらかじめ限定されている場合などに、画像 211 をみて取得するのではなく、あらかじめ取得しておくようにしておいてもよい。

30

【0034】

画像内位置算出部 202・領域特定部 203 は、対象物の 3 次元位置 214 に関する 3 次元位置情報を記憶した記憶部 205 を参照し、画像 211 の撮影位置および姿勢 212 を用いて、3 次元位置 214 に関する 3 次元位置情報から、画像 211 における対象物の領域を算出する。また、対象物の 3 次元位置 214 に関する 3 次元位置情報を記憶した記憶部 205 を参照し、画像 211 の撮影位置および姿勢 212 だけでなく、撮像装置のカメラパラメータ 213 も用いて、3 次元位置 214 に関する 3 次元位置情報から、画像 211 における対象物の領域を算出するようにしてもよい。

【0035】

たとえば、画像内位置算出部 202 が、対象物が画像 211 内に存在するかを判断する。そして、領域特定部 203 が、画像内位置算出部 202 による判断の結果、対象物が画像 211 内に存在する場合に、当該画像 211 における当該対象物の領域を算出するようにしてもよい。

40

【0036】

領域出力部 204 は、領域特定部 203 による画像 211 における対象物の領域の算出結果（後述する画像内位置結果 Ar511）に基づいて、当該対象物の画像の切り出しをおこなう。

【0037】

記憶部 205 には、複数種類の各種対象物（地物）の 3 次元位置 214 を格納している

50

。したがって、対象物の種類などを指定することで、指定された対象物の３次元位置 2 1 4 を記憶部 2 0 5 から抽出することができる。

【 0 0 3 8 】

図 2 においては、情報処理装置（サーバ）2 0 0 が、画像入力部 2 0 1、画像内位置算出部 2 0 2、領域特定部 2 0 3、領域出力部 2 0 4、記憶部 2 0 5 と、を有する構成とした。図示は省略するが、これらの各機能部の少なくとも一つを、情報処理装置 2 0 0 に加えて、あるいは、情報処理装置 2 0 0 に代えて、上述のように、移動体が備える車載機が有するようにしてもよい。

【 0 0 3 9 】

車載機が、各機能部 2 0 1、2 0 2、2 0 3、2 0 4、2 0 5 の少なくとも一つを有する場合は、情報処理装置 2 0 0 が実施する処理の内容と同じであってもよい。このように構成することによって、移動体に搭載された撮像装置が撮影した画像に基づいて、移動体に搭載された車載機が機械学習用の教師データを出力することができる。

10

【 0 0 4 0 】

（情報処理装置のハードウェア構成例）

図 3 は、情報処理装置のハードウェア構成の一例を示すブロック図である。情報処理装置 2 0 0 の一例であるサーバは、CPU（Central Processing Unit）3 0 1 と、メモリ 3 0 2 と、ネットワーク I / F（Interface）3 0 3 と、記録媒体 I / F 3 0 4 と、記録媒体 3 0 5 と、を有する。また、各構成部は、バス 3 0 0 によってそれぞれ接続される。

20

【 0 0 4 1 】

ここで、CPU 3 0 1 は、情報処理装置 2 0 0 の全体の制御を司る。メモリ 3 0 2 は、たとえば、ROM（Read Only Memory）、RAM（Random Access Memory）およびフラッシュ ROMなどを有する。具体的には、たとえば、フラッシュ ROMや ROMが各種プログラムを記憶し、RAMが CPU 3 0 1 のワークエリアとして使用される。メモリ 3 0 2 に記憶されるプログラムは、CPU 3 0 1 にロードされることで、コーディングされている処理を CPU 3 0 1 に実行させる。

【 0 0 4 2 】

ネットワーク I / F 3 0 3 は、通信回線を通じてネットワーク 3 1 0 に接続され、ネットワーク 3 1 0 を介して他の装置（たとえば、図示を省略する車載機や、他のサーバやシステム）に接続される。そして、ネットワーク I / F 3 0 3 は、ネットワーク 3 1 0 と自装置内部とのインターフェースを司り、他の装置からのデータの入出力を制御する。ネットワーク I / F 3 0 3 には、たとえば、モデムや LAN アダプタなどを採用することができる。

30

【 0 0 4 3 】

記録媒体 I / F 3 0 4 は、CPU 3 0 1 の制御にしたがって記録媒体 3 0 5 に対するデータのリード/ライトを制御する。記録媒体 3 0 5 は、記録媒体 I / F 3 0 4 の制御で書き込まれたデータを記憶する。記録媒体 3 0 5 としては、たとえば、磁気ディスク、光ディスクなどが挙げられる。

【 0 0 4 4 】

なお、情報処理装置 2 0 0 は、上述した構成部のほかに、たとえば、SSD（Solid State Drive）、キーボード、ポインティングデバイス、ディスプレイなどを有していてもよい。

40

【 0 0 4 5 】

また、図示は省略するが、車載機が、撮像装置のほか、情報処理装置 2 0 0 と同様のハードウェア構成を備えていてもよい。

【 0 0 4 6 】

（領域切り出し方法の一連の処理手順）

図 4 は、領域切り出し方法の一連の処理の手順の一例を示すフローチャートである。図 4 のフローチャートにおいて、情報処理装置 2 0 0 は、任意の画像 2 1 1 を読み込む（ス

50

テップ S 4 0 1)。このステップの処理は、たとえば、図 2 に示した画像入力部 2 0 1 によっておこなわれる。

【 0 0 4 7 】

つぎに、対象物の地物特徴点群の 3 次元位置 2 1 4 を、撮影位置・姿勢 2 1 2 を用いて透視変換をおこない、地物特徴点の 2 次元位置を推定（算出）する（ステップ S 4 0 2 ）。地物特徴点の 2 次元位置を推定するにあたり、カメラパラメータ 2 1 3 もあわせて用いるようにしてもよい。

【 0 0 4 8 】

そして、推定（算出）した各地物特徴点の 2 次元位置が、画像 2 1 1 の画像領域内、すなわち、カメラ撮像面 1 0 1 内にあるか否かを判断し、これら 2 次元位置が画像領域内にあるか否かをを用いて、対象物が画像 2 1 1 に写り込んでいる量を算出する（ステップ S 4 0 3 ）。

10

【 0 0 4 9 】

つぎに、算出した地物の写り込み量が、規定値、すなわち、システムなどで決定した条件に合致するかを判断する（ステップ S 4 0 4 ）。たとえば、写り込み量が 0（地物全体が画像領域外）～ 1（地物全体が画像領域内）の値の場合、写り込み量が規定値以上であるか否かを条件とし、写り込み量が規定値以上の場合に合致したとみなす。規定値が 1 に近いほど、対象物全体が画像に写っている可能性の高い画像のみを画像領域切り出しに使うことになる。完全に対象物が画像に写っている画像のみを使いたい場合は、条件の規定値を 1 に設定しておくといよい。また、対象物の約半分以上が写っている画像のみを使いたい場合は、規定値を 0 . 5 に設定しておくといよい。

20

【 0 0 5 0 】

これらのステップ（ステップ S 4 0 2 ～ S 4 0 4 ）の処理は、たとえば、図 2 に示した画像内位置算出部 2 0 2 によっておこなわれる。

【 0 0 5 1 】

ステップ S 4 0 4 において、推定（算出）した地物の写り込み量が条件に合致しない場合、たとえば写り込み量が規定値未満の場合（ステップ S 4 0 4 : N o ）は、当該画像 2 1 1 については何もせずに、一連の処理を終了する。これにより、画像 2 1 1 のうち、地物（対象物）が存在しない画像を排除し、地物（対象物）が存在する画像のみを抽出することができる。したがって、画像探索処理をより効率的かつ迅速におこなうことができる。

30

【 0 0 5 2 】

一方、ステップ S 4 0 4 において、推定した地物の写り込み量が条件に合致する場合、たとえば写り込み量が規定値以上の場合（ステップ S 4 0 4 : Y e s ）は、画像領域内にある 2 次元位置の画素群を含む画像領域を算出する（ステップ S 4 0 5 ）。このステップの処理は、たとえば、図 2 に示した領域特定部 2 0 3 によっておこなわれる。

【 0 0 5 3 】

その後、画像 2 1 1 から、ステップ S 4 0 4 において算出された画像領域を切り出す（ステップ S 4 0 6 ）。このステップの処理は、たとえば、図 2 に示した領域出力部 2 0 4 によっておこなわれる。そして、一連の処理を終了する。

【 0 0 5 4 】

40

これらの一連の処理を、画像ごとに行う。このようにして、情報処理装置 2 0 0 は、任意の画像 2 1 1 から対象物（地物）の画像領域の切り出しを、対象物がどれだけ画像に写っているかを考慮しながら、自動でおこなうことができる。切り出された対象物（地物）の画像領域にかかる情報は、教師データとして保存することができる。以下、これらの一連の処理を実行する各構成部 2 0 1 ～ 2 0 5 の具体的な内容について説明する。

【 0 0 5 5 】

（各構成部 2 0 1 ～ 2 0 5 の具体的な内容）

図 5 は、各構成部および各情報の内容の一例を示す説明図である。図 5 においては、1 つの画像（画像 A 2 1 1 ）に対し、当該画像を撮影した撮影カメラの撮影位置・姿勢情報と対象物の 3 次元位置に関する 3 次元位置情報を用いて、画像上の対象物が写る領域を推

50

定し、領域切り出しをおこなう例を示す。

【 0 0 5 6 】

図 5 において、各構成部は、画像入力部 2 0 1、画像内位置算出部 2 0 2、領域特定部 2 0 3、領域出力部 2 0 4、記憶部 2 0 5 からなる。また、各情報は、画像 A 2 1 1、撮影位置・姿勢 A a 2 1 2、カメラパラメータ A p 2 1 3、対象物の 3 次元位置 A o 2 1 4、画像内位置結果 A r 5 1 1、位置情報 A i 5 1 2、領域画像 A b 5 1 3 からなる。

【 0 0 5 7 】

(画像入力部 2 0 1 の処理内容)

図 5 において、画像入力部 2 0 1 は、任意の画像 A 2 1 1 を画像内位置算出部 2 0 2 に入力する。その際、画像 A 2 1 1 の撮影位置・姿勢 A a 2 1 2 および画像 A 2 1 1 を撮影したカメラのカメラパラメータ A p 2 1 3 が、画像内位置算出部 2 0 2 に入力される。画像入力部 2 0 1 は、具体的には、たとえば、図 3 に示した情報処理装置 2 0 0 において、メモリ 3 0 2 に記憶されたプログラムを C P U 3 0 1 が実行することによって、あるいは、ネットワーク I / F 3 0 3、記録媒体 I / F 3 0 4 によって、その機能を実現することができる。

10

【 0 0 5 8 】

(撮影位置・姿勢 A a 2 1 2 の内容)

撮影位置・姿勢 A a 2 1 2 は、画像 A 2 1 1 に対応する情報であり、たとえば、実座標上の 3 軸の位置情報と 3 次元ベクトル方向情報を持っている。図 6 は、撮影位置・姿勢 A a のデータ構成の一例を示す説明図である。図 6 に示すように、撮影位置・姿勢 A a 2 1 2 は、「動画 I D」、「フレーム番号」、「位置」、「姿勢」を含む各種情報を有している。

20

【 0 0 5 9 】

ここで、「動画 I D」は、画像 A 2 1 1 が含まれる動画を識別する一意の識別情報であり、「フレーム番号」は、当該動画の中の画像 A 2 1 1 が存在するフレームの番号に関する情報である。この「動画 I D」および「フレーム番号」によって、画像 A 2 1 1 を特定することができる。

【 0 0 6 0 】

「位置」は、撮影位置に関する実座標（世界座標）上の画像 A 2 1 1 の 3 軸の位置情報である。3 軸の位置情報は、たとえば、緯度（ P o s X）・経度（ P o s Y）・高さ（ P o s Z）によって示すことができる。「姿勢」は、撮影姿勢に関する 3 次元ベクトル方向に関する情報である。撮影姿勢に関する 3 次元ベクトル方向情報は、たとえば、回転を表す方法であるロール・ピッチ・ヨー（ R o l l・P i t c h・Y a w）によって示すことができる。

30

【 0 0 6 1 】

これらの画像 A 2 1 1 の 3 軸の位置情報および 3 次元ベクトル方向情報は、センサ（ G P S などの測位センサと I M U などの姿勢センサなど）から取得することができる。I M U は、運動を司る 3 軸の角度（または角速度）と加速度を検出する。I M U は、I N U（ i n e r t i a l n a v i g a t i o n u n i t）、I G U（ i n e r t i a l g u i d a n c e u n i t）、I R U（ i n e r t i a l r e f e r e n c e u n i t）とも呼ばれる。

40

【 0 0 6 2 】

基本的には、3 軸のジャイロと 3 方向の加速度計によって、3 次元の角速度と加速度が求められる。信頼性向上のために圧力計、流量計、 G P S など別種類のセンサが搭載されるようにしてもよい。I M U は、通常は、搭載する移動体の重心に置くようにするとよい。

【 0 0 6 3 】

また、任意の画像処理手法、たとえば、上述したように、 V - S L A M によって取得するようにしてもよい。これにより、各画像特徴の空間内での位置（画像特徴点群マップ）と、各画像の撮影位置および姿勢を求めることができる。 V - S L A M とともに G P S の位置座標を用いることで、実座標系での撮影位置・姿勢を算出し、算出した結果を撮影位

50

置・姿勢 A a 2 1 2 として用いるようにしてもよい。

【 0 0 6 4 】

(カメラパラメータ A p 2 1 3 の内容)

カメラパラメータ A p 2 1 3 は、画像 A 2 1 1 を撮影したカメラの固有内部パラメータを表す情報である。図 7 は、カメラパラメータ A p のデータ構成の一例を示す説明図である。図 7 に示すように、カメラパラメータ A p 2 1 3 は、「焦点距離」、「光学中心」、「歪み補正係数」を含む各種情報を有している。

【 0 0 6 5 】

「焦点距離」は、カメラの焦点距離に関する情報であり、 f_x および f_y によって示すことができる。また、「光学中心」は、レンズ面の光軸が通る位置に関する情報であり、中心座標 (c_x 、 c_y) によって示すことができる。また、「歪み補正係数」は、カメラレンズの特性によって生じる歪みを補正する係数に関する情報であり、レンズの半径方向歪み係数 (k_1 、 k_2) およびレンズの円周方向歪み係数 ($p_1 \sim p_3$) によって示すことができる。

【 0 0 6 6 】

(記憶部 2 0 5 の内容・対象物の 3 次元位置 A o 2 1 4 の内容)

また、図 5 に示すように、対象物の 3 次元位置 A o 2 1 4 が、画像内位置算出部 2 0 2 に入力される。対象物の 3 次元位置 A o 2 1 4 は、記憶部 2 0 5 に記憶されていてもよい。

【 0 0 6 7 】

対象物の 3 次元位置 A o 2 1 4 は、抽出したい対象物の位置を 1 点以上の 3 次元代表位置で表す情報である。図 8 は、対象物の 3 次元位置 A o のデータ構成の一例を示す説明図である。図 8 に示すように、「対象物 ID」は、対象物を識別する一意の識別情報であり、「対象物の種類」は、対象物の種類を示す情報である。

【 0 0 6 8 】

「対象物を表す点の数」は、対象物を表す点の数を示しており、たとえば、対象物の中心 1 点であってもよいし、対象物の外接する 4 点でもよい。また、「直角平面系番号」を有していてもよい。「実座標 3 次元」は、「対象物を表す点の数」が『 4 』なので、 $P o s [0] \sim P o s [3]$ の 4 つの座標が、それぞれ x 、 y 、 z の 3 次元で示される。

【 0 0 6 9 】

対象物を表す点を、対象物の外接する座標 4 点としたが、それには限定されない。たとえば、対象物の中心 1 点であってもよいし、4 点よりも少なくてもよいし、4 点よりも多くてもよい。

【 0 0 7 0 】

対象物の 3 次元位置 A o 2 1 4 における対象物の座標点は、図示を省略する地物 D B から取得するとよい。また、対象物の 3 次元位置 A o 2 1 4 における対象物の座標点は、三角測量によって取得するようにしてもよい。三角測量は、カメラの撮影位置・姿勢と、複数画像に写る同一対象物の画像内位置とから、対象物の 3 次元位置を推定する手法である。

【 0 0 7 1 】

図 9 は、三角測量の処理内容を示す説明図である。図 9 に示すように、対象物は、カメラ位置と対象物の画像内位置とを結ぶ直線 (対象物存在直線) 上にあるため、カメラ 1 による画像 1 の対象物存在直線 1 と、カメラ 2 による画像 2 の対象物存在直線 2 との交点が、対象物の 3 次元位置となる。これを、対象物の 3 次元位置 A o 2 1 4 とすることができる。

【 0 0 7 2 】

この三角測量において用いる撮影位置・姿勢に関する情報は、任意の既知の方法によって取得することができる。たとえば、撮影位置・姿勢 A a 2 1 2 と同様に、センサによって取得するようにしてもよいし、V - S L A M などの画像処理によって取得するようにしてもよい。

【 0 0 7 3 】

(画像内位置算出部 2 0 2 の処理内容・画像内位置結果 A r 5 1 1 の内容)

10

20

30

40

50

つぎに、画像内位置算出部 202 の処理の内容について説明する。図 5 に示すように、画像内位置算出部 202 は、これらの、画像 A 211 の撮影位置・姿勢 A a 212、画像 A 211 を撮影したカメラのカメラパラメータ A p 213、および、対象物の 3 次元位置 A o 214 に基づいて、対象物の各 3 次元座標点に対応する画像内位置（2 次元位置）を算出する。続けて、画像内位置算出部 202 は、各 2 次元位置が画像 A 211 内に含まれるかを判定し、その判定結果から対象物が画像 A 211 内に写り込む量を算出し、画像内位置結果 A r 511 として出力する。

【0074】

画像内位置算出部 202 は、具体的には、たとえば、図 3 に示した情報処理装置 200 において、メモリ 302 に記憶されたプログラムを C P U 301 が実行することによって、その機能を実現することができる。

10

【0075】

画像内位置結果 A r 511 は、対象物が画像 A 211 内に写り込む状態に関する情報であり、画像 A 211 についての画像内位置算出部 202 による算出結果に関する情報を持っている。図 10 は、画像内位置結果 A r のデータ構成の一例を示す説明図である。図 10 に示すように、画像内位置結果 A r 511 は、「動画 I D」、「フレーム番号」、「対象物 I D」、「対象物の種類」、「画像内位置結果」を含む各種情報を有している。また、画像内位置 A r 511 はさらに、「対象物を表す点の数」、「各点の画像内判定結果」、「各点の 2 次元位置」を有してもよい。

【0076】

20

ここで、「動画 I D」は、画像 A 211 が含まれる動画を識別する一意の識別情報であり、「フレーム番号」は、当該動画の中の画像 A 211 が存在するフレームの番号に関する情報である。これらは、図 6 に示した撮影位置・姿勢 A a 212 の「動画 I D」および「フレーム番号」である。「対象物 I D」は、対象物を識別する一意の識別情報であり、「対象物の種類」は、対象物の種類を示す情報であり、これらは、図 8 に示した対象物の 3 次元位置 A o 214 の「対象物 I D」および「対象物の種類」である。

【0077】

「画像内位置結果」は、対象物が画像 A 211 内に写り込む量に関する情報である。たとえば、対象物の 2 次元位置群が画像 A 211 内に存在する割合であり、2 次元位置群すべてが画像 A 211 内に存在する場合は、「画像内位置結果」が『1』となる。対象物の 2 次元位置群の一部が画像 A 211 内に存在する場合は、たとえば対象物を表す 4 点の 2 次元位置群のうち、3 つが画像 A 211 内に存在する場合は、「画像内位置結果」が『0.75』となる。一方、対象物の 2 次元位置群がすべて画像 A 211 外に存在する（画像 A 211 内に存在しない）場合は、「画像内位置結果」が『0』となる。

30

【0078】

「対象物を表す点の数」は、図 8 の対象物の 3 次元位置 A o 215 の「対象物を表す点の数」と同じであり、対象物の地物特徴点数である。

【0079】

「各点の画像内判定結果」は、対象物の各地物特徴点の 3 次元座標点、すなわち、図 8 の実座標 3 次元位置を保持する各代表点に相当する対象物の 2 次元位置群のそれぞれが、画像 A 211 内に存在するか否かの情報であり、「対象物を表す点の数」だけ保持し、図 8 の実座標 3 次元位置と同じ順番で保持する。2 次元位置が画像 A 211 内に存在する場合は『1』となり、存在しない場合は『0』となる。「各点の画像内判定結果」は、「対象物を表す点の数」が『4』なので、R e s u l t P o s [0] ~ R e s u l t P o s P o s [3] の 4 つの『0』または『1』の値で示される。

40

【0080】

「各点の 2 次元位置」は、対象物の各地物特徴点の画像内の 2 次元位置座標値であり、「対象物を表す点の数」だけ保持し、「各点の画像内判定結果」や図 8 の実座標 3 次元位置と同じ順番とする。「各点の 2 次元位置」は、「対象物を表す点の数」が『4』なので、2 D P o s [0] ~ 2 D P o s [3] の 4 つの座標が、それぞれ x、y の 2 次元で示さ

50

れる。

【 0 0 8 1 】

なお、「画像内位置結果」は、対象物の 2 次元位置群が画像内に存在する割合ではなく、対象物が存在しない『 0 』、存在する『 1 』、の 2 つの選択肢のどちらかに判定した結果であってもよい。たとえば、対象物の 2 次元位置群の一部が画像内に存在する場合でも、対象物の 2 次元位置群全てが存在しない場合と同様と見なして、「画像内位置結果」を『 0 』としてもよい。この場合、各 2 次元位置群が画像内か否かの「各点の画像内判定結果」は、画像内位置結果 A r 5 1 1 として保持せず、省略するようにしてもよい。

【 0 0 8 2 】

また、「画像内位置結果」は、厳密な画像内に存在する 2 次元位置群の割合ではなく、任意の階層値であってもよい。たとえば、割合の大小に従って 4 段階の値（0 ~ 3、割合が 0 ~ 0 . 2 5 未満を 0、0 . 2 5 以上 0 . 5 未満を 1、0 . 5 以上 0 . 7 5 未満を 2、0 . 7 5 以上を 3 とする、など）のいずれかを求めて、割合の代わりに用いてもよい。なお、割合を 2 段階の値にする場合は、前述した、割合を算出せずに対象物が存在しない『 0 』、存在する『 1 』、の 2 つの選択肢を求める場合と、実質的に同じになる。

【 0 0 8 3 】

また、対象物の地物特徴点の対象物の上部に多数存在する場合など、位置に偏りがある場合には、地物特徴点群の各 2 次元位置の画像領域内にある割合を求める際に、任意の重みを付けて算出し、数の少ない対象物の下部の特徴点画像領域内にあるか否かを重視する割合値にしてもよい。あるいは、割合による条件に加えて、あるいは条件の代わりに、重視する特徴点、たとえば下部にある特徴点の N 個以上が画像内に存在すること、という条件としてもよい。

【 0 0 8 4 】

さらに「画像内位置結果」は、対象物の地物特徴点の各 2 次元位置から求めた点、たとえば重心が、画像内に含まれるかを用いて算出してもよい。地物特徴点群の各 2 次元位置を用いて算出した重心位置が、画像内に含まれるかを判定し、含まれない = 『 0 』、含まれる = 『 1 』、の値とする。このとき、地物特徴点群の全 2 次元位置が画像内に含まれない場合は、重心位置も画像内に含まれないので、重心位置が画像内に含まれるか否かの算出処理を省略することができる。

【 0 0 8 5 】

また、「画像内位置結果」は、対象物の地物特徴の各 2 次元位置から求める画像内の対象物領域の面積と、その対象物領域が画像内に含まれる面積を求めて、面積比を用いてもよい。対象物領域の面積は、対象物の地物特徴の各 2 次元位置を頂点とする多角形の面積として求めることができる。対象物領域が画像内に含まれる面積は、概多角形の各辺と画面の左右上下端に相当する直線との交点を求めて、求めた交点と画面内に含まれる地物特徴の各 2 次元位置だけを頂点に用いた多角形の面積として算出することができる。たとえば、「画面内位置結果」は、（画面内に含まれる対象物領域の面積）÷（対象物領域の面積）として求めることができる。

【 0 0 8 6 】

画像内位置算出部 2 0 2 は、この「画像内位置結果」に関する情報に基づいて、画像 A 2 1 1 に対して以後の処理をおこなうか否かを判断する。すなわち、「画像内位置結果」が、画像内に含まれる 2 次元位置群の割合であり、0 ~ 1 の値である場合には、「画像内位置結果」があらかじめ決定しておいた規定値による条件（たとえば「画像内位置結果」の値範囲の中間である 0 . 5 以上）に合致する画像のみを用いることとし、「画像内位置結果」が規定値による条件に合致しない画像については、以後の処理に用いないようにし、処理対象となる画像を絞り込む。

【 0 0 8 7 】

なお、規定値は、図 2 の説明で前述したように、画像領域切り出しとして用いる画像 A 2 1 1 を、画像 A 2 1 1 内に写りこむ対象物の量で選定するための条件に用いる閾値であり、写りこむ量が大きい画像に限定したい場合は、「画像内位置結果」の値範囲の最大値

10

20

30

40

50

に近い値（２次元位置群の割合の場合は、より１に近い値）を設定しておく。また、前述したように、規定値による条件に加えて、さらに付随する条件、たとえば、重視する特徴点の２次元位置が画像内にあるか否か、を加えるようにしてもよい。

【００８８】

つぎに、画像内位置算出部２０２の具体的な処理手順について説明する。図１１は、画像内位置算出部の一連の処理の手順の一例を示すフローチャートである。図１１のフローチャートにおいて、画像内位置算出部２０２は、画像Ａ２１１の撮影位置・姿勢Ａ_aを入力する（ステップＳ１１０１）とともに、画像Ａ２１１を撮影したカメラパラメータＡ_pを入力する（ステップＳ１１０２）。また、画像内位置算出部２０２は、対象物の３次元位置Ａ_oを入力する（ステップＳ１１０３）。

10

【００８９】

つぎに、画像内位置算出部２０２は、対象物の各地物特徴点群が写り込む２次元位置を、対象物の３次元の各座標点に対する２次元位置群として、撮影位置・姿勢Ａ_aによる「透視変換（処理方法）」を用いて算出する（ステップＳ１１０４）。そして、推定した各地物特徴点群の２次元位置が画像Ａ２１１内に存在するか否かから、対象物が画像Ａ２１１内に写り込む量を算出する（ステップＳ１１０５）。

【００９０】

さらに、画像内位置結果Ａ_r５１１として、画像内位置結果：画像Ａ２１１内に対象物が写り込む量（写り込む割合０～１の値）、対象物を表す点の数：地物特徴点数、すなわち、対象物の特徴点数（２次元位置群の総数）、各点の画像内判定結果：各地物特徴点の２次元位置がそれぞれ画像Ａ２１１内にある（存在する）＝「１」か、ない（存在しない）＝「０」か、各点の２次元位置：各地物特徴点の２次元位置座標、を出力する（ステップＳ１１０６）。

20

【００９１】

つぎに、画像内位置算出部２０２は、画像内位置結果Ａ_r５１１の対象物が画像内に写り込む量である画像内位置結果が、あらかじめ設定してある規定値以上か否かを判定する（ステップＳ１１０７）。ここで、規定値以上の場合（ステップＳ１１０７：Ｙｅｓ）は、画像Ａ２１１を、以後の領域特定部以降の処理をおこなうものとして（ステップＳ１１０８）、画像内位置算出部２０２の一連の処理を終了する。

【００９２】

一方、規定値未満の場合（ステップＳ１１０７：Ｎｏ）は、画像Ａ２１１を、以後の領域特定部以降の処理をおこなわないものとして（ステップＳ１１０９）、画像内位置算出部２０２の一連の処理を終了する。この場合の画像内位置結果Ａ_r５１１（画像内位置結果：規定値未満の値）は、保管され、同じ対象物に対する別の探索に参照されるようにしてもよい。

30

【００９３】

ステップＳ１１０４における透視変換処理方法とは、３次元座標上の物体を２次元上に投影する変換処理に関する方法である。この透視変換処理方法により、複数カメラの撮像面の関係による、同一対象物が写る画像内位置を推定することが可能となる。透視変換処理は、具体的には、図１にも示したように、対象物の代表点の３次元位置と撮影位置を結んだ直線と撮影位置・姿勢とカメラパラメータ（焦点距離、光学中心）から推定する、カメラ撮像面との交点を求める。そして、その交点に相当する画像内の位置を２次元位置として算出するものである。

40

【００９４】

図１２は、透視変換の処理内容を示す説明図である。図１２において、物体Ｐは、実世界上の３次元座標位置を有する。具体的には、物体Ｐは、実座標原点０_wに対して、座標軸Ｘ_w、Ｙ_w、Ｚ_wの世界座標系における３次元座標点（ P_{xw} 、 P_{yw} 、 P_{zw} ）を有する。透視変換はまず、世界座標系における３次元座標点（ P_{xw} 、 P_{yw} 、 P_{zw} ）を、カメラの撮影位置・姿勢情報を用いてカメラ座標系の３次元座標点（ P_{xc} 、 P_{yc} 、 P_{zc} ）に変換する。

50

【 0 0 9 5 】

ここで、カメラ座標系とは、画像を撮影した撮像装置（カメラ）の位置を原点 O_c とし、画像の横方向を X_c 、縦方向を Y_c 、奥行き方向を Z_c とする座標系である。透視投影は具体的には、カメラ撮影位置を用いて実座標原点 O_w をカメラ原点 O_c に合わせる併進運動と、カメラ撮影姿勢を用いて世界座標系の各軸 X_w, Y_w, Z_w をカメラ座標系の各軸 X_c, Y_c, Z_c に変換する回転運動に相当する処理をおこなう。

【 0 0 9 6 】

つぎに、透視投影は、カメラの内部パラメータを用いて、カメラ座標系での実座標スケールでの座標値（たとえば $[m]$ など）を、画素単位のスケールでの座標値（ $[dot]$ ）に変換するとともに、画像内 2 次元位置の原点定義（画像の原点 O 、たとえば画像の左上）に基づく座標値へと変換する。

10

【 0 0 9 7 】

具体的には、物体 P のカメラ座標系の 3 次元座標点（ P_{xc}, P_{yc}, P_{zc} ）を、カメラ焦点距離によるスケール変換とカメラ原点が画像内に写りこんだ位置である画像中心位置による原点変更により、物体 P の位置にある画像上、すなわち画像座標系（原点 O 、各軸 x, y, z ）での物体 P の画像内位置（ P_x, P_y, P_z ）を求め、さらに画像とカメラの距離に依らない正規化画像座標系（原点 O' 、各軸 s, t ）での物体 P の画像内位置（ $P_s = P_x / P_z, P_t = P_y / P_z$ ）を求めることで、最終的な画像上の 2 次元画像内位置（ P_s, P_t ）を算出する。

【 0 0 9 8 】

20

これにより、物体 P の実世界上の 3 次元座標位置（ P_{xw}, P_{yw}, P_{zw} ）を、画像座標点（ P_s, P_t ）に変換し、3 次元座標上の物体 P を 2 次元の画面上に投影することができる。

【 0 0 9 9 】

透視投影で実世界上の 3 次元座標位置を画像座標点へ変換するには、下記式（1）～（3）を用いる。

【 0 1 0 0 】

【数 1】

$$\begin{bmatrix} P_{xc} \\ P_{yc} \\ P_{zc} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} P_{xw} \\ P_{yw} \\ P_{zw} \\ 1 \end{bmatrix} \quad \dots (1)$$

30

【 0 1 0 1 】

上記式（1）において、撮影位置・姿勢から求めた世界座標系からカメラ座標系への座標系変換行列である並進 - 回転の同次座標系変換行列として、 $r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}$ は、回転成分を表しており、 t_1, t_2, t_3 は、並進移動成分を表している。また、（ X_w, Y_w, Z_w ）は世界座標系の 3 次元座標を表している。

40

【 0 1 0 2 】

この式（1）の並進移動成分（ $t_1 \sim t_3$ ）によって、座標系原点を実座標原点 O_w からカメラの原点 O_c に合わせ、回転成分（ $r_{11} \sim r_{33}$ ）によって、世界座標系の座標軸定義 X_w, Y_w, Z_w による 3 次元座標値をカメラ座標系の座標軸定義 X_c, Y_c, Z_c による 3 次元座標値に変換することができる。

【 0 1 0 3 】

【数 2】

50

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} f_x & 0 & C_x \\ 0 & f_y & C_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} P_{xc} \\ P_{yc} \\ P_{zc} \end{bmatrix} \quad \dots (2)$$

【 0 1 0 4 】

上記式(2)は、式(1)で求めたカメラ座標系の値を、カメラの内部パラメータ行列を用いて、物体Pの位置にカメラ撮像面(画像面)のある画像座標系の値へと変換する。内部パラメータ行列の f_x 、 f_y は、図7に示したカメラパラメータAp213の画像横方向と縦方向の焦点距離を表しており、後述するカメラ中心からの距離を1とする画像スクリーン上で、実スケールによる位置を画素位置へ変換する値である。

10

【 0 1 0 5 】

また、 c_x 、 c_y は、カメラパラメータAp213の光学中心の横方向と縦方向の画素位置であり、画像内に写りこむカメラ原点Ocの画像内位置を表している。光学中心は、画像内位置の座標原点定義に依存する値であり、図12のように画像内位置を画像の左上を原点(画像の原点O)とする座標値で表す場合があるため、座標の原点定義を変更するために用いる。

20

【 0 1 0 6 】

式(2)は、内部パラメータ行列の焦点距離を用いたスケール変換(たとえば実スケール[m]から画素[dot]への変換)と、光学中心を用いた原点定義変更(たとえばカメラ原点Ocの画像座標系での投影位置に相当する画像中心から、画像左上の原点Oへの変更)により、カメラ座標系 X_c 、 Y_c 、 Z_c の3次元座標位置(P_{xc} 、 P_{yc} 、 P_{zc})を、物体Pの位置にある画像スクリーン上の画像、すなわち画像座標系の画像内位置(P_x 、 P_y 、 P_z)に変換する。

【 0 1 0 7 】

【数3】

30

$$\begin{bmatrix} P_s \\ P_t \\ 1 \end{bmatrix} = \begin{bmatrix} P_x/P_z \\ P_y/P_z \\ 1 \end{bmatrix} \quad \dots (3)$$

【 0 1 0 8 】

上記式(3)は、式(2)で求めた画像座標系の位置(P_x 、 P_y 、 P_z)を、一般的な3次元座標変換を用いて2次元の正規化画像座標系での画像内位置へと変換し、最終的な画像内位置座標(P_s 、 P_t)を求める。具体的には、画像座標系の位置(P_x 、 P_y 、 P_z)は、カメラ中心から奥行方向に P_z 離れた位置にある画像スクリーン上の画素位置であるため、距離 P_z 分だけ本来の画素位置よりも拡大した位置になっている。このため、 f_x 、 f_y が想定しているカメラ中心からの距離を1とする画像スクリーン上の画素位置へと、各座標成分を P_z で除算して正規化変換し、最終的な画像内位置座標($P_s = P_x / P_z$ 、 $P_t = P_y / P_z$)を得る。

40

【 0 1 0 9 】

なお、式(2)～(3)は一例であり、カメラ内部パラメータの焦点距離 f_x 、 f_y やカメラ光学中心 c_x 、 c_y の定義が異なる場合は、それらの定義に合わせて、適宜式(2)～(3)を変更してもよい。

50

【 0 1 1 0 】

式(2)では、省略するが、さらに、歪み補正係数を考慮するようにしてもよい。歪み補正係数は、カメラパラメータ A_{p213} に、焦点距離、光学中心とともに記憶されている。

【 0 1 1 1 】

このようにして、対象物の3次元位置 A_o から、撮影位置・姿勢 A_{a212} 、カメラパラメータ A_{p213} を用いて、透視変換処理をすることにより、対象物が画像内に写り込む2次元群を算出することができる。

【 0 1 1 2 】

(領域特定部203の処理内容・位置情報 A_{i512} の内容)

10

つぎに、領域特定部203の具体的な内容について説明する。図5に示すように、領域特定部203は、画像内位置算出部202によって領域特定部203以後の処理をおこなうと判定済みの画像 A_{211} に対して、出力された画像内位置結果 A_{r511} を入力し、画像 A から切り出す画像内領域を算出する。そして、領域特定部203は、算出した結果として、位置情報 A_{i512} を出力する。

【 0 1 1 3 】

領域特定部203は、具体的には、たとえば、図3に示した情報処理装置200において、メモリ302に記憶されたプログラムをCPU301が実行することによって、その機能を実現することができる。

【 0 1 1 4 】

20

領域特定部203は、たとえば、画像内位置結果 A_{r511} を参照し、推定した画像への写り込み量から画像内にあると判断された対象物の各2次元位置群に対し、概2次元位置群の外接矩形を画像内領域として算出し、矩形を定義する位置情報 A_{i512} を出力する。位置情報 A_i は、たとえば、矩形の左上、右下の2次元の画像内位置などであってもよい。

【 0 1 1 5 】

なお、画像内領域としては、矩形でなくてもよく、任意の直線または曲線による図形、たとえば多角形や、楕円や曲線図形でよい。2次元位置群を結んだ多角形でもよいし、2次元位置群を内包する楕円でもよい。また、画像内領域は2次元位置群の厳密な外接図形ではなく、外接図形より小さな図形でもよいし、外接図形よりも大きな図形であっても構わない。

30

【 0 1 1 6 】

図13は、位置情報 A_i のデータ構成の一例を示す説明図である。図13において、位置情報 A_{i512} は、「動画ID」、「フレーム番号」、「対象物ID」、「対象物の種類」、「対象物を表す点の数」、「画像内位置座標」を含む各種情報を有している。「動画ID」は、画像 A_{211} が含まれる動画を識別する一意の識別情報であり、「フレーム番号」は、当該動画の中の画像 A_{211} が存在するフレームの番号に関する情報である。これらは、図6に示した撮影位置・姿勢 A_{a212} 、図10に示した画像内位置結果 A_{r511} の「動画ID」および「フレーム番号」である。

【 0 1 1 7 】

40

「対象物ID」は、対象物を識別する一意の識別情報であり、「対象物の種類」は、対象物の種類を示す情報であり、これらは、図8に示した対象物の3次元位置 A_o の「対象物ID」および「対象物の種類」である。

【 0 1 1 8 】

「対象物を表す点の数」は、対象物、より具体的には、対象物が含まれる切り出す画像を表す点の数であり、図8に示した対象物の3次元位置 A_o と同数であってもよく、異なる数であってもよい。また、図10に示した対象物の画像内にあると判定された2次元位置群の数と同じであってもよく、異なる数であってもよい。

【 0 1 1 9 】

一般的には、切り出された画像は、矩形により管理されるため、「対象物を表す点の数

50

」は、4点とするのが好ましい。しかしながら、対象物の形状などによって、『4』には限定されるものではない。「画像内位置座標」は、「対象物を表す点の数」が『4』なので、Pos[0]～Pos[3]の4つの座標が、それぞれx、yの2次元で示される。

【0120】

図14は、領域特定部の一連の処理の手順の一例を示すフローチャートである。図14のフローチャートにおいて、領域特定部203は、まず、画像内位置算出部202によって出力された画像内位置結果Ar（画像内位置結果：「1」）511を入力する（ステップS1401）。

【0121】

そして、対象物の全代表点のうち、画像A211内にあると画像内位置算出部202で判定された代表点の2次元位置群に対し、外接矩形を画像内領域として定義し、画像内位置座標を特定する（ステップS1402）。その後、位置情報Ai（画像内位置座標）を作成し、その後、その位置情報Ai512を出力する（ステップS1403）。それにより、領域特定部203の一連の処理を終了する。

【0122】

（領域出力部204の処理内容・領域画像Abの内容）

つぎに、領域出力部204の具体的な内容について説明する。図5に示すように、領域出力部204は、領域特定部203によって出力された位置情報Ai512を入力し、領域画像Ab513を出力する。領域出力部204は、具体的には、たとえば、図3に示した情報処理装置200において、メモリ302に記憶されたプログラムをCPU301が実行することによって、あるいは、ネットワークI/F303、記録媒体I/F304によって、その機能を実現することができる。

【0123】

図15は、領域画像Abのデータ構成の一例を示す説明図である。図15において、領域画像Ab513は、「動画ID」、「フレーム番号」、「対象物ID」、「対象物の種類」、「保存先」、「ファイル名」を含む各種情報を有している。「動画ID」は、画像A211が含まれる動画を識別する一意の識別情報であり、「フレーム番号」は、当該動画の中の画像A211が存在するフレームの番号に関する情報である。これらは、図6に示した撮影位置・姿勢Aa212、図10に示した画像内位置結果Ar511の「動画ID」および「フレーム番号」である。

【0124】

「対象物ID」は、対象物を識別する一意の識別情報であり、「対象物の種類」は、対象物の種類を示す情報であり、これらは、図8、図13に示した対象物の3次元位置Ao214、位置情報Ai512の「対象物ID」および「対象物の種類」である。「保存先」は、切り出された領域画像が保存されるフォルダに関する情報であり、「ファイル名」は、切り出された領域画像のファイル名に関する情報である。

【0125】

図16は、切り出された領域画像の一例を示す説明図である。図16に示すように、切り出された領域画像1601は、地物の一例である、制限速度が50km/hであることを示す道路標識である。この領域画像1601は、たとえば、ビットマップ、Jpegなどの所定のフォーマットにより、データとして保存される。

【0126】

地物は、具体的には、標識、信号機、ガードレールのほか、道路の縁石、白線、横断歩道、停止線、路面ペイント（文字に加えて矢印などの標示物）などであってもよい。また、地下鉄、交番、任意店舗・設備などに対する看板、踏切、歩道橋、料金所、道路鋸、ゼブラゾーン、ポールなどの道路付帯施設であってもよい。

【0127】

また、電柱、バス停、電話・郵便ボックス、タクシー乗り場表示、自動販売機などであってもよい。また、側溝、マンホール、クッションドラムなどであってもよい。さらには、バリケード、パイロン、駐車スペース（路面標示枠）、駐車場の料金支払い機、ゴミ収

10

20

30

40

50

集ボックス、消火栓、任意店舗のショウウィンドウ、街路樹、などであってもよい。

【0128】

このように、地物は、撮影画像に写り込むことが可能で、かつ、その3次元位置が取得できる物体であれば、切り取りの対象とすることができる。

【0129】

図17は、領域出力部の一連の処理の手順の一例を示すフローチャートである。図17のフローチャートにおいて、領域出力部204は、領域特定部203によって出力された位置情報Ai512を入力する(ステップS1701)。つぎに、領域出力部204は、位置情報Ai512に相当する領域画像Abを切り出す(ステップS1702)。具体的には、図13に示した位置情報Ai512の画像内位置座標に基づいて、領域画像を切り出す。そして、切り出した領域画像について、図15に示した領域画像Ab513のファイルを作成する。

10

【0130】

その後、領域出力部204は、領域画像Ab513を出力する(ステップS1703)。具体的には、領域画像Ab513のファイルを、図3に示した記録媒体305に記録する、あるいは、ネットワーク310を介して、他の情報処理装置へ出力する。それによって、切り出した領域画像Abを、領域画像1601とともに、所定の記憶領域に蓄積することができ、機械学習(ディープラーニング)用の教師データとして活用に供することができる。

【0131】

20

以上説明したように、本実施の形態にかかる情報処理装置200によれば、画像から対象物の領域を切り出すにあたり、対象物の3次元位置に関する3次元位置情報を記憶した記憶部を参照し、画像の撮影位置および姿勢に関する情報を用いて、3次元位置に関する3次元位置情報から、画像における対象物の領域を算出することができる。

【0132】

これにより、算出して切り出した画像領域を機械学習用の教師データとすることで、人手を使わずに、自動で任意画像または映像からの対象物が写る画像領域の画像切り出しを実現することができる。このように、任意シーンを使った十分な機械学習用の教師データ収集が可能となり、収集した機械学習用の教師データを活用した学習ができるようになる。

【0133】

30

また、本実施の形態にかかる情報処理装置200によれば、画像は、撮像装置によって撮影された画像であってもよく、撮像装置によって移動しながら撮影された画像であってもよい。これにより、容易に大量の機械学習用の教師データにかかる画像を取得することができる。

【0134】

また、本実施の形態にかかる情報処理装置200によれば、画像の撮影位置および姿勢に関する情報および撮像装置のパラメータに関する情報を用いて、画像における前記対象物の領域を算出することができる。これにより、より正確な対象物の領域を算出することができる。

【0135】

40

また、本実施の形態にかかる情報処理装置200によれば、透視変換処理により、画像における対象物の領域を算出することができる。これにより、対象物の3次元位置に関する3次元位置情報から、対象物が画像内に写り込む2次元位置を含む領域を迅速に算出することができる。

【0136】

また、本実施の形態にかかる情報処理装置200によれば、対象物が画像内に写り込む量を算出して、写り込む量から対象物が画像内に存在するか否かを判断し、判断の結果、対象物が前記画像内に存在する場合に、画像における対象物の領域を算出することができる。これにより、対象物が写っていない、または、ごく一部しか写っていない画像を避けて画像を探索することができ、対象物が写っている画像のみにその後の画像切り出し処理

50

をおこなえばよく、機械学習用の教師データを効率的に収集することができる。

【0137】

また、本実施の形態にかかる情報処理装置200によれば、画像における対象物の領域の算出結果に基づいて、対象物の画像を切り出すことができる。これにより、画像の切り出し処理を自動的におこなうことができる。

【0138】

また、本実施の形態にかかる情報処理装置200によれば、画像の撮影位置および姿勢に関する情報は、画像を用いたV-SLAMによって取得することができる。これにより、画像の撮影位置および姿勢に関する情報を効率よく抽出することができる。

【0139】

また、本実施の形態にかかる情報処理装置200によれば、3次元位置に関する3次元位置情報は、三角測量を用いて決定することができる。これにより、より簡易に3次元位置に関する3次元位置情報を取得することができる。

【0140】

また、本実施の形態にかかる情報処理装置200によれば、対象物は、画像に表示される地物であってもよい。これにより、地物を機械学習用の教師データとして活用することができる。

【0141】

また、本実施の形態にかかる情報処理装置200によれば、撮像装置は、移動体に搭載されたカメラであってもよい。これにより、移動しながら撮影された画像から対象物（地物）の機械学習用の教師データを取得することができる。

【0142】

なお、本実施の形態で説明した領域切り出し方法は、あらかじめ用意されたプログラムをパーソナルコンピュータやワークステーションなどのコンピュータで実行することにより実現することができる。プログラム配信プログラムは、ハードディスク、フレキシブルディスク、CD (Compact Disc) - ROM、MO (Magnetooptical Disk)、DVD (Digital Versatile Disk)、USB (Universal Serial Bus) メモリなどのコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また、領域切り出しプログラムは、インターネットなどのネットワークを介して配布してもよい。

【0143】

また、本実施の形態で説明した情報処理装置200は、スタンダードセルやストラクチャードASIC (Application Specific Integrated Circuit) などの特定用途向けICやFPGAなどのPLD (Programmable Logic Device) によっても実現することができる。

【0144】

上述した実施の形態に関し、さらに以下の付記を開示する。

【0145】

(付記1) 画像から対象物の領域を切り出す領域切り出し方法であって、
情報処理装置が、
前記対象物の3次元位置情報を記憶した記憶部を参照し、
前記画像の撮影位置および姿勢に関する情報を用いて、前記3次元位置情報から、前記画像における前記対象物の領域を算出する、
ことを特徴とする領域切り出し方法。

【0146】

(付記2) 前記画像は、撮像装置によって撮影された画像であることを特徴とする付記1に記載の領域切り出し方法。

【0147】

(付記3) 前記画像は、前記撮像装置によって移動しながら撮影された画像であることを

10

20

30

40

50

特徴とする付記 2 に記載の領域切り出し方法。

【 0 1 4 8 】

(付記 4) 前記画像の撮影位置および姿勢に関する情報および前記撮像装置のパラメータに関する情報を用いて、前記画像における前記対象物の領域を算出することを特徴とする付記 2 または 3 に記載の領域切り出し方法。

【 0 1 4 9 】

(付記 5) 透視変換処理により、前記画像における前記対象物の領域を算出することを特徴とする付記 1 ~ 4 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 0 】

(付記 6) 前記対象物が前記画像内に存在するかを判断し、
前記判断の結果、前記対象物が前記画像内に存在する場合に、当該画像における当該対象物の領域を算出することを特徴とする付記 1 ~ 5 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 1 】

(付記 7) 前記画像における前記対象物の領域の算出結果に基づいて、当該対象物の画像を切り出すことを特徴とする付記 1 ~ 6 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 2 】

(付記 8) 前記画像の撮影位置および姿勢に関する情報は、当該画像を用いた V - S L A M によって取得することを特徴とする付記 1 ~ 7 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 3 】

(付記 9) 前記 3 次元位置情報は、三角測量を用いて決定することを特徴とする付記 1 ~ 8 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 4 】

(付記 1 0) 前記対象物は、前記画像に表示される地物であること特徴とする付記 1 ~ 9 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 5 】

(付記 1 1) 前記撮像装置は、移動体に搭載されたカメラであることを特徴とする付記 2 ~ 4 のいずれか一つに記載の領域切り出し方法。

【 0 1 5 6 】

(付記 1 2) 画像から対象物の領域を切り出す領域切り出しプログラムであって、
情報処理装置に、
前記対象物の 3 次元位置情報を記憶した記憶部を参照し、
前記画像の撮影位置および姿勢に関する情報を用いて、前記 3 次元位置情報から、前記画像における前記対象物の領域を算出する、
処理を実行させることを特徴とする領域切り出しプログラム。

【符号の説明】

【 0 1 5 7 】

- 1 0 0 カメラ
- 1 0 1 カメラ撮像面
- 1 0 2 地物 (標識) の 3 次元位置
- 1 0 3 対物存在直線
- 1 0 4 地物 (対象物) の画像内位置
- 1 0 5 地物画像 (機械学習用の教師データ)
- 2 0 0 情報処理装置 (サーバ)
- 2 0 1 画像入力部
- 2 0 2 画像内位置算出部
- 2 0 3 領域特定部
- 2 0 4 領域出力部
- 2 0 5 記憶部

10

20

30

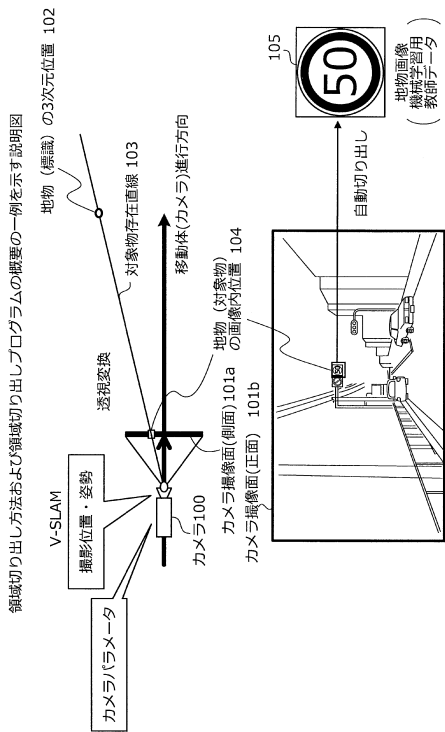
40

50

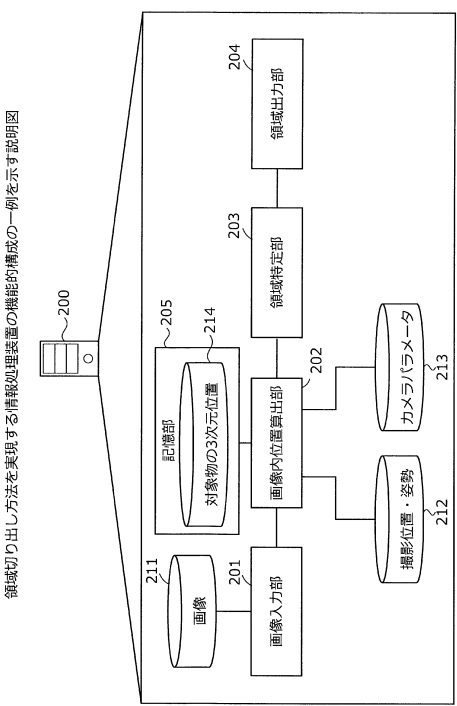
- 2 1 1 画像（画像 A）
- 2 1 2 撮影位置・姿勢（撮影位置・姿勢 A a）
- 2 1 3 カメラパラメータ（カメラパラメータ A p）
- 2 1 4 対象物の 3 次元位置（対象物の 3 次元位置 A o）
- 5 1 1 画像内位置結果（画像内位置結果 A r）
- 5 1 2 位置情報（位置情報 A i）
- 5 1 3、1 6 0 1 領域画像（領域画像 A b）

【図面】

【図 1】



【図 2】



10

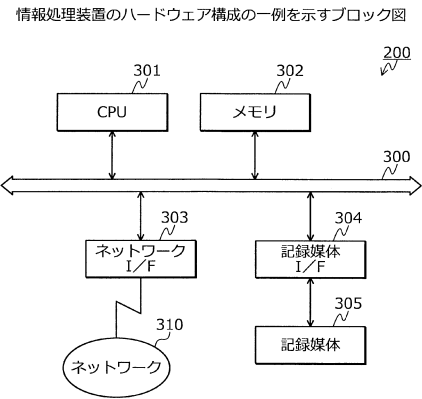
20

30

40

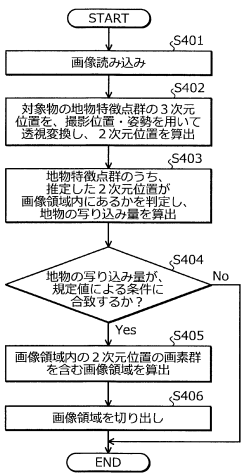
50

【図 3】



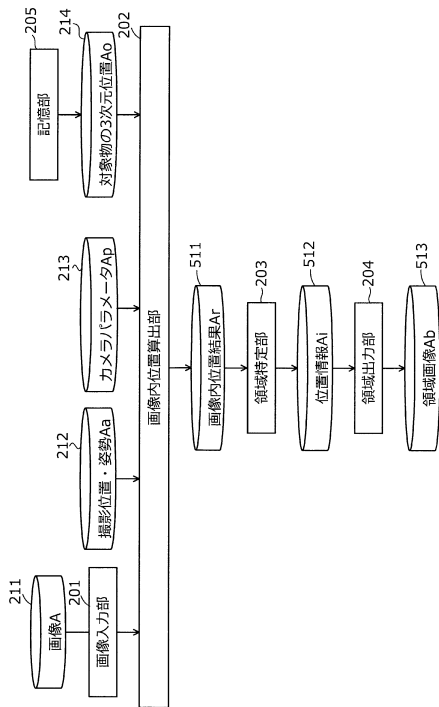
【図 4】

領域切り出し方法の一連の処理の手順の一例を示すフローチャート



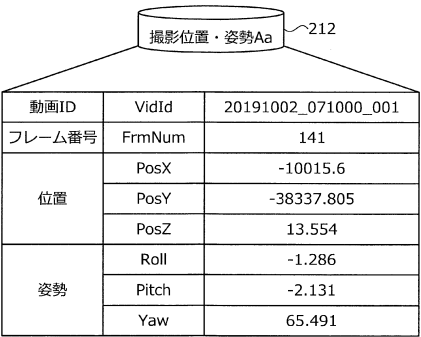
【図 5】

各構成部および各情報の内容の一例を示す説明図



【図 6】

撮影位置・姿勢 A a のデータ構成の一例を示す説明図



10

20

30

40

50

【 図 7 】

カメラパラメータ A p のデータ構成の一例を示す説明図

カメラパラメータAp213		
焦点距離	fx	968.7548
	fy	970.2972
光学中心	cx	982.918
	cy	548.1811
歪み補正係数	k1	-0.30387
	k2	0.092433
	p1	-0.0003
	p2	0.000648
	p3	-0.0127

【 図 8 】

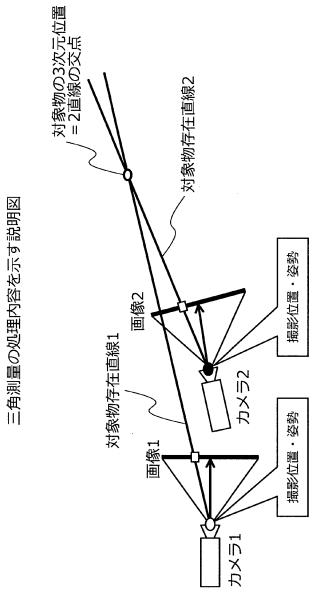
対象物の 3 次元位置 A o のデータ構成の一例を示す説明図

対象物の 3 次元位置Ao214		
対象物ID	objID	1
対象物の種類	objType	8
対象物を表す点の数	PosNum	4
直角平面系番号	CoordNum	9
実座標3次元[m]	Pos[0].x	-9946.05283
	Pos[0].y	-38173.5689
	Pos[0].z	19.556403
	Pos[1].x	-9946.05332
	Pos[1].y	-38173.5267
	Pos[1].z	19.226615
	Pos[2].x	-9945.022811
	Pos[2].y	-38173.9059
	Pos[2].z	19.186373
	Pos[3].x	-9945.05121
	Pos[3].y	-38173.9875
	Pos[3].z	19.53184

10

20

【 図 9 】



【 図 1 0 】

画像内位置結果 A r のデータ構成の一例を示す説明図

画像内位置結果Ar511		
動画ID	VidId	20191002_071000_001
フレーム番号	FrmNum	141
対象物ID	ObjID	1
対象物の種類	ObjType	8
画像内位置結果	Result	0.75
対象物を表す点の数	PosNum	4
各点の画像内判定結果	ResultPos[0]	1
	ResultPos[1]	0
	ResultPos[2]	1
	ResultPos[3]	1
各点の 2 次元位置 [dot]	2DPos[0].x	969.91
	2DPos[0].y	20.28
	2DPos[1].x	1015.24
	2DPos[1].y	-8.15
	2DPos[2].x	1033.71
	2DPos[2].y	120.88
	2DPos[3].x	953.12
	2DPos[3].y	129.32

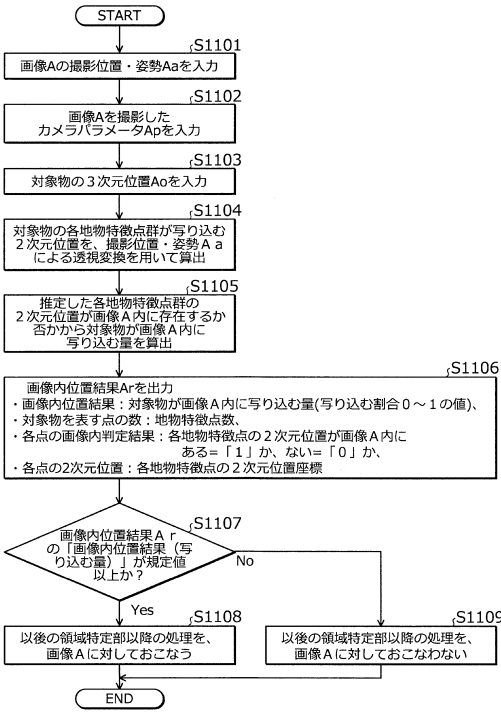
30

40

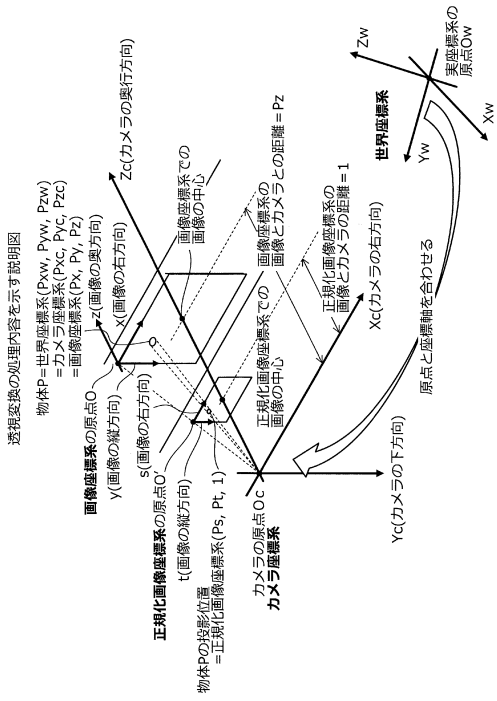
50

【図 1 1】

画像内位置算出部の一連の処理の手順の一例を示すフローチャート



【図 1 2】



10

20

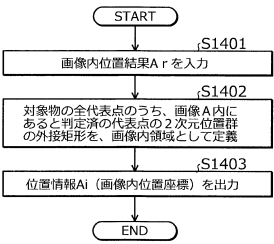
【図 1 3】

位置情報Aiのデータ構成の一例を示す説明図

位置情報Ai		
動画ID	VidId	20191002_071000_001
フレーム番号	FrmNum	141
対象物ID	ObjId	1
対象物の種類	ObjType	8
対象物を表す点の数	PosNum	4
画像内位置座標[dot]	Pos[0].x	964.393
	Pos[0].y	550.439
	Pos[1].x	964.343
	Pos[1].y	552.24
	Pos[2].x	970.296
	Pos[2].y	552.329
	Pos[3].x	970.285
	Pos[3].y	550.437

【図 1 4】

領域特定部の一連の処理の手順の一例を示すフローチャート



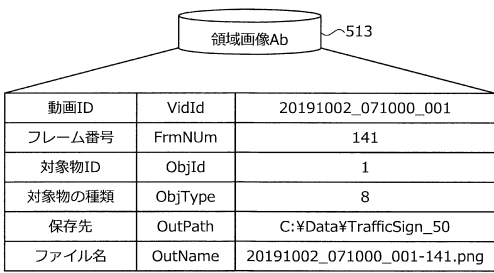
30

40

50

【図 15】

領域画像A bのデータ構成の一例を示す説明図



【図 16】

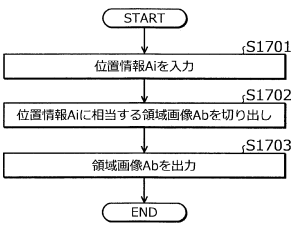
切り出された領域画像の一例を示す説明図



10

【図 17】

領域出力部の一連の処理の手順の一例を示すフローチャート



20

30

40

50

フロントページの続き

審査官 合田 幸裕

- (56)参考文献 特開 2 0 0 7 - 0 8 0 0 6 0 (J P , A)
特開 2 0 0 8 - 2 9 8 6 9 7 (J P , A)
特開 2 0 0 7 - 2 1 2 1 8 7 (J P , A)
特開 2 0 1 6 - 1 7 7 3 8 8 (J P , A)
特開 2 0 1 9 - 1 2 1 1 3 6 (J P , A)
国際公開第 2 0 0 8 / 0 9 9 9 1 5 (W O , A 1)
- (58)調査した分野 (Int.Cl., D B 名)
G 0 6 T 7 / 0 0
G 0 6 T 7 / 5 9 3
G 0 8 G 1 / 1 6
I E E E X p l o r e
J S T P l u s (J D r e a m I I I)