

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7383435号  
(P7383435)

(45)発行日 令和5年11月20日(2023.11.20)

(24)登録日 令和5年11月10日(2023.11.10)

|                 |      |           |         |              |
|-----------------|------|-----------|---------|--------------|
| (51)国際特許分類      | F I  |           |         |              |
| G 0 6 T         | 7/00 | (2017.01) | G 0 6 T | 7/00 6 6 0 B |
| G 0 6 T         | 1/00 | (2006.01) | G 0 6 T | 7/00 3 5 0 C |
| G 0 6 T         | 1/40 | (2006.01) | G 0 6 T | 1/00 3 4 0 B |
| H 0 4 N         | 7/18 | (2006.01) | G 0 6 T | 1/40         |
|                 |      |           | H 0 4 N | 7/18 K       |
| 請求項の数 20 (全22頁) |      |           |         |              |

|          |                             |          |                   |
|----------|-----------------------------|----------|-------------------|
| (21)出願番号 | 特願2019-168734(P2019-168734) | (73)特許権者 | 000001007         |
| (22)出願日  | 令和1年9月17日(2019.9.17)        |          | キヤノン株式会社          |
| (65)公開番号 | 特開2021-47538(P2021-47538A)  |          | 東京都大田区下丸子3丁目30番2号 |
| (43)公開日  | 令和3年3月25日(2021.3.25)        | (74)代理人  | 110003281         |
| 審査請求日    | 令和4年8月26日(2022.8.26)        |          | 弁理士法人大塚国際特許事務所    |
|          |                             | (72)発明者  | 牟田 元              |
|          |                             |          | 東京都大田区下丸子3丁目30番2号 |
|          |                             |          | キヤノン株式会社内         |
|          |                             | (72)発明者  | 馬場 康夫             |
|          |                             |          | 東京都大田区下丸子3丁目30番2号 |
|          |                             |          | キヤノン株式会社内         |
|          |                             | (72)発明者  | 矢野 光太郎            |
|          |                             |          | 東京都大田区下丸子3丁目30番2号 |
|          |                             |          | キヤノン株式会社内         |
|          |                             | 審査官      | 笠田 和宏             |
| 最終頁に続く   |                             |          |                   |

(54)【発明の名称】 画像処理装置、画像処理方法、及びプログラム

(57)【特許請求の範囲】

【請求項1】

画像中において対象物が存在すると推定される領域を表す密度分布を推定する密度推定手段と、

前記密度推定手段により推定される密度分布を表す密度マップの各位置について前記対象物の代表点を表すか否かを識別することにより、前記画像中における前記対象物の位置を推定する位置推定手段と、

前記位置推定手段により推定される複数の前記画像中の前記対象物の位置を対応付けする対応付け手段と、

前記対応付け手段により対応付けられる複数の前記画像中の前記対象物の位置に基づいて、計測線を通じた前記対象物の数を計測する計測手段と、

を備えることを特徴とする画像処理装置。

【請求項2】

前記位置推定手段は、前記密度マップの各位置について前記対象物の代表点である確率を表す位置マップを生成することを特徴とする、請求項1に記載の画像処理装置。

【請求項3】

前記位置推定手段は、前記密度マップを学習済みモデルに入力することにより生成される第1の位置マップと、前記第1の位置マップをプーリングすることにより生成される第2の位置マップを生成し、前記第1の位置マップと前記第2の位置マップとの間で値が一致している要素の座標を、前記対象物の位置として推定する、請求項1又は2に記載の画

像処理装置。

【請求項 4】

前記位置推定手段は、前記画像において互いに隣接する、前記対象物の位置を表すと推定された複数の位置に基づいて、前記対象物の位置を求めることを特徴とする、請求項 1 から 3 のいずれか 1 項に記載の画像処理装置。

【請求項 5】

前記画像から、前記画像の一部である 1 以上の部分画像を抽出する取得手段を更に有し、前記密度推定手段は、前記 1 以上の部分画像中の対象物の密度分布を推定することを特徴とする、請求項 1 から 4 のいずれか 1 項に記載の画像処理装置。

【請求項 6】

前記取得手段は、前記計測線が前記部分画像に含まれるように、前記画像から前記 1 以上の部分画像を抽出することを特徴とする、請求項 5 に記載の画像処理装置。

【請求項 7】

前記取得手段は、ユーザによる指定に基づいて、前記 1 以上の部分画像を抽出する領域を前記画像に設定することにより、前記 1 以上の部分画像を抽出することを特徴とする、請求項 5 又は 6 に記載の画像処理装置。

【請求項 8】

前記取得手段は、前記部分画像に中央領域とマージン領域を設定し、かつ、前記画像から複数の前記部分画像を抽出する場合に、隣り合う 2 つの前記部分画像の前記中央領域どうしが隣接するように前記部分画像を抽出し、

前記密度推定手段は、前記部分画像の前記中央領域に含まれる前記対象物の密度分布を推定する

ことを特徴とする、請求項 5 から 7 のいずれか 1 項に記載の画像処理装置。

【請求項 9】

前記取得手段は、前記対象物の大きさに基づいて前記マージン領域の大きさを設定することを特徴とする、請求項 8 に記載の画像処理装置。

【請求項 10】

前記計測手段は、前記計測線によって区切られる一方の領域から他方の領域へと移動した前記対象物の数を計測することを特徴とする、請求項 1 から 9 のいずれか 1 項に記載の画像処理装置。

【請求項 11】

前記計測手段は、前記一方の領域から前記計測線を超えて前記他方の領域へと移動した前記対象物の数を計測し、前記計測された対象物の数は、前記一方の領域から前記計測線へと移動したが前記他方の領域へと移動していない前記対象物の数を含まないことを特徴とする、請求項 10 に記載の画像処理装置。

【請求項 12】

前記計測手段は、前記一方の領域から前記他方の領域へと通過した前記対象物の数と、前記他方の領域から前記一方の領域へと通過した前記対象物の数と、の差分を算出することを特徴とする、請求項 10 又は 11 に記載の画像処理装置。

【請求項 13】

前記対応付け手段は、複数の前記画像のうちの第 1 の時刻に対応する第 1 の画像中の前記対象物の位置と、複数の前記画像のうちの第 2 の時刻に対応する第 2 の画像中の前記対象物の位置と、を対応付けることを特徴とする、請求項 1 から 12 のいずれか 1 項に記載の画像処理装置。

【請求項 14】

前記対応付け手段は、前記第 1 の画像中の前記対象物の位置と前記第 2 の画像中の前記対象物の位置とを対応付けることにより前記対象物毎の移動軌跡を出力し、

前記計測手段は、前記移動軌跡が前記計測線と交差する前記対象物の数を計測する、請求項 13 に記載の画像処理装置。

【請求項 15】

10

20

30

40

50

前記対応付け手段は、前記第 1 の画像中の前記対象物の位置と前記第 2 の画像中の前記対象物の位置とを結ぶ距離に基づいたコストを算出し、前記第 1 の画像中の前記対象物の位置と前記第 2 の画像中の前記対象物の位置との組合せのそれぞれの前記コストが最小となる組合せに基づいて対応付けを行う、ことを特徴とする、請求項 1 3 又は 1 4 に記載の画像処理装置。

【請求項 1 6】

前記対応付け手段は、

少なくとも前記第 2 の画像中の前記対象物の位置に基づいて、前記第 2 の時刻より後の第 3 の時刻における第 3 の画像中の前記対象物の位置を予測し、

前記第 3 の時刻における前記第 3 の画像中の予測された前記対象物の位置と、前記第 3 の画像中の前記対象物の位置と、に基づいて、前記第 2 の画像中の前記対象物の位置と、前記第 3 の画像中の前記対象物の位置と、を対応付ける

ことを特徴とする、請求項 1 3 から 1 5 の何れか 1 項に記載の画像処理装置。

【請求項 1 7】

前記密度推定手段は、ニューラルネットワークを用いて前記密度分布を推定することを特徴とする、請求項 1 から 1 6 のいずれか 1 項に記載の画像処理装置。

【請求項 1 8】

前記位置推定手段は、ニューラルネットワークを用いて前記対象物の位置を推定することを特徴とする、請求項 1 から 1 7 のいずれか 1 項に記載の画像処理装置。

【請求項 1 9】

画像処理装置が行う画像処理方法であって、

画像中において対象物が存在すると推定される領域を表す密度分布を推定する密度推定工程と、

前記密度推定工程において推定される密度分布を表す密度マップの各位置について前記対象物の代表点を表すか否かを識別することにより、前記画像中における前記対象物の位置を推定する位置推定工程と、

前記位置推定工程において推定される複数の前記画像中の前記対象物の位置を対応付けする対応付け工程と、

前記対応付け工程において対応付けられる複数の前記画像中の前記対象物の位置に基づいて、計測線を通じた前記対象物の数を計測する計測工程と、

を含むことを特徴とする画像処理方法。

【請求項 2 0】

コンピュータを、請求項 1 から 1 8 のいずれか 1 項に記載の画像処理装置として機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は画像処理装置、画像処理方法、及びプログラムに関し、特にビデオカメラ又は防犯カメラ等で撮像された動画像の解析技術に関する。

【背景技術】

【0 0 0 2】

近年、カメラ等で撮像された動画像から、撮影領域における対象物（例えば人）の流れ（例えば人数又は方向）、すなわち人流を解析する装置が提案されている。対象物の流れを測定するためには、動画像から対象物を検出することが必要である。特許文献 1 は、定点カメラで撮像された画像に含まれる人物領域を検出し、人物の数を計測する方法を開示している。

【先行技術文献】

【特許文献】

【0 0 0 3】

【文献】特開 2 0 0 7 - 2 0 1 5 5 6 号公報

10

20

30

40

50

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0004】

例えば群衆を撮像する場合のように、対象物が混雑しているシーン（混雑シーン）において対象物の流れを精度よく測定するためには、このようなシーンにおいても対象物の位置を精度よく検出することが求められる。一方、特許文献1では、画像が人物で占められている領域を検出することにより人物を検出している。この方法では、人物同士の重なりが大きくなると人物の検出精度が低下するため、混雑シーンにおいて流れの測定精度が低下する。

## 【0005】

本発明は、混雑シーンにおける対象物の流れの計測精度を向上させることを目的とする。

## 【課題を解決するための手段】

## 【0006】

本発明の目的を達成するために、例えば、本発明の画像処理装置は以下の構成を備える。すなわち、

画像中において対象物が存在すると推定される領域を表す密度分布を推定する密度推定手段と、

前記密度推定手段により推定される密度分布を表す密度マップの各位置について前記対象物の代表点を表すか否かを識別することにより、前記画像中における前記対象物の位置を推定する位置推定手段と、

前記位置推定手段により推定される複数の前記画像中の前記対象物の位置を対応付けする対応付け手段と、

前記対応付け手段により対応付けられる複数の前記画像中の前記対象物の位置に基づいて、計測線を通じた前記対象物の数を計測する計測手段と、

を備えることを特徴とする。

## 【発明の効果】

## 【0007】

混雑シーンにおける対象物の流れの計測精度を向上させる。

## 【図面の簡単な説明】

## 【0008】

【図1】画像処理装置のハードウェア構成の一例を示す図。

【図2】画像処理装置の機能構成の一例を示す図。

【図3】画像処理装置の処理の流れの一例を示す図。

【図4】入力画像から部分画像を抽出する方法の一例を示す図。

【図5】部分画像の中央領域及びマージン領域を説明する図。

【図6】入力画像における人物の密度分布推定の一例を示す図。

【図7】密度分布からの人物の位置座標推定の一例を示す図。

【図8】位置マップから人物の位置座標を求める処理の一例を示す図。

【図9】異なる時刻における人物の位置座標同士のマッチング方法を示す図。

【図10】マッチングで用いられるコスト行列及びマッチング結果の一例を示す図。

【図11】マッチングにより人物の移動軌跡を求める処理の一例を示す図。

【図12】人物の予測位置に基づくマッチング方法を示す図。

【図13】移動軌跡が途切れた場合に移動軌跡を補完する方法を示す図。

【図14】人物の移動軌跡から流量を計測する方法の一例を示す図。

【図15】人物の移動軌跡から流量を計測する方法の一例を示す図。

## 【発明を実施するための形態】

## 【0009】

以下、添付図面を参照して実施形態を詳しく説明する。なお、以下の実施形態は特許請求の範囲に係る発明を限定するものではない。実施形態には複数の特徴が記載されているが、これらの複数の特徴の全てが発明に必須のものとは限らず、また、複数の特徴は任意

10

20

30

40

50

に組み合わせられてもよい。さらに、添付図面においては、同一若しくは同様の構成に同一の参照番号を付し、重複した説明は省略する。

#### 【0010】

図1は、本発明の一実施形態に係る画像処理装置100のハードウェア構成例を示す。画像処理装置100は、制御装置11、記憶装置12、演算装置13、入力装置14、出力装置15、及びI/F装置16を有する。

#### 【0011】

制御装置11は、画像処理装置100の全体を制御する装置である。記憶装置12は、制御装置11の動作に必要なプログラム及びデータを保持する。演算装置13は、制御装置11からの制御に基づき、必要な演算処理を実行する。例えば、演算装置13は、後述するニューラルネットワーク演算を行ってもよい。入力装置14は、ヒューマンインターフェイスデバイス等であり、ユーザの操作による入力を取得する。出力装置15は、ディスプレイ等であり、画像処理装置100が生成した処理結果等をユーザに提示する。

#### 【0012】

I/F装置16は、ユニバーサルシリアルバス、イーサネット（登録商標）、若しくは光ケーブル等の有線インターフェース、又はWi-Fi若しくはBluetooth（登録商標）等の無線インターフェースである。I/F装置16を介して、他の装置を画像処理装置100に接続することができる。例えば、I/F装置16にはカメラ等の撮像装置を接続することができ、I/F装置16を介して画像処理装置100は撮像画像を取得することができる。別の例として、画像処理装置100は、I/F装置16を介して処理結果を外部に送信することができる。さらなる例として、画像処理装置100は、動作に必要なプログラム又はデータ等をI/F装置16を介して取得することができる。

#### 【0013】

後述する画像処理装置100の機能は、例えば、プロセッサ（例えば制御装置11）が、メモリ（例えば記憶装置12）上のプログラムに従って動作することにより実現することができる。上記の記憶装置12又はその他の記憶媒体は、このようなプログラムを格納することができる。もっとも、後述する画像処理装置100の少なくとも一部の機能は、専用のハードウェアによって実現されてもよい。また、本発明の一実施形態に係る画像処理装置は、例えばネットワークを介して接続された複数の装置によって構成されていてもよい。

#### 【0014】

図2は、本発明の一実施形態に係る画像処理装置100の機能構成例を示す。画像処理装置100は、機能構成として、取得部201、密度推定部202、座標推定部203、追跡部204、計測部205を有している。

#### 【0015】

取得部201は、複数の画像を含む動画像を取得する。取得部201が取得した動画像に対して、対象物の流れを計測するための解析が行われる。対象物の流れとは、画像上での対象物の流れであってもよいし、画像解析により推定される現実空間における対象物の流れであってもよい。なお、解析の対象物は特に限定されず、人物、自転車若しくはバイク等の乗り物、車若しくはトラック等の車両、又は家畜等の動物等が挙げられる。

#### 【0016】

動画像とは、例えば、ストリーミング、動画ファイル、フレーム毎に保存された一連の画像ファイル、又はメディアに保存された動画等であり、これらは複数のフレーム画像を含んでいる。複数の画像のそれぞれは、例えば、同一の場所にある撮像装置により異なる時刻に撮像されていてもよい。取得部201は、CMOSセンサー若しくはCCDセンサー等の固体撮像素子、又はこれらの固体撮像素子を備えるカメラ等の撮像装置から動画像を取得することができる。また、取得部201は、ハードディスク若しくはSSD等の記憶装置、又は記録メディア等から、動画像データを取得してもよい。

#### 【0017】

密度推定部202は、取得部201が取得した画像を用いて、流れの計測対象である対

10

20

30

40

50

象物の、画像中での密度分布を推定する。密度分布とは、画像において、流量計測の対象物（例えば人物の頭部等）が存在すると推定される箇所を表現したものである。密度分布は、画像中で対象物が存在すると推定された領域を示すことができる。密度分布は、例えば、対象物が存在する可能性が高い位置を中心とする、対象物の大きさに応じた一様分布であってもよい。また、密度分布は、対象物が存在する可能性が高い位置に極大を持つような任意の分布であってもよい。後者の例としては、例えば、ガウス分布が挙げられる。密度分布は、対象物が存在する可能性を示していてもよい。密度推定部 202 は、画像中の各位置における対象物の密度を表す密度マップを、密度分布を表す情報として出力することができる。なお、この密度マップの解像度は、画像と同じであってもよいし、画像より低くてもよい。

10

#### 【0018】

画像中で対象物が重なる場合に、密度推定部 202 は、それぞれの対象物に対応する密度分布が重ね合わされた密度分布を求めてもよい。この密度分布は、1つの対象物に対応する密度分布における密度の総和が1となるように規格化されていてもよい。この場合、密度推定部 202 が生成する密度分布における密度の総和は、画像中の対象物の数に対応する。

#### 【0019】

密度推定部 202 が密度分布を推定する方法としては、様々な方法を利用できる。例えば、画像、又は画像から抽出された特徴量、又はこれらの双方を、ニューラルネットワークに入力する方法が挙げられる。このようなニューラルネットワークとしては、例えば、畳み込みニューラルネットワーク、逆畳み込みニューラルネットワーク、これらのその双方が連結されたオートエンコーダー、及び U-Net 等のショートカットを持つネットワーク、等が挙げられる。密度推定部 202 が用いるニューラルネットワークは、画像が入力されると対象物の密度分布を出力するように、予め学習されている。このような学習は、学習用の画像と、上記のような画像の密度分布を表す教師データと、を用いて行うことができる。例えば、ニューラルネットワークからの出力値の総和が対象物の数となるように、すなわち出力値が画像の各位置の対象物の密度を表すように、学習を行うことができる。

20

#### 【0020】

座標推定部 203 は、密度推定部 202 により得られた密度分布から、画像中の対象物の位置を推定する。例えば、座標推定部 203 は、推定された対象物の位置の座標を出力することができる。対象物の位置の座標とは、画像における対象物の代表点の座標である。対象物の代表点は、例えば、対象物である人物の頭部中心であってもよい。

30

#### 【0021】

座標推定部 203 が位置を推定する方法は特に限定されず、例えば密度分布の重心を対象物の位置として推定してもよい。一方で、密度推定部 202 が出力する1つの対象物に対応する密度分布の境界はぼやけることがあり、また、この境界は様々な形状を有している。さらに、1つの対象物に対応する密度分布同士が重なり合うことにより、密度分布が複雑な形状を有する可能性がある。さらには、特定の状況で密度推定部 202 が想定外の出力をすることもあり得る。一実施形態においては、座標推定部 203 はニューラルネットワークを用いて対象物の位置を推定する。ニューラルネットワークを用いることにより、入力される密度分布にかかわらず、事前に定義された一定の演算量で位置の推定結果を得ることができる。このため、座標推定部 203 は、密度推定部 202 が出力する様々な密度分布に対応可能である。

40

#### 【0022】

座標推定部 203 が用いるニューラルネットワークは、例えば、密度分布を表す密度マップが入力されると、密度マップの各位置についての2クラス分類の推定結果を確率で表す、位置マップを出力することができる。座標推定部 203 は、このように、画像の各位置について対象物の代表点を表すかどうかを識別する2クラス分類を行うことができる。この2クラス分類においては、密度マップの各位置が対象物の位置座標を表すなら推定結

50

果は1に、対象物の位置座標を表さないなら推定結果は0となる。位置マップ701は、密度マップの各位置について、対象物の位置座標を表す確率を表すことができる。このような位置マップはほとんどの要素が0であるスパースなものであるため、スパースな出力を扱える2クラス分類問題に帰着させることで、このような位置マップを生成するニューラルネットワークの学習が容易となる。

#### 【0023】

このようなニューラルネットワークとしては、例えば、畳み込みニューラルネットワーク、逆畳み込みニューラルネットワーク、これらのその双方が連結されたオートエンコーダー、及びU-Net等のショートカットを持つネットワーク、等が挙げられる。また、座標推定部203が用いるニューラルネットワークは、0以上1以下の値域を持つ出力層を持つことができる。このような出力層としては、例えば、ステップ関数、シグモイド関数、又はソフトマックス関数等が挙げられる。このような出力層を持つニューラルネットワークは、2クラス分類問題に適した出力を与えることができる。

10

#### 【0024】

座標推定部203が用いるニューラルネットワークは、密度マップが入力されると対象物の位置を出力するように、予め学習されている。このような学習は、例えば、学習用の密度マップと、人物の位置に対応する要素が1であり、それ以外の要素が0である、密度マップに対応する位置マップを表す教師データと、を用いて行うことができる。

#### 【0025】

計測部205は、座標推定部203により推定された、複数の画像中の対象物の位置から、対象物の流れを計測する。計測部205は、例えば、対象物の流量、又は対象物の流れの向きなどを計測することができる。流量とは、例えば、あらかじめ設定された計測領域又は計測線を通して対象物の数であってもよい。例えば、計測部205は、計測領域によって区切られる一方の領域から他方の領域へと移動した対象物の数を計測することができる。計測領域として線分（計測線）を用いる場合、計測部205は、計測線を交差した対象物の数を計測することができる。

20

#### 【0026】

このような計測領域は、撮影シーンに基づき予め決められていてもよい。この場合、計測部205は、予め決められた計測領域の設定値を記憶装置12等のメモリから取得してもよい。また、計測部205は、出力装置15に表示された画像を参照しながら入力装置14に接続されたヒューマンインターフェースデバイス等を実行することによってユーザが入力した設定値を取得してもよい。

30

#### 【0027】

計測領域は、1つであってもよいし、複数であってもよい。また、計測領域の形状は特に限定されず、例えば、任意の折れ線、曲線、多角形、円形、若しくは楕円形、又は任意の閉曲線からなる図形であってもよい。また、計測領域は、太さのない線であってもよいし、太さを有する線（又は長い矩形）であってもよい。このような計測領域は、線又は矩形等の任意のガイド図形を基準に設定されてもよい。例えば、ガイド図形に従う形状の計測領域が、画像中の特定の位置に設けられてもよい。このようなガイド図形は1つであっても複数であってもよい。また、ガイド図形が、幅を大きくする等の方法により拡張されてもよい。

40

#### 【0028】

対象物の位置に基づいて対象物の流れを計測する方法は特に限定されないが、本実施形態において計測部205は、追跡部204による対象物の追跡結果を用いて計測を行う。追跡部204は、対象物の位置を追跡する。追跡部204は、座標推定部203により推定された対象物の位置に基づいて、対象物の位置を追跡することができる。例えば、追跡部204は、複数の画像のそれぞれにおいて推定された対象物の位置に基づいて、同一と推定される対象物について、時間にわたる位置の変化を追跡することができる。

#### 【0029】

一実施形態において追跡部204は、座標推定部203が出力した対象物の位置座標に

50

基づいて、対象物の位置座標を追跡することにより、対象物の移動軌跡を得る。追跡部 204 は、複数の画像のうちの第 1 の時刻における第 1 の画像中の対象物の位置と、複数の画像のうちの第 2 の時刻における第 2 の画像中の対象物の位置と、を対応付けることができる。このような対応付けを繰り返すことにより、追跡部 204 は対象物の移動軌跡を得ることができる。

#### 【0030】

例えば、追跡部 204 は、複数の異なる時刻における画像中の対象物の位置座標リストを用いて追跡を行うことができる。例えば、密度推定部 202 及び座標推定部 203 は、同じ場所を同じ画角で異なる時刻に撮像することにより得られた複数の画像のそれぞれに対して上記の処理を行うことにより、それぞれの画像に対応する対象物の位置座標リストを生成することができる。追跡部 204 は、異なる時刻間で、位置座標リストに示される対象物の位置座標をマッチングすることにより、対象物の移動軌跡を得ることができる。この場合、計測部 205 は、例えば移動軌跡が計測線と交差しているかどうかを判定することにより、対象物の流量を計測することができる。具体的な処理例については図 3 を参照しながら後述する。

#### 【0031】

本実施形態に係る画像処理装置 100 の処理例を、図 3 を参照して説明する。以下では、人物の流量計測を行う例について説明する。具体的には、あらかじめ設定された計測領域を通過する人数が計測される。

#### 【0032】

ステップ S301 において取得部 201 は動画像を取得する。また、取得部 201 は、流量計測に用いられる複数の画像を動画像から取得する。例えば、取得部 201 は、計測開始時刻から計測終了時刻までの間の複数回の撮像により得られた複数の画像を取得することができる。なお、取得部 201 は、撮像装置から各フレーム画像を順次取得してもよい。以下では、ステップ S301 で取得される複数の画像のそれぞれを入力画像と呼ぶ。ステップ S302 及び S303 の処理は、入力画像のそれぞれについて行われる。

#### 【0033】

取得部 201 は、入力画像から、1 つ又は複数の部分画像を抽出してもよい。取得部 201 は、複数の画像のそれぞれから、画像の一部である 1 以上の部分画像を抽出することができる。このように入力画像よりも小さいサイズを有する部分画像を用いることで、ステップ S302 以降の処理量を減らすことができる。このような構成により、リアルタイムでの流量計測が容易となる。入力画像から抽出される部分画像の領域は、あらかじめ設定されていてもよい。また、部分画像の領域はユーザにより設定されてもよい。例えば、ユーザが入力装置 14 に接続したヒューマンインターフェースデバイス等をユーザが操作することにより、出力装置 15 に表示された入力画像上の部分領域を設定することができる。部分画像の領域の設定方法及び具体的な形状は特に限定されない。例えば、取得部 201 は、ユーザが指定した領域に基づいて、抽出される部分領域を自動的に決定してもよい。取得部 201 は、既に説明したように計測領域のユーザ指定を取得することができる。この場合、取得部 201 は、指定された計測領域に基づいて部分画像の領域を自動的に設定してもよい。例えば、取得部 201 は、計測領域を含むように入力画像から部分画像を抽出することができる。

#### 【0034】

また、取得部 201 は、図 4 (A) のように、入力画像 400 の全体を複数の領域へと分割してもよい。この場合、複数の領域から選択された 1 以上の領域のそれぞれから部分画像を抽出してもよい。例えば、図 4 (A) に示すように入力画像 400 に計測領域 401 を設定した場合、計測領域 401 を包含する、太枠で囲まれた矩形領域から、流量計測に使用する部分画像を抽出することができる。また、図 4 (B) のように、入力画像 402 の一部に、計測領域 403 を包含する 1 つ又は複数の矩形領域を設定することができる。そして、こうして設定された矩形領域から、流量計測に使用する部分画像を抽出することができる。

10

20

30

40

50



## 【 0 0 3 5 】

図 4 ( A ) ( B ) において部分画像の形状は矩形であるが、部分画像の形状は、例えば、任意の多角形、円形、若しくは楕円形、又は任意の閉曲線からなる図形であってもよい。また、計測領域の形状が線分ではなく、例えば、任意の折れ線、曲線、多角形、円形、若しくは楕円形、又は任意の閉曲線からなる図形である場合にも、同様の手法で部分画像を抽出することができる。

## 【 0 0 3 6 】

ところで、カメラ等の撮像装置により得られた入力画像内では、位置により計測対象となる人物の映る大きさが異なることがある。したがって、画像内の位置に応じて部分画像の大きさを変えることができる。例えば、部分画像の大きさと人物の大きさとの比がほぼ一定となるように、部分画像の大きさを決定し、この大きさに応じて部分画像を抽出することができる。ここで、人物の大きさとは、例えば、人物の頭部と肩を包含する部分の大きさ、人物の頭部の大きさ、人物の全身の大きさ等でありうる。図 4 ( A ) ( B ) はこのように部分画像の大きさを変化させる例を示す。すなわち、図 4 ( A ) ( B ) においては、すなわち遠くの人物が映る領域、すなわち入力画像の上方において、部分画像の大きさが小さくなっている。

10

## 【 0 0 3 7 】

このように部分画像を設定した場合、ステップ S 3 0 2 及び S 3 0 3 の処理は、入力画像中の部分画像のそれぞれについて行うことができる。すなわち、密度推定部 2 0 2 は、それぞれの部分画像中の対象物の密度分布を推定し、座標推定部 2 0 3 は、それぞれの部分画像についての密度分布から対象物の位置を推定することができる。一方で、ステップ S 3 0 2 及び S 3 0 3 における推定結果の精度は、部分画像の境界領域（周辺領域）において低くなる傾向がある。このため、一実施形態においては、部分画像中に中央領域を規定し、この中央領域から検出された人物に関して、ステップ S 3 0 5 における流量計測を行うことができる。このような構成によれば、精度の高い人物検出結果を用いて流量計測が行われるため、計測精度が向上する。

20

## 【 0 0 3 8 】

部分画像に中央領域を規定する方法としては、例えば図 5 ( A ) に示すように、部分画像 5 0 0 の内側に、中央領域 5 0 1、及び部分画像 5 0 0 の境界線と中央領域 5 0 1 の境界線とに囲まれたマージン領域 5 0 2 を規定する方法が挙げられる。図 5 ( A ) の例において、中央領域 5 0 1 は矩形であるが、中央領域 5 0 1 の形状は特に限定されない。例えば、中央領域 5 0 1 の形状は、多角形、円形、楕円形、又は任意の閉曲線など、部分画像 5 0 0 の内部に含まれる任意の形状でありうる。

30

## 【 0 0 3 9 】

マージン領域の大きさは、流量計測の対象となる人物の大きさに合わせて設定することができる。例えば、マージン領域を、人物の検出に必要な部分を含むことができる大きさに設定することができる。図 5 ( B ) はその一例を示す。図 5 ( B ) では、部分画像 5 0 0 のマージン領域 5 0 2 が、流量計測の対象となる人物 5 0 3 又は人物 5 0 4 の検出に必要な部分を含むことができる大きさに設定されている。より具体的には、人物の頭部と肩を含むことができるように、マージン領域の大きさが設定されている。他の例としては、人物の頭部のみ、あるいは人物の全身がマージン領域に含まれるように、マージン領域を設定してもよい。

40

## 【 0 0 4 0 】

入力画像から複数の部分画像を抽出する場合、隣接する部分画像が互いに重なるように部分画像の領域を設定することができる。例えば、取得部 2 0 1 は、画像を複数の領域に分割することができる。そして、取得部 2 0 1 は、複数の領域から選択された 1 以上の領域のそれぞれについて、領域（中央領域）及びこの領域の周辺にあるマージン領域から、部分画像を抽出することができる。図 5 ( C ) はそのような分割の一例を示す。図 5 ( C ) においては、隣り合う 2 つの部分画像 5 0 5 及び部分画像 5 0 6 が互いに部分的に重なり合っている。より具体的には、部分画像 5 0 5 の中央領域 5 0 7 と、部分画像 5 0 6 の

50

中央領域 5 0 8 との間に隙間ができないように、部分画像 5 0 5 及び部分画像 5 0 6 が抽出されている。検出対象である人物 5 1 0 が部分画像 5 0 5 のマージン領域 5 1 2 にある場合であっても、部分画像 5 0 5 に隣接する部分画像 5 0 6 の中央領域 5 0 8 から人物 5 1 0 を検出できる。

#### 【 0 0 4 1 】

図 5 ( C ) に示されるように、各部分画像の中央領域同士を隙間なく隣接させることにより、いずれかの部分画像の中央領域で計測対象となる人物を検出することができるため、計測精度が向上する。このような部分画像の抽出方法としては、入力画像を複数の領域に分割し、それぞれの領域 ( 中央領域に相当 ) 及び周りのマージン領域から、1 つの部分画像を抽出する方法が挙げられる。さらに、各部分画像のマージン領域を確保しつつ、必要最小限の面積で計測領域をカバーできる分割方法を、最適化計算等で決定してもよい。図 5 ( D ) においては、入力画像 5 2 0 に設定された計測領域 5 2 1 に沿って、それぞれがマージン領域を有する複数の部分画像 5 2 2 が抽出されている。

10

#### 【 0 0 4 2 】

以下の説明において、入力画像は、入力画像そのもの、又は部分画像を指すものとする。すなわち、ステップ S 3 0 2 及び S 3 0 3 の処理は、入力画像に対して、又は入力画像中の部分画像のそれぞれに対して、行うことができる。

#### 【 0 0 4 3 】

ステップ S 3 0 2 において、密度推定部 2 0 2 は、上述のように、取得部 2 0 1 が取得した入力画像を元に計測対象となる人物の密度分布を推定する。図 6 ( A ) の例では、密度推定部 2 0 2 は、入力画像 6 0 0 における人物 6 0 1 の頭部位置の密度分布を推定する。密度推定部 2 0 2 は、図 6 ( B ) に示すように、人物 6 0 1 の頭部位置の密度分布 6 0 3 を示す密度マップ 6 0 2 を生成している。

20

#### 【 0 0 4 4 】

なお、密度推定部 2 0 2 が用いる入力画像は 1 つに限定されず、密度推定部 2 0 2 は複数の入力画像を用いて密度分布を推定してもよい。例えば、密度推定部 2 0 2 は、複数の入力画像をニューラルネットワークに入力することにより、密度分布を推定してもよい。密度推定部 2 0 2 が用いる複数の入力画像は、例えば、動画画像に含まれる、複数の異なるフレームの画像であってもよい。また、密度推定部 2 0 2 が出力する密度分布も 1 つには限定されない。例えば密度推定部 2 0 2 は、複数の入力画像にそれぞれ対応している、複数の密度分布を出力してもよい。密度推定部 2 0 2 が複数の密度分布を出力する場合、後続する座標推定部 2 0 3 には全ての密度分布が入力されてもよいし、選択された一部の密度分布が入力されてもよい。

30

#### 【 0 0 4 5 】

ステップ S 3 0 3 において座標推定部 2 0 3 は、上述のように、密度推定部 2 0 2 が出力した密度分布を元に人物の位置座標を推定する。例えば座標推定部 2 0 3 は、図 7 ( A ) に示す密度マップ 7 0 0 をニューラルネットワークに入力することにより、図 7 ( B ) に示す位置マップ 7 0 1 を出力することができる。位置マップ 7 0 1 は、人物の位置座標が、位置マップ 7 0 1 における要素の位置と一致している場合を示す。この場合、1 人の人物について 1 つの要素の値 ( 確率値 ) が大きくなるため、このような要素の位置を人物の位置座標として判定することができる。

40

#### 【 0 0 4 6 】

一方で、図 7 ( C ) に示す位置マップ 7 0 2 のように、人物の位置座標が、位置マップ 7 0 2 における要素の位置 ( すなわち格子点の位置 ) からずれる場合、1 人の人物について複数の要素の値 ( 確率値 ) が大きくなる。このように、入力画像の互いに隣接する複数の位置が、対象物の位置を表すと推定されることがある。この場合、座標推定部 2 0 3 は、対象物の位置を表すと推定された、入力画像の互いに隣接する複数の位置に基づいて、対象物の位置を求めることができる。例えば、座標推定部 2 0 3 は、複数の要素に分かれている確率値を 1 つにまとめることにより、人物の位置座標を決定することができる。

#### 【 0 0 4 7 】

50

例えば図8(A)~(C)に示すように、座標推定部203は、ニューラルネットワークから出力された位置マップ800に対して $3 \times 3$ のmax poolingを適用することによってマップ801を得ることができる。そして、座標推定部203は、位置マップ800と、マップ801とを、要素ごとに比較することができる。この場合、座標推定部203は、マップ802に示されるように、位置マップ800とマップ801との間で値が一致している要素の座標を、人物の位置座標として判定することができる。max poolingのカーネルサイズは $3 \times 3$ に限定されず、 $5 \times 5$ などの別のサイズであってもよい。また、max poolingの処理は、単独で行われてもよいし、座標推定部203が用いるニューラルネットワークの出力層の次にmax pooling層を追加することにより行われてもよい。

10

#### 【0048】

複数の確率値をまとめる方法はこの方法には限定されず、座標推定部203は、例えばwatershed等の、他の極大値検出法を用いてもよい。また、座標推定部203は、確率値が検出された複数の要素をクラスタリングし、各クラスタの中心座標を人物の位置座標とする方法を用いてもよい。クラスタリングの方法としては、例えば、階層クラスタリング又はK平均法などが挙げられる。クラスタの中心座標は、クラスタ内の要素座標の単純平均でもよいし、クラスタ内の要素のうち最大の確率値を持つ要素座標でもよいし、又は、各要素の要素座標の存在確率等による重みづけ平均値でもよい。

#### 【0049】

なお、座標推定部203が用いるニューラルネットワークが出力する位置マップにおいては、人物の位置座標に対応しない要素が、ノイズである微小な確率値を持っているかもしれない。このようなノイズがノイズが乗っている状態で、図8の例に示すように隣接する要素の確率値を1点にまとめる処理を行うと、微小なノイズを持つ要素の座標が人物の位置座標と判定されてしまうかもしれない。このため、座標推定部203は、ニューラルネットワークが出力した位置マップにおいて、閾値以下の、ノイズとみなせるような微小な確率値をゼロに修正してから、確率値を1点にまとめる処理、及び人物の位置座標の判定を行ってもよい。

20

#### 【0050】

ステップS304において追跡部204は、座標推定部203が出力した位置座標を元に人物の位置座標を追跡し、人物の移動軌跡を得る。上記のように、ステップS302及びS303の処理を、異なる時刻に得られた入力画像のそれぞれについて行うことにより、複数の異なる時刻における入力画像に含まれる人物の位置座標リストを得ることができる。図9は、時刻 $t_1$ における位置座標リスト900と、時刻 $t_1$ より後の時刻 $t_2$ における位置座標リスト901を示す。それぞれの位置座標リスト900、901には、入力画像から検出された人物の位置が示されている。追跡部204は、位置座標リスト900と位置座標リスト901との間で、人物の位置のマッチング902(対応付け)を行うことにより、時刻 $t_1$ から時刻 $t_2$ までの人物の移動軌跡903を得ることができる。

30

#### 【0051】

マッチング方法は特に限定されない。追跡部204は、例えば、マッチング結果に基づいて算出されるコストの総和が小さくなるように、マッチングを行うことができる。このようなマッチングを行うためのアルゴリズムとしては、例えば、ハンガリアンマッチング、線形計画法、又はニューラルネットワーク等が挙げられる。参照することができるコストの例としては、人物間の距離等が挙げられる。また、入力画像間での人物の画像情報の類似度に基づくコストを用いることもできる。さらに、1つの種類のコストが用いられてもよく、複数の書類のコストが組み合わせて用いられてもよい。例えば、人物間の距離と、入力画像間での人物の画像情報の類似度と、の双方に基づいてコストが算出されてもよい。追跡部204は、このようなコストが最小となるようにマッチングを行うことができる。

40

#### 【0052】

以下では、図10を参照して、人物間の距離をコストとして用いる例について説明する

50

。図10(A)に示される入力画像1000には、人物の位置座標A, B, Cが示されている。また、入力画像1000の前フレームの画像からは、人物の位置座標b, c, dが検出されている。なお、この例においては、位置座標bと位置座標Bは同じ人物の位置座標であり、位置座標cと位置座標Cも同じ別の人物の位置座標である。したがって、位置座標の組(b, B)及び(c, C)を含むマッチング結果が正解である。

#### 【0053】

一方で、追跡部204が、人物間の距離の総和が最小になるようにマッチングを行うと、位置座標の組(c, B)及び(b, C)が得られる可能性がある。そこで、追跡部204は、距離bBに対応するコストと距離cCに対応するコストとの和が、距離cBに対応するコストと距離Cbに対応するコストとの和より小さくなるように、コストを定義することができる。一実施形態において、距離dに対するコストは、距離2dに対応するコストの半分よりも小さい。例えば、長い距離Cbをより強調するような距離の尺度を用いてコストを定義することができる。このような距離の尺度としては、平方ユークリッド距離等が挙げられる。

#### 【0054】

また、入力画像1000の位置座標Aに映っている人物は、前フレームの撮像時には撮像範囲外の位置座標aに存在しており、撮像範囲内に移動してきた。さらに、前フレームの画像において位置座標dに映っていた人物は、入力画像1000の撮像時には撮像範囲外の位置座標Dに移動していた。このような場合、位置座標A, dとマッチングされるべき位置座標a, Dは、入力画像1000及び前フレームの入力画像からは検出されない。このような例において、入力画像1000の撮像範囲内から検出された全ての位置座標を、人物間の距離の総和が最小になるようにマッチングすると、正しくないマッチング結果が得られる可能性がある。例えば、図10(A)の例では、位置座標の組(A, b)、(B, c)、及び(C, d)が得られるかもしれない。

#### 【0055】

一実施形態において、追跡部204は、位置座標A, dのような孤立した位置座標がマッチングされないように、マッチングを行うことができる。例えば、マッチングの最大距離rを設定することができる。この場合、追跡部204は、最大距離rよりも離れている位置座標同士がマッチングしないように、マッチングを行うことができる。

#### 【0056】

図10(B)(C)は、孤立した位置座標がマッチングされないようにするための、ハンガリアンマッチングで用いるコスト行列及びマッチング結果の例を示す。図10(B)において、B, C, b, cは、それぞれ図10(A)に示す位置座標B, C, b, cに対応する。位置座標A, dは、自分以外の全ての位置座標との距離が最大距離rを越えている、孤立した位置座標である。このため、位置座標A, dが他の位置座標とマッチングされないように、位置座標A, dに関するコストは図10(B)のコスト行列から除外されている。追跡部204は、このようなコスト行列を、例えば平方ユークリッド距離に基づいて生成することができる。図10(B)の例では、コストの総和が最小となるように、太枠で示される位置座標の組(B, b)及び(C, c)が得られている。

#### 【0057】

図10(C)に示す別の例において、A, B, C, b, c, dは、それぞれ図10(A)に示す位置座標A, B, C, b, c, d(以下実座標と呼ぶ)に対応する。図10(C)の例では、追跡部204は、最初の実座標同士のコスト行列を、例えば平方ユークリッド距離に基づいて生成する。続いて、追跡部204は、入力画像1000から検出された実座標である位置座標A, B, Cと同じ数である3つのダミー、すなわちx1, x2、及びx3に対応する列を追加する。さらに、追跡部204は、前フレームの画像から検出された実座標である位置座標b, c, dと同じ数である数3つのダミー、すなわちX1, X2、及びX3に対応する行を追加する。そして、追跡部204は、ダミーと任意の実座標との間のコストを、最大距離rよりも大きい値r<sub>1</sub>を用いて設定し、距離がrよりも大きい実座標同士の間のコストを、r<sub>1</sub>よりもさらに大きいr<sub>2</sub>を用いて設定する。すなわち

10

20

30

40

50

、 $r < r_1 < r_2$ である。追跡部204は、このようなコスト行列を、例えば平方ユークリッド距離に基づいて生成することができる。このようなコスト行列を用いてマッチングを行うと、孤立した実座標は他の実座標とはマッチングせず、ダミーとマッチングする。また、余ったダミーはダミー同士でマッチングする。図10(C)では、コスト行列を用いたマッチングの結果が太枠で示されている。追跡部204は、これらのダミーとのマッチング結果を除外することにより、マッチング結果を得ることができる。

【0058】

さらなる別の例において、追跡部204は、マッチングを行う前に、最大距離 $r$ の範囲内にある位置座標同士をグルーピングしてもよい。この場合、追跡部204は、グループごとに、グループに含まれる位置座標同士のマッチングを行うことができる。例えば、追跡部204は、各グループについて個別に図10(B)又は図10(C)に示すようなコスト行列を生成し、マッチングを行うことができる。図10(A)に示されるような孤立した位置座標A、dは、このようなグルーピングにより取り除かれる。このような方法によれば、人物の数、すなわち検出された位置座標の数が多い場合に、処理速度を向上させることができる。グルーピングの方法は特に限定されないが、例えば、孤立した位置座標を含むコスト行列又は隣接グラフを用いて連結されていない独立したグラフを探索する方法が挙げられる。また、階層クラスタリング又はK平均法等のクラスタリング法を用いることもできる。

10

【0059】

上記のように、追跡部204は、入力画像から検出された人物の位置座標と、前の時刻に撮像された画像から検出された人物の位置座標と、をマッチングさせることができる。一方で、追跡部204は、以前に得られたマッチング結果を参照して、このマッチングを行ってもよい。

20

【0060】

一実施形態において、追跡部204は、時刻 $t_1$ における位置座標と、時刻 $t_2$ における位置座標と、のマッチング結果を参照して、時刻 $t_2$ における位置座標と、時刻 $t_2$ より後の時刻 $t_3$ における位置座標と、をマッチングさせる。例えば図11の例では、追跡部204は、時刻 $t_1$ と時刻 $t_2$ との間での位置座標のマッチング結果1100と、時刻 $t_3$ における位置座標リスト1101と、のマッチング1102を行っている。このような処理により、時刻 $t_1$ から時刻 $t_2$ を経て時刻 $t_3$ に至る人物の移動軌跡1103が得られる。このような処理をさらに後の時刻における位置座標に対して繰り返すことで、長時間にわたる同一人物の移動軌跡を得ることができる。

30

【0061】

この際に追跡部204は、時刻 $t_2$ における位置座標、及び時刻 $t_3$ における位置座標に加えて、時刻 $t_2$ における人物の移動方向又は移動速度を用いて、マッチングを行うことができる。すなわち、追跡部204は、時刻 $t_2$ における人物の移動方向又は移動速度と矛盾しない時刻 $t_3$ における位置座標を、時刻 $t_2$ における位置座標と優先的にマッチングさせることができる。

【0062】

例えば追跡部204は、少なくとも時刻 $t_2$ における入力画像中の対象物の位置に基づいて、時刻 $t_3$ における入力画像中の対象物の位置を予測することができる。図12(A)は、時刻 $t_2$ と時刻 $t_3$ の間での人物の位置座標の動きの例を示す。この例では、時刻 $t_2$ において位置座標b、cが、時刻 $t_3$ において位置座標B、Cが、それぞれ検出されている。この例では、人物同士が接近しているため、誤った(B、c)というマッチング結果が得られる可能性がある。追跡部204は、図12(B)に示すように、時刻 $t_3$ における人物の予測位置座標 $b^*$ 、 $c^*$ を用いてマッチングを行うことにより、マッチング精度を向上させることができる。

40

【0063】

追跡部204は、時刻 $t_3$ における人物の予測位置座標 $b^*$ 、 $c^*$ を、時刻 $t_3$ より前の時刻における移動軌跡から求めることができる。例えば、追跡部204は、時刻 $t_2$ まで

50

のマッチング結果に基づいて得られた移動軌跡を用いて、人物の予測位置座標  $b^*$  ,  $c^*$  を求めることができる。図 12 ( B ) の例では、時刻  $t_1$  における位置座標  $b'$  と時刻  $t_2$  における位置座標  $b$  とがマッチングされており、時刻  $t_1$  における位置座標  $c'$  と時刻  $t_2$  における位置座標  $c$  とがマッチングされている。この例において追跡部 204 は、時刻  $t_1$  から時刻  $t_2$  に至る移動軌跡 1201 から、時刻  $t_3$  の予測位置座標 1202 (  $b^*$  ,  $c^*$  ) を求めることができる。予測位置座標の計算方法は特に限定されない。例えば追跡部 204 は、時刻  $t_1$  及び時刻  $t_2$  の位置座標を用いて、線形外挿により時刻  $t_3$  の予測位置座標を求めることができる。

【0064】

そして、追跡部 204 は、上述の方法に従って、予測位置座標 1202 (  $b^*$  ,  $c^*$  ) と、検出された時刻  $t_3$  における位置座標 1203 (  $B$  ,  $C$  ) とのマッチング 1204 を行う。このように、追跡部 204 は、検出された時刻  $t_2$  における位置座標  $b$  ,  $c$  の代わりに、時刻  $t_3$  における予測位置座標  $b^*$  ,  $c^*$  を、検出された時刻  $t_3$  における位置座標  $B$  ,  $C$  とマッチングさせる。この方法によれば、位置座標  $b$  ,  $c$  よりも、予測位置座標  $b^*$  ,  $c^*$  の方が、位置座標  $B$  ,  $C$  と近接していることが期待されるので、マッチングの精度を向上させることができる。予測位置座標  $b^*$  ,  $c^*$  と位置座標  $B$  ,  $C$  とのマッチング結果に基づいて、追跡部 204 は、位置座標  $B$  ,  $C$  と位置座標  $b$  ,  $c$  とをマッチングさせることができる。こうして、時刻  $t_1$  から時刻  $t_2$  を経て時刻  $t_3$  に至る移動軌跡 1205 が得られる。

【0065】

なお、時刻  $t_3$  における予測位置座標の計算には、2つの時刻 ( 時刻  $t_1$  ,  $t_2$  ) における位置座標を用いる代わりに、時刻  $t_3$  より前の複数の時刻における位置座標を用いてもよい。

【0066】

何らかの理由で、取得部 201 が取得する入力画像が欠落する、密度推定部 202 による人物の密度分布が欠落する、又は座標推定部 203 が人物の位置座標を検出しないことがあるかもしれない。例えば図 13 ( A ) の例に示すように、時刻  $t_3$  における人物の位置 1304 が検出されず、時刻  $t_4$  で再検出されることがある。このように人物の移動軌跡が途切れている場合、人物は計測領域を表す線分 1301 及び 1302 を通過していないと判定されるかもしれない。このような事象を避けるため、図 13 ( B ) に示すように、追跡部 204 は、時刻  $t_3$  の予測位置座標 1305 を用いることで、途切れた移動軌跡を埋め合わせることができる。このような方法により、移動軌跡が途切れていても、人物が計測領域を通過したと判定することが可能となる。時刻  $t_3$  の予測位置座標 1305 は、上述のように求めることができる。

【0067】

例えば、追跡部 204 は、時刻  $t_3$  において人物が予測位置座標 1305 に存在すると予測し、予測位置座標 1305 を通る人物の移動軌跡を生成することができる。なお、追跡部 204 は、このように生成された移動軌跡が時刻  $t_3$  以降に検出された位置座標とマッチングしない場合、人物の移動軌跡から予測位置座標 1305 を除外してもよい。

【0068】

人物の移動軌跡が途切れた場合における、人物の位置座標の推定は上記の方法に限定されない。例えば、追跡部 204 は、テンプレートマッチング等の物体追跡方法を用いて、人物の位置座標を予測してもよい。さらに、人物の位置座標の不検出が 2 フレーム以上にわたる場合でも、予測位置座標の計算を繰り返すことで人物の移動軌跡を追跡することができる。

【0069】

さらに、追跡部 204 は、時刻  $t_2$  における人物の位置座標と、時刻  $t_4$  における人物の位置座標とをマッチングさせてもよい。この際に、追跡部 204 は、時刻  $t_2$  ( 及び時刻  $t_1$  等のより前の時刻 ) における人物の位置座標から、時刻  $t_4$  における人物の予測位置座標を求めてもよい。上記のように、この予測位置座標を時刻  $t_4$  において検出された

10

20

30

40

50

人物の位置座標とマッチングさせることにより、時刻  $t_2$  までの人物の移動軌跡を、時刻  $t_4$  において検出された人物の位置座標とマッチングさせることができる。この場合、時刻  $t_2$  において検出された位置座標から、時刻  $t_4$  において検出された位置座標へと向かう人物の移動軌跡を求めることができる。この場合において、追跡部 204 は、さらに時刻  $t_3$  における人物の予測位置座標をさらに求めてもよい。そして、追跡部 204 は、時刻  $t_2$  において検出された位置座標から、時刻  $t_3$  における予測位置座標を通して、時刻  $t_4$  において検出された位置座標へと向かう人物の移動軌跡を求めてもよい。

【0070】

なお、図9に示す例において、マッチングに用いられる位置座標リスト 900, 901 においては、それぞれの位置に対応する人物は特定されていない。もっとも、それぞれの位置に対応する人物が、画像識別処理等により識別されてもよい。この場合、追跡部 204 は、人物を識別する情報を参照して、異なる時刻における人物の位置座標をマッチングさせてもよい。

【0071】

S305において、計測部205は、追跡部204が出力した人物の移動軌跡を元に、人物の流量を計測する。例えば、計測部205は、追跡部204が出力した人物の移動軌跡が、設定した計測領域と交差したかどうかを判定することで、人物の流量を計測することができる。なお、計測領域の設定方法は上述のとおりである。図4(A)(B)を参照して説明したように、ステップS301において部分画像が抽出される領域を設定する際に、計測領域401, 403が設定されていてもよい。

【0072】

計測領域が線分である場合の人物の流量の計測方法例を、図14を参照して説明する。図14(A)には、時刻  $t_1$  における入力画像1400と、計測領域である線分1401が示されている。また、入力画像1400からは、人物1402の位置座標1403、及び人物1404の位置座標1405が検出されている。同様に、図14(B)には、時刻  $t_2$  における入力画像1406が示されている。入力画像1406からは、人物1402の位置座標1407が検出されており、位置座標1403から位置座標1407に至る移動軌跡1409が得られている。また、入力画像1406からは、人物1404の位置座標1408が検出されており、位置座標1405から位置座標1408に至る移動軌跡1410が得られている。

【0073】

図14(B)の例において計測部205は、移動軌跡1409は線分1401と交差していると判定することができ、したがって、人物1402は線分1401を通過したと判定することができる。同様に計測部205は、移動軌跡1410も線分1401と交差していると判定することができ、したがって、人物1404も線分1401を通過したと判定することができる。ここで、計測部205はさらに、人物1404は人物1402とは逆方向に線分1401を通過したと判定してもよい。移動軌跡と線分との交差判定は、例えば、線分と線分との交差判定等の幾何学的方法で行うことができる。

【0074】

一実施形態において計測部205は、人物が計測領域で分けられる一方の領域から他方の領域へと移動した場合に、人物は計測領域を通過したと判定する。一方で計測部205は、計測領域上に人物が存在する間は、人物は計測領域を通過したと判定しない。すなわち、計測部205は、一方の領域から計測領域を超えて他方の領域へと移動した人物の数を計測するが、こうして計測された人物の数は、一方の領域から計測領域へと移動したが他方の領域へと移動していない人物の数を含まない。

【0075】

例えば、図15には、計測領域である線分1500と、人物の移動軌跡1501とが示されている。この例において、計測部205は、a, d, fの3カ所において、人物が線分1500を下から上に通過したと判定することができ、c, eの2カ所において、人物が線分1500を上から下に通過したと判定することができる。一方で、計測部205は

、人物が線分 1 5 0 0 上にちょうど乗っている b , g においては、人物は線分 1 5 0 0 を通過していないと判定することができる。上述のように、線分 1 5 0 0 は太さを有する線分（すなわち矩形領域）であってもよく、この場合、人物が線分 1 5 0 0 上に存在する間は計測領域を通過したと判定されない。

【 0 0 7 6 】

このような判定を各人物の移動軌跡について行うことにより、計測部 2 0 5 は人物の流量を計測することができる。図 1 5 に示す例の場合、計測部 2 0 5 は、線分 1 5 0 0 の下から上への累積通過数と、上から下への累積通過数とを、それぞれの方向への人物の流量として計測することができる。ここで、図 1 5 を参照して説明したように、計測領域上に人物が存在する間には人物は計測領域を通過したと判定しないことにより、計測領域付近における人物の細かな動きに由来する、累積通過数の重複カウントを避けることができる。

10

【 0 0 7 7 】

計測部 2 0 5 は、各人物の計測領域の移動方向ごとの累積通過数の差分を、人物の流量として算出してもよい。具体的には、計測部 2 0 5 は、一方の領域から他方の領域へと通過した人物の数と、他方の領域から一方の領域へと通過した人物の数と、の差分を算出することができる。例えば、図 1 5 の場合、移動軌跡 1 5 0 1 で表される人物の、線分 1 5 0 0 の下から上への累積通過数は 3 回であり、上から下への累積通過数は 2 回である。計測部 2 0 5 は、移動方向ごとの累積通過数の差分を取ることで、線分 1 5 0 0 の下から上に 1 人の人物が通過したと判定することができる。

【 0 0 7 8 】

20

上述の通り、計測領域の形状は特に限定されない。例えば、計測領域が所定の領域を囲む閉曲線である場合、計測部 2 0 5 は、この領域へと入る人物の流量と、この領域から出る人物の流量とを計測することができる。

【 0 0 7 9 】

本実施形態によれば、画像中で対象物が重なっている場合であっても、まず密度分布が推定され、この密度分布から対象物の位置が推定される。このため、特許文献 1 のように画像からそれぞれの対象物を直接検出する場合と比較して、混雑したシーンにおける対象物の位置推定精度を向上させることができる。また、位置推定結果を示す位置マップはスパース性が高い（例えば、ほとんどの位置において値が 0 になる）ため、ニューラルネットワークを用いて画像から直接対象物の位置推定を行うことは容易ではない。一方、本実施形態によれば、密度マップの各位置についての 2 クラス分類問題を解くことで位置推定結果を得ることができるので、位置推定が容易となる。一実施形態においては、このような効果を得るために、取得部 2 0 1、密度推定部 2 0 2、及び座標推定部 2 0 3 がステップ S 3 0 1 ~ S 3 0 3 の処理を行う一方で、追跡部 2 0 4 及び計測部 2 0 5 が行うステップ S 3 0 4 ~ S 3 0 5 の処理を省略できる。

30

【 0 0 8 0 】

また、本実施形態によれば、対象物の位置座標がフレーム間で追跡されるため、精度よく対象物の流れを計測することができる。一実施形態においては、このような効果を得るために、追跡部 2 0 4 及び計測部 2 0 5 がステップ S 3 0 4 ~ S 3 0 5 の処理を行う一方で、取得部 2 0 1、密度推定部 2 0 2、及び座標推定部 2 0 3 が行うステップ S 3 0 1 ~ S 3 0 3 の処理を省略できる。

40

【 0 0 8 1 】

（その他の実施例）

本発明は、上述の実施形態の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1 以上の機能を実現する回路（例えば、ASIC）によっても実現可能である。

【 0 0 8 2 】

発明は上記実施形態に制限されるものではなく、発明の精神及び範囲から離脱することなく、様々な変更及び変形が可能である。従って、発明の範囲を公にするために請求項を

50



添付する。

【符号の説明】

【 0 0 8 3 】

1 0 0 : 画像処理装置、 2 0 1 : 取得部、 2 0 2 : 密度推定部、 2 0 3 : 座標推定部、  
2 0 4 : 追跡部、 2 0 5 : 計測部

10

20

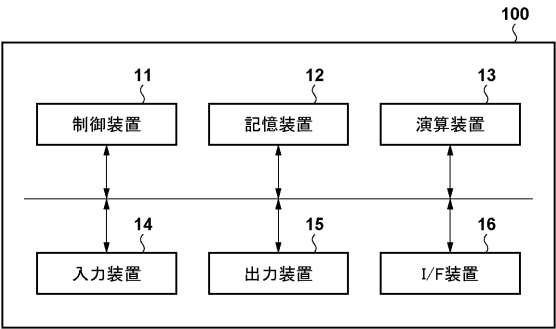
30

40

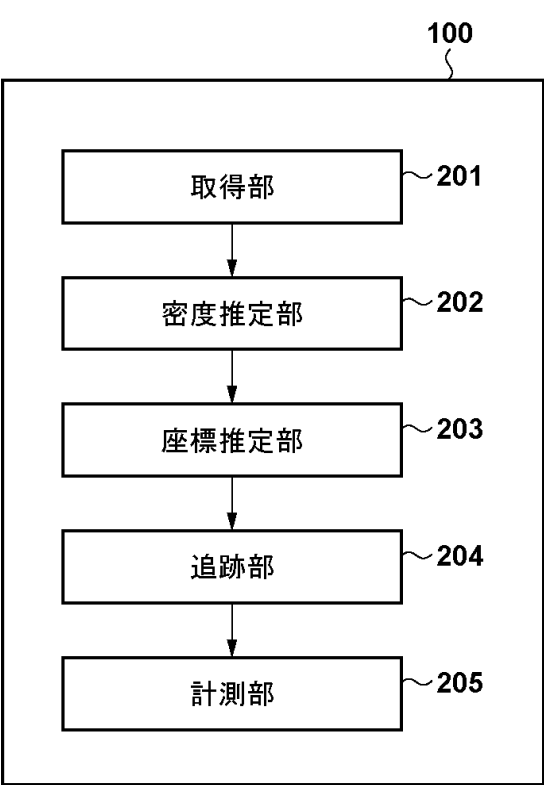
50

【 図 面 】

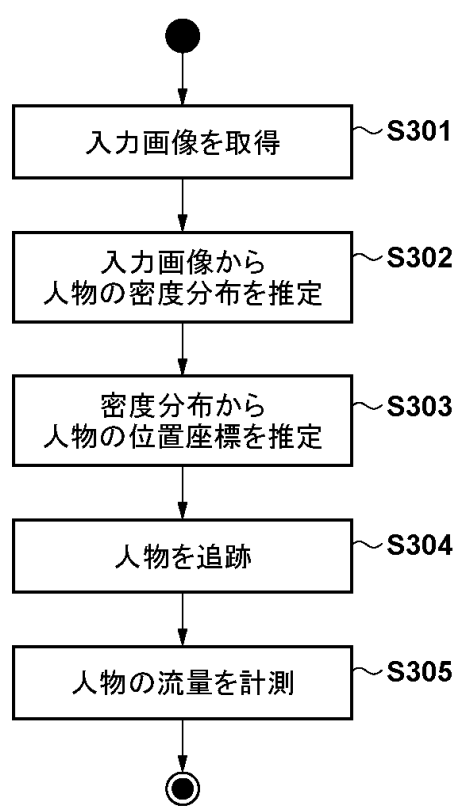
【 図 1 】



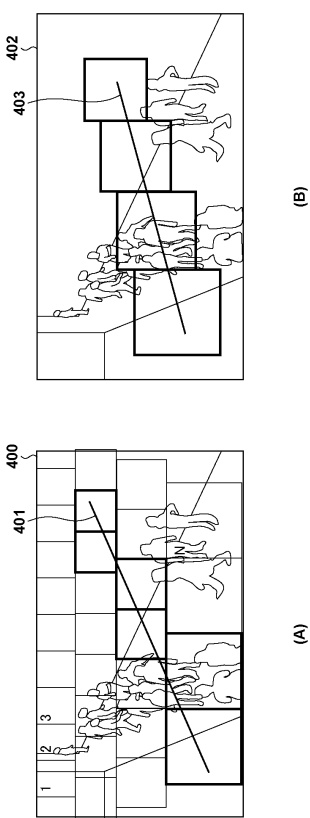
【 図 2 】



【 図 3 】



【 図 4 】



10

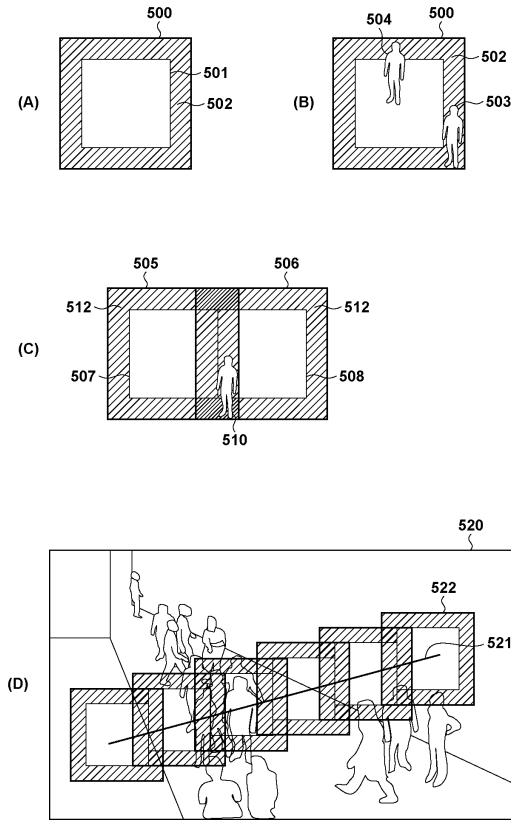
20

30

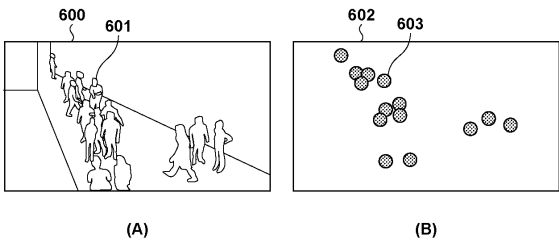
40

50

【図 5】



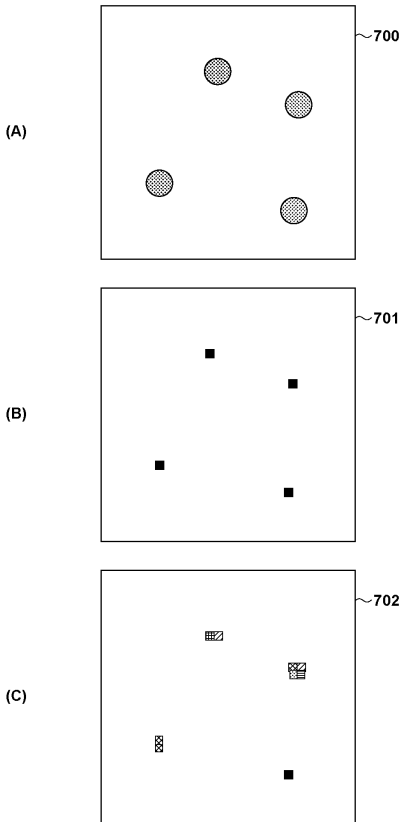
【図 6】



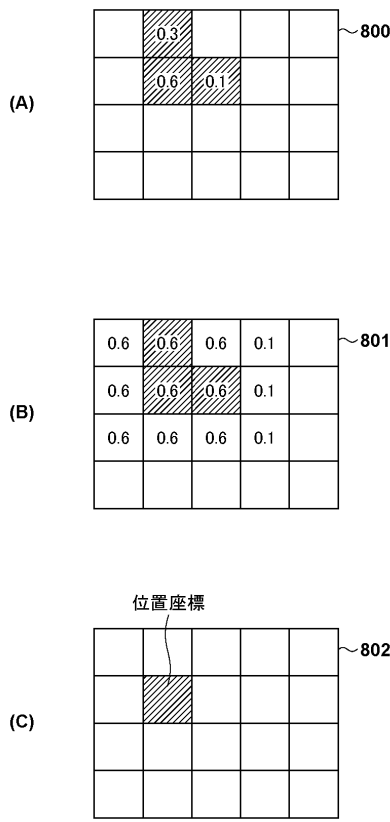
10

20

【図 7】



【図 8】

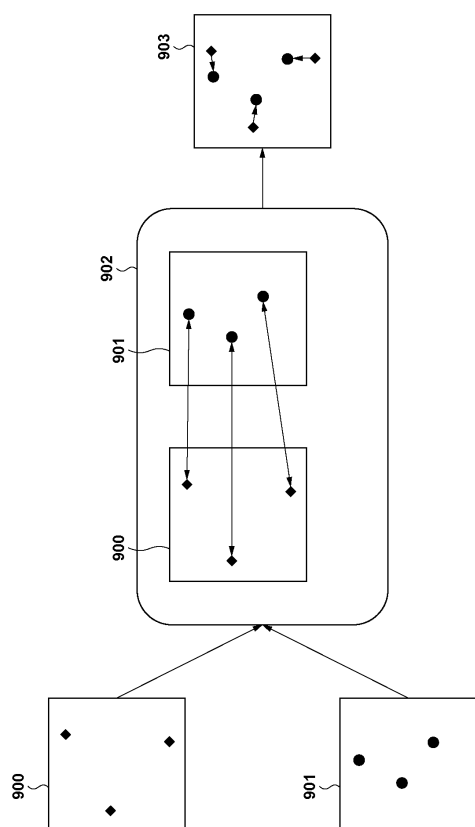


30

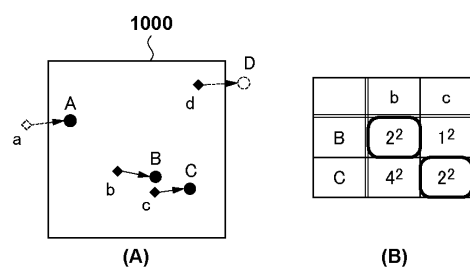
40

50

【 図 9 】



【 図 1 0 】



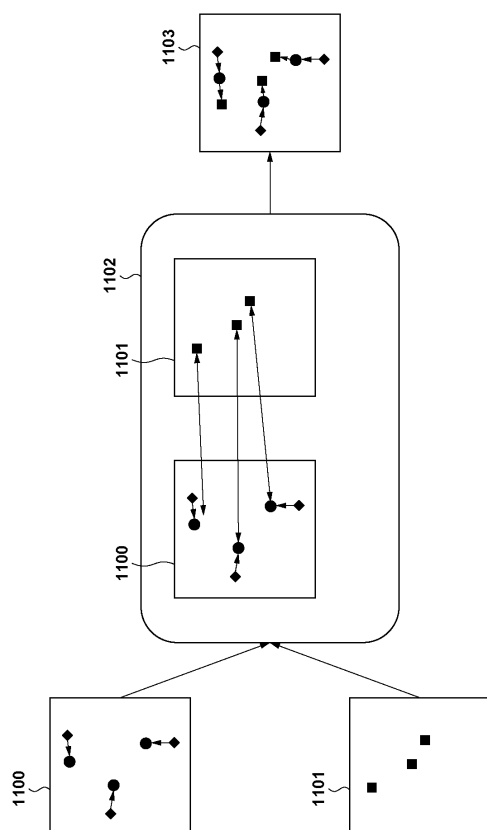
|   |       |       |
|---|-------|-------|
|   | b     | c     |
| B | $2^2$ | $1^2$ |
| C | $4^2$ | $2^2$ |

(C)

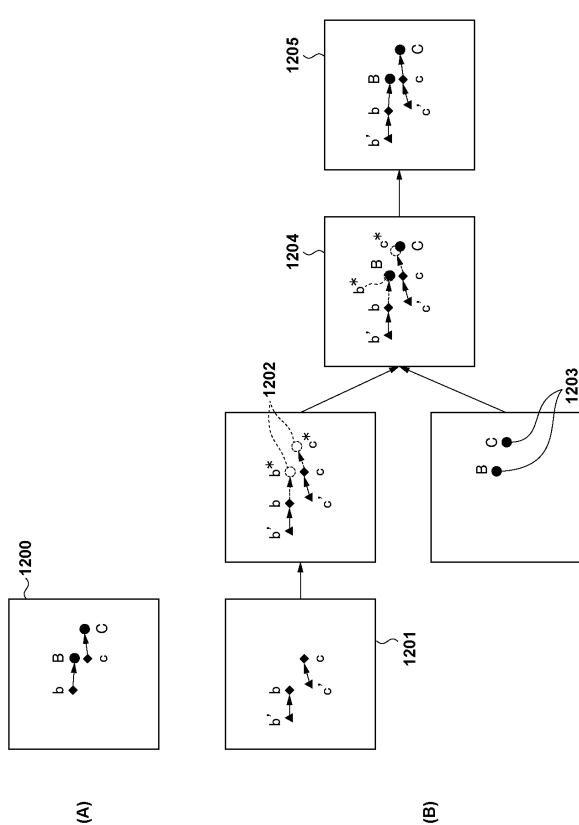
10

20

【 図 1 1 】



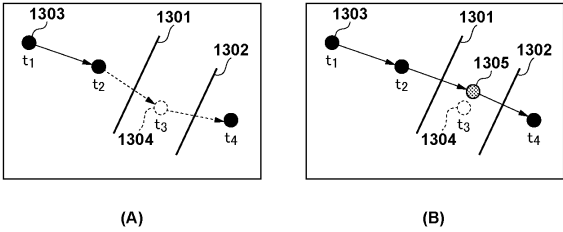
【图 1 2】



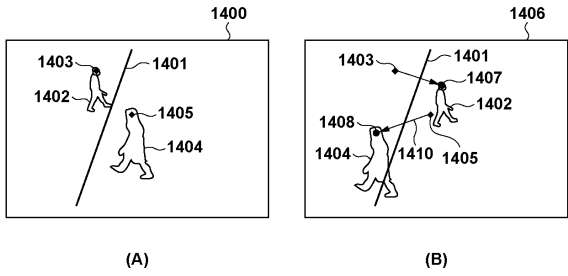
30

40

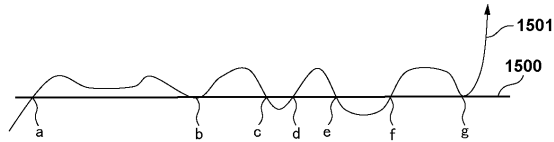
【 図 1 3 】



【 図 1 4 】



【 図 1 5 】



10

20

30

40

50

---

フロントページの続き

- (56)参考文献      特開 2 0 0 9 - 2 1 1 3 1 1 ( J P , A )  
                    中国特許出願公開第 1 1 0 2 1 0 6 0 3 ( C N , A )  
                    米国特許出願公開第 2 0 1 6 / 0 0 8 8 2 8 8 ( U S , A 1 )  
                    中国特許出願公開第 1 0 6 4 0 7 9 4 6 ( C N , A )  
                    Weihong Ren , 外 3 名 , Fusing Crowd Density Maps and Visual Object Trackers for People  
                    Tracking in Crowd Scenes , 2018 IEEE/CVF Conference on Computer Vision and Pattern R  
                    ecognition , 2018年06月18日 , pp. 5353-5362
- (58)調査した分野 (Int.Cl. , D B 名)
- |         |         |
|---------|---------|
| G 0 6 T | 7 / 0 0 |
| G 0 6 T | 1 / 0 0 |
| G 0 6 T | 1 / 4 0 |
| H 0 4 N | 7 / 1 8 |