

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-502516

(P2024-502516A)

(43)公表日 令和6年1月22日(2024.1.22)

(51)国際特許分類	F I	テーマコード(参考)
G 0 6 F 16/783(2019.01)	G 0 6 F 16/783	5 B 1 7 5
G 0 6 T 7/70(2017.01)	G 0 6 T 7/70	A 5 L 0 9 6

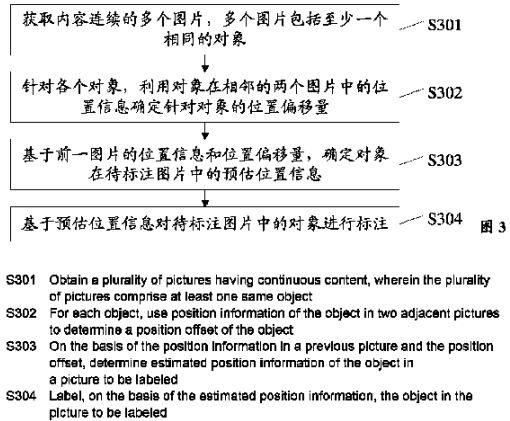
審査請求 有 予備審査請求 未請求 (全34頁)

(21)出願番号	特願2023-508558(P2023-508558)	(71)出願人	514322098
(86)(22)出願日	令和4年6月20日(2022.6.20)		ベイジン バイドゥ ネットコム サイエ
(85)翻訳文提出日	令和5年2月7日(2023.2.7)		ンス テクノロジー カンパニー リミテ
(86)国際出願番号	PCT/CN2022/099883		ッド
(87)国際公開番号	WO2023/103329		Beijing Baidu Netco
(87)国際公開日	令和5年6月15日(2023.6.15)		m Science Technolog
(31)優先権主張番号	202111489679.1		y Co., Ltd.
(32)優先日	令和3年12月8日(2021.12.8)		中華人民共和国 ベキン 100085,
(33)優先権主張国・地域又は機関	中国(CN)		ハイディアン ディストリクト, シャン
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		ディ テンス ストリート, 10番, バ イドゥ キャンパス 2階 2/F Baidu Campus, N o.10, Shangdi 10th Street, Haidian Dis trict, Beijing 1000 最終頁に続く

(54)【発明の名称】 データ注釈方法、装置、システム、デバイスおよび記憶媒体

(57)【要約】

本発明は、データ注釈方法、装置、システム、デバイスおよび記憶媒体を提供し、データ処理技術分野に関し、特に、人工知能、ビッグデータ、ディープラーニングなどの分野に関する。本発明の具体的な実現手段は、内容が連続する複数のピクチャーを取得し、複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含み、各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前において、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前において、1つ前のピクチャーに隣接するピクチャーであり、1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定し、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈することである。本発明によれば、複数のピクチャー



【特許請求の範囲】**【請求項 1】**

内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含むことと、

各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであることと、

前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定することと、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈することと、を含む、

データ注釈方法。

【請求項 2】

前記内容が連続する複数のピクチャーを取得することは、

少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものであること、を含む、

請求項 1 に記載の方法。

【請求項 3】

前記方法は、さらに、

少なくとも1つの動画データを取得することと、

各動画データについて、前記動画データをフレーム抽出し、前記動画データに対応するピクチャーを取得することと、

前記ピクチャーの時間順に前記ピクチャーを並べて、前記動画データに対応する時間順に並べた複数のピクチャーを取得することと、を含む、

請求項 2 に記載の方法。

【請求項 4】

前記の少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得することは、

注釈システムデータベースから少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、ここで、前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものであること、を含む、

請求項 2 に記載の方法。

【請求項 5】

前記方法は、さらに、

前記予測位置情報を展示することと、

前記予測位置情報を校正して得られた校正後の予測位置情報を受信することと、を含む、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈することは、

10

20

30

40

50

前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈すること、を含む、請求項 1 ~ 4 のいずれか一項に記載の方法。

【請求項 6】

内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも 1 つ含む、第 1 取得モジュールと、

各オブジェクトに対して、隣接する 2 つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する 2 つのピクチャーは 2 つ前のピクチャーと 1 つ前のピクチャーとを含み、前記 1 つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記 2 つ前のピクチャーは時系列で前記 1 つ前のピクチャーの前にあって、前記 1 つ前のピクチャーに隣接するピクチャーであり、前記 1 つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定する、確定モジュールと、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈する、注釈モジュールと、を含む、

データ注釈装置。

【請求項 7】

前記第 1 取得モジュールは、少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1 つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

請求項 6 に記載の装置。

【請求項 8】

少なくとも 1 つの動画データを取得する、第 2 取得モジュールと、

各動画データについて、前記動画データをフレーム抽出し、前記動画データに対応するピクチャーを取得する、フレーム抽出モジュールと、

前記ピクチャーの時間順に前記ピクチャーを並べて、前記動画データに対応する時間順に並べた複数のピクチャーを取得する、ソートモジュールとを、さらに含む、

請求項 7 に記載の装置。

【請求項 9】

前記第 1 取得モジュールは、注釈システムデータベースから少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、ここで、前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

請求項 7 に記載の装置。

【請求項 10】

前記予測位置情報を展示する、展示モジュールと、

前記予測位置情報を校正して得られた校正後の予測位置情報を受信する、受信モジュールと、をさらに含む、

前記注釈モジュールは、前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈する、

請求項 6 ~ 9 のいずれか一項に記載の装置。

【請求項 11】

10

20

30

40

50

注釈システムバックエンドと注釈システムフロントエンドとを含むデータ注釈システムであって、

前記注釈システムバックエンドは、内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含み、各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであり、前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定し、前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈し、

10

前記注釈システムフロントエンドは、前記注釈システムバックエンドが前記注釈対象ピクチャーにおける前記オブジェクトに対する注釈を展示する、

データ注釈システム。

【請求項12】

前記システムは、さらに、注釈システムデータベースを含み、

前記注釈システムバックエンドは、前記注釈システムデータベースから少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

20

請求項11に記載のシステム。

【請求項13】

前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

30

請求項12に記載のシステム。

【請求項14】

前記注釈システムバックエンドは、前記予測位置情報を展示し、前記予測位置情報を校正して得られた校正後の予測位置情報を受信し、前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈する、

請求項11～13のいずれか一項に記載のシステム。

【請求項15】

少なくとも1つのプロセッサと、

40

前記少なくとも1つのプロセッサと通信接続しているメモリと、を含み、

前記メモリには、前記少なくとも1つのプロセッサによって実行できる命令が記憶されており、前記少なくとも1つのプロセッサによって前記命令を実行することにより、前記少なくとも1つのプロセッサが請求項1～5のいずれか一項に記載の方法を実行できる、

電子デバイス。

【請求項16】

コンピューター命令が記憶されている非一時的なコンピューター可読記憶媒体であって、前記コンピューター命令は、前記コンピューターに請求項1～5のいずれか一項に記載の方法を実行させる、

50

非一時的なコンピューター可読記憶媒体。

【請求項 17】

コンピュータープログラムを含むコンピュータープログラム製品であって、前記コンピュータープログラムは、プロセッサによって実行されると、請求項 1 ~ 5 のいずれか一項に記載の方法を実行させる、

コンピュータープログラム製品。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、2021年12月8日に中国特許局に提出した、出願番号が202111489679.1であり、発明名称が「データ注釈方法、装置、システム、デバイスおよび記憶媒体」である中国特許出願を優先権として主張し、その内容の全ては、援用により本発明に組み込まれる。

【0002】

本発明は、データ処理技術分野に関し、特に、人工知能、ビッグデータ、およびディープラーニングなどの分野に関する。

【背景技術】

【0003】

データを注釈することは、コンピューター研究・応用分野で重要な過程である。例えば、大量の注釈されたデータに基づいてモデルのトレーニングなどを行う。

【発明の概要】

【0004】

本発明は、データ注釈方法、装置、システム、デバイスおよび記憶媒体を提供する。

【0005】

本発明の第1態様では、データ注釈方法を提供する。前記データ注釈方法は、内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含むことと、

各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであることと、

前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定することと、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈することと、を含む。

【0006】

本発明の第2態様では、データ注釈装置を提供する。前記データ注釈装置は、内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含む、第1取得モジュールと、

各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであり、前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定する、確定モジュールと、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈

10

20

30

40

50

する、注釈モジュールと、を含む。

【0007】

本発明の第3態様では、データ注釈システムを提供する。前記データ注釈システムは、注釈システムバックエンドと注釈システムフロントエンドとを含み、

前記注釈システムバックエンドは、内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含み、各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであり、前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定し、前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈し、

10

前記注釈システムフロントエンドは、前記注釈システムバックエンドが前記注釈対象ピクチャーにおける前記オブジェクトに対する注釈を展示する。

【0008】

本発明の第4態様では、電子デバイスを提供する。前記電子デバイスは、

少なくとも1つのプロセッサと、

20

前記少なくとも1つのプロセッサと通信接続しているメモリと、を含み、

前記メモリには、前記少なくとも1つのプロセッサによって実行される命令が記憶されており、前記少なくとも1つのプロセッサによって前記命令を実行することにより、前記少なくとも1つのプロセッサが第1態様に記載の方法を実行できる。

【0009】

本発明の第五態様では、コンピューター命令が記憶されている非一時的なコンピューター可読記憶媒体を提供する。ここで、前記コンピューター命令は、前記コンピューターに第1態様に記載の方法を実行させる。

【0010】

本発明の第六態様では、コンピュータープログラム製品を提供する。前記コンピュータープログラム製品は、コンピュータープログラムを含み、前記コンピュータープログラムは、プロセッサによって実行されると、第1態様に記載の方法を実現する。

30

【発明の効果】

【0011】

本発明は、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、1つ前のピクチャーの位置情報および位置シフト量に基づいて注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定し、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈することができる。これにより、人間の記憶によってピクチャーにおける各オブジェクトを注釈することを回避し、複数のピクチャーに同じオブジェクトを注釈する速度を高めることができる。

40

【0012】

この部分で記述された内容は、本発明実施例の要旨又は重要な特徴を特定することを意図しておらず、本発明の範囲に対する限定にならないことを理解すべきである。本発明の他の特徴は、以下の明細書によって容易に理解される。

【図面の簡単な説明】

【0013】

以下の図面は、本技術案をより理解するためのものであり、本発明に対する限定にならない。

【図1】図1は本発明実施例におけるブロック選択の模式図である。

【図2】図2は本発明実施例におけるピクチャーにオブジェクトタグを注釈する模式図で

50

ある。

【図 3】図 3 は本発明実施例が提供するデータ注釈方法のフローチャートである。

【図 4】図 4 は本発明実施例における動画データに対応する時間順に並べた複数のピクチャーを確定するフローチャートである。

【図 5】図 5 は本発明実施例におけるシステムのインタラクションの模式図である。

【図 6】図 6 は本発明実施例が提供するデータ注釈方法の応用の模式図である。

【図 7】図 7 は本発明実施例が提供するデータ注釈装置の一構成模式図である。

【図 8】図 8 は本発明実施例が提供するデータ注釈装置の別の構成模式図である。

【図 9】図 9 は本発明実施例が提供するデータ注釈装置のさらに別の構成模式図である。

【図 10】図 10 は本発明実施例が提供するデータ注釈システムの一構成模式図である。

10

【図 11】図 11 は本発明実施例が提供するデータ注釈システムの別の構成模式図である。

【図 12】図 12 は本発明実施例のデータ注釈方法を実現するための電子デバイスのブロック図である。

【発明を実施するための形態】

【0014】

本発明の目的、技術案、及び利点をより明らかに説明するためには、以下、図面を参照し実施例を挙げて、本発明をさらに詳しく説明する。勿論、説明される実施例は単に本発明の一部の実施例に過ぎず、すべての実施例ではない。本発明の実施例に基づき、当業者が進歩性を有する努力をせずに想到し得るほかの実施例はすべて本発明の保護範囲に属する。

20

【0015】

以下、附図を参照しながら、本発明の例示的な実施例に対して説明する。理解の便宜上、本発明実施例の様々な詳細が含まれるが、これらは例示的なものに過ぎないとみなされるべきである。そのため、当業者であれば、本発明の範囲及び精神から逸脱することなく、本明細書に記載の実施例に対して様々な変更及び修正を行うことができることを認識するべきである。また、以下の説明では、明確や簡潔のために、公知の機能や構成に対する説明を省略する。

【0016】

現在、スマートシティという概念での人工知能 (Artificial Intelligence、AI) アプリケーションは、多くの場合にはカメラが収集した画面に対して注釈する必要がある。例えば、オブジェクトが撮像範囲に入ってから、当該オブジェクトが撮像範囲から離れるまで、オブジェクトを注釈する。これは、カメラが撮ったピクチャーに対して注釈することと理解してよい。

30

【0017】

通常、ブロック選択することで注釈を行い、ブロック選択する同時に、異なるオブジェクトに対して、例えば、船、車両などの異なる物体に対して、異なる ID (タグ) をマークすることで、同じ ID によって、異なる画面において同一オブジェクトが識別される。ここで、ブロック選択とは、ピクチャーにおけるあるオブジェクトを中心として、当該オブジェクトの最大カバー範囲を注釈することである。図 1 に示すように、矩形の枠で 1 つの船をブロック選択した。オブジェクトに対して ID をマークした。オブジェクトの付近に 1 つの ID をマークすることができる。図 2 のように、ピクチャーにおいて、船の ID を 1 としてマークした。その後のピクチャーには、当該船の位置が変化することがあり、例えば、左下にずれている場合、位置が変化した後の当該船に対して、1 つ前のピクチャーと同じ ID : 1 をマークし続ける必要がある。

40

【0018】

本発明実施例が提供するデータ注釈方法は、電子デバイスに用いられることができる。具体的には、電子デバイスは、端末装置、サーバーなどであってよい。

【0019】

本発明実施例は、データ注釈方法を提供する。前記データ注釈方法は、

50

内容が連続する複数のピクチャーを取得し、複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含むことと、

各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前において、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前において、1つ前のピクチャーに隣接するピクチャーであることと、

1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定することと、

予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈することと、を含む。

【0020】

本発明は、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、1つ前のピクチャーの位置情報および位置シフト量に基づいて注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定し、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトに対して注釈を行うことができる。これにより、人間の記憶によってピクチャーにおける各オブジェクトを注釈することを回避し、複数のピクチャーに同じオブジェクトを注釈する速度を高めることができる。

【0021】

なお、注釈者は記憶に基づいて同じオブジェクトを注釈すると、人為的なミスが発生しやすい。例えば、オブジェクトが何度も撮像範囲内に入り、異なる注釈のニーズがある場合、ミスが発生する可能性が大きくなる。例えば、同一オブジェクトが何度も撮像範囲に入った場合は同じIDを用いる要求がある。本発明実施例では、人間の記憶によってピクチャーにおける各オブジェクトを注釈することを回避でき、さらに、注釈の正確度を向上させることができる。

【0022】

図3は本発明実施例が提供するデータ注釈方法のフローチャートである。図3を参照して、本発明実施例が提供するデータ注釈方法は、以下のステップを含んでよい。

【0023】

S301において、内容が連続する複数のピクチャーを取得する。

複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含む。

【0024】

オブジェクトとは、ピクチャーにおけるターゲット、例えば、船、車両などの物体と理解してもよい。

【0025】

内容が連続するとは、複数のピクチャーにおける内容が連続していることを指す。具体的には、内容が連続する複数のピクチャーは、動画における連続フレームであってもよく、時間順に並べた複数のピクチャー、例えば、動画データからフレーム抽出して得られた複数のピクチャーなどであってもよい。

【0026】

1つの実現可能な形態において、画像収集デバイスが収集したピクチャーをリアルタイムで受けてよい。内容が連続する複数のピクチャーは画像収集デバイスがリアルタイムで収集した連続フレームである。ここで、画像収集デバイスは、ビデオカメラ、カメラなどを含んでもよい。

【0027】

別の実現可能な形態において、データベースから、格納されている動画データに対応する時間順に並べた複数のピクチャーを取得してもよい。内容が連続する複数のピクチャーとして、動画データにおけるすべてのフレームを取得してもよく、動画における一部のフレームを取得してもよい。例えば、画像収集デバイスがデバイスデータを収集した後、動

10

20

30

40

50

画データをデータベースに格納し、動画データをフレーム抽出して並べて、動画データに対応する時間順に並べた複数のピクチャーを取得し、動画データに対応する時間順に並べた複数のピクチャーをデータベースに格納する。

【0028】

S302において、各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定する。

【0029】

ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前において、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前において、1つ前のピクチャーに隣接するピクチャーである。

10

【0030】

1つの実現可能な形態において、当該位置シフト量は、1つ前のピクチャーにおけるオブジェクトの位置情報から当該2つ前のピクチャーにおける当該オブジェクトの位置情報を減算して得られたものであってよい。

【0031】

別の実現可能な形態において、当該位置シフト量は、2つ前のピクチャーにおけるオブジェクトの位置情報から当該1つ前のピクチャーにおける当該オブジェクトの位置情報を減算して得られたものであってもよい。

【0032】

S303において、1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定する。

20

【0033】

通常、内容が連続する複数のピクチャー（例えば、動画における連続フレーム）における同一オブジェクトの位置変化は規律があり、例えば、同一オブジェクトについて、1つのピクチャーにおける位置が当該ピクチャーの内容と連続する1つ前のピクチャーにおける位置に対する変化は、変化1と理解してよく、当該ピクチャーの内容と連続する1つ後のピクチャーにおける当該オブジェクトの位置が当該ピクチャーにおける当該オブジェクトの位置に対する変化は、変化2と理解してよく、変化1と変化2との間の差は、一般に小さく、ひいては変化1と変化2は同じである可能性もある。本発明実施例では、隣接する2つのピクチャーにおける同一オブジェクトの位置情報によって注釈対象ピクチャーにおける当該オブジェクトの予測位置情報を確定することができる。簡単に理解すると、隣接する2つのピクチャーにおける同一オブジェクトの位置情報によって注釈対象ピクチャーにおける当該オブジェクトの位置情報を予測する。

30

【0034】

1つの実現可能な形態において、位置シフト量が1つ前のピクチャーにおけるオブジェクトの位置情報から2つ前のピクチャーにおける当該オブジェクトの位置情報を減算して得られたものである場合、S303は、1つ前のピクチャーにおけるオブジェクトの位置情報に当該位置シフト量を加算し、即ち、1つ前のピクチャーにおけるオブジェクトの位置情報を元に、当該位置シフト量を増加することと理解してもよく、そして、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を取得すること、を含んでよい。

40

【0035】

1つの実現可能な形態において、位置シフト量が2つ前のピクチャーにおけるオブジェクトの位置情報から1つ前のピクチャーにおける当該オブジェクトの位置情報を減算して得られたものである場合、S303は、1つ前のピクチャーにおけるオブジェクトの位置情報から当該位置シフト量を減算し、即ち、1つ前のピクチャーにおけるオブジェクトの位置情報元に、当該位置シフト量を減算することと理解してよく、そして、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を取得すること、を含んでよい。

【0036】

注釈対象ピクチャーは、内容が連続する複数のピクチャーのうち、1つ目のピクチャー

50

および2つ目のピクチャー以外の他のピクチャーを含んでよい。

【0037】

複数のピクチャーにおける時間順に並べる1つ目のピクチャーおよび2つ目のピクチャーを注釈し、1つ目のピクチャーおよび2つ目のピクチャーにおけるオブジェクトの位置情報を後続の注釈のベースとして、即ち、1つ目のピクチャーおよび2つ目のピクチャーにおけるオブジェクトの位置情報に基づいて、後続の各ピクチャーにおける各オブジェクトのそれぞれの位置情報を予測すると理解してよい。

【0038】

簡単に理解すると、隣接する2つのピクチャーによって次のピクチャーの位置を予測することである。例えば、時系列を有する5つのピクチャーについて、5つのピクチャーは1つの動画における連続フレームであってよく、例えば、1つの動画の一部である。または、5つのピクチャーは、動画からフレーム抽出し、並べて得られたピクチャーである。まず、1つ目のピクチャーおよび2つ目のピクチャーを注釈し、具体的に、1つ目のピクチャーにおける各オブジェクトの位置を注釈し、各オブジェクトの位置に当該オブジェクトのタグを注釈する。そして、2つ目のピクチャーにおいて、各オブジェクトに対して1つ目のピクチャーにおける同じタグを注釈し、2つ目のピクチャーにおける各オブジェクトの位置を注釈する。次に、各オブジェクトに対して、1つ目のピクチャーおよび2つ目のピクチャーによって3つ目のピクチャーにおける当該オブジェクトの位置を予測し、即ち、1つ目のピクチャーおよび2つ目のピクチャーにおける当該オブジェクトの位置情報によって当該オブジェクトの位置シフト量を確定し、2つ目のピクチャーにおける当該オブジェクトの座標から1つ目のピクチャーにおける当該オブジェクトの座標を減算して、座標の差を取得し、その座標の差は当該位置シフト量として理解されることができ、2つ目のピクチャーの位置情報に当該位置シフト量を加算して、3つ目のピクチャーにおける当該オブジェクトの位置情報を取得する。同様に、2つ目のピクチャーおよび3つ目のピクチャーによって4つ目のピクチャーにおける当該オブジェクトの位置を予測し、即ち、2つ目のピクチャーおよび3つ目のピクチャーにおける当該オブジェクトの位置情報によって当該オブジェクトの位置シフト量を確定し、3つ目のピクチャーにおける当該オブジェクトの座標から2つ目のピクチャーにおける当該オブジェクトの座標を減算して、座標の差を取得し、その座標の差は当該位置シフト量として理解されることができ、3つ目のピクチャーの位置情報に当該位置シフト量を加算して、4つ目のピクチャーにおける当該オブジェクトの位置情報を取得する。3つ目のピクチャーおよび4つ目のピクチャーによって5つ目のピクチャーにおける当該オブジェクトの位置を予測し、即ち、3つ目のピクチャーおよび4つ目のピクチャーにおける当該オブジェクトの位置情報によって当該オブジェクトの位置シフト量を確定し、4つ目のピクチャーにおける当該オブジェクトの座標から3つ目のピクチャーにおける当該オブジェクトの座標を減算して、座標の差を取得し、その座標の差は当該位置シフト量として理解されることができ、4つ目のピクチャーの位置情報に当該位置シフト量を加算して、5つ目のピクチャーにおける当該オブジェクトの位置情報を取得する。

【0039】

ここで、5つのピクチャーは、隣接するピクチャーに基づいて注釈対象ピクチャーの位置を予測する過程を説明するためのものであり、実際応用の過程において、ピクチャーの数量がさらに多くなる可能性がある。

【0040】

S304において、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈する。

【0041】

1つの実現可能な形態において、注釈対象ピクチャーに、当該予測位置情報を直接に注釈し、当該予測位置情報に基づいて確定された位置に、オブジェクトのタグを注釈することができる。1つの形態において、当該予測位置情報が存在する位置は、注釈対象ピクチャーにおける当該オブジェクトの位置と理解してよく、かつ、当該位置に当該オブジェ

10

20

30

40

50

トのタグを注釈することができる。これにより、注釈対象ピクチャーにおける当該オブジェクトの注釈を完成する。別の形態において、予測位置情報が1つの座標であり、当該座標を矩形の中心点として、当該中心点に基づいて1つの矩形を書くことができる。当該矩形が当該オブジェクトに対する矩形の枠であると理解してよく、当該矩形の枠の位置に当該オブジェクトのタグをマークすることができ、例えば、当該矩形の枠の左側に当該オブジェクトのタグをマークする。

【0042】

このような形態は、各オブジェクトのタグおよび位置を人為的に記憶する難しさを回避し、マークの速度を高めることができ、かつ、各オブジェクトのタグおよび位置を人為的に記憶することによる注釈ミス回避し、注釈の正確度を高めることができる。

10

【0043】

別の実現可能な形態において、当該予測位置情報を展示することによって、注釈者は当該予測位置情報を校正し、校正後の予測位置情報を入力する。これにより、当該校正後の予測位置情報を受信し、注釈対象ピクチャーに校正後の予測位置情報を注釈することができる。注釈対象ピクチャーに校正後の予測位置情報を注釈し、校正後の予測位置情報に基づいて確定された位置に、オブジェクトのタグを注釈する。1つの形態において、当該校正後の予測位置情報は、注釈対象ピクチャーにおける当該オブジェクトの位置と理解してよく、かつ、当該位置に当該オブジェクトのタグを注釈することができる。これにより、注釈対象ピクチャーにおける当該オブジェクトのマークを完成する。別の形態において、当該校正後の予測位置情報を矩形の中心点として、当該中心点に基づいて1つの矩形を書くことができる。当該矩形は当該オブジェクトに対する矩形の枠であると理解してよく、当該矩形の枠の位置に当該オブジェクトのタグをマークすることができ、例えば、当該矩形の枠の左側に当該オブジェクトのタグをマークする。

20

【0044】

この方法は、人間の記憶によってピクチャーにおける各オブジェクトを注釈することを回避し、注釈の速度を高めることができる。かつ、注釈者は、注釈の過程において、予測位置に対して少量の校正を行うことだけで、隣接するピクチャーにおける同一オブジェクトの位置記録の作業を完成でき、注釈者がIDを記憶する難しさを大幅に下げている、人間の記憶が注釈効率および正確性に与える影響を下げている、注釈速度および正確度を向上することができる。なお、得られた予測位置情報を校正することは、注釈の正確性をさらに向上することができる。

30

【0045】

1つの好ましい実施例において、内容が連続する複数のピクチャーを取得することは、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得すること、を含む。

【0046】

1つの動画データについて、動画データに対応する時間順に並べた複数のピクチャーは、動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順に従って、ピクチャーを並べて得られたものである。

【0047】

時間順に並べた複数のピクチャーは、内容が連続する複数のピクチャーと理解してよい。即ち、異なる動画データにそれぞれ対応する、内容が連続する複数のピクチャーに対して注釈することを実現できる。動画の注釈において、連続画面における同じオブジェクトに対して継続的に注釈する効率を高めることを実現できる。

40

【0048】

1つの実現可能な形態において、動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、電子デバイスが動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順に従って、ピクチャーを並べて得られたものであつてよい。

【0049】

50

図 4 に示すように、本発明実施例が提供するデータ注釈方法は、以下のステップを含む。

【 0 0 5 0 】

S 4 0 1 において、少なくとも 1 つの動画データを取得する。

【 0 0 5 1 】

1 つの画像収集デバイスが収集した動画データを取得してよい。例えば、1 つのビデオカメラは複数のカメラを備える場合、複数のカメラのそれぞれが収集した動画データを取得してよい。または、1 つのカメラが異なる時間帯で収集した動画データをそれぞれ取得してよい。少なくとも 2 つの画像収集デバイスが収集した動画データのそれぞれを取得してもよい。

10

【 0 0 5 2 】

本発明実施例において、電子デバイスは直接に画像収集デバイスから動画データを取得することができ、または、画像収集デバイスは収集された動画データをデータベースに格納し、電子デバイスはデータベースから動画データを取得することなどができる。

【 0 0 5 3 】

1 つの実現可能な形態において、少なくとも 2 つのビデオカメラによってデータをそれぞれ収集し、カメラにより得られたデータに対して番号を付け、各カメラがそれぞれ唯一の識別コードを有し、当該識別コードは、異なるカメラにより得られたデータを区別するために用いられる。例えば、番号 1 のカメラにより得られたすべてのデータに対して、「0 0 1 - 」をデータファイル名のプレフィックスとして、番号 2 のカメラにより得られた

20

【 0 0 5 4 】

データファイルパッケージによってビデオカメラが収集したデータを格納することができる。データファイルのデータファイル名にはカメラに対応する標識コードを含むことで、当該標識コードは、当該データファイルに格納されているデータがどのカメラにより収集したデータであるかを標識するために用いられる。

【 0 0 5 5 】

1 つの実現可能な形態において、異なるカメラにより得られたデータは、異なるデータファイルにそれぞれ格納することができ、即ち、1 つのカメラが収集したすべてのデータ

30

【 0 0 5 6 】

S 4 0 2 において、各動画データについて、動画データをフレーム抽出し、動画データに対応するピクチャーを取得する。

【 0 0 5 7 】

本発明実施例において、フレーム抽出する方法は制限されなく、実際の要求に応じて選択すればよい。例えば、予め設定された時間の間隔ごとに 1 つのフレームの画像を抽出するように、動画データ全体に対してフレーム抽出してよい。ここで、予め設定された時間の間隔は、実際の要求や経験などに応じて確定してよく、例えば、1 秒ごとに 1 つのフレームの画像を抽出してもよい。即ち、抽出された隣接する 2 つのフレームの画像の間隔が 1 秒であることと理解してもよい。または、予め設定された時間範囲の動画ごとに 1 つのフレームの画像を抽出するように、例えば、1 秒の動画ごとに 1 つのフレームの画像を抽出するように、フレーム抽出してよい。抽出された当該フレームの画像は、この 1 秒の動画におけるいずれか 1 つのフレームの画像であってもよい。

40

【 0 0 5 8 】

フレーム抽出して得られたピクチャーには、時間タグを含むことができる。例えば、ビデオカメラがデータを集める過程において、1 つのフレームの画像のそれぞれの撮影時

50

間を記録することができる。当該時間タグは、当該撮影時間を示すことができる。

【0059】

または、時間タグは、シフト時間を示すことができる。シフト時間とは、ある1フレームの画像に対して、抽出された当該フレームの画像のシフト時間を指す。当該フレームの画像は、動画データにおけるいずれか1つのフレームの画像であってよく、例えば、抽出された1番目のフレームの画像であってよい。

【0060】

本発明実施例において、フレーム抽出して得られたピクチャーにおける時間タグは、制限されなく、同一組のピクチャーが同一種類の時間の時間タグを使用すればよい。例えば、1つの動画データをフレーム抽出して得られたすべてのピクチャーの時間は、いずれも撮影時間を標識するためのものであり、または、いずれもシフト時間を標識するためのものである。

10

【0061】

本発明実施例において、フレーム抽出して得られた動画データに対応するピクチャーを格納する。ここで、動画データのタグおよびピクチャーの時間タグで当該ピクチャーを命名することができる。例えば、動画データファイル名のプレフィックス（動画データを生成した画像収集デバイス、例えばビデオカメラを標識する）を画像データファイルのファイル名のプレフィックスとして保存する。画像データファイル名は、動画データプレフィックス+当該フレームを生成した時刻のタイムスタンプに構成される。動画データプレフィックスは動画データを格納するデータファイルのデータファイル名のプレフィックスであり、当該フレームを生成した時刻のタイムスタンプは上記した時間タグである。

20

【0062】

1つの場合では、フレーム抽出する過程において、同じ秒の動画から複数フレームのピクチャーを抽出した可能性がある。この場合、時間タグが標識した時間の単位が秒であれば、これらの複数フレームの画像を区別することができない。そのため、この場合は、時間タグが標識した時間の単位がミリ秒であってよく、即ち、ピクチャーを命名する過程において、当該フレームを生成した時刻のタイムスタンプにミリ秒フィールドを増加してよい。

【0063】

S403において、ピクチャーの時間順にピクチャーを並べて、動画データに対応する時間順に並べた複数のピクチャーを取得する。

30

【0064】

ピクチャーの時間タグに基づいて、各動画データからフレーム抽出して得られたピクチャーのそれぞれを時間順に並べることができる。例えば、時間の前後順に従って、即ち、時間的に前のピクチャーを前に、時間的に後のピクチャーを後に並べることができる。これにより、各動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得することができる。

【0065】

フレーム抽出して得られた複数のピクチャーを時間が自然に流れる順に並べることで、複数のピクチャーに連続性を持たせることができ、各ピクチャーは、ピクチャーの内容について、前後のピクチャーと連続性を生じる。時間順に並べた複数のピクチャーは、内容が連続する複数のピクチャーであると理解してよい。

40

【0066】

異なるカメラが生成したデータを組に分けて、各組に同一カメラが生成した画像データを格納することができる。例えば、異なるカメラにより得られた動画データにそれぞれ対応する時間順に並べた複数のピクチャーをそれぞれ分けて格納する。例えば、全部で、3つのカメラが収集した動画データが3つある場合、3つの動画データから、それらに対応する3組の時間順に並べた複数のピクチャーをそれぞれ取得でき、これらの3組のピクチャーはそれぞれ分けて格納されることができる。

【0067】

50

これにより、内容が連続する複数のピクチャーを取得することは、時間順に並べた複数のピクチャーを取得することを含む。内容が連続する複数のピクチャーは、上記した時間順に並べた複数のピクチャーを含む。

【0068】

本発明実施例において、動画データを取得した後、まず、動画データをフレーム抽出し、そして、フレーム抽出して得られたピクチャーを時間順に並べることができる。並べた複数のピクチャーは、内容が連続する複数のピクチャーであると理解してよい。具体的に、複数の動画データにおける各動画データについて、それぞれに対応する内容が連続する複数のピクチャーを取得することができる。異なる動画データにそれぞれ対応する内容が連続する複数のピクチャーに対して、上記したステップS302～S304をそれぞれ実行することで、異なる動画データにそれぞれ対応する内容が連続する複数のピクチャーを注釈することができる。

10

【0069】

本発明実施例において、異なるカメラが撮った動画データを取得することができ、各カメラが撮った動画データをそれぞれフレーム抽出し、異なるカメラに対応する画像データを取得し、画像データには、各ピクチャーのカメラタグおよび時間タグを含む。これにより、注釈する過程において、時間順に並べた複数のピクチャーを取得し、2つ前のピクチャーにおけるオブジェクトを注釈し（各オブジェクトのオブジェクトタグおよび位置を注釈する）、各オブジェクトに対して、後続のピクチャーにおける同一オブジェクトの位置を予測することができる。具体的に、予測することは、各ピクチャーに対して、当該ピクチャーと当該ピクチャーの1つ前のピクチャーにおける同一オブジェクトの位置との差を算出し、オブジェクトの位置シフト量を取得し、当該ピクチャーにおけるオブジェクトの位置に当該位置シフト量を加算して、次のピクチャーにおける当該オブジェクトの位置を取得し、当該位置に基づいて次のピクチャーに当該オブジェクトを注釈することを含む。予測した位置に基づいて当該オブジェクトを注釈する前に、予測して得られた位置を校正することもできる。

20

【0070】

別の実現可能な形態において、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、電子デバイスがデータベース、例えば注釈システムデータベースから取得したものであってよく、各動画データに対応する時間順に並べた複数のピクチャーは、動画データ処理システムが注釈システムデータベースに格納したものである。

30

【0071】

各動画データを処理し、各動画データに対応する時間順に並べた複数のピクチャーを取得することは、

各動画データについて、動画データをフレーム抽出し、動画データに対応するピクチャーを取得することと、ピクチャーの時間順にピクチャーを並べて、動画データに対応する時間順に並べた複数のピクチャーを取得することと、を含んでもよい。

【0072】

具体的に、動画データ処理システムは、上記したS401～S403を実行して、各動画データに対応する時間順に並べた複数のピクチャーを取得し、各動画データに対応する時間順に並べた複数のピクチャーを注釈システムデータベースに保存する。これにより、注釈する過程において、注釈システムデータベースから、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得することができる。

40

【0073】

動画データ処理システムによって動画データを処理して、時間順に並べた複数のピクチャーを取得し、注釈システムデータベースによって複数の動画データにそれぞれ対応する時間順に並べた複数のピクチャーを保存し、注釈する過程において、注釈システムデータベースからこれらの時間順に並べた複数のピクチャーを直接に取得して注釈することによって、注釈する過程において、電子デバイスの処理負荷を低減し、電子デバイスの計算リ

50

ソースなどを節約することができる。

【0074】

図5に示すように、複数のカメラがデータをそれぞれ収集し、複数のカメラのそれぞれに対応する動画データを取得する。各カメラが動画データを取得した後、当該カメラの動画データを動画データ処理システムに送信することができる。動画データ処理システムは、各動画データをそれぞれ処理し、各動画データに対応する時間順に並べた複数のピクチャーを取得し、各動画データに対応する時間順に並べた複数のピクチャーを注釈システムデータベースに格納する。ここで、時間順に並べた複数のピクチャーは、内容が連続するピクチャーであると理解してもよい。

【0075】

注釈システムバックエンドは、アプリケーションプログラムインタフェース (Application Programming Interface、API) を介して注釈システムデータベースからこれらの内容が連続するピクチャーを取得し、これらの内容が連続するピクチャーを注釈する。注釈システムフロントエンドは、注釈されたピクチャーを展示する。

【0076】

ここで、注釈システムバックエンドおよび注釈システムフロントエンドは、1つの電子デバイスにおける2つの部分であってよく、または、2つの独立した電子デバイスなどであってよい。

【0077】

注釈システムバックエンドは内容が連続するピクチャーを取得し、これらの内容が連続するピクチャーを注釈することは、

内容が連続する複数のピクチャーを取得し、複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含むことと、各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前にあって、1つ前のピクチャーに隣接するピクチャーであることと、1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定することと、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈することと、を含んでよい。具体的な過程は、上記した実施例に詳しく説明した。

【0078】

1つの具体的な例において、図6に示すように、本発明実施例が提供するデータ注釈方法は、以下のステップを含む。

【0079】

ステップAにおいて、動画データを収集し、カメラによって番号を付ける。

【0080】

複数のカメラが動画データをそれぞれ収集し、異なるカメラが収集した動画データに対してカメラの番号によって標識する。

【0081】

各カメラは唯一の識別コードをそれぞれ有し、当該識別コードは、異なるカメラにより得られた動画データを区別するために用いられる。例えば、番号1のカメラにより得られた動画データは、いずれも「001-」をデータファイル名のプレフィックスとし、即ち、番号1のカメラにより得られた動画データは、1つのデータファイルに格納されており、「001-」をデータファイル名のプレフィックスとする。番号2のカメラにより得られたデータは、いずれも「002-」データファイル名のプレフィックスとし、即ち、番号2のカメラにより得られた動画データは、他のデータファイルに格納されており、「002-」をデータファイル名のプレフィックスとする。このようにして、異なるカメラに

10

20

30

40

50

より得られた動画データを、異なるデータファイルにそれぞれ格納し、異なるデータファイル名で区別する。

【0082】

ステップBにおいて、動画データをフレーム抽出し、動画シフト時間に従ってピクチャーに対して番号を付ける。

【0083】

本発明実施例において、動画データを取得した後、まず動画データを処理し、即ち、まず動画データをフレーム抽出してよい。

【0084】

動画データを動画データファイルに格納してよく、各動画データファイルから動画データをそれぞれ取得し、動画データを処理する。

10

【0085】

シフト時間に従ってピクチャーに対して番号を付けることは、フレーム抽出して得られた各ピクチャーに当該ピクチャーのシフト時間を標識することと理解してもよい。シフト時間は、ある1つのフレームのピクチャーに対して、抽出された当該フレームのピクチャーのシフト時間であってよい。当該フレームの画像は、動画データにおけるいずれか1つのフレームの画像であってよく、例えば、抽出された1番目のフレームの画像であってよい。

【0086】

1つの例において、1秒の動画ごとに1つのフレームのピクチャーを抽出するように、すべての動画データをフレーム抽出するとともに、各ピクチャーのシフト時間を計算する。シフト時間は、ミリ秒まで精確なものであってよい。シフト時間をピクチャーの作成時間として格納する。ピクチャーは、データベースに格納されてよく、例えば、データベースにおける画像データファイルに格納されてよい。

20

【0087】

ピクチャーを格納する過程において、シフト時間を格納することのほか、動画データファイル名のプレフィックスを画像データファイルのファイル名のプレフィックスとして保存することもできる。画像データファイル名は、元プレフィックス(動画データファイル名のプレフィックス)+当該フレームを生成した時刻のタイムスタンプ(時間タグ、例えば、撮影時間、シフト時間など)+ミリ秒数に構成される。即ち、画像データファイル名は、動画データプレフィックス+当該フレームを生成した時刻のタイムスタンプに構成されることと理解してもよい。動画データプレフィックスは動画データを格納するデータファイルのデータファイル名のプレフィックスであり、当該フレームを生成した時刻のタイムスタンプは上記した時間タグである。時間タグが標識した時間の単位はミリ秒であってよく、即ち、ピクチャーを命名する過程において、当該フレームを生成した時刻のタイムスタンプにミリ秒フィールドを追加してよい。ミリ秒数を追加することは、あるフレーム抽出方法において1秒の動画から複数のピクチャーを抽出した場合があることを考慮するためである。

30

【0088】

ステップCにおいて、ピクチャーを組に分けて並べる。

40

【0089】

異なるカメラが生成した画像データを組に分けて、各組に同一カメラが生成した画像データを格納することができる。ここで、各カメラが生成した画像データは、当該カメラの動画データをフレーム抽出して得られた複数のピクチャーである。

【0090】

各ピクチャーはいずれも時間属性を有し、即ち、各ピクチャーにはいずれも時間タグが標識されているため、同一組にあるデータ、即ち複数のピクチャーを、ピクチャーの時間タグに従って並べることができ、複数のピクチャーに連続性を持たせ、ピクチャーを時間が自然に流れる順に並べさせることを目的とする。これにより、各ピクチャーは、ピクチャーの内容について、前後のピクチャーと連続性を生じる。

50

【0091】

1つの実現可能な形態において、各カメラが動画データを取得した後、当該カメラの動画データを動画データ処理システムに送信することができる。動画データ処理システムは、各動画データをそれぞれ処理する。具体的に、動画データ処理システムがカメラによって番号を付け、動画データをフレーム抽出し、動画シフト時間に従ってピクチャーに対して番号を付け、ピクチャーを組に分けて並べる。これにより、各動画データに対応する時間順に並べた複数のピクチャーを取得する。

【0092】

なお、各動画データに対応する時間順に並べた複数のピクチャーを注釈システムデータベースに格納することができる。

【0093】

ステップDにおいて、ピクチャーの組の1つ目のピクチャーおよび2つ目のピクチャーにおける各オブジェクトに対して、IDおよび位置を注釈する。

【0094】

ステップEにおいて、次のピクチャーにおける各オブジェクトの位置を予測する。

【0095】

注釈システムデータベースから少なくとも1つの動画データに対応する時間順に並べた複数のピクチャーを取得することができる。例えば、複数のカメラにそれぞれ対応する複数組のピクチャーを取得することができる。例えば、注釈システムバックエンドは、注釈システムデータベースから少なくとも1つの動画データに対応する時間順に並べた複数のピクチャーを取得する。

【0096】

並べた各組のピクチャーに対して、当該組のピクチャーのうち1つ目のピクチャーおよび2つ目のピクチャーにおけるオブジェクトを注釈し、各オブジェクトのIDおよび座標を記録する。IDは当該オブジェクトの唯一の標識であり、座標はピクチャーにおける当該オブジェクトの位置である。

【0097】

当該組のピクチャーのうち残りのピクチャー（1つ目のピクチャーおよび2つ目のピクチャー以外の他のピクチャー）におけるオブジェクトの位置を逐次に予測する。予測方法は、隣接する2つのピクチャーにおける同じIDのオブジェクトの位置シフト量を計算し、例えば、隣接する2つのピクチャーの位置（1つの座標であってよい）の差を計算し、各オブジェクトに対して、ピクチャーにおける当該オブジェクトの現在の位置に当該オブジェクトの位置シフト量を加算することで、次のピクチャーにおける各オブジェクトの座標を計算することである。

【0098】

当該ピクチャーと当該ピクチャーの1つ前のピクチャーにおける同一オブジェクトの位置の差を計算し、オブジェクトの位置シフト量を取得し、当該ピクチャーにおけるオブジェクトの位置に当該位置シフト量を加算して、次のピクチャーにおける当該オブジェクトの位置を取得する。

【0099】

ステップFにおいて、各オブジェクトの位置を校正する。

【0100】

次のピクチャーにおける各オブジェクトの予測位置情報を取得した後、注釈対象ピクチャーに当該予測位置情報を直接に注釈し、当該予測位置情報に基づいて確定された位置にオブジェクトのタグを注釈することができる。

【0101】

または、当該予測位置情報を展示することによって、注釈者は当該予測位置情報を校正し、校正後の予測位置情報を入力することもできる。これにより、当該校正後の予測位置情報を受信し、注釈対象ピクチャーに当該校正後の予測位置情報を注釈することができる。注釈対象ピクチャーに校正後の予測位置情報を注釈し、校正後の予測位置情報に基づい

10

20

30

40

50

て確定された位置に、オブジェクトのタグを注釈する。

【0102】

注釈者は、注釈する過程において、予測位置に対して少量の座標校正を行うことだけで、隣接するピクチャーにおける同一オブジェクトの位置記録の作業を完成でき、注釈者がIDを記憶する難しさを大幅に下げるとともに、注釈の正確度を向上することができる。

【0103】

ステップGにおいて、注釈結果を整理する。

【0104】

注釈された結果をファイル名に従って並べて整理してから、交付する。

【0105】

ファイル名は、上記した画像データファイル名である。

【0106】

各カメラが生成した画像データは1組のピクチャーであり、即ち、各カメラの動画データに対して得られた時間順に並べた複数のピクチャーが1組のピクチャーであると理解されてよい。複数のカメラに対応する複数組のピクチャーをそれぞれ注釈し、複数組の注釈されたピクチャーを取得する。

【0107】

注釈が完成された後にも、注釈結果を展示することができる。具体的に、各ピクチャーに注釈された各オブジェクトのタグおよび位置を展示してもよい。例えば、注釈システムフロントエンドは、注釈された各ピクチャーを展示する。

【0108】

本発明実施例は、動画中の連続画面における同じオブジェクトを注釈すること実現でき、動画中の連続画面における同じオブジェクトを注釈する速度を向上し、注釈の正確度を高めることができる。かつ、注釈過程は注釈者の能力に制限されない。

【0109】

本発明実施例は、さらに、データ注釈装置を提供する。図7に示すように、前記データ注釈装置は、

内容が連続する複数のピクチャーを取得し、複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含む、第1取得モジュール701と、

各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前であって、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前であって、1つ前のピクチャーに隣接するピクチャーであり、1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定する、確定モジュール702と、

予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈する、注釈モジュール703と、を含む。

【0110】

好ましくは、第1取得モジュール701は、具体的に、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得する。1つの動画データについて、動画データに対応する時間順に並べた複数のピクチャーは、動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順にピクチャーを並べて得られたものである。

【0111】

好ましくは、図8に示すように、当該装置は、さらに、

少なくとも1つの動画データを取得する、第2取得モジュール801と、

各動画データについて、動画データをフレーム抽出し、動画データに対応するピクチャーを取得する、フレーム抽出モジュール802と、

10

20

30

40

50

ピクチャーの時間順にピクチャーを並べて、動画データに対応する時間順に並べた複数のピクチャーを取得する、ソートモジュール 803 と、を含む。

【0112】

好ましくは、第1取得モジュール701は、具体的に、注釈システムデータベースから少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得する。少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが注釈システムデータベースに格納されているものであり、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各動画データのそれぞれについて、動画データ処理システムが動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順にピクチャーを並べて得られたものである。

10

【0113】

好ましくは、図9に示すように、当該装置は、さらに、

予測位置情報を展示する、展示モジュール901と、

予測位置情報を校正して得られた校正後の予測位置情報を受信する、受信モジュール902と、

注釈対象ピクチャーに校正後の予測位置情報を注釈し、校正後の予測位置情報に基づいて確定された位置にオブジェクトのタグを注釈する、注釈モジュール703と、を含み。

【0114】

本発明実施例は、さらに、データ注釈システムを提供する。図10に示すように、前記データ注釈システムは、注釈システムバックエンド1001と注釈システムフロントエンド1002とを含み、

20

注釈システムバックエンド1001は、内容が連続する複数のピクチャーを取得し、複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含み、各オブジェクトに対して、隣接する2つのピクチャーにおけるオブジェクトの位置情報によってオブジェクトの位置シフト量を確定し、ここで、隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、注釈対象ピクチャーに隣接するピクチャーであり、2つ前のピクチャーは時系列で1つ前のピクチャーの前にあって、1つ前のピクチャーに隣接するピクチャーであり、1つ前のピクチャーの位置情報および位置シフト量に基づいて、注釈対象ピクチャーにおけるオブジェクトの予測位置情報を確定し、予測位置情報に基づいて注釈対象ピクチャーにおけるオブジェクトを注釈し、

30

注釈システムフロントエンド1002は、注釈システムバックエンド1001による注釈対象ピクチャーにおけるオブジェクトに対する注釈を展示する。

【0115】

好ましくは、図11に示すように、当該データ注釈システムは、さらに、注釈システムデータベース1101と、

注釈システムデータベース1101から少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1つの動画データについて、動画データに対応する時間順に並べた複数のピクチャーは、動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順にピクチャーを並べて得られたものである、注釈システムバックエンド1001と、を含む。

40

【0116】

好ましくは、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが注釈システムデータベース1101に格納するものであり、少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各動画データのそれぞれについて、動画データ処理システムが動画データをフレーム抽出し、フレーム抽出して得られた動画データに対応するピクチャーの時間順にピクチャーを並べて得られたものである。

【0117】

50

好ましくは、注釈システムバックエンド1001は、予測位置情報を展示し、予測位置情報を校正して得られた校正後の予測位置情報を受信し、注釈対象ピクチャーに校正後の予測位置情報を注釈し、校正後の予測位置情報に基づいて確定された位置にオブジェクトのタグを注釈する。

【0118】

装置、システム実施例は、方法実施例と基本的に同様であるため、簡単に説明したが、関連する内容について、方法実施例の説明部分に参照すればよい。

【0119】

本発明の技術案に関するユーザーの個人情報に対する収集、保存、使用、加工、伝送、提供及び公開等の処理は、いずれも関連する法律および規制に従い、公序良俗に反しないものである。

【0120】

本発明実施例によれば、本発明は、デバイス、可読記憶媒体、及びコンピュータープログラム製品をさらに提供する。

【0121】

図12は、本発明実施例を実施できる例示である電子デバイス1200の模式的なブロック図を示す。電子デバイスとしては、例えば、ラップトップパソコン、デスクトップパソコン、ワークステーション、パーソナルデジタルアシスタント、サーバー、ブレードサーバー、メインフレームコンピューター、及び他の適当なコンピューターなど、様々な形態のデジタルコンピューターである。電子デバイスとしては、さらに、例えば、パーソナルデジタルアシスタント、セルフォン、スマートフォン、ウェアラブル機器、及び他の類似の計算装置など、様々な形態のモバイル装置である。本明細書に記載のコンポーネント、それらの接続関係、及びそれらの機能は例示的なものに過ぎず、本発明の実施に関して本明細書に記載及び/又は主張された内容に限定するものではない。

【0122】

図12に示すように、デバイス1200は、リードオンリーメモリ(ROM)1202に記憶されているコンピュータープログラムによって、又は記憶手段1208からランダムアクセスメモリ(RAM)1203にロードされたコンピュータープログラムによって、各種の適当な動作や処理を実行する、計算手段1201を含む。RAM1203には、さらに、デバイス1200の動作に必要な各種のプログラムやデータが記憶されていることもできる。計算手段1201、ROM1202及びRAM1203は、それぞれ、バス1204によって接続する。入力/出力(I/O)インターフェース1205も、バス1204に接続されている。

【0123】

デバイス1200における複数のコンポーネントは、I/Oインターフェース1205に接続されている。前記複数のコンポーネントは、例えばキーボード、マウス等の入力手段1206と、例えば各種の型のディスプレイ、スピーカー等の出力手段1207と、例えばディスク、光ディスク等の記憶手段1208と、例えばネットワークカード、モデム、無線通信トランシーバ等の通信手段1209と、を含む。通信手段1209によって、デバイス1200は、例えばインターネット等のコンピューターネットワーク、及び/又は各種の通信ネットワークを介して、他のデバイスと情報/データを交換することが可能である。

【0124】

計算手段1201は、処理及び計算能力を有する、各種の汎用な及び/又は専用の処理コンポーネントであってよい。計算手段1201のいくつかの例示は、中央処理ユニット(CPU)、画像処理ユニット(GPU)、各種の専用の人工知能(AI)演算チップ、機械学習モデルアルゴリズムを実行する各種の演算ユニット、デジタルシグナルプロセッサ(DSP)、及び任意の適当なプロセッサ、コントローラ、マイクロコントローラ等を含むが、これらに限定されない。計算手段1201は、上述した各方法や処理、例えばデータ注釈方法を実行する。例えば、いくつかの実施例では、データ注釈方法は、コン

ピューターソフトウェアプログラムとして実現されることができ、例えば記憶手段 1 2 0 8 等の機械可読媒体に有形的に含まれている。いくつかの実施例では、コンピュータープログラムの一部又は全部は、ROM 1 2 0 2 及び / 又は通信手段 1 2 0 9 を介して、デバイス 1 2 0 0 にロード及び / 又はインストールされることができ、コンピュータープログラムは、RAM 1 2 0 3 にロードされ、計算手段 1 2 0 1 によって実行されると、上述したデータ注釈方法の 1 つ又は複数のステップを実行することができる。予備的に、他の実施例において、計算手段 1 2 0 1 は、他の任意の適当な方式によって（例えば、ファームウェアによって）、実行データ注釈方法を実行するように構成されてよい。

【 0 1 2 5 】

本明細書において、上述したシステムや技術の各種実施形態は、デジタル電子回路システム、集積回路システム、フィールドプログラマブルゲートアレイ（FPGA）、特定用途向け集積回路（ASIC）、特定用途向け標準製品（ASSP）、システムオンチップ（SOC）、複合プログラマブルロジックデバイス（CPLD）、コンピューターハードウェア、ファームウェア、ソフトウェア、及び / 又はそれらの組み合わせにおいて実現できる。これらの各種実施形態は、1 つ以上のコンピュータープログラムにおいて実施されており、当該 1 つ以上のコンピュータープログラムは、少なくとも 1 つのプログラマブルプロセッサを含むプログラマブルシステムにおいて実行及び / 又は解釈されてよく、当該プログラマブルプロセッサは、専用又は汎用のプログラマブルプロセッサであってよく、記憶システム、少なくとも 1 つの入力装置、及び少なくとも 1 つの出力装置からデータ及び命令を受信し、データ及び命令を当該記憶システム、当該少なくとも 1 つの入力装置、及び当該少なくとも 1 つの出力装置に伝送することができる、ことを含んでもよい。

10

20

【 0 1 2 6 】

本発明の方法を実施するためのプログラムコードは、1 つまたは複数のプログラミング言語の任意の組み合わせを用いて編集してよい。これらのプログラムコードは、汎用コンピューター、専用コンピューター又は他のプログラマブルデータ処理装置のプロセッサ又はコントローラに提供されることができ、これによって、プログラムコードがプロセッサ又はコントローラによって実行される時、フローチャート図及び / 又はブロック図で規定された機能 / 動作を実施させられることができる。プログラムコードは、完全に機械で実行されることができ、部分的に機械で実行されることができ、スタンドアロンパッケージとして部分的に機械で実行され且つ部分的にリモート機械で実行され、又は完全にリモート機械やサーバーで実行されることができ、

30

【 0 1 2 7 】

本発明の文脈では、機械可読媒体は、命令実行システム、装置又はデバイスが使用されるプログラム、又は、命令実行システム、装置又はデバイスと結合して使用されるプログラムを含み又は記憶されている有形の媒体であってよい。機械可読媒体は、機械可読シグナル媒体、又は機械可読記憶媒体であってよい。機械可読媒体は、電子、磁気、光学、電磁、赤外線、又は半導体のシステム、装置又はデバイス、又はこれらの任意の適当な組み合わせを含んでよいが、これに限定されない。機械可読記憶媒体のより具体的な例示としては、1 本以上のワイヤに基づく電気接続、携帯型コンピューターディスク、ハードディスク、ランダムアクセスメモリ（RAM）、リードオンリーメモリ（ROM）、消去可能なプログラマブルリードオンリーメモリ（EPROM 又はフラッシュメモリ）、光ファイバー、携帯型コンパクトディスクリードオンリーメモリ（CD-ROM）、光学記憶デバイス、磁気記憶デバイス、又はこれらの任意の適当な組み合わせを含んでよい。

40

【 0 1 2 8 】

ユーザーとのインタラクションを提供するために、上述したシステム及び技術は、コンピューターで実施されることができ、当該コンピューターは、ユーザーに情報を表示することに用いられる表示装置（例えば、CRT（ブラウン管）又はLCD（液晶ディスプレイ）モニター）と、キーボード及びポインティングデバイス（例えば、マウス又はトラックボール）とを有し、ユーザーは、当該キーボード及び当該ポインティングデバイスに

50

よって、コンピューターに入力を提供することができる。ユーザーとのインタラクションを提供するために、他の種類の装置も使える。例えば、ユーザーに提供されるフィードバックは、任意の形式の感覚フィードバック（例えば、視覚フィードバック、聴覚フィードバック、又は触覚フィードバック）であってよい。また、任意の形式（音入力、音声入力、触覚入力を含む）によって、ユーザーからの入力を受信してよい。

【 0 1 2 9 】

ここで記述したシステム及び技術は、バックグラウンドコンポーネントを含む計算システム（例えば、データサーバーとして）、又は中間コンポーネントを含む計算システム（例えば、アプリケーションサーバー）、又はフロントエンドコンポーネントを含む計算システム（例えば、グラフィカルユーザーインターフェース又はウェブブラウザを備えるユーザーコンピューターであって、ユーザーは、当該グラフィカルユーザーインターフェース又は当該ウェブブラウザによって上記したシステム及び技術の実施形態とインタラクションを行うことができる）、又はこのようなバックグラウンドコンポーネント、中間コンポーネント、又はフロントエンドコンポーネントの任意の組み合わせを含む計算システムで、実施されることができる。任意の形式又は媒体によるデジタルデータ通信（例えば、通信ネットワーク）を介して、システムのコンポーネントを互いに接続することができる。通信ネットワークの例示としては、ローカルエリアネットワーク（LAN）と、ワイドエリアネットワーク（WAN）と、インターネットとを含む。

10

【 0 1 3 0 】

コンピューターシステムは、クライアントとサーバーとを含んでよい。クライアントとサーバーとは、一般的に互いに離れており、通常、通信ネットワークを介してインタラクションを行う。対応するコンピューターで実行され、互いにクライアント-サーバー関係を有するコンピュータープログラムによって、クライアントとサーバーとの関係を確立する。サーバーは、クラウドサーバー、分散システムのサーバー、又はブロックチェーンを結合したサーバーであってよい。

20

【 0 1 3 1 】

なお、上記の様々なプロセスを使用して、ステップを並べ替え、追加、又は削除することができる。例えば、本発明に記載の各ステップは同時に実行されてもよいし、順に実行されてもよいし、他の順番で実行されてもよく、本開示の技術的解決手段の所望の結果を得られるものであれば、本明細書では特に限定しない。

30

【 0 1 3 2 】

上記具体的な実施形態は、本開示の保護範囲を限定するものと見なされない。当業者が理解したように、設計上の要件や他の要素に基づいて、様々な修正や、組み合わせ、置き換えを行うことができる。本開示の趣旨においてなされた修正、同等な置き換えや改善等は、いずれも本開示の保護範囲に含まれる。

【 図面 】

【 図 1 】



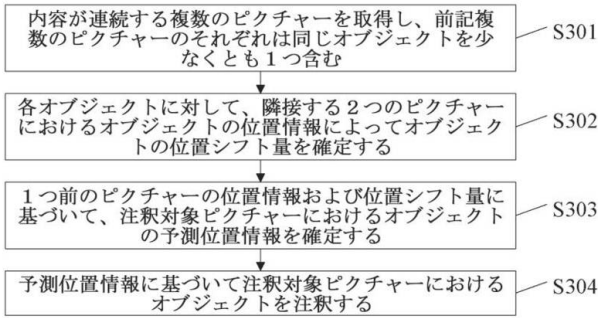
【 図 2 】



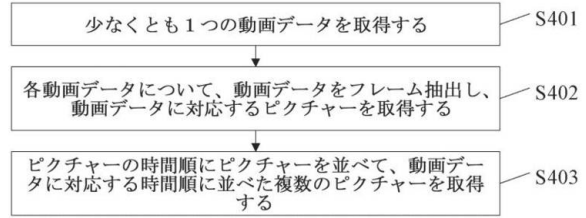
40

50

【 図 3 】

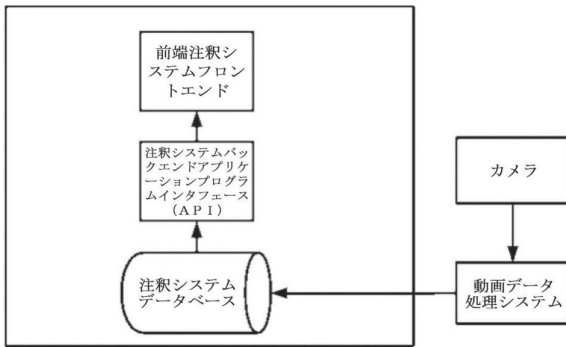


【 図 4 】

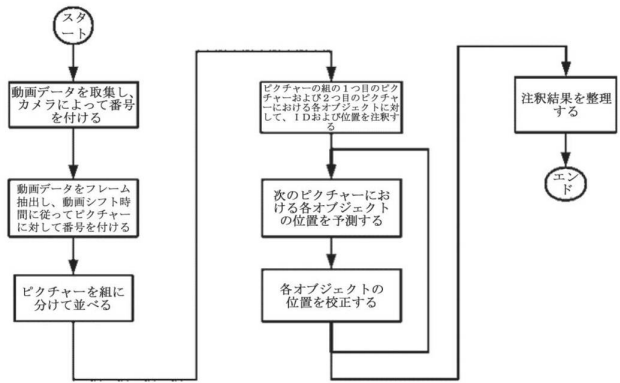


10

【 図 5 】



【 図 6 】



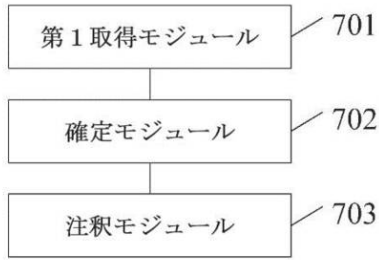
20

30

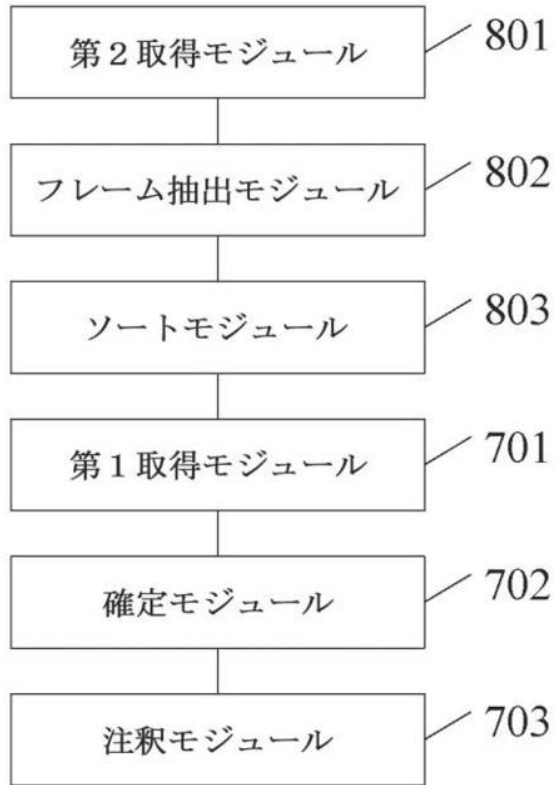
40

50

【図7】



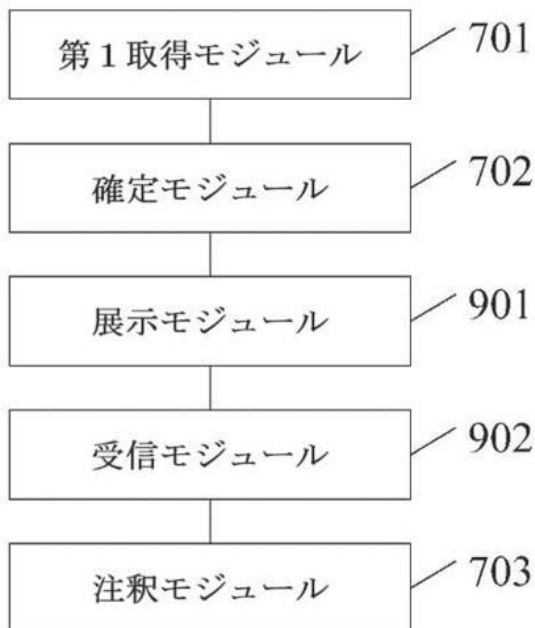
【図8】



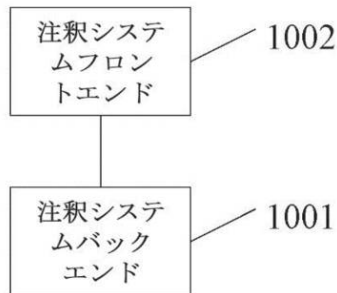
10

20

【図9】



【図10】

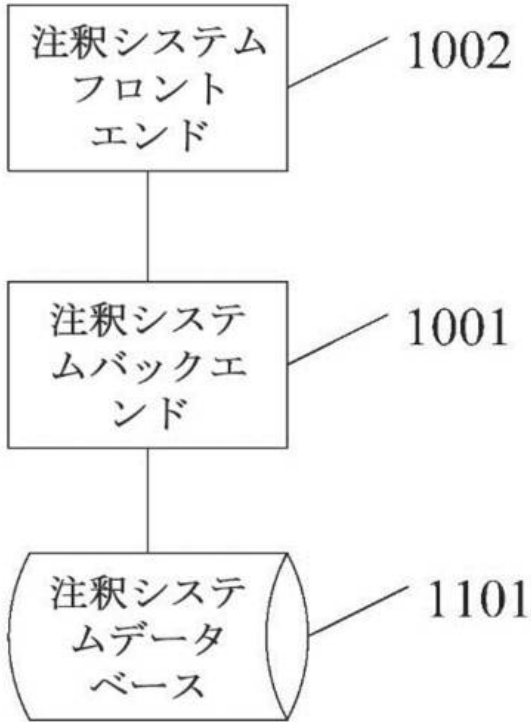


30

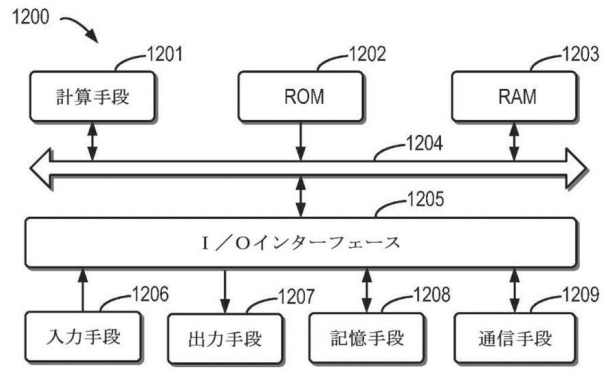
40

50

【図 1 1】



【図 1 2】



10

20

30

40

50

【 手続補正書 】

【 提出日 】 令和5年2月7日(2023.2.7)

【 手続補正 1 】

【 補正対象書類名 】 特許請求の範囲

【 補正対象項目名 】 全文

【 補正方法 】 変更

【 補正の内容 】

【 特許請求の範囲 】

【 請求項 1 】

内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも1つ含むことと、

各オブジェクトに対して、隣接する2つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する2つのピクチャーは2つ前のピクチャーと1つ前のピクチャーとを含み、前記1つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記2つ前のピクチャーは時系列で前記1つ前のピクチャーの前にあって、前記1つ前のピクチャーに隣接するピクチャーであることと、

前記1つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定することと、

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈することと、を含む、

データ注釈方法。

【 請求項 2 】

前記内容が連続する複数のピクチャーを取得することは、

少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものであること、を含む、

請求項1に記載の方法。

【 請求項 3 】

前記方法は、さらに、

少なくとも1つの動画データを取得することと、

各動画データについて、前記動画データをフレーム抽出し、前記動画データに対応するピクチャーを取得することと、

前記ピクチャーの時間順に前記ピクチャーを並べて、前記動画データに対応する時間順に並べた複数のピクチャーを取得することと、を含む、

請求項2に記載の方法。

【 請求項 4 】

前記の少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得することは、

注釈システムデータベースから少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、ここで、前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも1つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものであること、を含む、

請求項2に記載の方法。

10

20

30

40

50

【請求項 5】

前記方法は、さらに、
前記予測位置情報を展示することと、
前記予測位置情報を校正して得られた校正後の予測位置情報を受信することと、を含み

、
前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈することは、

前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈すること、を含む、

請求項 1 ~ 4 のいずれか一項に記載の方法。

10

【請求項 6】

内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも 1 つ含む、第 1 取得モジュールと、

各オブジェクトに対して、隣接する 2 つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する 2 つのピクチャーは 2 つ前のピクチャーと 1 つ前のピクチャーとを含み、前記 1 つ前のピクチャーは時系列で注釈対象ピクチャーの前において、前記注釈対象ピクチャーに隣接するピクチャーであり、前記 2 つ前のピクチャーは時系列で前記 1 つ前のピクチャーの前において、前記 1 つ前のピクチャーに隣接するピクチャーであり、前記 1 つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定する、確定モジュールと、

20

前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈する、注釈モジュールと、を含む、

データ注釈装置。

【請求項 7】

前記第 1 取得モジュールは、少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1 つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

30

請求項 6 に記載の装置。

【請求項 8】

少なくとも 1 つの動画データを取得する、第 2 取得モジュールと、

各動画データについて、前記動画データをフレーム抽出し、前記動画データに対応するピクチャーを取得する、フレーム抽出モジュールと、

前記ピクチャーの時間順に前記ピクチャーを並べて、前記動画データに対応する時間順に並べた複数のピクチャーを取得する、ソートモジュールとを、さらに含む、

請求項 7 に記載の装置。

【請求項 9】

前記第 1 取得モジュールは、注釈システムデータベースから少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、ここで、前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

40

請求項 7 に記載の装置。

【請求項 10】

前記予測位置情報を展示する、展示モジュールと、

50

前記予測位置情報を校正して得られた校正後の予測位置情報を受信する、受信モジュールと、をさらに含み、

前記注釈モジュールは、前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈する、

請求項 6 ~ 9 のいずれか一項に記載の装置。

【請求項 1 1】

注釈システムバックエンドと注釈システムフロントエンドとを含むデータ注釈システムであって、

前記注釈システムバックエンドは、内容が連続する複数のピクチャーを取得し、前記複数のピクチャーのそれぞれは同じオブジェクトを少なくとも 1 つ含み、各オブジェクトに対して、隣接する 2 つのピクチャーにおける前記オブジェクトの位置情報によって前記オブジェクトの位置シフト量を確定し、ここで、前記隣接する 2 つのピクチャーは 2 つ前のピクチャーと 1 つ前のピクチャーとを含み、前記 1 つ前のピクチャーは時系列で注釈対象ピクチャーの前にあって、前記注釈対象ピクチャーに隣接するピクチャーであり、前記 2 つ前のピクチャーは時系列で前記 1 つ前のピクチャーの前にあって、前記 1 つ前のピクチャーに隣接するピクチャーであり、前記 1 つ前のピクチャーの位置情報および前記位置シフト量に基づいて、前記注釈対象ピクチャーにおける前記オブジェクトの予測位置情報を確定し、前記予測位置情報に基づいて前記注釈対象ピクチャーにおける前記オブジェクトを注釈し、

前記注釈システムフロントエンドは、前記注釈システムバックエンドが前記注釈対象ピクチャーにおける前記オブジェクトに対する注釈を展示する、

データ注釈システム。

【請求項 1 2】

前記システムは、さらに、注釈システムデータベースを含み、

前記注釈システムバックエンドは、前記注釈システムデータベースから少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーを取得し、1 つの動画データについて、前記動画データに対応する時間順に並べた複数のピクチャーは、前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

請求項 1 1 に記載のシステム。

【請求項 1 3】

前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、動画データ処理システムが前記注釈システムデータベースに格納するものであり、前記少なくとも 1 つの動画データにそれぞれ対応する時間順に並べた複数のピクチャーは、各前記動画データのそれぞれについて、前記動画データ処理システムが前記動画データをフレーム抽出し、フレーム抽出して得られた前記動画データに対応するピクチャーの時間順に前記ピクチャーを並べて得られたものである、

請求項 1 2 に記載のシステム。

【請求項 1 4】

前記注釈システムバックエンドは、前記予測位置情報を展示し、前記予測位置情報を校正して得られた校正後の予測位置情報を受信し、前記注釈対象ピクチャーに前記校正後の予測位置情報を注釈し、前記校正後の予測位置情報に基づいて確定された位置に前記オブジェクトのタグを注釈する、

請求項 1 1 ~ 1 3 のいずれか一項に記載のシステム。

【請求項 1 5】

少なくとも 1 つのプロセッサと、

前記少なくとも 1 つのプロセッサと通信接続しているメモリと、を含み、

前記メモリには、前記少なくとも 1 つのプロセッサによって実行できる命令が記憶されており、前記少なくとも 1 つのプロセッサによって前記命令を実行することにより、

10

20

30

40

50

前記少なくとも1つのプロセッサが請求項1に記載の方法を実行できる、
電子デバイス。

【請求項16】

コンピューター命令が記憶されている非一時的なコンピューター可読記憶媒体であって、前記コンピューター命令は、前記コンピューターに請求項1に記載の方法を実行させる

、
非一時的なコンピューター可読記憶媒体。

【請求項17】

コンピュータープログラムを含むコンピュータープログラム製品であって、前記コンピュータープログラムは、プロセッサによって実行されると、請求項1に記載の方法を実行させる、

コンピュータープログラム製品。

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/099883

A. CLASSIFICATION OF SUBJECT MATTER G06F 16/51(2019.01)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) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNPAT, WPI, EPODOC, CNKI, IEEE: 图片, 标注, 对象, 偏移, 位置, 时序, 相邻, picture, tag, object, deviation, location, sequence, adjacent		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 114168767 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 11 March 2022 (2022-03-11) claims 1-17	1-17
Y	CN 110032914 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 19 July 2019 (2019-07-19) description, paragraphs [0064]-[0150]	1-17
Y	CN 113034541 A (BEIJING GRIDSUM TECHNOLOGY CO., LTD. et al.) 25 June 2021 (2021-06-25) description, paragraphs [0011]-[0093]	1-17
A	CN 113378958 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 10 September 2021 (2021-09-10) entire document	1-17
A	CN 111860302 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 30 October 2020 (2020-10-30) entire document	1-17
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 25 August 2022		Date of mailing of the international search report 21 September 2022
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China Facsimile No. (86-10)62019451		Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/099883

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 114168767 A	11 March 2022	None	
CN 110032914 A	19 July 2019	None	
CN 113034541 A	25 June 2021	None	
CN 113378958 A	10 September 2021	None	
CN 111860302 A	30 October 2020	None	

10

20

30

40

50

国际检索报告

国际申请号

PCT/CN2022/099883

A. 主题的分类		
G06F 16/51 (2019.01)i		
按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
G06F		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CNPAT, WPI, EPDOC, CNKI, IEEE: 图片, 标注, 对象, 偏移, 位置, 时序, 相邻, picture, tag, object, deviation, location, sequence, adjacent		
C. 相关文件		
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 114168767 A (北京百度网讯科技有限公司) 2022年3月11日 (2022 - 03 - 11) 权利要求1-17	1-17
Y	CN 110032914 A (北京京东尚科信息技术有限公司 等) 2019年7月19日 (2019 - 07 - 19) 说明书第[0064]-[0150]段	1-17
Y	CN 113034541 A (北京国双科技有限公司 等) 2021年6月25日 (2021 - 06 - 25) 说明书第[0011]-[0093]段	1-17
A	CN 113378958 A (北京百度网讯科技有限公司) 2021年9月10日 (2021 - 09 - 10) 全文	1-17
A	CN 111860302 A (北京百度网讯科技有限公司) 2020年10月30日 (2020 - 10 - 30) 全文	1-17
<input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2022年8月25日		2022年9月21日
ISA/CN的名称和邮寄地址		受权官员
中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088		李娜
传真号 (86-10)62019451		电话号码 86-(10)-53961346

PCT/ISA/210 表(第2页) (2015年1月)

10

20

30

40

50

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/099883

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	114168767	A	2022年3月11日	无	
CN	110032914	A	2019年7月19日	无	
CN	113034541	A	2021年6月25日	无	
CN	113378958	A	2021年9月10日	无	
CN	111860302	A	2020年10月30日	无	

10

20

30

40

PCT/ISA/210 表(同族专利附件) (2015年1月)

50

フロントページの続き

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

85, China

(74)代理人 100147485

弁理士 杉村 憲司

(74)代理人 230118913

弁理士 杉村 光嗣

(74)代理人 100163511

弁理士 辻 啓太

(72)発明者 王 磊

中華人民共和国 ペキン 100085, ハイディアン ディストリクト, シャンディ テンス ストリート, 10番, バイドウ キャンパス 2階

(72)発明者 王 迎

中華人民共和国 ペキン 100085, ハイディアン ディストリクト, シャンディ テンス ストリート, 10番, バイドウ キャンパス 2階

(72)発明者 ザン 暁 ティン

中華人民共和国 ペキン 100085, ハイディアン ディストリクト, シャンディ テンス ストリート, 10番, バイドウ キャンパス 2階

Fターム(参考) 5B175 DA04 FA01 FB02

5L096 AA06 CA04 CA05 DA01 FA18 FA69

【要約の続き】

ーにおいて、同じオブジェクトを注釈する速度を高めることができる。

【選択図】図3