

明 細 書

発明の名称： 情報処理装置

技術分野

[0001] 本発明は、情報処理装置にかかり、特に、複数の認識エンジンによる認識結果を利用した新たな出力を行う情報処理装置に関する。

背景技術

[0002] 情報処理技術の発達に伴い、様々な認識処理を行う認識エンジンの開発が行われている。例えば、静止画像データから人物を特定する認識エンジン、動画画像データから人間の動線をトレースする位置情報を生成する認識エンジン、音声データからテキストデータを生成する認識エンジンなど、様々な認識エンジンが存在している。一例として、特許文献1には、画像認識や音声認識等の認識器を搭載しているロボットが開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2004-283959号公報

発明の概要

発明が解決しようとする課題

[0004] しかしながら、上述したように様々な認識エンジンが開発されているものの、これら各認識エンジンは個別のアプリケーションに依存して特定のシステムに組み込まれているため、他の用途向けに再利用することが困難である。仮に再利用しようとした場合であっても、各認識エンジンの出力形式は当該認識エンジン毎に異なっているため、これらを利用可能な新たなシステムの開発に多大なコストがかかる。従って、複数の認識エンジンからの出力結果を利用した要求がアプリケーションからあった場合であっても、柔軟に対応することができず、また、そのようなシステム開発には多大なコストがかかる、という問題があった。

[0005] このため、本発明の目的は、上述した課題である、低コストにて複数の認

識エンジンからの出力結果を再利用することが困難であること、を解決することにある。

課題を解決するための手段

[0006] 上記目的を達成するために、本発明の一形態である情報処理装置は、
認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、
前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を備え、
前記認識結果取得手段は、前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、
前記統合認識結果出力手段は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、
という構成を取る。

[0007] また、本発明の他の形態であるプログラムは、
情報処理装置に、
認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、
前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を実現させると共に、
前記認識結果取得手段は、前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、
前記統合認識結果出力手段は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、
ことを実現させるためのプログラムである。

[0008] また、本発明の他の形態である情報処理方法は、
認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力され、前記複数の認識エンジンにて共通するデータ形式の各認識

結果情報を取得し、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報に基づ

いて当該各認識結果情報を統合して新たな認識結果を出力する、

という構成を取る。

発明の効果

[0009] 本発明は、以上のような構成を取ることで、低コストにて複数の認識

エンジンからの出力結果を再利用することができる情報処理装置を提供する

ことができる。

図面の簡単な説明

[0010] [図1] 本発明における中間層装置の構成を示すブロック図である。

[図2] 本発明の実施形態 1 における中間層装置に登録された認識結果の構造の

一例を示す図である。

[図3] 本発明の実施形態 1 における中間層装置に登録された認識結果の構造の

一例を示す図である。

[図4] 本発明の実施形態 1 における中間層装置に登録された認識結果の構造の

一例を示す図である。

[図5] 本発明の実施形態 1 における中間層装置の動作を示すフローチャートで

ある。

[図6] 本発明の実施形態 1 における中間層装置にて認識エンジンから取得した

認識結果の一例を示す図である。

[図7] 本発明の実施形態 1 における中間層装置にて認識エンジンから取得した

認識結果の一例を示す図である。

[図8] 本発明の実施形態 1 における中間層装置にて認識エンジンから取得した

認識結果を統合するときの様子を示す図である。

[図9] 本発明の実施形態 1 における中間層装置にて認識エンジンから取得した

認識結果を統合したときの様子を示す図である。

[図10] 本発明の実施形態 2 における中間層装置に登録された認識結果の構造

の一例を示す図である。

[図11]本発明の実施形態2における中間層装置に登録された認識結果の構造の一例を示す図である。

[図12]本発明の実施形態2における中間層装置にて認識エンジンから取得した認識結果を統合するときの様子を示す図である。

[図13]本発明の実施形態2における中間層装置にて認識エンジンから取得した認識結果を統合したときの様子を示す図である。

[図14]本発明の付記1における情報処理装置の構成を示すブロック図である。

発明を実施するための形態

[0011] <実施形態1>

本発明の第1の実施形態を、図1乃至図9を参照して説明する。図1乃至図4は、本実施形態における中間層装置の構成を説明するための図であり、図5乃至図9は、動作を説明するための図である。

[0012] [構成]

実施形態1における情報処理システムは、図1に示すように、複数の認識エンジン21-23と、これら認識エンジン21-23に接続された中間層装置10と、この中間層装置10に接続された情報処理端末に実装されたアプリケーション部30と、により構成されている。

[0013] まず、上記認識エンジン21-23は、認識対象データに対してそれぞれ異なる認識処理を行う情報処理装置である。認識エンジン21-23としては、例えば、認識対象データである静止画像データや動画データ、音声データなどから人物を特定する認識エンジンや、動画データから人間の動線をトレースする位置情報を生成する認識エンジン、認識対象データである音声データからテキストデータを生成する認識エンジンなどがある。そして、本実施形態における認識エンジンは、動画データ内から予め設定された人物を特定する人物認識エンジン（以下、符号21で表す）と、動画データ内から物体の移動軌跡を表す動線を認識する動線認識エンジン（以下、符号22で表す）と、である。但し、認識エンジンは、上述した認識処理を行う

ものに限定されず、また、その数も上述した2つや図1で示す3つであることに限定されない。例えば、上記人物認識エンジンは、人物のみならず、予め設定された物体を特定する物体認識エンジンであってもよい。

[0014] 具体的に、上記人物認識エンジン21は、認識処理を要求された「特定の人物」の予め登録された特徴量データを、動画像データの各日時におけるフレーム画像内から検出する。そして、検出された場合には、その日時の動画像データを取得したカメラが設置された位置に「特定の人物」が所在していたことを認識し、当該認識した「特定の人物」に設定された人物識別情報 (person ID) と、検出した日時及び位置を表す所在情報とを、認識結果情報として出力する。なお、本実施形態における人物認識エンジン21は、上述した認識結果情報を、後述するように中間層装置10に予め登録されているデータ形式に対応して当該中間層装置10に出力するが、そのデータ形式については後述する。

[0015] また、動線認識エンジン22は、具体的に、動画像データ内で移動している物体を検出し、その物体を時系列的な移動軌跡を認識する。そして、移動している物体を識別する物体識別情報 (object ID) と、検出された一連の移動軌跡を識別する移動軌跡識別情報 (trajectory ID) と、検出された各日時とそのときの位置 (座標) からなる動線情報とを、認識結果情報として出力する。なお、本実施形態における動線認識エンジン22は、上述した認識結果情報を、後述するように中間層装置10に予め登録されているデータ形式にて当該中間層装置10に出力するが、そのデータ形式については後述する。

[0016] 次に、アプリケーション部30について説明する。アプリケーション部30は、中間層装置10に接続された情報処理端末に実装されており、上述した認識エンジン21-23による認識結果を中間層装置10に要求する機能を有する。本実施形態では、特に、人物認識エンジン21による認識結果と、動線認識エンジン22による認識結果と、を再利用して統合した新たな認識結果を要求する。具体的に、本実施形態で説明する例では、アプリケーシ

ョン部30は、「特定の人物の動線」を中間層装置10に要求する。

[0017] 次に、中間層装置10について説明する。中間層装置10は、演算装置と記憶装置とを備えた情報処理装置である。そして、中間層装置10は、図1に示すように、演算装置にプログラムが組み込まれることで構築された、認識エンジン出力取得部11と、インスタンス・クラス対応判断部12と、出力インスタンス保持部13と、クラス構造登録部15と、を備えている。また、中間層装置10は、記憶装置にクラス構造保持部14とクラスID保持部16とを備えている。以下、各構成について詳述する。

[0018] まず、認識エンジン出力取得部11（認識結果取得手段）は、上述したように、人物認識エンジン21と動線認識エンジン22から、それぞれが認識処理により生成した認識結果情報を取得する。すると、インスタンス・クラス対応判断部12（認識結果取得手段）は、取得した各認識結果情報を、予め定義されたデータ形式を表すクラス構造にて保持する。

[0019] ここで、人物認識エンジン21に対応する認識結果情報のデータ形式であり、例えばグラフ構造又は木構造で定義された認識結果定義情報であるクラス構造は、クラス構造登録部15からクラス構造保持部14（定義情報記憶手段）及びクラスID保持部16に保存され定義されている。具体的に、人物認識エンジン21に対応するクラス構造C1は、図2に示すように、検出された特定の人物を識別する人物識別情報（person ID）C11の下位に、当該人物が検出された画像（フレーム）が撮影された位置の座標（X座標値、Y座標値）及び日時（検知時刻）にて表される所在情報C12を配置したグラフ構造にて定義されている。これにより、インスタンス・クラス対応判断部12は、図6に示すように、人物認識エンジン21に対応するクラス構造C1の人物識別情報（person ID）や位置等（X座標値、Y座標値、検知時刻）に、実際に人物認識エンジン21から取得した認識結果情報に含まれる値を当てはめて保持する。

[0020] また、動線認識エンジン22に対応する認識結果情報のデータ形式であり、例えばグラフ構造又は木構造で定義する認識結果定義情報であるクラス構

造C 2は、図3に示すように、検出された移動する物体を識別する物体識別情報 (object ID) C 2 1の下位に、その物体の一連の移動軌跡を識別する移動軌跡識別情報 (trajectory ID) C 2 2を配置すると共に、さらにこの下位に、物体が検出された移動軌跡を形成する各日時 (検知時刻) 及びそのときの位置の座標 (X座標値、Y座標値) にて表される動線情報C 2 3を配置したグラフ構造にて定義されている。これにより、インスタンス・クラス対応判断部1 2は、図7に示すように、動線認識エンジン2 2に対応するクラス構造C 2の物体識別情報 (object ID) や位置等 (X座標値、Y座標値、検知時刻) に、実際に動線認識エンジン2 2から取得した認識結果情報に含まれる値を当てはめて保持する。

[0021] そして、中間層装置1 0が備える出力インスタンス保持部1 3 (統合認識結果出力手段) は、上述したように予め定義されたクラス構造C 1, C 2に当てはめて保持した人物認識エンジン2 1からの認識結果情報と、動線認識エンジン2 2からの認識結果情報と、を統合して新たな認識結果として保持する。具体的に、出力インスタンス保持部1 3は、予めクラス構造保持部1 4 (定義情報記憶手段) 及びクラスID保持部1 6に記憶されており、各認識結果情報を統合する条件と統合後の新たな認識結果とを表すデータ形式である統合形式をグラフ構造又は木構造で表した統合定義情報であるクラス構造C 3に従って、各認識結果情報の統合を行う。

[0022] ここで、各認識結果情報を統合する条件と統合後の新たな認識結果のデータ形式を表すクラス構造は、図4に示すように、人物認識エンジン2 1からの認識結果C 1の一部である検出された特定の人物を識別する人物識別情報 (person ID) C 1 1に、動線認識エンジン2 2からの認識結果C 2の一部である移動軌跡識別情報 (trajectory ID) C 2 2及び動線情報C 2 3を関連付けて統合する構造C 3を表している。

[0023] このとき、上述したように各認識結果情報を統合する条件は、人物認識エンジン2 1から出力された人物識別情報 (person ID) C 1 1と、動線認識エンジン2 2から出力された物体識別情報 (object ID) C 2 1の一部と、が一

致することである。これを判定するために、出力インスタンス保持部 13 は、各エンジン 21, 22 にて検出された各認識結果情報内に、予め定められた基準により相互に対応すると判断される情報があるか否かを調べる。具体的には、動線認識エンジン 22 から出力された動線情報 C 23 内に、人物認識エンジン 21 から出力された所在情報 C 12 つまり特定の人物が検出された日時及び位置と一致する情報があるかどうかを調べる。そして、出力インスタンス保持部 13 は、所在情報 C 12 と動線情報 C 23 の一部とが一致した場合には、当該一致した所在情報 C 12 に対応する人物識別情報 (person ID) C 11 に、一致した動線情報 C 23 及びその上位の移動軌跡識別情報 C 22 を関連付けて、図 9 に示すように、新たな認識結果を保持する。

[0024] 但し、各認識結果情報を統合する条件は、上述したように、各エンジン 21, 22 にて検出された各認識結果情報内に相互に一致する情報があることに限定されない。例えば、動線認識エンジン 22 から出力された動線情報 C 23 内に、人物認識エンジン 21 から出力された所在情報 C 12 つまり特定の人物が検出された日時及び位置の「近傍」の情報がある場合に、上記動線情報 C 23 と所在情報 C 12 とが相互に対応すると判断して、その所在情報 C 12 に対応する人物識別情報 C 11 に、動線情報 C 23 及びその上位の移動軌跡識別情報 C 22 を関連付けて統合してもよい。このとき、特定の人物が検出された日時及び位置の「近傍」とは、例えば、検出された日時及び位置に対して予め設定された許容範囲内の日時及び位置を含むこととする。

[0025] なお、上記では、各認識エンジン 21, 22 による認識結果情報のデータ形式や、出力インスタンス保持部 13 による各認識結果情報の統合形式を、グラフ構造又は木構造で定義している場合を例示したが、別のデータ形式にて定義してもよい。

[0026] また、出力インスタンス保持部 13 は、上述のように生成して保持した新たな認識結果を、アプリケーション部 30 に出力する。

[0027] [動作]

次に、上述した情報処理システム、特に、中間層装置 10 の動作を、図 5

のフローチャート及び図6乃至図9のデータ遷移図を参照して説明する。

[0028] まず、中間層装置10は、アプリケーション部30から認識結果の要求を受ける（ステップS1）。ここでは、「特定の人物の動線」といった認識結果の要求を受けたこととする。

[0029] 続いて、中間層装置10は、アプリケーション部30から要求された認識結果を出力するために利用する各認識エンジンからの認識結果情報を取得する。本実施形態では、人物認識エンジン21と、動線認識エンジン22から、それぞれが認識処理により生成した各認識結果情報を取得する（ステップS2）。そして、中間層装置10は、取得した各認識結果情報を、予め定義されたクラス構造に従って保持する。具体的に、人物認識エンジン21から取得した認識結果情報を、図2に示す人物認識エンジンに対応したクラス構造C1に当てはめ、図6に示すように保持する。また、動線認識エンジン22から取得した認識結果情報を、図2に示す動線認識エンジンに対応したクラス構造C2に当てはめ、図7に示すように保持する。

[0030] 続いて、中間層装置10は、図6、図7に示すように、各クラス構造C1'、C2'に当てはめて保持した人物認識エンジン21及び動線認識エンジン22からの各認識結果情報に含まれる人物識別情報（person ID）C11'と物体識別情報（object ID）C21'との同一性を調べる（ステップS3、図8の矢印Y2参照）。具体的に、中間層装置10は、図8の矢印Y1に示すように、動線認識エンジン22から出力された動線情報C23'内に、人物認識エンジン21から出力された所在情報C12'と一致する情報があるかどうかを調べる。

[0031] 中間層装置10は、動線情報C23'内に所在情報C12'と一致する情報が存在した場合には、これら一致した各情報を含む人物識別情報（person ID）C11'と物体識別情報（object ID）C21'とが同一であると判断する。そして、中間層装置10は、図9に示すように、人物識別情報（person ID）C11'に、これと同一であると判断した物体識別情報（object ID）C21'の下位に位置する移動軌跡識別情報C22'及びさらにその下位に位

置する動線情報C 2 3' を関連付けて統合し、新たな認識結果として保持する（ステップS 4）。

[0032] その後、中間層装置1 0は、上述したように生成して保持した新たな認識結果を、アプリケーション部3 0に出力する（ステップS 5）。

[0033] なお、上記では、人物認識エンジン2 1から出力された所在情報C 1 2' と動線認識エンジン2 2から出力された動線情報C 2 3' の一部とが一致する場合に、各エンジン2 1, 2 2からの各認識結果を統合したが、統合する条件は上述したことに限定されない。例えば、人物認識エンジン2 1からの認識結果情報に含まれる人物識別情報（person ID）C 1 1' と、動線認識エンジン2 2からの認識結果情報に含まれる物体識別情報（object ID）C 2 1' と、が一致する場合に、上述したように各認識結果情報を統合してもよい。あるいは、上述したように、動線認識エンジン2 2から出力された動線情報C 2 3' 内に、人物認識エンジン2 1から出力された所在情報C 1 2' の「近傍」の情報がある場合に、各認識結果情報を統合してもよい。

[0034] また、上記では、アプリケーション部3 0から認識結果の要求を受けてから（ステップS 1）、各認識エンジンから各認識結果を取得して（ステップS 2）、当該各認識結果を合成する（ステップS 4）、という手順で処理を行っているが、かかる手順で処理を実行することに限定されない。例えば、図5に示すステップS 2, S 3, S 4の処理、つまり、各認識エンジンから各認識結果を取得して当該各認識結果を合成する処理を予め行って、合成された結果を中間層装置1 0内に蓄積しておく。その後、アプリケーション部3 0からの認識結果の要求を受けて、かかる要求に対応する合成結果を、蓄積されているデータ中から抽出して、アプリケーション部3 0に対して出力する、という手順で処理を行ってもよい。

[0035] 以上のように、本実施形態における中間層装置1 0を設け、複数の認識エンジンで共通のデータ形式で認識結果情報を保持することで、各認識エンジンの認識結果を他の用途向けに再利用することが容易となる。従って、認識エンジンや認識結果を要求するアプリケーションの開発が容易となり、低コ

ストにて認識エンジンの汎用性が向上する。

[0036] <実施形態 2>

次に、本発明の第 2 の実施形態を、図 1、図 5、図 10 乃至図 13 を参照して説明する。図 1 は、本実施形態における中間層装置の構成を説明するための図であり、図 5、図 10 乃至図 13 は、動作を説明するための図である。

[0037] [構成]

実施形態 2 における情報処理システムは、上述した実施形態 1 と同様に、図 1 に示すように、複数の認識エンジン 21-23 と、これら認識エンジン 21-23 に接続された中間層装置 10 と、この中間層装置 10 に接続された情報処理端末に実装されたアプリケーション部 30 と、により構成されている。

[0038] そして、本実施形態における各認識エンジンは、いずれも認識対象データから人物を特定する人物認識エンジンであるが、それぞれ認識手法が異なる。例えば、第 1 の人物認識エンジン（以下、符号 21 で表す）は、動画像データから予め設定された人物の顔の特徴データに基づいて、特定の人物を認識するものである。また、第 2 の人物認識エンジン（以下、符号 22 で表す）は、音声データから予め設定された人物の声の特徴データに基づいて、特定の人物を認識するものである。さらに、第 3 の人物認識エンジン（以下、符号 23 で表す）は、音声データから予め設定された発話内容（キーワード）に基づいて、特定の人物を認識するものである。但し、認識エンジンは、処理を行うものに限定されず、また、その数も上述した数に限定されない。例えば、上記人物認識エンジンは、人物のみならず、予め設定された物体を特定する物体認識エンジンであってもよい。

[0039] そして、本実施形態における各人物認識エンジン 21, 22, 23 は、認識結果情報を中間層装置 10 に予め登録されている共通のデータ形式にて当該中間層装置 10 に出力する。具体的に、各人物認識エンジン 21, 22, 23 は、各認識手法により認識した結果、候補となる単数あるいは複数の人

物を特定する情報からなる人物候補情報を、認識結果情報として出力する。

[0040] これに対応して、中間層装置 10 は、各人物認識エンジン 21, 22, 23 に対応する各認識結果情報のデータ形式を定義した、例えばグラフ構造又は木構造の認識結果定義情報であるクラス構造 C101 を、図 10 に示すようにクラス構造保持部 14 等に記憶している。具体的に、各人物認識エンジン 21, 22, 23 に対応するクラス構造 C101 は、図 10 に示すように、認識処理を行った人物認識エンジンを識別するエンジン識別情報 C110 の下位に、認識候補となる単数あるいは複数の人物を識別する人物識別情報 C111, C112, C113 を配置したグラフ構造にて定義されている。なお、ここでは、認識処理によって候補となる可能性が高い順から 3 人の人物の人物識別情報が、上から順番に配置されて格納されるよう定義されている。これにより、中間層装置 10 のインスタンス・クラス対応判断部 12 は、認識エンジン出力取得部 11 にて各人物認識エンジン 21, 22, 23 から取得した認識結果情報を、図 12 の符号 C101-1, C101-2, C101-3 に示すように、各人物認識エンジン 21, 22, 23 毎にクラス構造 C101 に当てはめて保持する。

[0041] そして、中間層装置 10 が備える出力インスタンス保持部 13 は、上述したように予め定義されたクラス構造 C101 に当てはめて保持した各人物認識エンジン 21, 22, 23 からの認識結果情報を統合して、新たな認識結果として保持する。具体的に、出力インスタンス保持部 13 は、予めクラス構造保持部 14 等に定義されており、各認識結果情報を統合する条件と統合後の新たな認識結果を表すデータ形式である統合形式をグラフ構造又は木構造で定義する統合定義情報であるクラス構造 C102 に従って、各認識結果情報の統合を行う。

[0042] ここで、各認識結果情報を統合する条件と統合後の新たな認識結果のデータ形式を表すクラス構造は、図 11 に示すように、各人物認識エンジン 21, 22, 23 からの認識結果情報 C120 である候補とされた人物の人物識別情報のうち、一人の人物識別情報 (person ID) C121 を特定するよう統

合する構造を表している。

[0043] このとき、上述したように各認識結果情報を統合する条件は、異なる人物認識エンジン 21, 22, 23 から出力された各人物識別情報 (person ID) の相互間で一致する人物識別情報が存在することである。これを判定するために、出力インスタンス保持部 13 は、各エンジン 21, 22 にてそれぞれ出力された各認識結果情報内に、予め定められた基準により相互に対応すると判断される人物識別情報があるか否かを調べる。具体的には、3つの人物認識エンジン 21, 22, 23 からそれぞれ出力された3つの認識結果情報の全てに、同一の人物識別情報があるかどうかを調べ、同一の人物識別情報 (person ID) を特定し、この人物識別情報を新たな認識結果として保持する。

[0044] 但し、各認識結果情報を統合する条件は、上述したように、各エンジン 21, 22, 23 にて検出された各認識結果情報内に相互に一致する情報があることに限定されない。例えば、まず、検出した各人物識別情報 (person ID) に対応する人物のデータから抽出された特徴値データである各属性情報の類似性を、予め定められた基準に従って判定する。そして、かかる基準にて類似性度合が予め設定された範囲内である場合には、各人物識別情報の人物が同一人物であると判断する。

[0045] なお、上記では、各認識エンジン 21, 22, 23 による認識結果情報のデータ形式や、出力インスタンス保持部 13 による各認識結果情報の統合形式を、グラフ構造又は木構造で定義している場合を例示したが、別のデータ形式にて定義してもよい。

[0046] また、出力インスタンス保持部 13 は、上述のように生成して保持した新たな認識結果を、アプリケーション部 30 に出力する。

[0047] [動作]

次に、上述した情報処理システム、特に、中間層装置 10 の動作を、図 5 のフローチャート及び図 10 乃至図 13 のデータ遷移図を参照して説明する。

[0048] まず、中間層装置 10 は、アプリケーション部 30 から認識結果の要求を受ける（ステップ S1）。例えば、「人物を特定」といった認識結果の要求を受けたこととする。

[0049] 続いて、中間層装置 10 は、アプリケーション部 30 から要求された認識結果を出力するために利用する各認識エンジンからの認識結果情報を取得する（ステップ S2）。そして、中間層装置 10 は、取得した各認識結果情報を、予め定義されたクラス構造に従って保持する。

[0050] 具体的に、本実施形態では、動画像データから人物の顔の特徴データに基づいて人物を認識する第 1 の人物認識エンジン 21 から取得した候補となる 3 つの人物識別情報を、図 12 の符号 C101-1 に示すようにクラス構造に当てはめて保持する。また、音声データから人物の声の特徴データに基づいて人物を認識する第 2 の人物認識エンジン 22 から取得した候補となる 3 つの人物識別情報を、図 12 の符号 C101-2 に示すようにクラス構造に当てはめて保持する。さらに、音声データから発話内容（キーワード）に基づいて人物を認識する第 3 の人物認識エンジン 23 から取得した候補となる 3 つの人物識別情報を、図 12 の符号 C101-3 に示すようにクラス構造に当てはめて保持する。

[0051] 続いて、中間層装置 10 は、図 12 に示すように、定義されたクラス構造に当てはめて保持した各人物認識エンジン 21, 22, 23 からのそれぞれ候補となる人物識別情報（person ID）C111-1～C113-1, C111-2～C113-2, C111-3～C113-3 の中で、相互に対応するもの、ここでは共通する人物識別情報を調べる（ステップ S3）。すると、図 12 に示す例では、全ての認識結果情報で、符号 C111-1, C112-2, C111-3 に示すように「Person-A」の人物識別情報が共通するため、図 13 に示すように「Person-A」で表される 1 人の人物識別情報 C121' に、認識結果を統合する（ステップ S4）。あるいは、各人物識別情報に付随する各属性情報の類似性を判定し、類似性度合が基準よりも高い場合に、各人物識別情報の一の人物が同一人物であると判断して、認識結果を

統合してもよい。

[0052] なお、上記では、アプリケーション部30から認識結果の要求を受けてから（ステップS1）、各認識エンジンから各認識結果を取得して（ステップS2）、当該各認識結果を合成する（ステップS4）、という手順で処理を行っているが、かかる手順で処理を実行することに限定されない。例えば、まず、各認識エンジンから各認識結果を取得して当該各認識結果を合成する処理を予め行って、合成された結果を中間層装置10内に蓄積しておく。その後、アプリケーション部30からの認識結果の要求を受けて、かかる要求に対応する合成結果を蓄積されているデータの中から抽出して、アプリケーション部30に対して出力する、という手順で処理を行ってもよい。

[0053] その後、中間層装置10は、上述したように生成して保持した新たな認識結果を、アプリケーション部30に出力する（ステップS5）。

[0054] 以上のように、本実施形態における中間層装置10を設け、複数の認識エンジンで共通のデータ形式で認識結果情報を保持することで、各認識エンジンの認識結果を他の用途向けに再利用することが容易となる。従って、認識エンジンや認識結果を要求するアプリケーションの開発が容易となり、低コストにて認識エンジンの汎用性が向上する。

[0055] なお、上記では、各人物認識エンジン21、22、23から出力された候補となる人物のうち、全てに共通して含まれる人物を特定するよう認識結果を統合したが、統合する条件は上述したことに限定されない。例えば、認識エンジンの数に対して過半数の認識エンジンからの認識結果に含まれている人物を特定するよう認識結果を統合するなど、他の方法にて認識結果を統合してもよい。また、上記では、各人物認識エンジン21、22、23が3人の候補となる人物を認識結果として出力している場合を例示したが、1人の候補のみを認識結果としてもよく、他の複数の候補を認識結果としてもよい。

[0056] <付記>

上記実施形態の一部又は全部は、以下の付記のようにも記載されうる。以

下、本発明における情報処理装置の構成の概略を、図 14 を参照して説明する。但し、本発明は、以下の構成に限定されない。

[0057] (付記 1)

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジン 211, 212, 213 から出力された各認識結果情報を取得する認識結果取得手段 201 と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段 202 と、を備え、

前記認識結果取得手段 201 は、前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、

前記統合認識結果出力手段 202 は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、
情報処理装置 200。

[0058] (付記 2)

付記 1 に記載の情報処理装置であって、

前記統合認識結果出力手段は、前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、
情報処理装置。

[0059] (付記 3)

付記 1 又は 2 に記載の情報処理装置であって、

前記認識結果取得手段にて前記複数の認識エンジンからそれぞれ取得する前記認識結果情報のデータ形式を定義する認識結果定義情報と、前記統合認識結果出力手段にて前記各認識結果情報を統合する際の統合形式を定義する統合定義情報と、を記憶した定義情報記憶手段を備え、

前記認識結果取得手段は、前記認識結果定義情報にて定義されたデータ形式にて前記各認識結果情報を取得し、

前記統合認識結果出力手段は、前記統合定義情報にて定義された統合形式にて前記各認識結果情報を統合して前記新たな認識結果を出力する、
情報処理装置。

[0060] (付記 4)

付記 1 乃至 3 のいずれかに記載の情報処理装置であって、
前記認識結果取得手段にて前記複数の認識エンジンからそれぞれ取得する前記認識結果情報は、グラフ構造又は木構造のデータ形式である、
情報処理装置。

[0061] (付記 5)

付記 2 に記載の情報処理装置であって、
前記認識結果取得手段は、前記認識対象データから物体を認識する物体認識エンジンと、前記認識対象データから所定の物体の移動軌跡を表す動線を認識する動線認識エンジンと、から前記認識結果情報をそれぞれ取得し、
前記統合認識結果出力手段は、前記物体認識エンジンから取得した前記認識結果情報である物体情報に、前記動線認識エンジンから取得した前記認識結果情報である所定の物体の動線情報を関連付けて統合し、前記新たな認識結果として出力する、
情報処理装置。

[0062] (付記 6)

付記 5 に記載の情報処理装置であって、
前記認識結果取得手段は、前記物体認識エンジンが前記認識対象データから認識した物体の位置とその日時を表す所在情報を含む前記認識結果情報を前記物体認識エンジンから取得すると共に、前記動線認識エンジンが前記認識対象データから認識した所定の物体の動線情報を前記動線認識エンジンから取得し、
前記統合認識結果取得手段は、前記物体認識エンジンから取得した前記所在情報に予め定められた基準により対応する情報が前記動線認識エンジンから取得した前記動線情報に含まれている場合に、前記所在情報に対応する情

報が含まれた前記物体認識エンジンからの前記認識結果情報にて特定される物体情報に、前記動線認識エンジンからの前記認識結果である前記動線情報を関連付けて統合し、前記新たな認識結果として出力する、
情報処理装置。

[0063] (付記 7)

付記 2 に記載の情報処理装置であって、

前記認識結果取得手段は、前記認識対象データからそれぞれ異なる認識処理にて物体を認識する複数の物体認識エンジンから前記認識結果情報をそれぞれ取得し、

前記統合認識結果出力手段は、前記複数の物体認識エンジンからそれぞれ取得した前記各認識結果情報である各物体情報に基づいて一の物体情報を特定して統合し、前記新たな認識結果を出力する、
情報処理装置。

[0064] (付記 8)

付記 7 に記載の情報処理装置であって、

前記認識結果取得手段は、前記複数の物体認識エンジンが前記認識対象データから認識した物体の候補を表す物体候補情報を含む前記認識結果情報を前記複数の物体認識エンジンからそれぞれ取得し、

前記統合認識結果取得手段は、前記複数の物体認識エンジンからそれぞれ取得した前記各物体候補情報で予め定められた基準により相互に対応する物体情報を特定して、前記新たな認識結果として出力する、
情報処理装置。

[0065] (付記 9)

情報処理装置に、

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を實現させると共

に、

前記認識結果取得手段は、前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、

前記統合認識結果出力手段は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、
ことを実現させるためのプログラム。

[0066] (付記 10)

付記 9 に記載のプログラムであって、

前記統合認識結果出力手段は、前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、
ことを実現させるためのプログラム。

[0067] (付記 11)

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力され、前記複数の認識エンジンにて共通するデータ形式の各認識結果情報を取得し、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報に基づいて当該各認識結果情報を統合して新たな認識結果を出力する、
情報処理方法。

[0068] (付記 12)

付記 11 に記載の情報処理方法であって、

前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、
情報処理方法。

[0069] (付記 13)

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を備え、

前記認識結果取得手段は、前記認識対象データから物体を認識する物体認識エンジンが前記認識対象データから認識した物体の位置とその日時を表す所在情報を含む前記認識結果情報であり前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を、前記物体認識エンジンから取得すると共に、前記認識対象データから所定の物体の移動軌跡を表す動線を認識する動線認識エンジンが前記認識対象データから認識した所定の物体の動線情報を含む前記認識結果情報であり前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を、前記動線認識エンジンから取得し、

前記統合認識結果出力手段は、前記物体認識エンジンから取得した前記所在情報に予め定められた基準により対応する情報が前記動線認識エンジンから取得した前記動線情報に含まれている場合に、前記所在情報に対応する情報が含まれた前記物体認識エンジンからの前記認識結果情報にて特定される物体情報に、前記動線認識エンジンからの前記認識結果である前記動線情報を関連付けて統合し、前記新たな認識結果として出力する、
情報処理装置。

[0070] (付記 14)

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を備え、

前記認識結果取得手段は、前記認識対象データからそれぞれ異なる認識処理にて物体を認識する複数の物体認識エンジンが前記認識対象データから認識した物体の候補を表す物体候補情報を含む前記認識結果情報であり前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を、前記複数の物体認識エンジンからそれぞれ取得し、

前記統合認識結果出力手段は、前記複数の物体認識エンジンからそれぞれ

取得した前記各物体候補情報で予め定められた基準により相互に対応する一の物体情報を特定して統合し、前記新たな認識結果として出力する、
情報処理装置。

[0071] なお、上記各実施形態においてプログラムは、記憶装置に記憶されていたり、コンピュータが読み取り可能な記録媒体に記録されている。例えば、記録媒体は、フレキシブルディスク、光ディスク、光磁気ディスク、及び、半導体メモリ等の可搬性を有する媒体である。

[0072] 以上、上記各実施形態を参照して本願発明を説明したが、本願発明は、上述した実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明の範囲内で当業者が理解しうる様々な変更をすることができる。

[0073] なお、本発明は、日本国にて2011年2月17日に特許出願された特願2011-31613の特許出願に基づく優先権主張の利益を享受するものであり、当該特許出願に記載された内容は、全て本明細書に含まれるものとする。

符号の説明

- [0074] 10 中間層装置
- 11 認識エンジン出力取得部
 - 12 インスタンス・クラス対応判断部
 - 13 出力インスタンス保持部
 - 14 クラス構造保持部
 - 15 クラス構造登録部
 - 16 クラスID保持部
 - 21, 22, 23 認識エンジン
 - 30 アプリケーション部
 - 200 情報処理装置
 - 201 認識結果取得手段
 - 202 統合認識結果出力手段
 - 211, 212, 213 認識エンジン

請求の範囲

[請求項1] 認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を備え、

前記認識結果取得手段は、前記複数の認識エンジンにて共通するデータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、

前記統合認識結果出力手段は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、情報処理装置。

[請求項2] 請求項 1 に記載の情報処理装置であって、

前記統合認識結果出力手段は、前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、情報処理装置。

[請求項3] 請求項 1 又は 2 に記載の情報処理装置であって、

前記認識結果取得手段にて前記複数の認識エンジンからそれぞれ取得する前記認識結果情報のデータ形式を定義する認識結果定義情報と、前記統合認識結果出力手段にて前記各認識結果情報を統合する際の統合形式を定義する統合定義情報と、を記憶した定義情報記憶手段を備え、

前記認識結果取得手段は、前記認識結果定義情報にて定義されたデータ形式にて前記各認識結果情報を取得し、

前記統合認識結果出力手段は、前記統合定義情報にて定義された統合形式にて前記各認識結果情報を統合して前記新たな認識結果を出力

する、
情報処理装置。

[請求項4] 請求項 1 乃至 3 のいずれかに記載の情報処理装置であって、
前記認識結果取得手段にて前記複数の認識エンジンからそれぞれ取得する前記認識結果情報は、グラフ構造又は木構造のデータ形式である、
情報処理装置。

[請求項5] 請求項 2 に記載の情報処理装置であって、
前記認識結果取得手段は、前記認識対象データから物体を認識する物体認識エンジンと、前記認識対象データから所定の物体の移動軌跡を表す動線を認識する動線認識エンジンと、から前記認識結果情報をそれぞれ取得し、
前記統合認識結果出力手段は、前記物体認識エンジンから取得した前記認識結果情報である物体情報に、前記動線認識エンジンから取得した前記認識結果情報である所定の物体の動線情報を関連付けて統合し、前記新たな認識結果として出力する、
情報処理装置。

[請求項6] 請求項 5 に記載の情報処理装置であって、
前記認識結果取得手段は、前記物体認識エンジンが前記認識対象データから認識した物体の位置とその日時を表す所在情報を含む前記認識結果情報を前記物体認識エンジンから取得すると共に、前記動線認識エンジンが前記認識対象データから認識した所定の物体の動線情報を前記動線認識エンジンから取得し、
前記統合認識結果取得手段は、前記物体認識エンジンから取得した前記所在情報に予め定められた基準により対応する情報が前記動線認識エンジンから取得した前記動線情報に含まれている場合に、前記所在情報に対応する情報が含まれた前記物体認識エンジンからの前記認識結果情報にて特定される物体情報に、前記動線認識エンジンからの

前記認識結果である前記動線情報を関連付けて統合し、前記新たな認識結果として出力する、

情報処理装置。

[請求項7]

請求項2に記載の情報処理装置であって、

前記認識結果取得手段は、前記認識対象データからそれぞれ異なる認識処理にて物体を認識する複数の物体認識エンジンから前記認識結果情報をそれぞれ取得し、

前記統合認識結果出力手段は、前記複数の物体認識エンジンからそれぞれ取得した前記各認識結果情報である各物体情報に基づいて一の物体情報を特定して統合し、前記新たな認識結果を出力する、
情報処理装置。

[請求項8]

請求項7に記載の情報処理装置であって、

前記認識結果取得手段は、前記複数の物体認識エンジンが前記認識対象データから認識した物体の候補を表す物体候補情報を含む前記認識結果情報を前記複数の物体認識エンジンからそれぞれ取得し、

前記統合認識結果取得手段は、前記複数の物体認識エンジンからそれぞれ取得した前記各物体候補情報で予め定められた基準により相互に対応する物体情報を特定して、前記新たな認識結果として出力する、
情報処理装置。

[請求項9]

情報処理装置に、

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力された各認識結果情報を取得する認識結果取得手段と、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報を統合した新たな認識結果を出力する統合認識結果出力手段と、を實現させると共に、

前記認識結果取得手段は、前記複数の認識エンジンにて共通するデ

ータ形式の前記認識結果情報を前記複数の認識エンジンからそれぞれ取得し、

前記統合認識結果出力手段は、前記各認識結果情報に基づいて当該各認識結果情報を統合して前記新たな認識結果として出力する、
ことを実現させるためのプログラム。

[請求項10]

請求項9に記載のプログラムであって、

前記統合認識結果出力手段は、前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、
ことを実現させるためのプログラム。

[請求項11]

認識対象データに対してそれぞれ異なる認識処理を行う複数の認識エンジンから出力され、前記複数の認識エンジンにて共通するデータ形式の各認識結果情報を取得し、

前記複数の認識エンジンからそれぞれ取得した前記各認識結果情報に基づいて当該各認識結果情報を統合して新たな認識結果を出力する、
情報処理方法。

[請求項12]

請求項11に記載の情報処理方法であって、

前記各認識結果情報に予め定められた基準により相互に対応する情報が含まれている場合に、当該各認識結果情報を統合して前記新たな認識結果として出力する、
情報処理方法。