

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4613284号
(P4613284)

(45) 発行日 平成23年1月12日(2011.1.12)

(24) 登録日 平成22年10月29日(2010.10.29)

(51) Int.Cl. F I
G06T 7/20 (2006.01) G O 6 T 7/20 3 0 0 Z
G06T 1/00 (2006.01) G O 6 T 1/00 3 4 0 B

請求項の数 4 (全 17 頁)

<p>(21) 出願番号 特願2005-87651 (P2005-87651) (22) 出願日 平成17年3月25日(2005.3.25) (65) 公開番号 特開2006-268607 (P2006-268607A) (43) 公開日 平成18年10月5日(2006.10.5) 審査請求日 平成20年2月1日(2008.2.1)</p> <p>(出願人による申告)平成16年度独立行政法人情報通信研究機構、研究テーマ「超高速知能ネットワーク社会に向けた新しいインタラクション・メディアの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p> <p>特許権者において、実施許諾の用意がある。</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2 (74) 代理人 100090181 弁理士 山田 義人 (72) 発明者 篠沢 一彦 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内 (72) 発明者 小暮 潔 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>審査官 松尾 俊介</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

最終頁に続く

(54) 【発明の名称】 コミュニケーションロボットおよびそれを用いた動作識別システム

(57) 【特許請求の範囲】

【請求項1】

人間を撮影する撮影手段、
 前記人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、
 前記識別情報検出手段によって検出された識別情報に基づいて前記人間の動作を予測する動作予測手段、および
 前記撮影手段によって撮影された撮影画像が示す前記人間の実動作と前記動作予測手段によって予測された予測動作とから当該実動作を特定する動作特定手段を備える、コミュニケーションロボット。

【請求項2】

前記撮影画像が示す前記人間の実動作と前記予測動作の各々の類似度を算出する類似度算出手段をさらに備え、
 前記動作特定手段は、前記類似度算出手段によって算出された類似度が最も高い前記予測動作を前記実動作として特定する、請求項1記載のコミュニケーションロボット。

【請求項3】

コミュニケーションロボットと、このコミュニケーションロボットと通信可能に設けられるサーバとを備える動作識別システムであって、
 前記コミュニケーションロボットは、
 人間を撮影する撮影手段、
 前記人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、

検出した識別情報を前記サーバに送信する識別情報送信手段、
前記サーバから前記人間の予測動作情報を受信する予測動作情報受信手段、および
前記撮影手段によって撮影された撮影画像が示す前記人間の実動作と前記予測動作情
報受信手段によって受信された予測動作情報とから当該実動作を特定する動作特定手段を
備え、

前記サーバは、

前記識別情報送信手段によって送信された識別情報を受信する識別情報受信手段、お
よび

前記識別情報受信手段によって受信された識別情報に基づいて取得した前記人間の予
測動作情報を前記コミュニケーションロボットに送信する予測動作情報送信手段を備える
、動作識別システム。

10

【請求項 4】

コミュニケーションロボットと、このコミュニケーションロボットと通信可能に設けら
れるサーバとを備える動作識別システムであって、

前記コミュニケーションロボットは、

人間を撮影する撮影手段、

前記人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、お
よび

前記撮影手段によって撮影された撮影画像と前記識別情報検出手段によって検出され
た識別情報とを前記サーバに送信する送信手段を備え、

20

前記サーバは、

前記送信手段によって送信された撮影画像と識別情報とを受信する受信手段、

前記受信手段によって受信された識別情報に基づいて前記人間の動作を予測する動作
予測手段、および

前記識別情報受信手段によって受信された撮影画像が示す前記人間の実動作と前記動
作予測手段によって予測された予測動作とから当該実動作を特定する動作特定手段を備え
る、動作識別システム。

【発明の詳細な説明】

【技術分野】

【0001】

30

この発明はコミュニケーションロボットおよびそれをを用いた動作識別システムに関し、
特にたとえば、人間の動作を識別する、コミュニケーションロボットおよびそれをを用いた
動作識別システムに関する。

【背景技術】

【0002】

本件出願人は、特許文献 1 に代表されるように、人間と相互作用するコミュニケーショ
ンロボットを提案してきた。

【特許文献 1】特開 2002 - 355783 号公報

【発明の開示】

【発明が解決しようとする課題】

40

【0003】

背景技術のコミュニケーションロボットを用いて、その周囲に存在する人間の動作を識
別する場合には、たとえば、人間のあらゆる動作（行動）についての映像を予め記録して
おき、撮影した映像と予め記録しておいた映像とを比較し、いずれか 1 つの動作を特定す
る。これでは、計算量が膨大である。また、映像のみで動作を判断した場合には、異なる
動作であっても映像が似ている動作では、誤った動作として特定してしまう可能性がある
。

【0004】

それゆえに、この発明の主たる目的は、新規な、コミュニケーションロボットおよびそ
れを用いた動作識別システムを提供することである。

50

【 0 0 0 5 】

この発明の他の目的は、計算処理を低減して人間の行動を正確に特定できる、コミュニケーションロボットおよびそれを用いた動作識別システムを提供することである。

【課題を解決するための手段】

【 0 0 0 6 】

請求項1の発明は、人間を撮影する撮影手段、人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、識別情報検出手段によって検出された識別情報に基づいて人間の動作を予測する動作予測手段、および撮影手段によって撮影された撮影画像が示す人間の実動作と動作予測手段によって予測された予測動作とから当該実動作を特定する動作特定手段を備える、コミュニケーションロボットである。

10

【 0 0 0 7 】

請求項1の発明では、コミュニケーションロボットは、撮影手段、識別情報検出手段、動作予測手段、および動作特定手段を備える。撮影手段は、たとえば、人間を撮影する。識別情報検出手段は、ロボット自身が存在する環境において、その周辺に存在する物体（物品）、厳密には、人間の近傍に存在する物品についての識別情報を検出する。実施例では、物品に装着した無線タグから発信されるタグIDを検出する。動作予測手段は、識別情報検出手段によって検出された識別情報に基づいて人間の動作を予測する。たとえば、人間が鉛筆を所持したり、人間の近傍に鉛筆が存在する場合には、人間が文字等を書いたり、書いた文字等を読んだりするなどしていると、その動作を予測することができる。そして、動作特定手段は、撮影手段によって撮影された撮影画像が示す人間の実動作と動作予測手段によって予測された予測動作とから当該実動作を特定する。

20

【 0 0 0 8 】

請求項1の発明によれば、人間の近傍に存在する物体に基づいて動作を予測し、予測した動作から実動作を特定するので、あらゆる動作から1の動作を特定する場合よりも、大幅に特定処理を軽減することができる。また、物体と実動作との組み合わせにより、1の動作を特定するので、動作を正確に特定することができる。

【 0 0 0 9 】

請求項2の発明は請求項1に従属し、撮影画像が示す人間の実動作と予測動作の各々の類似度を算出する類似度算出手段をさらに備え、動作特定手段は、類似度算出手段によって算出された類似度が最も高い予測動作を実動作として特定する。

30

【 0 0 1 0 】

請求項2の発明では、コミュニケーションロボットは、類似度算出手段をさらに備える。この類似度算出手段は、撮影画像が示す人間の実動作と予測動作の各々の類似度を算出する。動作特定手段は、類似度が最も高い予測動作を実動作として特定する。

【 0 0 1 1 】

請求項2の発明では、識別情報に基づいて絞り込んだ予測動作との類似度を算出するだけなので、すべての動作との類似度を算出する場合に比べて処理負担を軽減することができる。また、絞り込んだ予測動作から実動作を特定するので、すべての動作について類似度を算出する場合よりも、識別精度を高くすることができる。

【 0 0 1 2 】

40

請求項3の発明は、コミュニケーションロボットと、このコミュニケーションロボットと通信可能に設けられるサーバとを備える動作識別システムであって、コミュニケーションロボットは、人間を撮影する撮影手段、人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、検出した識別情報をサーバに送信する識別情報送信手段、サーバから人間の予測動作情報を受信する予測動作情報受信手段、および撮影手段によって撮影された撮影画像が示す人間の実動作と予測動作情報受信手段によって受信された予測動作情報とから当該実動作を特定する動作特定手段を備え、サーバは、識別情報送信手段によって送信された識別情報を受信する識別情報受信手段、および識別情報受信手段によって受信された識別情報に基づいて取得した人間の予測動作情報をコミュニケーションロボットに送信する予測動作情報送信手段を備える、動作識別システムである。

50

【 0 0 1 3 】

請求項 3 の発明は、コミュニケーションロボットと、このコミュニケーションロボットと通信可能に設けられるサーバとを備える動作識別システムである。この動作識別システムでは、コミュニケーションロボットは、人間の近傍に存在する物体の識別情報を検出すると、その識別情報をサーバに送信する。サーバは、識別情報に基づいて動作を予測し、予測動作情報をコミュニケーションロボットに送信する。つまり、請求項 2 の発明では、動作を予測する手段ないし役割としてのサーバが設けられる点が請求項 1 の発明と異なる。

【 0 0 1 4 】

請求項 3 の発明によれば、請求項 1 の発明と同様に、動作特定についての計算処理を大幅に低減できる。また、ロボットにおける計算処理の負担を軽減できる。

10

【 0 0 1 5 】

請求項 4 の発明は、コミュニケーションロボットと、このコミュニケーションロボットと通信可能に設けられるサーバとを備える動作識別システムであって、コミュニケーションロボットは、人間を撮影する撮影手段、人間の近傍に存在する物体についての識別情報を検出する識別情報検出手段、および撮影手段によって撮影された撮影画像と識別情報検出手段によって検出された識別情報とをサーバに送信する送信手段を備え、サーバは、送信手段によって送信された撮影画像と識別情報とを受信する受信手段、受信手段によって受信された識別情報に基づいて人間の動作を予測する動作予測手段、および識別情報受信手段によって受信された撮影画像が示す人間の実動作と動作予測手段によって予測された

20

【 0 0 1 6 】

請求項 4 の発明では、コミュニケーションロボットは、識別情報と人間の撮影画像とをサーバに送信し、サーバが識別情報に基づいて行動を予測し、予測行動の各々と撮影画像とから、撮影画像が示す人間の動作を特定するようにしてある以外は、請求項 3 の発明と同じである。

【 0 0 1 7 】

請求項 4 の発明においても、請求項 1 の発明と同様に、動作特定についての計算処理を大幅に低減できる。また、ロボットにおける計算処理の負担を軽減できる。

【 発明の効果 】

30

【 0 0 1 8 】

この発明によれば、人間の近傍に存在する物体から人間の動作を推定し、推定した動作のいずれか 1 つに特定するので、計算量を大幅に低減することができる。また、映像のみならず、物体との組み合わせにより、動作を特定するので、正確に動作を特定することができる。

【 0 0 1 9 】

この発明の上述の目的、その他の目的、特徴および利点は、図面を参照して行う以下の実施例の詳細な説明から一層明らかとなろう。

【 発明を実施するための最良の形態 】

【 0 0 2 0 】

40

< 第 1 実施例 >

図 1 を参照して、この第 1 実施例の動作識別システム（以下、単に「システム」という。）10 は、コミュニケーションロボット（以下、単に「ロボット」という。）12 および複数のタグ 14 を含む。ロボット 12 は、人間との間で、身振り手振りのような身体動作ないし行動（以下、「コミュニケーション行動」ということがある。）を取ることができる。ただし、コミュニケーション行動としては、ロボット 12 と人間との間における会話が含まれる場合もある。タグ 14 は、たとえばパッシブタイプの無線タグ（RF タグ、IC タグなど）であり、少なくともロボット 12 が存在する環境に存在する物体（図 1 では省略する。）に装着される（貼り付けられる）。

【 0 0 2 1 】

50

ロボット12は、人間のような身体を有し、その身体を用いてコミュニケーションのために必要な複雑な身体動作を生成する。具体的には、図2を参照して、ロボット12は台車32を含み、この台車32の下面には、このロボット12を自律移動させる車輪34が設けられる。この車輪34は、車輪モータ(ロボット12の内部構成を示す図3において参照番号「36」で示す。)によって駆動され、台車32すなわちロボット12を前後左右任意の方向に動かすことができる。

【0022】

なお、図2では示さないが、この台車32の前面には、衝突センサ(図3において参照番号「38」で示す。)が取り付けられ、この衝突センサ38は、台車32への人や他の障害物の接触を検知する。そして、ロボット12の移動中に障害物との接触を検知すると、直ちに車輪34の駆動を停止してロボット12の移動を急停止させる。

【0023】

また、ロボット12の背の高さは、この第1実施例では、人、特に子供に威圧感を与えることがないように、100cm程度とされている。ただし、この背の高さは任意に変更可能である。

【0024】

台車32の上には、多角形柱のセンサ取付パネル40が設けられ、このセンサ取付パネル40の各面には、超音波距離センサ42が取り付けられる。この超音波距離センサ42は、取付パネル40すなわちロボット12の周囲の主として人との間の距離を計測するものである。

【0025】

台車32の上には、さらに、ロボット12の胴体が、その下部が上述の取付パネル40に囲まれて、直立するように取り付けられる。この胴体は下部胴体44と上部胴体46とから構成され、これら下部胴体44および上部胴体46は、連結部48によって連結される。連結部48には、図示しないが、昇降機構が内蔵されていて、この昇降機構を用いることによって、上部胴体46の高さすなわちロボット12の高さを変化させることができる。昇降機構は、後述のように、腰モータ(図3において参照番号「50」で示す。)によって駆動される。上で述べたロボット12の身長100cmは、上部胴体46をその最下位置にしたときの値である。したがって、ロボット12の身長は100cm以上にする事ができる。

【0026】

上部胴体46のほぼ中央には、1つの全方位カメラ52と、1つのマイク16とが設けられる。全方位カメラ52は、ロボット12の周囲を撮影するもので、後述の眼カメラ54と区別される。マイク16は、周囲の音、とりわけ人の声を取り込む。

【0027】

上部胴体46の両肩には、それぞれ、肩関節56Rおよび56Lによって、上腕58Rおよび58Lが取り付けられる。肩関節56Rおよび56Lは、それぞれ3軸の自由度を有する。すなわち、右肩関節56Rは、X軸、Y軸およびZ軸の各軸廻りにおいて上腕58Rの角度を制御できる。Y軸は、上腕58Rの長手方向(または軸)に平行な軸であり、X軸およびZ軸は、そのY軸に、それぞれ異なる方向から直交する軸である。左肩関節56Lは、A軸、B軸およびC軸の各軸廻りにおいて上腕58Lの角度を制御できる。B軸は、上腕58Lの長手方向(または軸)に平行な軸であり、A軸およびC軸は、そのB軸に、それぞれ異なる方向から直交する軸である。

【0028】

上腕58Rおよび58Lのそれぞれの先端には、肘関節60Rおよび60Lを介して、前腕62Rおよび62Lが取り付けられる。肘関節60Rおよび60Lは、それぞれ、W軸およびD軸の軸廻りにおいて、前腕62Rおよび62Lの角度を制御できる。

【0029】

なお、上腕58Rおよび58Lならびに前腕62Rおよび62L(いずれも図2)の変位を制御するX、Y、Z、W軸およびA、B、C、D軸では、「0度」がホームポジショ

10

20

30

40

50

ンであり、このホームポジションでは、上腕 5 8 R および 5 8 L ならびに前腕 6 2 R および 6 2 L は下方向に向けられる。

【 0 0 3 0 】

また、図 2 では示さないが、上部胴体 4 6 の肩関節 5 6 R および 5 6 L を含む肩の部分や上述の上腕 5 8 R および 5 8 L ならびに前腕 6 2 R および 6 2 L を含む腕の部分には、それぞれ、タッチセンサ (図 3 において参照番号 6 4 で包括的に示す。) が設けられていて、これらのタッチセンサ 6 4 は、人がロボット 1 2 のこれらの部位に接触したかどうかを検知する。

【 0 0 3 1 】

前腕 6 2 R および 6 2 L のそれぞれの先端には、手に相当する球体 6 6 R および 6 6 L がそれぞれ固定的に取り付けられる。なお、この球体 6 6 R および 6 6 L に代えて、この第 1 実施例のロボット 1 2 と異なり指の機能が必要な場合には、人の手の形をした「手」を用いることも可能である。また、球体 6 6 R には、右手用のタグリーダ 1 0 2 が設けられ、球体 6 6 L には、左手用のタグリーダ 1 0 4 が設けられる。ただし、いずれか一方の球体ないし手のみタグリーダを設けるようにしてもよい。

【 0 0 3 2 】

上部胴体 4 6 の中央上方には、首関節 6 8 を介して、頭部 7 0 が取り付けられる。この首関節 6 8 は、3 軸の自由度を有し、S 軸、T 軸および U 軸の各軸廻りに角度制御可能である。S 軸は首から真上に向かう軸であり、T 軸および U 軸は、それぞれ、この S 軸に対して異なる方向で直交する軸である。頭部 7 0 には、人の口に相当する位置に、スピーカ 7 2 が設けられる。スピーカ 7 2 は、ロボット 1 2 が、その周囲の人に対して音声または声によってコミュニケーションを図るために用いられる。ただし、スピーカ 7 2 は、ロボット 1 2 の他の部位たとえば胴体に設けられてもよい。

【 0 0 3 3 】

また、頭部 7 0 には、目に相当する位置に眼球部 7 4 R および 7 4 L が設けられる。眼球部 7 4 R および 7 4 L は、それぞれ眼カメラ 5 4 R および 5 4 L を含む。なお、右の眼球部 7 4 R および左の眼球部 7 4 L をまとめて眼球部 7 4 といい、右の眼カメラ 5 4 R および左の眼カメラ 5 4 L をまとめて眼カメラ 5 4 ということもある。眼カメラ 5 4 は、ロボット 1 2 に接近した人の顔や他の部分ないし物体等を撮影してその映像信号を取り込む。

【 0 0 3 4 】

なお、上述の全方位カメラ 5 2 および眼カメラ 5 4 のいずれも、たとえば C C D や C M O S のような固体撮像素子を用いるカメラであってよい。

【 0 0 3 5 】

たとえば、眼カメラ 5 4 は眼球部 7 4 内に固定され、眼球部 7 4 は眼球支持部 (図示せず) を介して頭部 7 0 内の所定位置に取り付けられる。眼球支持部は、2 軸の自由度を有し、 α 軸および β 軸の各軸廻りに角度制御可能である。 α 軸および β 軸は頭部 7 0 に対して設定される軸であり、 α 軸は頭部 7 0 の上へ向かう方向の軸であり、 β 軸は α 軸に直交しかつ頭部 7 0 の正面側 (顔) が向く方向に直交する方向の軸である。この第 1 実施例では、頭部 7 0 がホームポジションにあるとき、 α 軸は S 軸に平行し、 β 軸は U 軸に平行するように設定されている。このような頭部 7 0 において、眼球支持部が α 軸および β 軸の各軸廻りに回転されることによって、眼球部 7 4 ないし眼カメラ 5 4 の先端 (正面) 側が変位され、カメラ軸すなわち視線方向が移動される。

【 0 0 3 6 】

なお、眼カメラ 5 4 の変位を制御する α 軸および β 軸では、「0 度」がホームポジションであり、このホームポジションでは、図 2 に示すように、眼カメラ 5 4 のカメラ軸は頭部 7 0 の正面側 (顔) が向く方向に向けられ、視線は正視状態となる。

【 0 0 3 7 】

図 3 には、ロボット 1 2 の内部構成を示すブロック図が示される。この図 3 に示すように、ロボット 1 2 は、全体の制御のためにマイクロコンピュータまたは C P U 7 6 を含み

10

20

30

40

50

、このCPU76には、バス78を通して、メモリ80，モータ制御ボード82，センサ入力/出力ボード84および音声入力/出力ボード86が接続される。

【0038】

メモリ80は、図示しないが、ROMやHDD、RAM等を含み、ROMまたはHDDにはこのロボット12の制御プログラムおよびデータ等が予め格納されている。CPU76は、このプログラムに従って処理を実行する。具体的には、ロボット12の身体動作を制御するための複数のプログラム（行動モジュールと呼ばれる。）が記憶される。たとえば、行動モジュールが示す身体動作としては、「握手」、「抱っこ」、「万歳」...などがある。行動モジュールが示す身体動作が「握手」である場合には、当該行動モジュールを実行すると、ロボット12は、たとえば、右手を前に差し出す。また、行動モジュールが示す身体動作が「抱っこ」である場合には、当該行動モジュールを実行すると、ロボット12は、たとえば、両手を前に差し出す。さらに、行動モジュールが示す身体動作が「万歳」である場合には、当該行動モジュールを実行すると、ロボット12は、たとえば、両手を数回（たとえば、2回）上下させる。また、RAMは、一時記憶メモリとして用いられるとともに、ワーキングメモリとして利用され得る。

10

【0039】

モータ制御ボード82は、たとえばDSP(Digital Signal Processor)で構成され、右腕、左腕、頭および眼等の身体部位を駆動するためのモータを制御する。すなわち、モータ制御ボード82は、CPU76からの制御データを受け、右肩関節56RのX，YおよびZ軸のそれぞれの角度を制御する3つのモータと右肘関節60Rの軸Wの角度を制御する1つのモータを含む計4つのモータ（図3ではまとめて、「右腕モータ」として示す。）88の回転角度を調節する。また、モータ制御ボード82は、左肩関節56LのA，BおよびC軸のそれぞれの角度を制御する3つのモータと左肘関節60LのD軸の角度を制御する1つのモータとを含む計4つのモータ（図3ではまとめて、「左腕モータ」として示す。）90の回転角度を調節する。モータ制御ボード82は、また、首関節68のS，TおよびU軸のそれぞれの角度を制御する3つのモータ（図3ではまとめて、「頭部モータ」として示す。）92の回転角度を調節する。モータ制御ボード82は、また、腰モータ50、および車輪34を駆動する2つのモータ（図3ではまとめて、「車輪モータ」として示す。）36を制御する。さらに、モータ制御ボード82は、右眼球部74Rの軸および軸のそれぞれの角度を制御する2つのモータ（図3ではまとめて、「右眼球モータ」として示す。）94の回転角度を調節し、また、左眼球部74Lの軸および軸のそれぞれの角度を制御する2つのモータ（図3ではまとめて、「左眼球モータ」として示す。）96の回転角度を調節する。

20

30

【0040】

なお、この第1実施例の上述のモータは、車輪モータ36を除いて、制御を簡単化するためにそれぞれステッピングモータまたはパルスモータであるが、車輪モータ36と同様に、直流モータであってよい。

【0041】

センサ入力/出力ボード84も、同様に、DSPで構成され、各センサやカメラからの信号を取り込んでCPU76に与える。すなわち、超音波距離センサ42の各々からの反射時間に関するデータがこのセンサ入力/出力ボード84を通して、CPU76に入力される。また、全方位カメラ52からの映像信号が、必要に応じてこのセンサ入力/出力ボード84で所定の処理が施された後、CPU76に入力される。眼カメラ54からの映像信号も、同様にして、CPU76に与えられる。また、タッチセンサ64からの信号がセンサ入力/出力ボード84を介してCPU76に与えられる。

40

【0042】

スピーカ72には音声入力/出力ボード86を介して、CPU76から、合成音声データが与えられ、それに従って、スピーカ72からはそのデータに従った音声または声が発音される。また、マイク24からの音声入力、音声入力/出力ボード86を介してCPU76に取り込まれる。

50

【 0 0 4 3 】

また、CPU 76には、バス78を通して、通信LANボード98が接続される。この通信LANボード98も、同様に、DSPで構成され、CPU 76から与えられた送信データを無線通信装置100に与え、無線通信装置100から送信データを送信させる。また、通信LANボード98は無線通信装置100を介してデータを受信し、受信データをCPU 76に与える。

【 0 0 4 4 】

さらに、CPU 76には、バス78を介して、タグリーダ102、タグリーダ104およびデータベース106が接続される。ただし、データベース106は、ロボット12内部に設ける必要はなく、ロボット12と通信可能に、その外部に設けるようにすることもできる。タグリーダ102およびタグリーダ104は、上述したように、ロボット12の手(球体64Rおよび64L)に設けられ、ロボット12の周辺に存在する物体等に装着されたタグ14が発信するタグ情報(タグID)を受信して、CPU 76に与える。

10

【 0 0 4 5 】

また、データベース106には、図4および図5に示すように、物品データ106a、動作候補データ106bおよび動作テンプレートデータ106c等のデータが記憶される。図4(A)に示すように、物品データ106aは、タグ14の識別情報(タグID)に対応して、該当するタグ14が装着された物品(物体)の名称が記述される。つまり、物品データ106aは、タグIDから物品を特定するためのテーブルデータである。また、図4(B)に示すように、動作候補データ106bは、物品の名称に対応して、動作の候補(動作候補)が記述される。たとえば、物品が「鉛筆」である場合には、動作候補として、「持つ」、「書く」、「見る」、「置く」などの動作が記述される。また、物品が「包丁」である場合には、「持つ」、「置く」、「切る」、「研ぐ」、「洗う」などの動作が記述される。たとえば、人間の近くに鉛筆がある場合には、人間は当該鉛筆を持ったり(握ったり)、当該鉛筆で文字等を書いたり、当該鉛筆で書いた内容を見たり、当該鉛筆を置いたりするなどの動作をすると考えられる。また、人間の近くに包丁がある場合には、人間は当該包丁を持ったり(握ったり)、当該包丁を置いたり、当該包丁で食物(肉、魚、野菜、果物など)を切ったり、当該包丁を研いだり、当該包丁を洗ったりするなどの動作をすると考えられる。つまり、動作候補データ106bは、物品(物体)から予測(推測)され得る人間の動作(行動)の候補を決定するためのテーブルデータである。

20

30

【 0 0 4 6 】

また、図5に示すように、動作テンプレートデータ106cは、動作に対応して、当該動作を行う人間を撮影したときの映像(画像)ファイルのファイル名が記述される。図示は省略するが、データベース106には、実際の画像ファイルも記憶される。ただし、各動作に対応して、画像ファイルそのものを記述するようにしてもよい。

【 0 0 4 7 】

たとえば、図6に示すように、或る部屋200に、システム10は適用される。ただし、図6では、簡単のため、ロボット12および一部のタグ14を示してある。また、図6に示すように、部屋200には、人間202が存在し、その近傍にロボット12が存在する。また、部屋200には、机204が配置され、机204の上には、ノート206が載置される。たとえば、人間202は、鉛筆208を持って、ノート206に文字等を書いている。さらに、部屋200の隅には、ごみ箱210が置いてある。

40

【 0 0 4 8 】

また、上述したように、部屋200に存在する物品(物体)には、それぞれタグ14が装着される。図6では、机204のみにタグ14が装着されている様子を示してあるが、実際には、ノート206、鉛筆208およびごみ箱210にもタグ14は装着される。さらに、部屋200自体(または、屋内や屋外)を識別する場合には、つまり場所を識別する場合には、当該部屋200の入り口や壁等にタグ14を装着するようにしてもよい。

【 0 0 4 9 】

たとえば、システム10は、人間202の動作(行動)を識別(特定)する。具体的に

50

は、図3に示したCPU76が図7および図8示す動作識別処理を実行する。図7に示すように、CPU76は動作識別処理を開始すると、ステップS1で、タグリーダ102およびタグリーダ104を動かし、タグIDを検出する。具体的には、眼カメラ54の撮影画像（映像）から人間202を検出し、人間202に接近し、ロボット12の近傍に存在する物品（厳密には、タグID）を検出する。ただし、この第1実施例では、人間202の動作を識別するようにしてあるため、厳密には、その近傍（たとえば、1～2m程度で、人間の手の届く範囲）に存在する物品を検出するようにしてある。

【0050】

また、人間202の近傍に存在する物品および人間202の撮影画像（映像）から人間202の動作を識別（特定）するようにしてあるため、ロボット12を人間202に近づけるようにしてある。詳細な説明は省略するが、ロボット12は、眼カメラ54の撮影画像（映像）から人間202をパターンマッチングの手法により検出（推定）し、その方向に進行する。或いは、撮影画像（映像）に含まれる肌色領域を検出すると、その肌色領域が大きくなる方向に進行（または後退）或いは回転する。このようにして、ロボット12は、人間202に近づくことができる。

10

【0051】

ただし、人間202に近づいたか否かの判断には、超音波距離センサ42の検出結果（距離）を用いてもよく、眼カメラ54の撮影画像と超音波距離センサ42の検出結果とを用いるようにしてもよい。

【0052】

ロボット12は、人間202に近づくと、その両手、すなわち球体66Rおよび球体66Lに設けられる、タグリーダ102およびタグリーダ104によって受信されるタグIDを検出する。この第1実施例では、両手（腕）すなわち肩関節56R、56Lおよび肘関節60R、60Lを動かすことにより、タグリーダ102およびタグリーダ104を自在に動かし、タグリーダ102またはタグリーダ104によってタグIDを検出する。

20

【0053】

続くステップS3では、タグIDを検出したかどうかを判断する。ステップS3で“NO”であれば、つまりタグIDを検出していなければ、ステップS1に戻って、タグリーダ102またはタグリーダ104或いはその両方を動かし、タグIDを検出する。一方、ステップS3で“YES”であれば、つまりタグIDを検出すれば、ステップS5で、当該タグIDを検出したタグリーダ102または104（両方の場合には、いずれか一方でよい。）と眼カメラ54との距離を算出し、メモリ80に記憶（一時記憶）する。ここで、眼カメラ54の位置および両手（上腕58R、58Lおよび前腕62R、62Lなど）の長さは固定であり、タグリーダ102およびタグリーダ104は球体66Rおよび66Lに装着されているため、肩関節56R、56Lおよび肘関節60R、60Lの角度を考慮して、上記距離は算出される。

30

【0054】

また、眼カメラ54とタグリーダ102、104との距離を計測するのは、上述したように、タグ14はパッシブタイプのものであり、眼カメラ54と物品（物体）とのおおよその距離を計測し、最終的に、人間202と物品との距離を推定するためである。この結果と、上述したように、肌色領域が大きくなるように前進等することにより、人間202が所持（または装着）する、または近傍に存在する物品のタグIDを検出することができるのである。したがって、図示は省略するが、タグIDを検出した場合であっても、ステップS1において移動したときに、肌色領域が検出されなかったり、肌色領域が検出されたが、比較的その領域が小さく、ロボット12と人間202との距離が遠いと判断されたりした場合には、検出したタグIDは検出結果から排除（リジェクト）するようにしてある。これは、人間202の動作を正確に識別するためである。

40

【0055】

続いて、ステップS7では、両手（肩関節56R、56L、肘関節60R、60L）を動かして、タグリーダ102およびタグリーダ104を、眼カメラ54に写らない位置に

50

移動させる。ただし、タグリーダ102、104のみならず、両手も写らない位置に移動させる方が好ましい。これは、後述するように、眼カメラ54で人間202を撮影し、人間202の動作を識別するようにしてあるためである。つまり、できる限り、眼カメラ54によって人間202を撮影させるためである。

【0056】

続くステップS9では、検出したタグID、検出時刻および検出場所を記憶する。図3では省略したが、検出時刻は、ロボット12内部に設けられる時計回路(タイマ)から取得される。また、検出場所は、場所毎に予め割り当てたタグを設置しておき、このタグIDを検出すれば、物品同様に、場所を特定することができる。続くステップS11では、撮影画像から肌色が検出されたかどうかを判断する。ステップS11で“NO”であれば、つまり撮影画像から肌色が検出されなければ、そのままステップS15に進む。一方、ステップS11で“YES”であれば、つまり撮影画像から肌色(領域)が検出されると、ステップS13で、当該肌色領域が撮影画像の中心にくるように、眼カメラ54を方向転換する。これは、人間202(の動作)を撮影して、正確に動作を識別するためである。ただし、眼カメラ54のみならず、ロボット12全体の方向も変化(旋回)させるようにしてもよい。このようにすれば、人間202およびその動作を撮影することができる。

【0057】

なお、この第1実施例では、撮影画像から肌色領域が検出された場合には、何ら処理を施さないようにしてあるが、人間202(肌色領域)を検出すべく、ロボット12を旋回させるようにしてもよい。

【0058】

続くステップS15では、現在の場所(現在位置)において、タグIDをすべて読み取ったかどうかを判断する。つまり、ロボット12を移動させたり、旋回させたり、両手を動かしたりして、人間202の周囲に存在する物品(タグID)をすべて検出したかどうかを判断する。ステップS15で“NO”であれば、つまり、現在の場所において、タグIDをすべて読み取っていなければ、そのままステップS1に戻る。しかし、ステップS15で“YES”であれば、つまり現在の場所において、タグIDをすべて読み取れば、図8に示すステップS17で、検出した1または2以上のタグIDに基づいて、動作候補パターンを決定する。

【0059】

具体的には、CPU76は、物品データ106aを参照して、検出したタグIDに対応する1または2以上の物品を特定する。続いて、CPU76は、動作候補データ106bを参照して、特定した物品に対応する動作候補パターンを決定する。ただし、2以上の物品が特定された場合には、各物品についての動作候補パターンであって、すべての物品において重複する動作候補パターンのみが決定される。たとえば、物品として、「包丁」と「まな板」とが検出された場合には、動作候補としては、「持つ」、「置く」、「切る」、「洗う」が決定される。つまり、2つの物品で重複していない「研ぐ」についての動作候補パターンが排除される。

【0060】

続いて、ステップS19では、各動作候補パターンと撮影画像との類似度を計算する。ここでは、CPU76は、各動作候補パターンと撮影画像との類似度距離を計算する。たとえば、人間202に取り付けられたマーカの軌跡と、各動作候補パターンにおいて対応するマーカの軌跡について、「動的計画法(Dynamic Programming)」によって距離を計算する。

【0061】

なお、この「動的計画法(Dynamic Programming)」は、既に周知であり、また、本願発明の本質的部分ではないため、詳細な説明は省略するが、その内容については、たとえば、「高橋勝彦、関進、岡隆一、ジェスチャ動画像のスポッティング認識、信学技報 PR U92-157, pp. 9-16, 3 1993.」に開示されている。

【0062】

続くステップS 2 1では、動作を特定する。具体的には、ステップS 1 9で計算した類似度距離が最も小さい（最も類似度が高い）動作候補パターンについての動作を、人間2 0 2の動作として特定する。次のステップS 2 3では、特定した動作、時刻および場所を記憶する。たとえば、メモリ8 0またはデータベース1 0 6に行動履歴のテーブルデータを記憶しておき、これに特定した動作、時刻および場所を記憶するようにすればよい。ただし、ステップS 2 3で記憶する時間および場所は、ステップS 9で記憶した検出時間および検出場所である。そして、ステップS 2 5で、動作識別結果を出力して、動作識別処理を終了する。たとえば、ステップS 2 5では、ロボット1 2のスピーカ7 2から識別した動作を音声で出力することができる。ただし、ロボット1 2と通信可能にコンピュータを接続しておき、識別結果を当該コンピュータに送信するようにすれば、当該コンピュータに接続されるモニタ（図示せず）等の画像表示装置に識別した動作をテキスト表示したり、当該コンピュータに接続されるスピーカ（図示せず）から識別した動作を音声で出力したりすることもできる。

10

【0063】

この第1実施例によれば、人間の近傍に存在する物品を検出し、物品から動作候補を決定して、動作候補の中から人間の動作を識別するので、すべての動作候補の中から人間の動作を識別する場合よりも大幅に処理を低減することができる。

【0064】

また、この第1実施例によれば、映像のみならず、物品に基づいて動作を識別するため、正確に識別することができる。つまり、動作候補を絞り込むので、識別制度を高くすることができる。

20

【0065】

なお、この第1実施例では、特定した動作とともに、その動作を検出（撮影）した時間および場所とともに記録しておくため、その記録内容は、或る人間についての行動メモとして用いることもできる。

<第2実施例>

図9に示す第2実施例のシステム1 0は、人間の動作候補をロボット1 2と通信可能に設けたサーバで検出するようにした以外は、上述の実施例と同様であるため、重複した説明は省略する。

【0066】

30

図9を参照して、第2実施例のシステム1 0では、ロボット1 2はネットワーク2 0を介してサーバ2 2と通信可能に接続される。サーバ2 2としては、汎用のサーバを用いることができ、サーバに代えて、汎用のパーソナルコンピュータやワークステーションを用いることもできる。また、ネットワーク2 0は、有線または無線のいずれで構築されてもよい。また、システム1 0では、サーバ2 2にデータベース2 4が接続される。このデータベース2 4に、第1実施例で示した物品データ1 0 6 aおよび動作候補データ1 0 6 bが記憶され、サーバ2 2は、ロボット1 2からの問い合わせに応じて動作候補を検出し、検出した動作候補をロボット1 2に通知（送信）するのである。

【0067】

具体的なロボット1 2（CPU 7 6）の動作識別処理は、図7および図8のフロー図で示した動作識別処理とほぼ同じであるため、異なる処理についてのみ説明することにする。また、図7に示した処理は同じであるため、図示は省略してある。

40

【0068】

図10に示すように、CPU 7 6は、ステップS 1 7'で、検出した1または2以上のタグIDをサーバ2 2に送信する。つまり、動作候補を問い合わせる。ここで、図示は省略するが、サーバ2 2は、ロボット1 2から1または2以上のタグIDを受信すると、タグIDに基づいて、動作候補パターンを決定し、決定した動作候補パターンをロボット1 2に送信する。具体的には、サーバ1 2は、物品データ1 0 6 aを参照して、ロボット1 2から受信したタグIDに対応する1または2以上の物品を特定する。続いて、サーバ2 2は、動作候補データ1 0 6 bを参照して、特定した物品に対応する動作候補パターンを

50

決定する。ただし、2以上の物品が特定された場合には、各物品についての動作候補パターンであって、すべての物品において重複する動作候補パターンのみが決定される。たとえば、物品として、「包丁」と「まな板」とが検出された場合には、動作候補としては、「持つ」、「置く」、「切る」、「洗う」が決定される。つまり、2つの物品で重複していない「研ぐ」についての動作候補パターンが排除される。

【0069】

図10に戻って、ステップS18では、動作候補パターンを受信したかどうかを判断する。ステップS18で“NO”であれば、つまり動作候補パターンを受信していなければ、同じステップS18に戻って、動作候補パターンの受信を待機する。一方、ステップS18で“YES”であれば、つまり動作候補パターンを受信すれば、ステップS19で、各動作候補パターンと撮影画像との類似度を計算する。これ以降の処理は、図8を用いて説明した場合と同じである。

10

【0070】

この第2実施例によれば、人間の動作を少ない処理で確実に識別することができ、さらに、サーバ側で動作候補を検出するので、ロボットの処理負担を低減することができる。

<第3実施例>

第3実施例のシステム10は、サーバ22側で人間202の動作を識別するようにした以外は、第2実施例のシステム10と同じであるため、重複した説明は省略する。図示は省略するが、この第3実施例のシステム10では、サーバ22に接続されるデータベース24に、物品データ106a、動作候補データ106bおよび動作テンプレートデータ106cを記憶する。したがって、ロボット12内部に設けるデータベース106は削除することができる。

20

【0071】

具体的には、第3実施例のシステム10では、ロボット12は、人間202が所持するまたは人間202の近傍に存在する物品のタグIDを検出し、また、そのときの人間202の撮影画像を取得する。そして、検出したタグIDと撮影画像とをサーバ22に送信する。サーバ22では、上述の第2実施例で説明したように、タグIDに基づいて動作候補パターンを検出し、検出した動作候補パターンのそれぞれと撮影画像との類似度を計算し、当該撮影画像が示す人間202の動作を特定する。

【0072】

したがって、図7および図10で示される動作識別処理において、ステップS1～ステップS17'までの処理をロボット12側で実行し、ステップS18～ステップS25までの処理をサーバ22側で実行するようにすればよい。ただし、ステップS17'では、検出したタグIDおよび撮影画像をサーバ22に送信する。

30

【0073】

なお、詳細な説明は省略するが、第1実施例で説明したように、ステップS23で記憶する時間および場所は、ステップS9で記憶した検出時間および検出場所であるため、この第3実施例では、ステップS17'では、ステップS9で記憶した検出時間および検出場所もサーバ22に送信される。

【0074】

この第3実施例においても、人間の動作を少ない処理で確実に識別することができ、さらに、サーバ側で動作識別処理を実行するので、ロボットの処理負担を低減することができる。

40

【0075】

なお、上述の実施例では、いずれも、パッシブタイプのタグを用いたが、他の実施例として、アクティブタイプのタグ(赤外線タグなど)を用いることができる。かかる場合には、たとえば、各物品に赤外線タグを設けておき、一方、2台のカメラ(赤外線カメラ(センサ))をロボット12の頭部(眼カメラ54の上部)に設ける。そして、2台のカメラの検出結果(撮影画像)から三角測量を行うことにより、ロボット12と物品との距離またはロボット12から見た物品の位置(3次元位置)を計測(算出)する。同様に、眼

50

カメラ 5 4 の検出結果（撮影画像）から三角測量を行うことにより、ロボット 1 2 と人間 2 0 2 との距離またはロボット 1 2 から見た人間 2 0 2 の 3 次元位置を算出する。これにより、人間 2 0 2 と物品との距離（位置）を測定できる。つまり、人間 2 0 2 の近傍に存在する物品を検出することができる。なお、物品を検出した後の処理は、上述の実施例と同様である。

【図面の簡単な説明】

【 0 0 7 6 】

【図 1】図 1 はこの発明のコミュニケーションロボットを用いた動作識別システムの一例を示す図解図である。

【図 2】図 2 は図 1 実施例に示すロボットの外観を説明するための図解図である。

10

【図 3】図 3 は図 1 および図 2 に示すロボットの電気的な構成を示す図解図である。

【図 4】図 4 はロボットに内蔵されるデータベースに記憶される物品データおよび動作候補データの例を示す図解図である。

【図 5】図 5 はロボットに内蔵されるデータベースに記憶される動作テンプレートデータの例を示す図解図である。

【図 6】図 6 は図 1 に示すシステムの適用例を示す図解図である。

【図 7】図 7 は図 3 に示す CPU の動作識別処理の一部を示すフロー図である。

【図 8】図 8 は図 3 に示す CPU の動作識別処理の他の一部であり、図 7 のフロー図に後続するフロー図である。

【図 9】図 9 はこの発明のコミュニケーションロボットを用いた動作識別システムの他の例を示す図解図である。

20

【図 1 0】図 1 0 は他の実施例の動作識別処理の一部を示すフロー図である。

【符号の説明】

【 0 0 7 7 】

1 0 ... コミュニケーションロボットを用いた動作識別システム

1 2 ... コミュニケーションロボット

1 4 ... タグ

1 6 ... マイク

2 0 ... ネットワーク

2 2 ... サーバ

30

2 4 , 1 0 6 ... データベース

3 8 ... 衝突センサ

4 2 ... 超音波距離センサ

5 2 ... 全方位カメラ

5 4 ... 眼カメラ

6 4 ... タッチセンサ

7 6 ... CPU

8 0 ... メモリ

8 2 ... モータ制御ボード

8 4 ... センサ入力 / 出力ボード

40

8 6 ... 音声入力 / 出力ボード

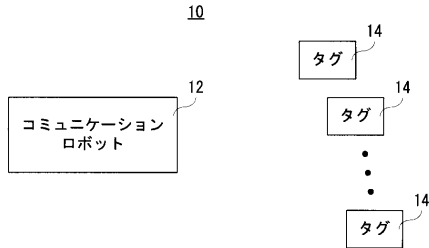
8 8 - 9 6 ... モータ

9 8 ... 通信 LAN ボード

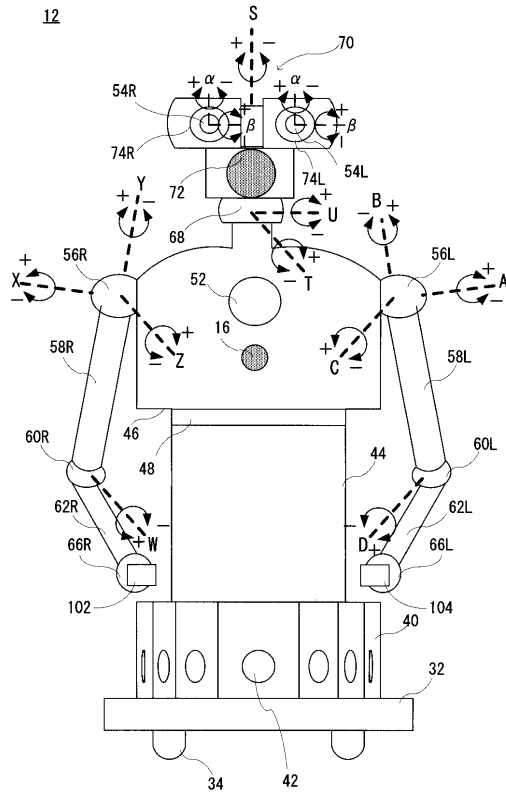
1 0 0 ... 無線通信装置

1 0 2 , 1 0 4 ... タグリーダー

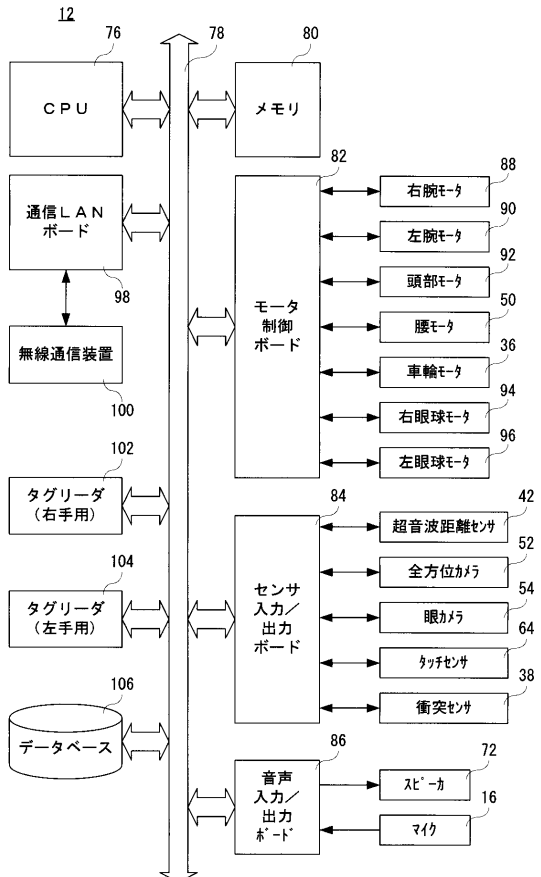
【図1】



【図2】



【図3】



【図4】

(A) 物品データ106a

タグID	物品
A	鉛筆
B	ノート
C	消しゴム
D	机
⋮	⋮
Y	包丁
Z	まな板
⋮	⋮

(B) 動作候補データ106b

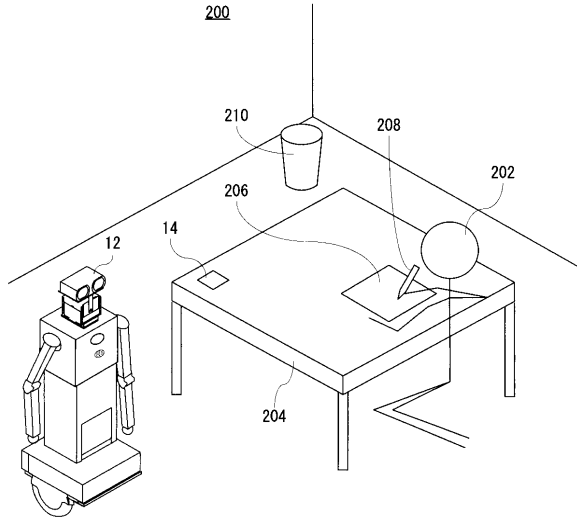
物品	動作候補
鉛筆	持つ、書く、見る、置く、...
ノート	持つ、書く、見る、置く、消す、...
消しゴム	持つ、置く、消す、...
机	置く、書く、見る、消す、...
⋮	⋮
包丁	持つ、置く、切る、研ぐ、洗う、...
まな板	持つ、置く、切る、洗う、...
⋮	⋮

【図5】

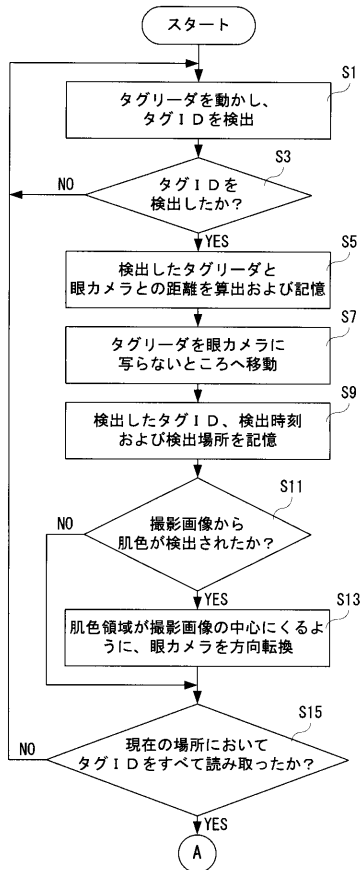
動作テンプレートデータ106c

動作	画像ファイル名
書く	ファイルA
見る	ファイルB
置く	ファイルC
持つ	ファイルD
⋮	⋮
切る	ファイルY
投げる	ファイルZ
⋮	⋮

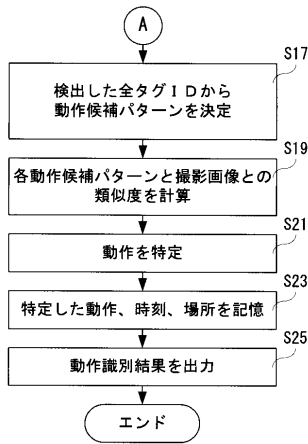
【図6】



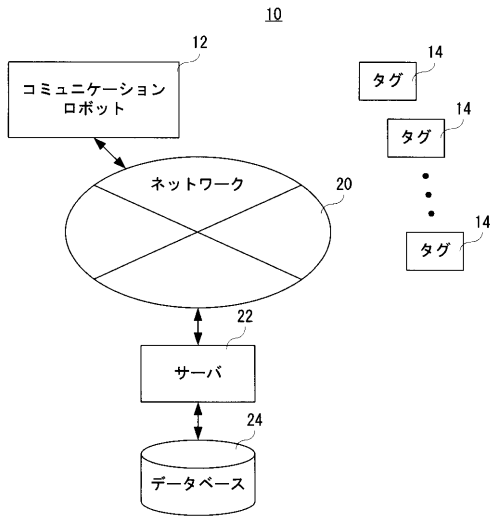
【図7】



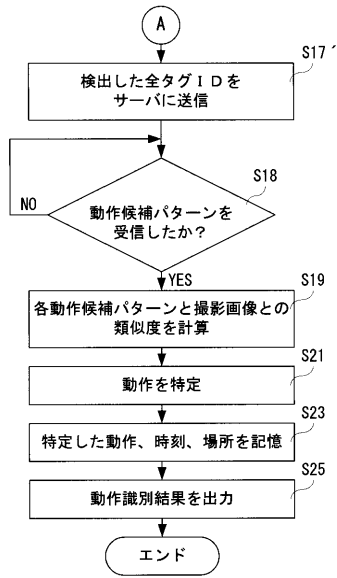
【図8】



【図9】



【図10】



フロントページの続き

(56)参考文献 特開2004-227208(JP,A)
特開2000-105596(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00~7/60

G06T 1/00