

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2019年10月3日(03.10.2019)



(10) 国際公開番号

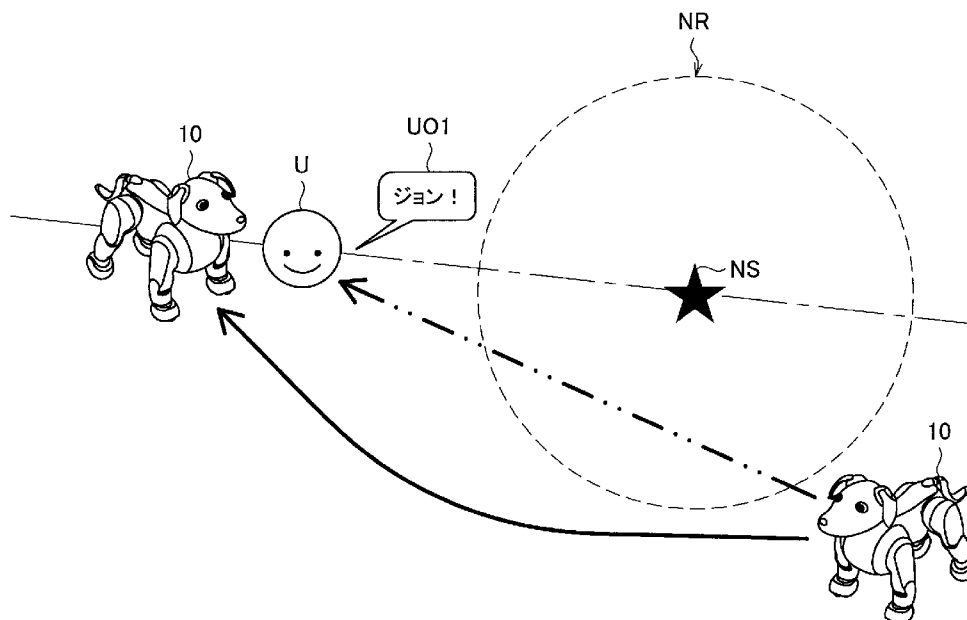
WO 2019/187834 A1

- (51) 国際特許分類:
B25J 13/00 (2006.01) *A63H 30/02* (2006.01)
A63H 3/33 (2006.01) *G10L 15/00* (2013.01)
A63H 11/00 (2006.01) *G10L 25/84* (2013.01)
A63H 11/18 (2006.01)
- (21) 国際出願番号: PCT/JP2019/006580
- (22) 国際出願日: 2019年2月21日(21.02.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-069788 2018年3月30日(30.03.2018) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 沢田 充 奨 (SAWATA, Ryosuke); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 小山 裕一郎 (KOYAMA, Yuichiro); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 特許業務法人酒井国際特許事務所 (SAKAI INTERNATIONAL PATENT OFFICE); 〒1000013 東京都千代田区霞が関3丁目8番1号 虎の門三井ビルディング Tokyo (JP).

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、およびプログラム

[図10]



U01 "John!"

(57) Abstract: [Problem] To cause an autonomous mobile body to execute a motion for further improving the accuracy of speech recognition. [Solution] Provided is an information processing device equipped with a motion control unit for controlling the motion of an autonomous mobile body so as to act on the basis of a recognition process, wherein when a target sound, which is sound to be subjected to a speech recognition process, has been detected, the motion control unit causes the autonomous mobile body to move to a position where the input level of a non-target sound, being a sound other



WO 2019/187834 A1

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): 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, IR, IS, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

than the subject sound, becomes lower, at the periphery of an approaching object determined on the basis of the target sound. Also provided is an information processing method comprising a step performed by a processor, the step being controlling the motion of an autonomous mobile body so as to act on the basis of a recognition process. The controlling further comprises, when a target sound, which is sound to be subjected to a speech recognition process, has been detected, moving the autonomous mobile body to a position where the input level of a non-target sound, being a sound other than the subject sound, becomes lower, at the periphery of an approaching object determined on the basis of the target sound.

(57) 要約: 【課題】自律移動体に音声認識の精度をより向上させる動作を実行させる。【解決手段】認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、を備え、前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、情報処理装置が提供される。また、プロセッサが、認識処理に基づいて行動を行う自律移動体の動作を制御すること、を含み、前記制御することは、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させること、をさらに含む、情報処理方法が提供される。

明 細 書

発明の名称： 情報処理装置、情報処理方法、およびプログラム
技術分野

[0001] 本開示は、情報処理装置、情報処理方法、およびプログラムに関する。

背景技術

[0002] 近年、ユーザの発話などの音源の方向を推定し、当該音源の方向に応じた動作を実行する種々の装置が開発されている。上記の装置には、例えば、推定した音源方向に基づいて、自律移動を実行する自律移動体が含まれる。例えば、特許文献1には、ユーザの発話や顔が認識された方向にロボット装置を移動させる技術が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2004-130427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、特許文献1に記載の技術では、ユーザの発話以外の音、すなわちノイズの存在が考慮されていない。このため、推定されたユーザの方向にロボット装置を単に接近させた場合、ノイズの入力レベルが上昇し、ユーザの発話の認識が困難となる可能性がある。

[0005] そこで、本開示では、自律移動体に音声認識の精度をより向上させる動作を実行させることが可能な、新規かつ改良された情報処理装置、情報処理方法、およびプログラムを提案する。

課題を解決するための手段

[0006] 本開示によれば、認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、を備え、前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、

前記自律移動体を移動させる、情報処理装置が提供される。

[0007] また、本開示によれば、プロセッサが、認識処理に基づいて行動を行う自律移動体の動作を制御すること、を含み、前記制御することは、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させること、をさらに含む、情報処理方法が提供される。

[0008] また、本開示によれば、コンピュータを、認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、を備え、前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、情報処理装置、として機能させるためのプログラムが提供される。

発明の効果

[0009] 以上説明したように本開示によれば、自律移動体に音声認識の精度をより向上させる動作を実行させることが可能となる。

[0010] なお、上記の効果は必ずしも限定的なものではなく、上記の効果とともに、または上記の効果に代えて、本明細書に示されたいずれかの効果、または本明細書から把握され得る他の効果が奏されてもよい。

図面の簡単な説明

[0011] [図1]本開示の一実施形態に係る自律移動体のハードウェア構成例を示す図である。

[図2]同実施形態に係る自律移動体が備えるアクチュエータの構成例である。

[図3]同実施形態に係る自律移動体が備えるアクチュエータの動作について説明するための図である。

[図4]同実施形態に係る自律移動体が備えるアクチュエータの動作について説明するための図である。

[図5]同実施形態に係る自律移動体が備えるディスプレイの機能について説明

するための図である。

[図6]同実施形態に係る自律移動体の動作例を示す図である。

[図7]同実施形態に係る自律移動体の機能構成例を示す図である。

[図8A]同実施形態に係る自律移動体の動作概要について説明するための図である。

[図8B]同実施形態に係る自律移動体の動作概要について説明するための図である。

[図9]同実施形態に係る目的音の未検出時における自律移動体の動作制御について説明するための図である。

[図10]同実施形態に係る目的音が検出された場合におけるノイズマップに基づく動作制御について説明するための図である。

[図11]同実施形態に係る目的音が検出された場合におけるノイズマップに基づく動作制御について説明するための図である。

[図12]同実施形態に係る接近対象が発話ユーザではない場合における動作制御について説明するための図である。

[図13]同実施形態に係る音源方向推定に基づくノイズマップの生成について説明するための図である。

[図14]同実施形態に係る音源方向推定に基づくノイズマップの生成について説明するための図である。

[図15]同実施形態に係るノイズ源の種別情報を含むノイズマップの一例を示す図である。

[図16]同実施形態に係るノイズマップの生成処理および更新処理に係る実行条件の設定例である。

[図17]同実施形態に係るノイズマップの統合処理について説明するための図である。

[図18]同実施形態に係るユーザ入力に基づくノイズマップの生成および更新について説明するための図である。

[図19]同実施形態に係るユーザ入力に基づくノイズマップの生成および更新

について説明するための図である。

[図20]同実施形態に係るノイズ領域の回避が困難な状況の一例を示す図である。

[図21]同実施形態に係る音声らしさの度合いを示す指標 α と定常性の度合いを示す指標 β の算出について説明するための図である。

[図22]同実施形態に係る音声らしさの度合いを示す指標 α と定常性の度合いを示す指標 β の算出について説明するための図である。

[図23]同実施形態に係るノイズマップ更新の流れを示すフローチャートである。

[図24]同実施形態に係る動作制御の流れを示すフローチャートである。

発明を実施するための形態

[0012] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0013] なお、説明は以下の順序で行うものとする。

1. 構成

1. 1. 自律移動体10の概要

1. 2. 自律移動体10のハードウェア構成例

1. 3. 自律移動体10の機能構成例

2. 実施形態

2. 1. 概要

2. 2. 動作制御の詳細

2. 3. ノイズマップの生成と更新

2. 4. ノイズ源の回避優先度に基づく動作制御

2. 5. 動作の流れ

3. まとめ

[0014] <1. 構成>

<< 1. 1. 自律移動体 10 の概要 >>

上述したように、近年においては、ユーザの発話などを認識し、認識結果に基づく動作を実行する種々の装置が開発されている。上記のような装置には、例えば、ユーザの発話や、周囲の環境などに応じて振る舞いを変化させる自律移動体が挙げられる。

[0015] ここで、一般に精度の高い音声認識を実現するためには、マイクロフォンを通じて得られる音の信号において、音声認識の対象音声である目的音（例えば、ユーザの発話音声）と対象音声ではない非目的音のパワーの比、すなわちSN比（Signal-to-Noise Ratio）を高めることが重要となる。特に、音声認識機能を有する自律移動体にあつては、SN比が向上する位置に移動を行うことで音声認識精度を高めることが望ましい。

[0016] しかし、特許文献1に記載の技術では、非目的音を考慮しておらず、ユーザの発話や顔が認識された方向にロボット装置を移動させるに留まっている。このため、特許文献1に記載の技術では、ユーザに接近すると共に非目的音を発するノイズ源にも同時に接近した結果、SN比が低下し、音声認識精度が低下する状況も想定される。

[0017] また、特許文献1に記載の技術では、ユーザの発話や顔が認識されたことをトリガーとし、ユーザに接近するようロボット装置の動作を制御している。このため、特許文献1に記載のロボット装置は、周辺に存在するユーザに常に追従する可能性が高く、ユーザに煩わしさを感じさせてしまうことも想定される。

[0018] 本開示の一実施形態に係る情報処理装置、情報処理方法、およびプログラムは、上記の点に着目して発想されたものであり、自律移動体に音声認識の精度をより向上させる動作を実行させることを可能とする。

[0019] ここで、まず、本開示の一実施形態に係る自律移動体10の概要について説明する。本開示の一実施形態に係る自律移動体10は、収集したセンサ情報に基づく状況推定を実行し、状況に応じた種々の動作を自律的に選択し実行する情報処理装置である。自律移動体10は、単にユーザの指示コマンド

に従った動作を行うロボットとは異なり、状況ごとに最適であると推測した動作を自律的に実行することを特徴の一つとする。

[0020] 本開示の一実施形態に係る自律移動体 10 は、例えば、音声認識処理の対象音声である目的音、すなわちユーザの発話を検出されない場合には、上記対象音声ではない非目的音の入力が回避されるよう自律動作を行ってよい。当該動作によれば、ユーザに常につきまとうことなく、またユーザの発話を検出された際には、当該発話に係る音声認識の精度が向上する可能性を効果的に高めることが可能である。

[0021] また、本開示の一実施形態に係る自律移動体 10 は、目的音を検出された場合には、当該目的音に基づいて定まる接近対象の周辺において非目的音の入力レベルがより低下する位置に移動してよい。すなわち、本開示の一実施形態に係る自律移動体 10 は、非目的音を考慮した移動動作を行うことで、S/N 比を向上させ、ユーザの発話に係る音声認識精度を効果的に向上させることが可能である。

[0022] このように、本開示の一実施形態に係る自律移動体 10 は、ヒトを含む動物と同様に、自己の状態や周囲の環境などを総合的に判断して自律動作を決定、実行する。上記の点において、本開示の一実施形態に係る自律移動体 10 は、指示に基づいて対応する動作や処理を実行する受動的な装置とは明確に相違する。

[0023] 本開示の一実施形態に係る自律移動体 10 は、空間内において自律的な姿勢制御を行い、種々の動作を実行する自律移動型ロボットであってよい。自律移動体 10 は、例えば、ヒトやイヌなどの動物を模した形状や、動作能力を有する自律移動型ロボットであってもよい。また、自律移動体 10 は、例えば、ユーザとのコミュニケーション能力を有する車両や無人航空機などの装置であってもよい。本開示の一実施形態に係る自律移動体 10 の形状、能力、また欲求などのレベルは、目的や役割に応じて適宜設計され得る。

[0024] <<1. 2. 自律移動体 10 のハードウェア構成例>>

次に、本開示の一実施形態に係る自律移動体 10 のハードウェア構成例に

ついて説明する。なお、以下では、自律移動体 10 がイヌ型の四足歩行ロボットである場合を例に説明する。

[0025] 図 1 は、本開示の一実施形態に係る自律移動体 10 のハードウェア構成例を示す図である。図 1 に示すように、自律移動体 10 は、頭部、胴部、4 つの脚部、および尾部を有するイヌ型の四足歩行ロボットである。また、自律移動体 10 は、頭部に 2 つのディスプレイ 510 を備える。

[0026] また、自律移動体 10 は、種々のセンサを備える。自律移動体 10 は、例えば、マイクロフォン 515、カメラ 520、ToF (Time of Flight) センサ 525、人感センサ 530、測距センサ 535、タッチセンサ 540、照度センサ 545、足裏ボタン 550、慣性センサ 555 を備える。

[0027] (マイクロフォン 515)

マイクロフォン 515 は、周囲の音を収集する機能を有する。上記の音には、例えば、ユーザの発話や、周囲の環境音が含まれる。自律移動体 10 は、例えば、頭部に 4 つのマイクロフォンを備えてもよい。複数のマイクロフォン 515 を備えることで、周囲で発生する音を感度高く収集すると共に、音源の定位を実現することが可能となる。

[0028] (カメラ 520)

カメラ 520 は、ユーザや周囲環境を撮像する機能を有する。自律移動体 10 は、例えば、鼻先と腰部に 2 つの広角カメラを備えてもよい。この場合、鼻先に配置される広角カメラは、自律移動体 10 の前方視野（すなわち、イヌの視野）に対応した画像を撮像し、腰部の広角カメラは、上方を中心とする周囲領域の画像を撮像する。自律移動体 10 は、例えば、腰部に配置される広角カメラにより撮像された画像に基づいて、天井の特徴点などを抽出し、SLAM (Simultaneous Localization and Mapping) を実現することができる。

[0029] (ToF センサ 525)

ToF センサ 525 は、頭部前方に存在する物体との距離を検出する機能

を有する。ToFセンサ525は、頭部の鼻先に備えられる。ToFセンサ525によれば、種々の物体との距離を精度高く検出することができ、ユーザを含む対象物や障害物などとの相対位置に応じた動作を実現することが可能となる。

[0030] (人感センサ530)

人感センサ530は、ユーザやユーザが飼育するペットなどの所在を検知する機能を有する。人感センサ530は、例えば、胸部に配置される。人感センサ530によれば、前方に存在する動物体を検知することで、当該動物体に対する種々の動作、例えば、興味、恐怖、驚きなどの感情に応じた動作を実現することが可能となる。

[0031] (測距センサ535)

測距センサ535は、自律移動体10の前方床面の状況を取得する機能を有する。測距センサ535は、例えば、胸部に配置される。測距センサ535によれば、自律移動体10の前方床面に存在する物体との距離を精度高く検出することができ、当該物体との相対位置に応じた動作を実現することができる。

[0032] (タッチセンサ540)

タッチセンサ540は、ユーザによる接触を検知する機能を有する。タッチセンサ540は、例えば、頭頂、あご下、背中など、ユーザが自律移動体10に対し触れる可能性が高い部位に配置される。タッチセンサ540は、例えば、静電容量式や感圧式のタッチセンサであってよい。タッチセンサ540によれば、ユーザによる触れる、撫でる、叩く、押すなどの接触行為を検知することができ、当該接触行為に応じた動作を行うことが可能となる。

[0033] (照度センサ545)

照度センサ545は、自律移動体10が位置する空間の照度を検出する。照度センサ545は、例えば、頭部背面において尾部の付け根などに配置されてもよい。照度センサ545によれば、周囲の明るさを検出し、当該明るさに応じた動作を実行することが可能となる。

[0034] (足裏ボタン 550)

足裏ボタン 550 は、自律移動体 10 の脚部底面が床と接触しているか否かを検知する機能を有する。このために、足裏ボタン 550 は、4 つの脚部の肉球に該当する部位にそれぞれ配置される。足裏ボタン 550 によれば、自律移動体 10 と床面との接触または非接触を検知することができ、例えば、自律移動体 10 がユーザにより抱き上げられたことなどを把握することが可能となる。

[0035] (慣性センサ 555)

慣性センサ 555 は、頭部や胴部の速度や加速度、回転などの物理量を検出する 6 軸センサである。すなわち、慣性センサ 555 は、X 軸、Y 軸、Z 軸の加速度および角速度を検出する。慣性センサ 555 は、頭部および胴部にそれぞれ配置される。慣性センサ 555 によれば、自律移動体 10 の頭部および胴部の運動を精度高く検出し、状況に応じた動作制御を実現することが可能となる。

[0036] 以上、本開示の一実施形態に係る自律移動体 10 が備えるセンサの一例について説明した。なお、図 1 を用いて説明した上記の構成はあくまで一例であり、自律移動体 10 が備え得るセンサの構成は係る例に限定されない。自律移動体 10 は、上記の構成のほか、例えば、温度センサ、地磁気センサ、GNSS (Global Navigation Satellite System) 信号受信機を含む各種の通信装置などをさらに備えてよい。自律移動体 10 が備えるセンサの構成は、仕様や運用に応じて柔軟に変形され得る。

[0037] 続いて、本開示の一実施形態に係る自律移動体 10 の関節部の構成例について説明する。図 2 は、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の構成例である。本開示の一実施形態に係る自律移動体 10 は、図 2 に示す回転箇所に加え、耳部と尾部に 2 つずつ、口に 1 つの合計 22 の回転自由度を有する。

[0038] 例えば、自律移動体 10 は、頭部に 3 自由度を有することで、頷きや首を

傾げる動作を両立することができる。また、自律移動体 10 は、腰部に備えるアクチュエータ 570 により、腰のスイング動作を再現することで、より現実のイヌに近い自然かつ柔軟な動作を実現することが可能である。

[0039] なお、本開示の一実施形態に係る自律移動体 10 は、例えば、1 軸アクチュエータと 2 軸アクチュエータを組み合わせることで、上記の 22 の回転自由度を実現してもよい。例えば、脚部における肘や膝部分においては 1 軸アクチュエータを、肩や大腿の付け根には 2 軸アクチュエータをそれぞれ採用してもよい。

[0040] 図 3 および図 4 は、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の動作について説明するための図である。図 3 を参照すると、アクチュエータ 570 は、モータ 575 により出力ギアを回転させることで、可動アーム 590 を任意の回転位置および回転速度で駆動させることができる。

[0041] 図 4 を参照すると、本開示の一実施形態に係るアクチュエータ 570 は、リアカバー 571、ギアBOXカバー 572、制御基板 573、ギアBOXベース 574、モータ 575、第 1 ギア 576、第 2 ギア 577、出力ギア 578、検出用マグネット 579、2 個のベアリング 580 を備える。

[0042] 本開示の一実施形態に係るアクチュエータ 570 は、例えば、磁気式 s v GMR (spin-valve Giant Magnetoresistive) であってもよい。制御基板 573 が、メインプロセッサによる制御に基づいて、モータ 575 を回転させることで、第 1 ギア 576 および第 2 ギア 577 を介して出力ギア 578 に動力が伝達され、可動アーム 590 を駆動させることが可能である。

[0043] また、制御基板 573 に備えられる位置センサが、出力ギア 578 に同期して回転する検出用マグネット 579 の回転角を検出することで、可動アーム 590 の回転角度、すなわち回転位置を精度高く検出することができる。

[0044] なお、磁気式 s v GMR は、非接触方式であるため耐久性に優れるとともに、GMR 飽和領域において使用することで、検出用マグネット 579 や位

置センサの距離変動による信号変動の影響が少ないという利点を有する。

[0045] 以上、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の構成例について説明した。上記の構成によれば、自律移動体 10 が備える関節部の屈伸動作を精度高く制御し、また関節部の回転位置を正確に検出することが可能となる。

[0046] 続いて、図 5 を参照して、本開示の一実施形態に係る自律移動体 10 が備えるディスプレイ 510 の機能について説明する。図 5 は、本開示の一実施形態に係る自律移動体 10 が備えるディスプレイ 510 の機能について説明するための図である。

[0047] (ディスプレイ 510)

ディスプレイ 510 は、自律移動体 10 の目の動きや感情を視覚的に表現する機能を有する。図 5 に示すように、ディスプレイ 510 は、感情や動作に応じた眼球、瞳孔、瞼の動作を表現することができる。ディスプレイ 510 は、文字や記号、また眼球運動とは関連しない画像などを敢えて表示しないことで、実在するイヌなどの動物に近い自然な動作を演出する。

[0048] また、図 5 に示すように、自律移動体 10 は、右眼および左眼にそれぞれ相当する 2 つのディスプレイ 510 r および 510 l を備える。ディスプレイ 510 r および 510 l は、例えば、独立した 2 つの OLED (Organic Light Emitting Diode) により実現される。OLED によれば、眼球の曲面を再現することが可能となり、1 枚の平面ディスプレイにより一対の眼球を表現する場合や、2 枚の独立した平面ディスプレイにより 2 つの眼球をそれぞれ表現する場合と比較して、より自然な外装を実現することができる。

[0049] 以上述べたように、ディスプレイ 510 r および 510 l によれば、図 5 に示すような自律移動体 10 の視線や感情を高精度かつ柔軟に表現することが可能となる。また、ユーザはディスプレイ 510 に表示される眼球の動作から、自律移動体 10 の状態を直観的に把握することが可能となる。

[0050] 以上、本開示の一実施形態に係る自律移動体 10 のハードウェア構成例に

ついて説明した。上記の構成によれば、図6に示すように、自律移動体10の関節部や眼球の動作を精度高くまた柔軟に制御することで、より実在の生物に近い動作および感情表現を実現することが可能となる。なお、図6は、本開示の一実施形態に係る自律移動体10の動作例を示す図であるが、図6では、自律移動体10の関節部および眼球の動作について着目して説明を行うため、自律移動体10の外部構造を簡略化して示している。同様に、以下の説明においては、自律移動体10の外部構造を簡略化して示す場合があるが、本開示の一実施形態に係る自律移動体10のハードウェア構成および外装は、図面により示される例に限定されず、適宜設計され得る。

[0051] <<1. 3. 自律移動体10の機能構成例>>

次に、本開示の一実施形態に係る自律移動体10の機能構成例について説明する。図7は、本開示の一実施形態に係る自律移動体10の機能構成例を示す図である。図7を参照すると、本開示の一実施形態に係る自律移動体10は、入力部110、認識部120、周辺環境推定部130、周辺環境保持部140、動作制御部150、駆動部160、出力部170を備える。

[0052] (入力部110)

入力部110は、ユーザや周囲環境に係る種々の情報を収集する機能を有する。入力部110は、例えば、ユーザの発話や周囲で発生する環境音、ユーザや周囲環境に係る画像情報、および種々のセンサ情報を収集する。このために、入力部110は、図1に示す各種のセンサを備える。

[0053] (認識部120)

認識部120は、入力部110が収集した種々の情報に基づいて、ユーザや周囲の物体、また自律移動体10の状態に係る種々の認識を行う機能を有する。一例としては、認識部120は、人識別、顔認識、表情や視線の認識、音声認識、物体認識、色認識、形認識、マーカー認識、障害物認識、段差認識、明るさ認識などを行ってよい。

[0054] (周辺環境推定部130)

周辺環境推定部130は、入力部110が収集したセンサ情報や、認識部

120による認識結果に基づいて、非目的音の発生状況を示すノイズマップを生成、更新する機能を有する。周辺環境推定部130が有する機能については別途詳細に説明する。

[0055] (周辺環境保持部140)

周辺環境保持部140は、周辺環境推定部130が生成、更新したノイズマップを保持する機能を有する。

[0056] (動作制御部150)

動作制御部150は、認識部120による認識結果や周辺環境保持部140が保持するノイズマップに基づいて行動計画を行い、当該行動計画に基づいて、駆動部160および出力部170の動作を制御する。動作制御部150は、例えば、上記の行動計画に基づいて、アクチュエータ570の回転制御や、ディスプレイ510の表示制御、スピーカによる音声出力制御などを行う。本開示の一実施形態に係る動作制御部150が有する機能については別途詳細に説明する。

[0057] (駆動部160)

駆動部160は、動作制御部150による制御に基づいて、自律移動体10が有する複数の関節部を屈伸させる機能を有する。より具体的には、駆動部160は、動作制御部150による制御に基づき、各関節部が備えるアクチュエータ570を駆動させる。

[0058] (出力部170)

出力部170は、動作制御部150による制御に基づいて、視覚情報や音情報の出力を行う機能を有する。このために、出力部170は、ディスプレイ510やスピーカを備える。

[0059] 以上、本開示の一実施形態に係る自律移動体10の機能構成について説明した。なお、図7に示す構成はあくまで一例であり、本実施形態に係る自律移動体10の機能構成は係る例に限定されない。本開示の一実施形態に係る自律移動体10は、例えば、情報処理サーバや他の自律移動体と通信を行う通信部などを備えてよい。また、認識部120、周辺環境推定部130、動

作制御部 150 などが有する機能は、上記の情報処理サーバの機能として実現されてもよい。この場合、情報処理サーバは、自律移動体 10 の入力部 110 が収集したセンサ情報に基づいて各種の認識処理、ノイズマップの生成または更新、行動計画を実行し、自律移動体 10 の駆動部 160 と出力部 170 の制御を行うことが可能である。本開示の一実施形態に係る自律移動体 10 の機能構成は仕様や運用に応じて柔軟に変形され得る。

[0060] <2. 実施形態>

<<2. 1. 概要>>

次に、本開示の実施形態の概要について説明する。上述したように、本開示の一実施形態に係る自律移動体 10 は、目的音に係る音声認識の精度を向上させるために、目的音と非目的音とに係る S/N 比が向上するように自律動作を行う。

[0061] ここで、S/N 比を向上させるための手法としては、入力信号に信号処理を施す手法（マルチマイク信号処理、シングルマイク信号処理）や、指向性マイクロフォンなどを用いる手法も想定される。しかし、S/N 比に対しては、目的音源や非目的音源（以下、ノイズ源、とも称する）との物理的な距離の影響が最も強いといえる。

[0062] このために、本実施形態に係る自律移動体 10 は、単純に目的音源へ近づくのではなく、目的音源へ近づきながらも非目的音源から可能な限り遠ざかることで、S/N 比を効果的に向上させることができる。

[0063] 図 8A および図 8B は、本実施形態に係る自律移動体 10 の動作概要について説明するための図である。図 8A には、本実施形態に係る比較手法による移動制御の一例が示されている。図 8A に示す一例の場合、比較装置 90 は、ユーザ U の音声、すなわち目的音が検出された場合、当該目的音の入力レベル（音圧レベル）が上昇するように、ユーザ U に最短距離を以って接近している。しかし、この際、比較装置 90 は、非目的音を発するノイズ源 N の存在を考慮していないため、ユーザ U に接近すると同時にノイズ源 N にも接近している。このように、比較手法の場合、目的音の入力レベルと共

に非目的音の入力レベルも上昇し、結果としてS/N比の向上効果が薄れ、ユーザUの発話に係る音声認識精度が低下する可能性が生じる。

[0064] 一方、図8Bは、本実施形態に係る情報処理方法による移動制御の一例が示されている。図8Bに示すように、本実施形態に係る自律移動体10は、ユーザUの音声、すなわち目的音が検出された場合、非目的音を発するノイズ源の存在を考慮して移動を行う。具体的には、本実施形態に係る動作制御部150は、目的音を検出された場合、目的音に基づいて定まる接近対象の周辺において非目的音の入力レベルがより低下する位置に、自律移動体10を移動させてよい。

[0065] ここで、上記の接近対象は、目的音、すなわち発話音声を発したユーザUであってよい。すなわち、本実施形態に係る動作制御部150は、ユーザUの発話を検出された場合、接近対象であるユーザUの周辺においてユーザUの発話に係る入力レベルがより上昇し、ノイズ音源NSが発する非目的音の入力レベルがより低下する位置に、自律移動体10を移動させることができる。

[0066] 図8Bに示す一例の場合、本実施形態に係る動作制御部150は、接近対象であるユーザUを中心に非目的音の発生位置、すなわちノイズ源NSとは反対側に自律移動体10を移動させている。このように、本実施形態に係る動作制御部150によれば、ノイズ音源NSからより遠ざかり、かつ接近対象であるユーザUにより近づくように、自律移動体10を移動させることができる。また、自律移動体10をユーザUを挟んでノイズ源NSとは反対側に移動させる場合、ユーザUが壁の役割を果たしノイズ源NSが発する非目的音の入力レベルがさらに効果的に低下する効果が期待される。

[0067] 本実施形態に係る動作制御部150が有する上記の機能によれば、目的音の入力レベルを上昇させると共に、非目的音の入力レベルを低下、または入力レベルの上昇率を軽減することができ、結果としてS/N比を効果的に向上させることが可能となる。このように、本実施形態に係る動作制御部150によれば、入力信号に対する信号処理や指向性マイクロフォンを用いたビー

ムフォーミングを行わずとも、自律移動体 10 が本来有する移動機能だけを利用して S/N 比を大幅に改善することができる。

[0068] <<2. 2. 動作制御の詳細>>

次に、本実施形態に係る動作制御部 150 による自律移動体 10 の動作制御についてより詳細に説明する。上述したように、本実施形態に係る動作制御部 150 は、目的音に加え非目的音の存在を考慮した動作を自律移動体 10 に実行させることで、S/N 比を改善し目的音に係る音声認識精度を効果的に向上させることができる。

[0069] ここで、本実施形態に係る目的音とは、認識部 120 による音声認識の対象音声と定義できる。上記対象音声は、音声全般であってもよい。例えば、自律移動体 10 がテレビジョン装置やラジオなどから出力される音声や、ユーザまたは第三者の発話音声のすべてを音声認識の対象とする場合、上記のような音声は、すべて目的音といえる。この場合、認識部 120 は、例えば、人の声の倍音構造などに着目しピッチを比較するなどして、上記のような音声全般を目的音として検出することが可能である。なお、上記のような音声をすべて目的音とする場合、自律移動体 10 が例えば、テレビジョン装置から出力される音声に応じて何らかのアクションを行うことなどが可能となる。

[0070] 一方、本実施形態に係る目的音は、上記のような音声のうち、例えば、事前に登録された所定のユーザの音声のみであってもよい。この場合、認識部 120 は、予め登録されたユーザの音声特徴に基づく話者識別や、入力信号の到来方向に存在する人物の顔識別を行うことで、所定のユーザの発話音声のみを目的音として検出可能である。

[0071] 他方、本実施形態に係る目的音は、所定ユーザが発する発話音声のうち、特定のキーワードや、行動指示などに係る特定の文言のみであってもよい。この場合、認識部 120 は、入力信号に基づく音声認識を行うことで特定のキーワードや文言を含む所定のユーザの発話音声のみを目的音として検出することが可能である。

[0072] また、本実施形態に係る非目的音とは、目的音以外のすべての音として定義できる。本実施形態に係る非目的音には、例えば、キッチンにおける作業音や、換気扇や冷蔵庫、車などの装置が発する種々の非音声が挙げられる。

[0073] 以上、本実施形態に係る目的音と非目的音について詳細に説明した。続いて、図9を参照して、目的音が検出されない場合における動作制御について説明する。図9は、本実施形態に係る目的音の未検出時における自律移動体10の動作制御について説明するための図である。図9には、周辺環境保持部140が保持するノイズマップの一例が示されている。

[0074] ここで、本実施形態に係るノイズマップとは、周辺環境推定部130により生成、更新される、非目的音の発生状況を示すマップである。本実施形態に係るノイズマップには、例えば、自律移動体10が存在する空間に存在するノイズ源や、当該ノイズ源が発する非目的音の入力レベルが強い（例えば、閾値以上）領域であるノイズ領域に係る情報が含まれる。図9に示す一例の場合、ノイズマップには、ノイズ源NS1およびNS2と、ノイズ源NS1およびNS2のそれぞれに対応するノイズ領域NR1およびNR2の情報が含まれている。

[0075] 本実施形態に係る動作制御部150は、上記のような情報を含むノイズマップに基づいて、自律移動体10の動作を制御することを特徴の一つとする。例えば、本実施形態に係る動作制御部150は、目的音が検出されない場合、ノイズマップに基づいて非目的音の入力を回避するように自律移動体10の動作を制御してよい。

[0076] より具体的には、本実施形態に係る動作制御部150は、目的音が検出されない場合、ノイズマップに基づいて非目的音の入力レベルが閾値以下となるエリア内に自律移動体10の移動範囲を限定することができる。例えば、図9に示す一例の場合、動作制御部150は、目的音が検出されていない場合、ノイズ領域NR1およびNR2に自律移動体10が進入しないよう、両領域を除いたエリアに自律移動体10の移動範囲を限定してよい。この際、動作制御部150は、例えば、上記エリア内において自律移動体10をラン

ダムに移動させたり、非目的音の音圧が最小であると予想される位置 P_{min} などに自律移動体 10 を移動させてよい。

[0077] 本実施形態に係る動作制御部 150 による上記の制御によれば、目的音が検出されていない場合であっても、非目的音の入力になるべく抑えられるように自律移動体 10 を動作させることで、ユーザからの呼びかけなどがあった場合、すなわち目的音が検出された際に、当該目的音に係る音声認識の精度を効果的に向上させることが可能である。

[0078] 続いて、本実施形態に係る目的音が検出された場合におけるノイズマップに基づく動作制御について詳細に説明する。上述したように、本実施形態に係る動作制御部 150 は、目的音が検出された場合、接近対象の周辺において目的音の入力レベルがより上昇し非目的音の入力レベルがより低下する位置に、自律移動体 10 を移動させることを特徴の一つとする。この際、本実施形態に係る動作制御部 150 は、ノイズマップを参照することで上記の動作制御を精度高く実現することが可能である。

[0079] 図 10 は、本実施形態に係る目的音が検出された場合におけるノイズマップに基づく動作制御について説明するための図である。図 10 に示す一例の場合、動作制御部 150 は、自律移動体 10 の名前を呼ぶユーザ U の発話音声 U01、すなわち目的音が検出されたことに基づいて、自律移動体 10 をユーザ U に接近させている。この際、本実施形態に係る動作制御部 150 は、ノイズマップを参照し、ノイズマップに含まれるノイズ源 NS およびノイズ領域 NR を考慮して自律移動体 10 の移動を制御する。

[0080] 例えば、図 10 が示す状況において、図中に二点鎖線で示すように、自律移動体 10 を最短距離でユーザ U に接近させた場合、自律移動体 10 がノイズ領域 NR 内を移動することとなる。しかし、本実施形態に係る動作制御部 150 は、ノイズマップを参照することで、自律移動体 10 がノイズ領域 NR に進入またはノイズ領域 NR 内で停止することなく、かつノイズ源 NS からより遠ざかる位置に自律移動体 10 を移動させることが可能である。より具体的には、動作制御部 150 は、図中に実線で示すようにノイズ領域 NR

を迂回させ、接近対象であるユーザUを中心にノイズ源NSとは反対側に自律移動体10を移動させてよい。

[0081] このように、本実施形態に係る動作制御部150は、ノイズマップを参照することにより、ノイズ源やノイズ領域を正確に把握し、目的音の入力レベルが上昇し、かつ非目的音の入力レベルが低下する位置に自律移動体10を移動させることが可能である。本実施形態に係る動作制御部150による上記の動作制御によれば、SN比を向上させ、目的音に係る音声認識精度を効果的に向上させることが可能となる。

[0082] なお、本実施形態に係る動作制御部150は、必ずしも、接近対象を中心にノイズ源とは反対側に自律移動体10を移動させなくてもよい。例えば、図11に示す一例の場合、ノイズ源NSとユーザUを結ぶ直線上には壁が存在している。このように、ノイズ源と接近対象を結ぶ直線上に障害物がある場合、動作制御部150は、上記反対側に近い位置かつノイズ領域NRに進入しない位置で自律移動体10を停止させてよい。この場合であっても、目的音の入力レベル上昇と非目的音の入力レベル低下を両立し、SN比を向上させる効果を得ることが可能である。なお、動作制御部150は、ノイズマップに含まれる壁や家具などの情報に基づいて障害物の存在を把握してもよい、認識部120が認識した障害物の存在に基づいて上記のような動作制御を行ってもよい。

[0083] 次に、本実施形態に係る接近対象が発話ユーザではない場合の動作制御について説明する。図12は、本実施形態に係る接近対象が発話ユーザではない場合における動作制御について説明するための図である。

[0084] 図12に示す一例の場合、ユーザU1は、図10や図11に示した一例とは異なり、自身のところではなく、ユーザU2のもとへ移動を指示する旨の音声発話U02を行っている。この際、本実施形態に係る動作制御部150は、認識部120が認識した発話音声U02に基づいて、接近対象をユーザU2に設定し、ユーザU2に接近するよう自律移動体10の動作を制御する。

[0085] このように、本実施形態に係る接近対象は、音声発話を発した発話ユーザのみではなく、当該発話音声に基づく音声認識処理により特定された他のユーザなどの動体、充電ステーションなどの固定物体、または任意の位置であってもよい。

[0086] 本実施形態に係る動作制御部150は、接近対象が発話ユーザではない場合であっても、同様にノイズマップを参照し、ノイズ源NSやノイズ領域NRを考慮した移動を自律移動体10に実行させることができる。図12に示す一例の場合、動作制御部150は、ユーザU2を中心にノイズ源NSとは反対側に自律移動体10を移動させている。本実施形態に係る動作制御部150による上記の動作制御によれば、この後発せられることが予想されるユーザU2の発話音声、すなわち目的音に係る音声認識の精度を効果的に向上させることが可能となる。

[0087] 以上、本実施形態に係るノイズマップに基づく動作制御について説明した。上述したように、本実施形態に係る動作制御部150は、周辺環境保持部140が保持するノイズマップを参照することで、目的音の入力レベルのみではなく非目的音の入力レベルを考慮した自律移動体10の移動を実現し、SM比を向上させることで精度の高い音声認識を実現することができる。

[0088] なお、上記では、本実施形態に係る動作制御部150が、目的音が検出されたことに基づいて、接近対象へ移動するよう自律移動体10を制御する場合を主な例に述べたが、本実施形態に移動のトリガーに係る例に限定されない。本実施形態に係る動作制御部150は、例えば、ユーザの顔が認識されたことや、ユーザによる移動指示に係るジェスチャが認識されたことに基づいて、自律移動体10を接近対象に移動させる制御を行ってもよい。この場合であっても、ノイズマップを参照し、非目的音の入力レベルがより低下する位置に自律移動体を移動させることで、後に発せられると予想される目的音に係る音声認識の精度を高めることが可能である。

[0089] <<2. 3. ノイズマップの生成と更新>>

次に、本実施形態に係るノイズマップの生成と更新について詳細に説明す

る。本実施形態に係る周辺環境推定部 130 は、例えば、音源方向推定や音圧測定の結果に基づいて上述したようなノイズマップを生成することができる。

[0090] まず、本実施形態に係る音源方向推定に基づくノイズマップの生成について説明する。図 13 は、本実施形態に係る音源方向推定に基づくノイズマップの生成について説明するための図である。

[0091] 音源方向推定に基づくノイズマップの生成において、周辺環境推定部 130 は、まず任意の地点において音源定位を行い音源の方向を推定する。図 13 に示す一例の場合、周辺環境推定部 130 は、地点 P1 において、ノイズ源 NS1 および NS2 の方向をそれぞれ推定している。なお、この時点においては、ノイズ源 NS1 および NS2 の方向は推定可能であるが、自律移動体 10 とノイズ源 NS1 および NS2 との距離は不明である。

[0092] 続いて、周辺環境推定部 130 は、前回推定された音源方向とは異なる地点に移動し、再び音源定位を行い音源の方向を推定する。図 13 に示す一例の場合、周辺環境推定部 130 は、地点 P2 において、再びノイズ源 NS1 および NS2 の方向をそれぞれ推定している。この際、周辺環境推定部 130 は、自律移動体 10 移動距離と、地点 P1 および P2 において推定した方向の交点とから、空間上におけるノイズ源 NS1 および NS2 の位置を推定することができる。

[0093] この後、周辺環境推定部 130 は、また別の地点において音源方向推定を繰り返すことで、音源位置に係る推定精度を向上させることが可能である。図 13 に示す一例の場合、周辺環境推定部 130 は、地点 P3 において再度ノイズ源 NS1 および NS2 の方向を推定している。

[0094] このように、本実施形態に係る周辺環境推定部 130 は、複数地点において音源方向の推定を繰り返すことで、空間上におけるノイズ源 NS1 および NS2 の位置を精度高く推定し、例えば、それぞれの推定位置から所定距離にある領域をノイズ領域 NR1 および NR2 として設定したノイズマップを生成することが可能である。

[0095] 続いて、本実施形態に係る音圧測定に基づくノイズマップの生成について説明する。上述した音源方向推定に基づくノイズマップの生成は、自律移動体 10 が、同時に発生している音源数よりも多い数のマイクロフォンを備えていない場合、困難である。一方、本実施形態に係る周辺環境推定部 130 は、自律移動体 10 が単一のマイクロフォンのみしか備えていない場合であっても、下記に説明する音圧測定に基づいてノイズマップを生成することが可能である。図 14 は、本実施形態に係る音源方向推定に基づくノイズマップの生成について説明するための図である。

[0096] 音源方向推定に基づくノイズマップの生成において、周辺環境推定部 130 は、まず任意の地点において音圧レベルの測定を実行する。図 14 に示す一例の場合、周辺環境推定部 130 は、地点 P4 において音圧レベルを測定している。続いて、周辺環境推定部 130 は、測定済の地点とは異なる他の地点において音圧レベルの測定を繰り返し実行する。図 14 に示す一例の場合、周辺環境推定部 130 は、地点 P5 および P6 において音圧レベルの測定を実行している。

[0097] このように、本実施形態に係る周辺環境推定部 130 は、複数地点において音圧測定を繰り返すことで、図 14 に示すように音圧レベルに係る等圧線を推定し、音圧レベルが閾値以上であるエリアをノイズ領域として設定することができる。また、周辺環境推定部 130 は、ノイズ領域において最も音圧レベルが高い地点をノイズ源の位置として推定することも可能である。図 14 に示す一例の場合、周辺環境推定部 130 は、推定した等圧線に基づいてノイズ領域 NR1 および NR2 を設定し、またノイズ源 NS1 および NS2 の位置を推定している。

[0098] このように、本実施形態に係る周辺環境推定部 130 によれば、自律移動体 10 が単一のマイクロフォンのみを備える場合であっても、複数地点において音圧測定を繰り返し実行することで、精度の高いノイズマップを生成することが可能である。なお、音圧測定に基づくノイズマップ生成を行う場合には、目的音と非目的音を切り分けることが必要となるが、当該切り分けは

、上述した認識部 120 の機能により実現可能である。

[0099] また、本実施形態に係るノイズマップは、ノイズ源の種別などの情報を含んでもよい。図 15 は、本実施形態に係るノイズ源の種別情報を含むノイズマップの一例を示す図である。図 15 に示す一例の場合、ノイズ源 NS1 および NS2 がそれぞれキッチンおよびテレビジョン装置であることがノイズマップの情報として含まれていることがわかる。

[0100] 本実施形態に係る周辺環境推定部 130 は、例えば、認識部 120 による物体認識の結果に基づいて、図 15 に示すようなノイズ源の種別情報を含んだノイズマップを生成することが可能である。本実施形態に係る周辺環境推定部 130 が有する上記の機能によれば、動作制御部 150 が、ノイズ源の特定に応じてより精度の高い動作制御を行うことが可能となる。

[0101] 以上、本実施形態に係るノイズマップ生成の一例について説明した。続いて、本実施形態に係るノイズマップの生成または更新に係るタイミングについて説明する。

[0102] 例えば、ノイズマップの更新については、常時動的に行われる場合も想定される。この場合、周囲で発生する非目的音を漏れなく検出できる一方、動作制御部 150 による動作制御に対し有用ではない突発音などの情報もすべてノイズマップの情報として含まれることとなる。また、ノイズマップの更新を常時動的に実行する場合、計算量が膨大となるため、高性能のプロセッサなどが必要となる。

[0103] このため、本実施形態に係る周辺環境推定部 130 は、非目的音の収集に関し効率が良い条件においてのみノイズマップの生成処理や更新処理を実行してよい。ここで、上記の効率が良い条件とは、非目的音が多数発生し得る状況である。また、非目的音が多数発生し得る状況とは、すなわちユーザが空間上において活動を行っている状況という。このため、本実施形態に係る周辺環境推定部 130 は、ユーザが自律移動体 10 が設置される空間に存在するタイミングで、ノイズマップの生成処理や更新処理を実行してよい。

[0104] この際、本実施形態に係る周辺環境推定部 130 は、ユーザの予定や各種

のセンサ情報に基づいて、ユーザの不在あるいは存在を推定し、ユーザが存在する可能性が高い条件においてのみノイズマップの生成処理や更新処理を実行することができる。

[0105] 図16は、本実施形態に係るノイズマップの生成処理および更新処理に係る実行条件の設定例である。図16に示す一例の場合、ユーザの予定（不在または在宅）、鍵音の検出有無、ドア開閉音の検出有無、「ただいま」などの発話検出有無、人感センサによる動体の検出有無などの要素の組み合わせごとに、生成処理や更新処理を実施するか否かが設定されている。なお、図16では、上記すべての要素において、ユーザが自律移動体10と同一の空間に存在すると判定された場合にのみ、処理を実行する場合の一例が示されている。

[0106] なお、上記のような実行可否は、自律移動体10の特性や状況などに応じて動的に設定可能であってもよい。このように、本実施形態に係る周辺環境推定部130によれば、周辺環境にユーザが存在する時間帯において収集された非目的音に基づいてノイズマップの生成や更新を行うことで、より精度の高いノイズマップを保持することが可能となる。

[0107] 次に、本実施形態に係るノイズマップの更新処理について詳細に説明する。上述したように、本実施形態に係る周辺環境推定部130は、周辺環境にユーザが存在する時間帯において収集された非目的音に基づいてノイズマップを動的に更新することが可能である。

[0108] しかし、この場合、集音が行われるタイミングでは、発生する非目的音が異なる場合も想定される。このため、単純に最新の集音データに基づいてノイズマップを上書きした場合、突発音など本来影響の少ない非目的音の情報がノイズマップに含まれることとなり、動作制御部150による動作制御の精度が低下する要因ともなりかねない。

[0109] このため、本実施形態に係る周辺環境推定部130は、最新の集音データに基づいて既存のノイズマップを上書きするのではなく、最新の集音データを既存のノイズマップに統合することで、当該ノイズマップの更新を行って

よい。

[0110] 図 17 は、本実施形態に係るノイズマップの統合処理について説明するための図である。図 17 に示す一例の場合、本実施形態に係る周辺環境推定部 130 は、3 回の集音に基づいて、ノイズマップの統合を行っている。ここで、1 回目の集音において、キッチンおよびテレビジョン装置に係る非目的音が検出され、2 回目の集音において、窓およびテレビジョン装置に係る非目的音が検出され、3 回目の集音において、キッチンおよびテレビジョン装置に係る非目的音が検出された場合を想定する。

[0111] この場合、本実施形態に係る周辺環境推定部 130 は、3 回分の集音データを、例えば平均化するなどして統合することで、ノイズマップの更新を行ってよい。本実施形態に係る周辺環境推定部 130 によれば図 17 に示すように、各非目的音の発生頻度などをノイズマップに反映することができ、突発音など発生頻度の少ない非目的音の影響を抑えることが可能となる。なお、図 17 に示す一例の場合、非目的音の発生頻度は、ハッチングの濃度の濃淡（濃度が濃いほど発生頻度が高い）により示されている。この場合、動作制御部 150 は、非目的音を発する頻度が高いテレビジョン装置を、より重点的に回避するように、自律移動体 10 の移動を制御してよい。

[0112] 以上、本実施形態に係るノイズマップの生成および更新について説明した。なお、上記で示した手法はあくまで一例であり、本実施形態に係るノイズマップの生成および更新は、係る例に限定されない。

[0113] 本実施形態に係る周辺環境推定部 130 は、例えば、ユーザが入力した情報に基づいてノイズマップの生成や更新を行ってもよい。図 18 および図 19 は、本実施形態に係るユーザ入力に基づくノイズマップの生成および更新について説明するための図である。

[0114] 例えば、本実施形態に係る周辺環境推定部 130 は、図 18 に示すように、情報処理端末 20などを介してユーザが入力した家具の配置情報などに基づいて、ノイズマップの生成や更新を行ってもよい。この際、周辺環境推定部 130 は、情報処理端末 20 が備える表示部において、ユーザの部屋に見

立てた入力領域 1 A に、各家具に対応するアイコン 1 C を配置するように依頼し、入力された情報に基づいて、ノイズマップの生成や更新を実行することができる。

[0115] また、例えば、本実施形態に係る周辺環境推定部 130 は、図 19 に示すように、ユーザ U が行う指差しなどのジェスチャやノイズ源の教示に係る発話音声 UO3 に基づいて、ノイズ源 NS を特定し、ノイズマップに反映することが可能である。

[0116] <<2. 4. ノイズ源の回避優先度に基づく動作制御>>

次に、本実施形態に係るノイズ源の回避優先度に基づく動作制御について説明する。上記では、本実施形態に係る動作制御部 150 が、ノイズマップを参照し、ノイズ領域を避けるように自律移動体 10 の移動を制御することについて述べた。

[0117] しかし、状況によっては、ノイズ領域を回避して移動することが困難な場合も想定される。図 20 は、本実施形態に係るノイズ領域の回避が困難な状況の一例を示す図である。図 20 の左側には、音源方向推定に基づくノイズマップが、図 20 の右側には音圧測定に基づくノイズマップがそれぞれ例示されている。ここで、両ノイズマップに着目すると、いずれのノイズマップにおいても、自律移動体 10 がノイズ領域 NR1 ~ NR4 に囲まれており、他の場所への移動が困難であることがわかる。

[0118] このような場合、本実施形態に係る動作制御部 150 は、ノイズ源 NS1 ~ NS4 に係る回避優先度に基づいて、より回避優先度が低いノイズ源に対応するノイズ領域に移動するよう自律移動体 10 を制御してよい。

[0119] ここで、本実施形態に係る回避優先度は、例えば、ノイズ源が発する非目的音の種別や特性により決定され得る。上述したように、本実施形態に係る非目的音は、目的音以外の種々の音を含む。一方、非目的音が目的音に係る音声認識の精度に与える影響は、非目的音の特性に応じて異なる。

[0120] このため、本実施形態に係る周辺環境推定部 130 は、音声認識精度に対する影響度の大きさに基づいて、非目的音を分類し、当該影響度の大きい順

に回避優先度を高く設定したノイズマップを生成してよい。

[0121] ここでは、非目的を4つのカテゴリ1～4に分類する例を示す。例えば、カテゴリ1は、比較的音量が大きく、人間の音声でありながら目的音ではない非目的音であってよい。カテゴリ1には、例えば、テレビジョン装置やラジオ、その他の装置が出力する音声や、ヴォーカルを含む楽曲、ユーザではない第三者同士による会話などが含まれる。カテゴリ1は、4つのカテゴリ中、音声認識精度に与える影響が最も高く、かつ回避優先度が最も高い非目的音であってよい。

[0122] また、カテゴリ2は、発生が非定常的かつ比較的音量が大きいためノイズ抑圧の効果が十分に得にくい非目的音であってよい。カテゴリ2には、例えば、食器洗いや調理などの作業音、窓を解放した際に流入する屋外音などが含まれる。カテゴリ2は、4つのカテゴリ中、音声認識精度に与える影響が2番目に高く、かつ回避優先度が2番目に高い非目的音であってよい。

[0123] また、カテゴリ3は、発生が定常的であり、比較的ノイズ抑圧の効果が得やすい非目的音であってよい。カテゴリ3には、例えば、エアコンディショナーや換気扇、PCのファンなどが発する音が含まれる。カテゴリ2は、4つのカテゴリ中、音声認識精度に与える影響が3番目に高く、かつ回避優先度が3番目に高い非目的音であってよい。

[0124] また、カテゴリ4は、突発的に発せられるため、瞬時的にしか影響のない非目的音であってよい。カテゴリ4には、例えば、ドアの開扉音、足音、電子レンジが発する音などが含まれる。カテゴリ4は、4つのカテゴリ中、音声認識精度に与える影響が最も低く、かつ回避優先度が最も低い非目的音であってよい。

[0125] このように、本実施形態に係る周辺環境推定部130は、非目的音の特性に応じた回避優先度を設定したノイズマップを生成することが可能である。

[0126] また、本実施形態に係るノイズ源の回避優先度は、非目的音に係る音響的な何らかの定量的な指標に基づいて設定されてもよい。上記の定量的な指標には、例えば、音声らしさの度合いを示す指標や、定常性の度合いを示す指

標が挙げられる。

[0127] 一般に、音声認識の対象となる目的音、すなわちユーザの発話音声は、「非定常的」な「音声」である。一方、「非定常的」な「音声」には、例えば、第三者同士による会話や、テレビジョン装置やラジオが出力する音声などの非目的音も含まれる。このため、「非定常的」な「音声」である目的音に係る音声認識の精度を向上させるためには、当該目的音との分離が困難である「非定常的」な「音声」である非目的音を回避することが重要となる。

[0128] 一方、「非定常的」な「非音声」としては、食器洗いや調理に係る作業音などが、「定常的」な「非音声」には、エアコンディショナー、換気扇、PCのファンなどが発する音が挙げられるが、このような非目的音は、目的音との分離が比較的容易であるため、上記の「非定常的」な「音声」である非目的音と比べ回避優先度は低いといえる。

[0129] また、「定常的」な「音声」としては、例えば、「あーーーー」、など同じ音を長く発話し続ける場合が該当するが、このような音は日常において生じる可能性が著しく低いため、無視してもよい。

[0130] 以上のことから、本実施形態に係る周辺環境推定部130は、音声らしさの度合いを示す指標 α と定常性の度合いを示す指標 β とに基づいて、非目的音が目的音に係る音声認識に与える影響度を算出し、算出した値に基づいて回避優先度を設定してよい。

[0131] 図21および図22は、本実施形態に係る音声らしさの度合いを示す指標 α と定常性の度合いを示す指標 β の算出について説明するための図である。図21には、本実施形態に係る自律移動体10が、複数のマイクロフォンを備える場合における算出の流れが示されている。

[0132] 自律移動体10が複数のマイクロフォンを備える場合、周辺環境推定部130は、図21に示すように、ノイズ源NS1～NS4の方向へ順にビームフォーミングを実行し、指標 α および指標 β を算出する。係る手法によれば、ノイズ領域NR1～NR4に進入することなくノイズ源NR1～NR4に係る指標 α および指標 β を算出することが可能であることから、算出中に目

的音が検出された場合であっても、当該目的音に係る音声認識の精度を維持することができる。

[0133] また、図22は、本実施形態に係る自律移動体10が、単一のマイクロフォンを備える場合における算出の流れが示されている。自律移動体10が単一のマイクロフォンを備える場合、周辺環境推定部130は、図22に示すように、ノイズ源NS1～NS4の周辺において、それぞれ指標 α および指標 β を算出してよい。なお、動作制御部150は、各ノイズ源に係る指標 α および指標 β の算出が完了した場合、速やかにノイズ領域NR1～NR4から脱出するよう自律移動体10を制御してよい。

[0134] 以上、本実施形態に係る指標 α および指標 β の算出の流れについて説明した。本実施形態に係る周辺環境推定部130は、上記のように算出した指標 α および指標 β に基づいて、各ノイズ源の影響度を算出し、当該影響度に基づいて回避優先度を設定することが可能である。周辺環境推定部130は、例えば、指標 α および指標 β の合計値を影響度の高さとし、当該合計値が高い順に回避優先度を高く設定してもよい。

[0135] なお、本実施形態に係る周辺環境推定部130は、音声らしさの度合いを示す指標 α を、例えば、音のスペクトロエントロピーに基づいて算出してもよい。音のスペクトロエントロピーは、VAD (Voice Activity Detection) 技術にも用いられる指標であり、人間の音声は他の音と比べて低い値を示す傾向にある。

[0136] 本実施形態に係る周辺環境推定部130は、音のスペクトロエントロピー、すなわち指標 α を例えば、下記の数式(1)により算出することができる。なお、数式(1)における f は、ある周波数を示し、 S_f は、観測信号の周波数 f の振幅スペクトルを示す。また、数式(1)における P_f は、下記の数式(2)により定義される。

[0137]

[数1]

$$\alpha = -\sum_f P_f \cdot \log P_f \quad \cdot \cdot \cdot (1)$$

$$P_f = \frac{S_f}{\sum_f S_f} \quad \cdot \cdot \cdot (2)$$

[0138] また、本実施形態に係る周辺環境推定部 130 は、定常性の度合いを示す指標 β を、例えば、音の尖度に基づいて算出してもよい。音の尖度とは、音の定常性、非定常性を判別するためにしばしば用いられる指標であり、下記の数式 (3) により算出され得る。なお、下記の数式 (3) における T は、尖度計算を行う音区間長を示し、例えば、3～5 秒などの長さが設定されてよい。また、数式 (3) における t は、ある時刻を示し、 $x(t)$ は、時刻 t における音声波形を示す。

[0139] [数2]

$$\beta = \frac{\sum_{t=1}^T x(t)^4}{(\sum_{t=1}^T x(t)^2)^2} - 3 \quad \cdot \cdot \cdot (3)$$

[0140] 以上、本実施形態に係るノイズ源の回避優先度の設定について説明した。本実施形態に係る回避優先度の設定によれば、自律移動体 10 が、目的音に係る音声認識に影響を与える非目的音を優先的に回避することが可能となる。

[0141] <<2. 5. 動作の流れ>>

次に、本実施形態に係る自律移動体 10 の動作の流れについて詳細に説明する。まず、本実施形態に係るノイズマップの更新の流れについて説明する。図 23 は、本実施形態に係るノイズマップ更新の流れを示すフローチャートである。

[0142] 図 23 を参照すると、まず、周辺環境推定部 130 が、入力部 110 が収集したセンサ情報や認識部 120 による認識結果に基づいて、周辺環境の推定を行う (S1101)。具体的には、周辺環境推定部 130 は、ノイズ源やノイズ領域の推定を行う。

[0143] 次に、周辺環境推定部 130 は、周辺環境保持部 140 に既存のノイズマ

ップが保持されているか否かを判定する（S 1 1 0 2）。

[0144] ここで、周辺環境保持部 1 4 0 に保持されているノイズマップが存在しない場合（S 1 1 0 2 : N O）、周辺環境推定部 1 3 0 は、ステップ S 1 1 0 1 において推定した周辺環境に基づいてノイズマップを生成し、周辺環境保持部 1 4 0 に記憶させる（S 1 1 0 7）。

[0145] 一方、周辺環境保持部 1 4 0 に既存のノイズマップが存在する場合（S 1 1 0 2 : Y E S）、周辺環境推定部 1 3 0 は、続いて、推定した周辺環境と既存のノイズマップとでノイズ源の数が増加しているか否かを判定する（S 1 1 0 3）。

[0146] ここで、ノイズ源の数が増加している場合（S 1 1 0 3 : Y E S）、周辺環境推定部 1 3 0 は、ステップ S 1 1 0 1 において推定した周辺環境に基づいてノイズマップを統合し（S 1 1 0 6）、統合後のノイズマップを周辺環境保持部 1 4 0 に記憶させる（S 1 1 0 7）。

[0147] 一方、ノイズ源の数が増加していない場合（S 1 1 0 3 : N O）、周辺環境推定部 1 3 0 は、続いて、推定した周辺環境と既存のノイズマップとでノイズ源の位置が増加しているか否かを判定する（S 1 1 0 4）。

[0148] ここで、ノイズ源の位置が増加している場合（S 1 1 0 4 : Y E S）、周辺環境推定部 1 3 0 は、ステップ S 1 1 0 1 において推定した周辺環境に基づいてノイズマップを統合し（S 1 1 0 6）、統合後のノイズマップを周辺環境保持部 1 4 0 に記憶させる（S 1 1 0 7）。

[0149] 一方、ノイズ源の位置が増加していない場合（S 1 1 0 4 : N O）、周辺環境推定部 1 3 0 は、続いて、推定した周辺環境と既存のノイズマップとでノイズ源が発する非目的音の音圧が増加しているか否かを判定する（S 1 1 0 5）。

[0150] ここで、ノイズ源が発する非目的音の音圧が増加している場合（S 1 1 0 5 : Y E S）、周辺環境推定部 1 3 0 は、ステップ S 1 1 0 1 において推定した周辺環境に基づいてノイズマップを統合し（S 1 1 0 6）、統合後のノイズマップを周辺環境保持部 1 4 0 に記憶させる（S 1 1 0 7）。

- [0151] 一方、ノイズ源が発する非目的音の音圧が変化していない場合（S 1 1 0 5 : N O）、周辺環境推定部 1 3 0 は、ノイズマップを更新せず、周辺環境保持部 1 4 0 に既存のノイズマップを維持させる。
- [0152] 次に、本実施形態に係る動作制御の流れについて詳細に説明する。図 2 4 は、本実施形態に係る動作制御の流れを示すフローチャートである。
- [0153] 図 2 4 を参照すると、動作制御部 1 5 0 は、まず、周辺環境保持部 1 4 0 が保持するノイズマップの読み込みを行う（S 1 2 0 1）。
- [0154] 続いて、動作制御部 1 5 0 は、ステップ S 1 2 0 1 において読み込んだノイズマップに基づいて、自律移動体 1 0 にノイズ領域を回避した自律行動を行わせる（S 1 2 0 2）。
- [0155] また、動作制御部 1 5 0 は、ステップ S 1 2 0 2 における自律行動中に、目的音が検出されたか否かを継続的に判定する（S 1 2 0 3）。
- [0156] ここで、目的音が検出された場合（S 1 2 0 3 : Y E S）、動作制御部 1 5 0 は、ステップ S 1 2 0 1 において読み込んだノイズマップに基づいて、接近対象の周辺において非目的音の入力レベルがより低下する位置に自律移動体 1 0 を移動させる（S 1 2 0 4）。
- [0157] 次に、動作制御部 1 5 0 は、目的音の音声認識結果に基づいて、対応する動作を自律移動体 1 0 に実行させる（S 1 2 0 5）。
- [0158] 以上、本実施形態に係る自律移動体 1 0 の動作の流れについて説明した。なお、上記では、本実施形態に係る自律移動体 1 0 が S N 比を向上させるために、目的音と非目的音の入力レベルを考慮した移動を行うことを中心に述べた。しかし、本実施形態に係る S N 比の向上手法は係る例に限定されず、例えば、信号処理やビームフォーミング技術が併用して用いられてもよい。
- [0159] 例えば、本実施形態に係る動作制御部 1 5 0 は、自律移動体 1 0 を、接近対象とノイズ源の間に移動させ、接近対象の方向に対しビームフォーミングを張るよう制御してもよい。自律移動体 1 0 が犬型のロボット装置である場合、動作制御部 1 5 0 は、接近対象であるユーザの顔の高さに応じた仰角を以ってビームフォーミングが張られるよう制御してよい。この場合、水平方

向から到達する非目的音を効果的に排除し、S/N比を効果的に向上させる効果が期待される。

[0160] また、本実施形態に係る動作制御部150は、ノイズ領域を回避するために、例えば、ユーザを誘導する動作を自律移動体10に行わせてもよい。動作制御部150は、例えば、接近対象であるユーザがノイズ領域に居る場合、ユーザがノイズ領域から離れ自律移動体10に近づくよう誘導する動作を自律移動体10に行わせることで、ノイズ領域に進入することなく目的音の入力レベルを上昇させることが可能である。上記の誘導は、例えば、吠える、ノイズ領域手前で止まる、うろつくなどの動作により実現され得る。また、自律移動体10が、例えば、人型のロボット装置など、言語によるコミュニケーション機能を有する場合には、ノイズ領域から離れてほしい旨を明示的に音声で伝えてもよい。

[0161] <3. まとめ>

以上説明したように、本開示の一実施形態に係る情報処理装置の一例である自律移動体10は、認識処理に基づいて自律移動体10の動作を制御する動作制御部150を備える。また、本開示の一実施形態に係る動作制御部150は、音声認識処理の対象音声である目的音が検出された場合、当該目的音に基づいて定まる接近対象の周辺において対象音声ではない非目的音の入力レベルがより低下する位置に、自律移動体10を移動させること、を特徴の一つとする。係る構成によれば、自律移動体に音声認識の精度をより向上させる動作を実行させることが可能となる。

[0162] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0163] また、本明細書に記載された効果は、あくまで説明的または例示的なもの

であって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

[0164] また、本明細書の自律移動体 10 の処理に係る各ステップは、必ずしもフローチャートに記載された順序に沿って時系列に処理される必要はない。例えば、自律移動体 10 の処理に係る各ステップは、フローチャートに記載された順序と異なる順序で処理されても、並列的に処理されてもよい。

[0165] なお、以下のような構成も本開示の技術的範囲に属する。

(1)

認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、
を備え、

前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、

情報処理装置。

(2)

前記動作制御部は、前記目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記目的音の入力レベルがより上昇し、前記非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、

前記(1)に記載の情報処理装置。

(3)

前記動作制御部は、前記目的音が検出された場合、前記非目的音を発するノイズ源からより遠ざかり、前記接近対象により近づく位置に、前記自律移動体を移動させる、

前記(1)または(2)に記載の情報処理装置。

(4)

前記動作制御部は、前記目的音が検出された場合、前記接近対象を中心に前記非目的音を発するノイズ源とは反対側に前記自律移動体を移動させる、前記（１）～（３）のいずれかに記載の情報処理装置。

（５）

前記目的音は、ユーザの発話音声であり、

前記接近対象は、前記発話音声を発した発話ユーザである、前記（１）～（４）のいずれかに記載の情報処理装置。

（６）

前記接近対象は、ユーザの発話音声に基づく前記音声認識処理により特定された動体、固定物体、または位置である、前記（１）～（５）のいずれかに記載の情報処理装置。

（７）

前記動作制御部は、周囲環境における前記非目的音の発生状況を示すノイズマップに基づいて、前記自律移動体の動作を制御する、前記（１）～（６）のいずれかに記載の情報処理装置。

（８）

前記ノイズマップは、前記非目的音を発するノイズ源の情報を含み、

前記動作制御部は、前記ノイズ源に係る回避優先度に基づいて、前記自律移動体の動作を制御する、前記（７）に記載の情報処理装置。

（９）

前記ノイズ源に係る回避優先度は、前記ノイズ源の種別に基づいて決定される、前記（８）に記載の情報処理装置。

（１０）

前記ノイズ源に係る回避優先度は、前記ノイズ源が発する前記非目的音の前記音声認識処理に対する影響度に基づいて決定される、前記（８）に記載の情報処理装置。

(11)

前記影響度は、前記非目的音の音声らしさの度合いを示す指標または定常性の度合いを示す指標のうち少なくともいずれかに基づいて算出される、
前記(10)に記載の情報処理装置。

(12)

前記動作制御部は、前記目的音が検出されない場合、前記ノイズマップに基づいて、前記非目的音の入力を回避するよう前記自律移動体の動作を制御する、
前記(7)～(10)のいずれかに記載の情報処理装置。

(13)

前記動作制御部は、前記目的音が検出されない場合、前記ノイズマップに基づいて、前記非目的音の入力レベルが閾値以下となるエリア内に前記自律移動体の移動範囲を制限する、
前記(7)～(12)のいずれかに記載の情報処理装置。

(14)

前記ノイズマップを生成する周辺環境推定部、
をさらに備える、
前記(7)～(13)のいずれかに記載の情報処理装置。

(15)

前記周辺環境推定部は、前記非目的音を発するノイズ源に係る方向推定または音圧測定に基づいて、前記ノイズマップを生成する、
前記(14)に記載の情報処理装置。

(16)

前記周辺環境推定部は、収集された前記非目的音に基づいて、前記ノイズマップを動的に更新する、
前記(14)または(15)に記載の情報処理装置。

(17)

前記周辺環境推定部は、前記非目的音を発するノイズ源の数、位置、また

は音圧の変化に基づいて、前記ノイズマップを動的に更新する、
前記（１６）に記載の情報処理装置。

（１８）

前記周辺環境推定部は、周辺環境にユーザが存在する時間帯において収集された前記非目的音に基づいて、前記ノイズマップの生成または更新を行う、
前記（１６）または（１７）に記載の情報処理装置。

（１９）

プロセッサが、認識処理に基づいて行動を行う自律移動体の動作を制御すること、

を含み、

前記制御することは、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させること、

をさらに含む、

情報処理方法。

（２０）

コンピュータを、

認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、
を備え、

前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、

情報処理装置、

として機能させるためのプログラム。

符号の説明

[0166]	1 0	自律移動体
	1 1 0	入力部
	1 2 0	認識部
	1 3 0	周辺環境推定部
	1 4 0	周辺環境保持部
	1 5 0	動作制御部
	1 6 0	駆動部
	1 7 0	出力部

請求の範囲

- [請求項1] 認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、
- を備え、
- 前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、
- 情報処理装置。
- [請求項2] 前記動作制御部は、前記目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記目的音の入力レベルがより上昇し、前記非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、
- 請求項1に記載の情報処理装置。
- [請求項3] 前記動作制御部は、前記目的音が検出された場合、前記非目的音を発するノイズ源からより遠ざかり、前記接近対象により近づく位置に、前記自律移動体を移動させる、
- 請求項1に記載の情報処理装置。
- [請求項4] 前記動作制御部は、前記目的音が検出された場合、前記接近対象を中心に前記非目的音を発するノイズ源とは反対側に前記自律移動体を移動させる、
- 請求項1に記載の情報処理装置。
- [請求項5] 前記目的音は、ユーザの発話音声であり、
- 前記接近対象は、前記発話音声を発した発話ユーザである、
- 請求項1に記載の情報処理装置。
- [請求項6] 前記接近対象は、ユーザの発話音声に基づく前記音声認識処理により特定された動体、固定物体、または位置である、
- 請求項1に記載の情報処理装置。

- [請求項7] 前記動作制御部は、周囲環境における前記非目的音の発生状況を示すノイズマップに基づいて、前記自律移動体の動作を制御する、請求項1に記載の情報処理装置。
- [請求項8] 前記ノイズマップは、前記非目的音を発するノイズ源の情報を含み、
前記動作制御部は、前記ノイズ源に係る回避優先度に基づいて、前記自律移動体の動作を制御する、請求項7に記載の情報処理装置。
- [請求項9] 前記ノイズ源に係る回避優先度は、前記ノイズ源の種別に基づいて決定される、
請求項8に記載の情報処理装置。
- [請求項10] 前記ノイズ源に係る回避優先度は、前記ノイズ源が発する前記非目的音の前記音声認識処理に対する影響度に基づいて決定される、
請求項8に記載の情報処理装置。
- [請求項11] 前記影響度は、前記非目的音の音声らしさの度合いを示す指標または定常性の度合いを示す指標のうち少なくともいずれかに基づいて算出される、
請求項10に記載の情報処理装置。
- [請求項12] 前記動作制御部は、前記目的音が検出されない場合、前記ノイズマップに基づいて、前記非目的音の入力を回避するよう前記自律移動体の動作を制御する、
請求項7に記載の情報処理装置。
- [請求項13] 前記動作制御部は、前記目的音が検出されない場合、前記ノイズマップに基づいて、前記非目的音の入力レベルが閾値以下となるエリア内に前記自律移動体の移動範囲を制限する、
請求項7に記載の情報処理装置。
- [請求項14] 前記ノイズマップを生成する周辺環境推定部、
をさらに備える、

請求項 7 に記載の情報処理装置。

[請求項15] 前記周辺環境推定部は、前記非目的音を発するノイズ源に係る方向推定または音圧測定に基づいて、前記ノイズマップを生成する、
請求項 1 4 に記載の情報処理装置。

[請求項16] 前記周辺環境推定部は、収集された前記非目的音に基づいて、前記ノイズマップを動的に更新する、
請求項 1 4 に記載の情報処理装置。

[請求項17] 前記周辺環境推定部は、前記非目的音を発するノイズ源の数、位置、または音圧の変化に基づいて、前記ノイズマップを動的に更新する、
請求項 1 6 に記載の情報処理装置。

[請求項18] 前記周辺環境推定部は、周辺環境にユーザが存在する時間帯において収集された前記非目的音に基づいて、前記ノイズマップの生成または更新を行う、
請求項 1 6 に記載の情報処理装置。

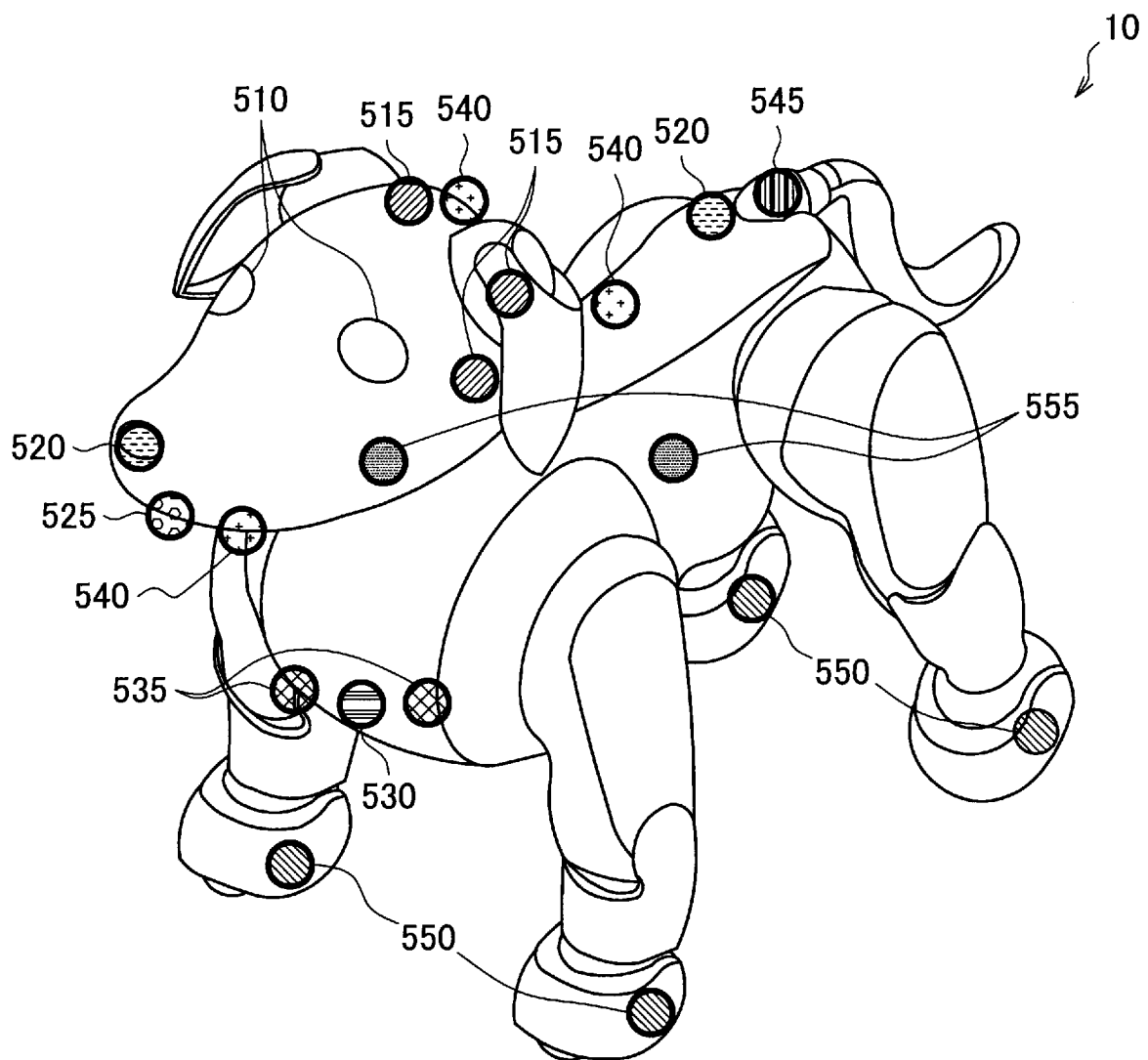
[請求項19] プロセッサが、認識処理に基づいて行動を行う自律移動体の動作を制御すること、
を含み、
前記制御することは、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させること、
をさらに含む、
情報処理方法。

[請求項20] コンピュータを、
認識処理に基づいて行動を行う自律移動体の動作を制御する動作制御部、
を備え、

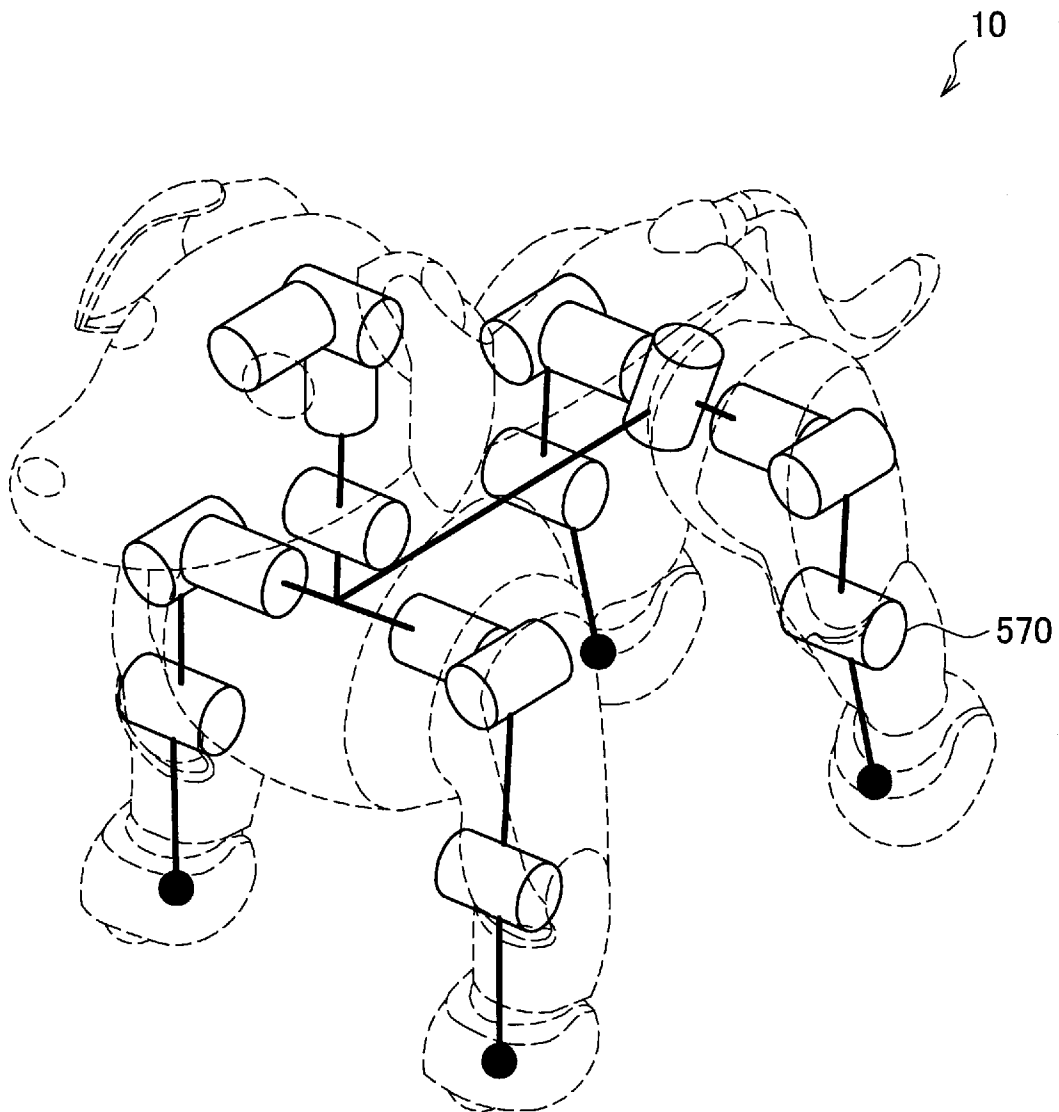
前記動作制御部は、音声認識処理の対象音声である目的音が検出された場合、前記目的音に基づいて定まる接近対象の周辺において前記対象音声ではない非目的音の入力レベルがより低下する位置に、前記自律移動体を移動させる、

情報処理装置、
として機能させるためのプログラム。

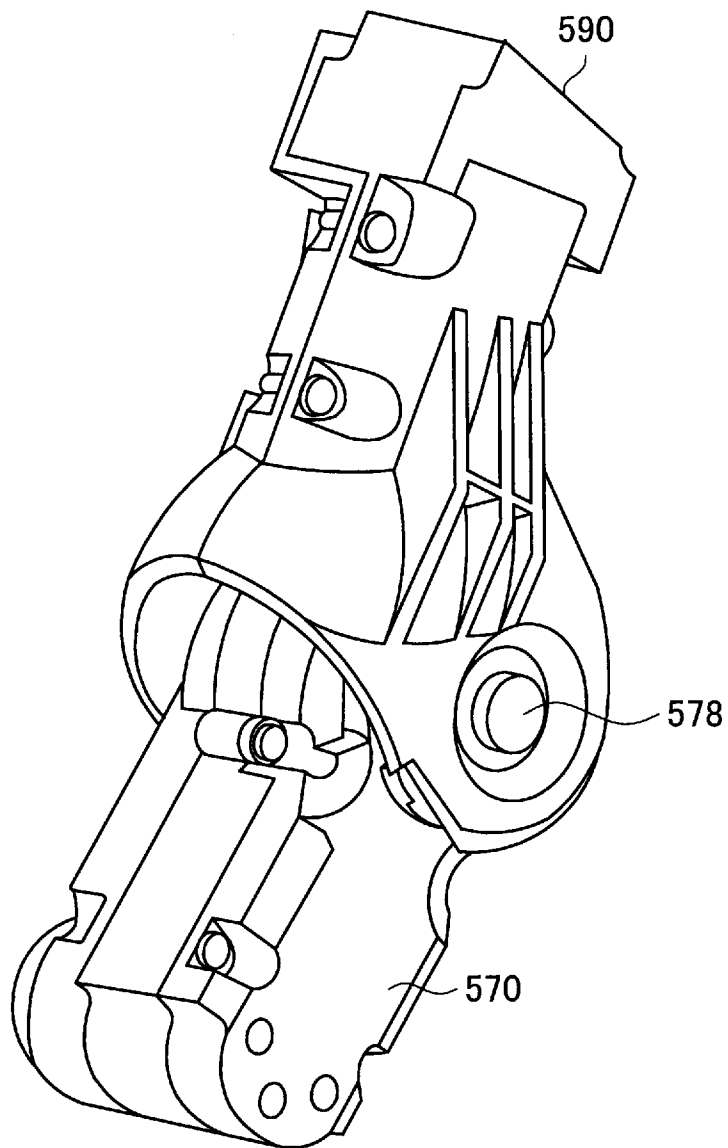
[図1]



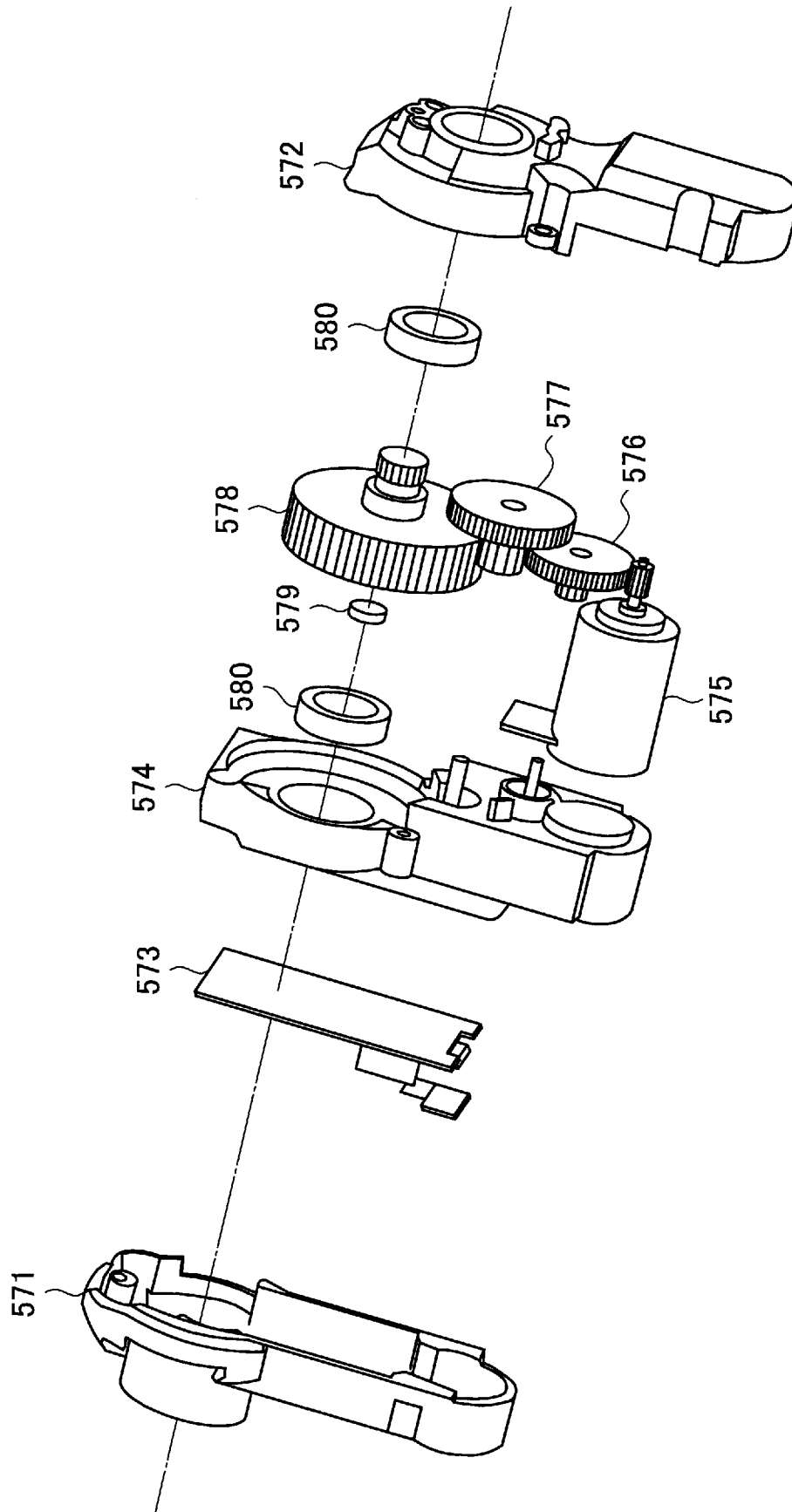
[図2]



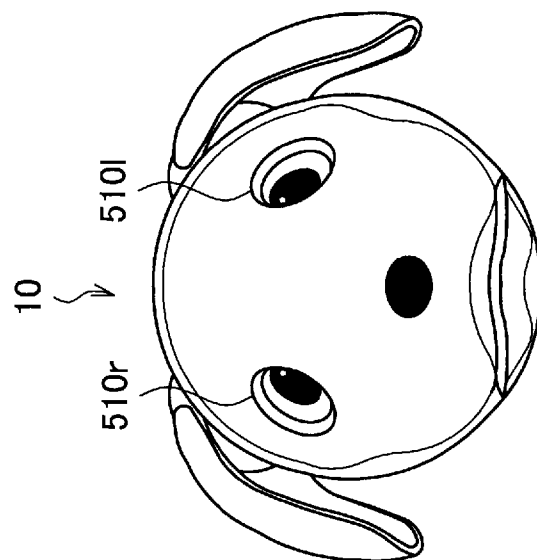
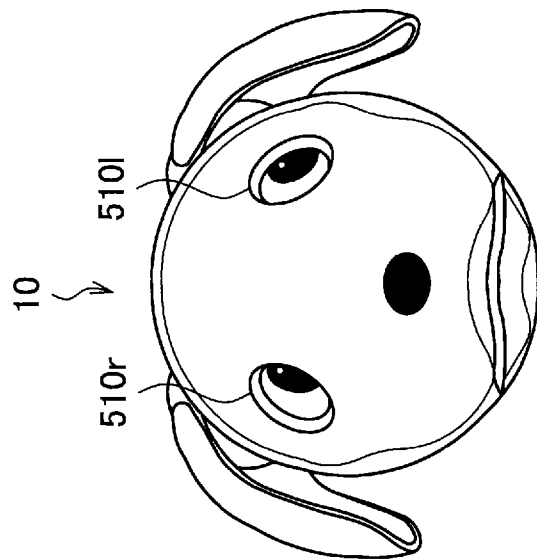
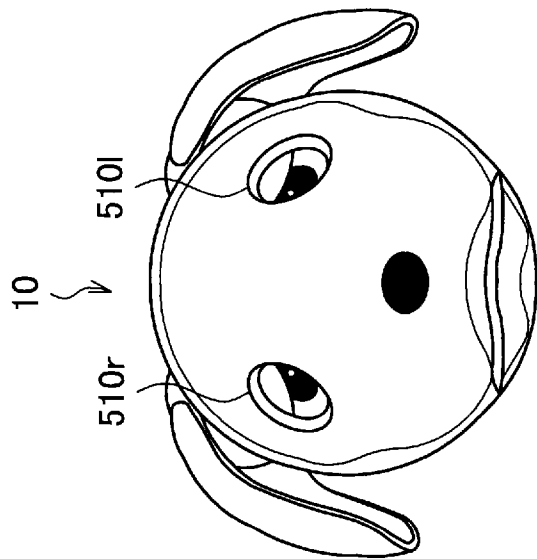
[図3]



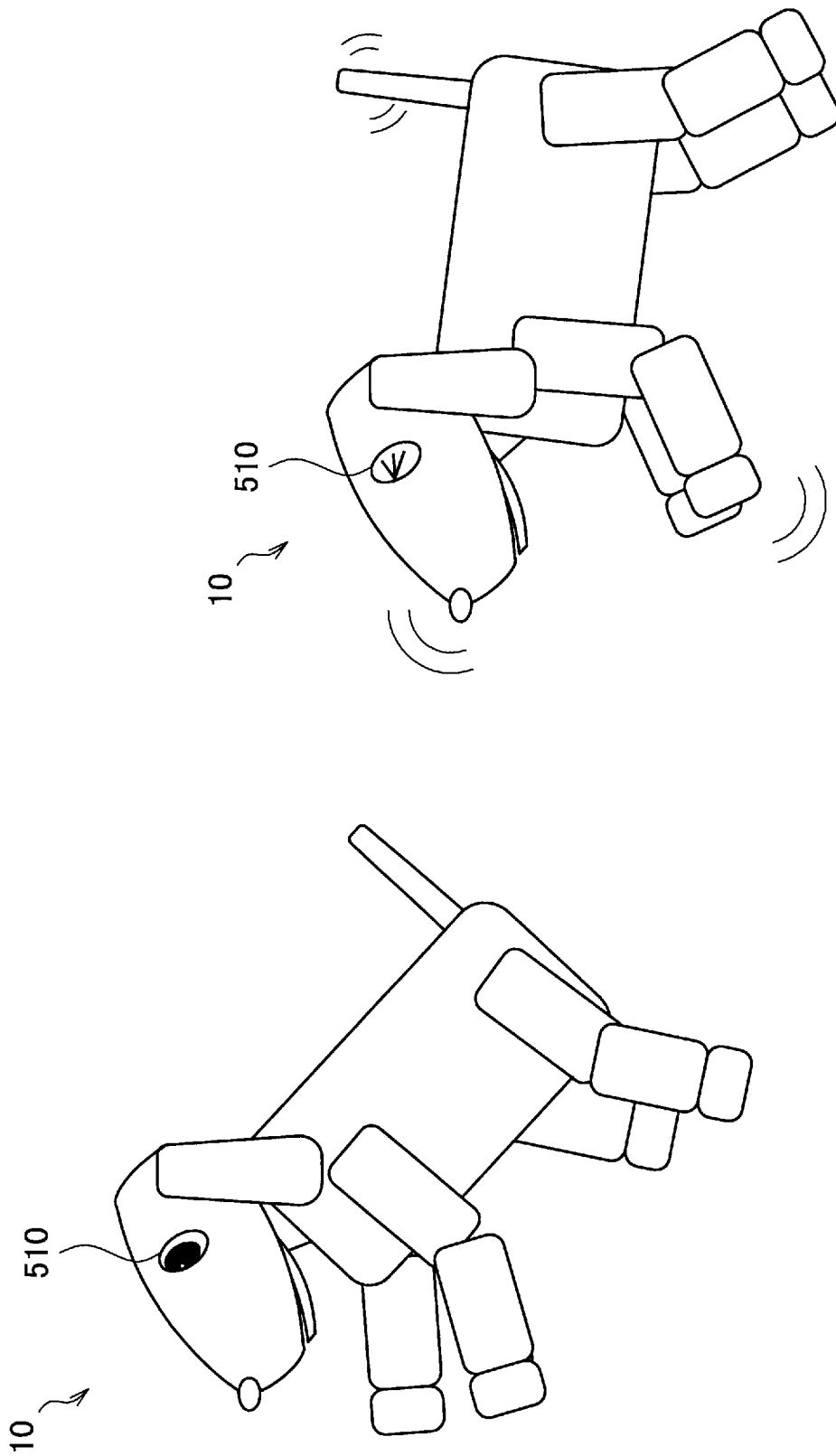
[図4]



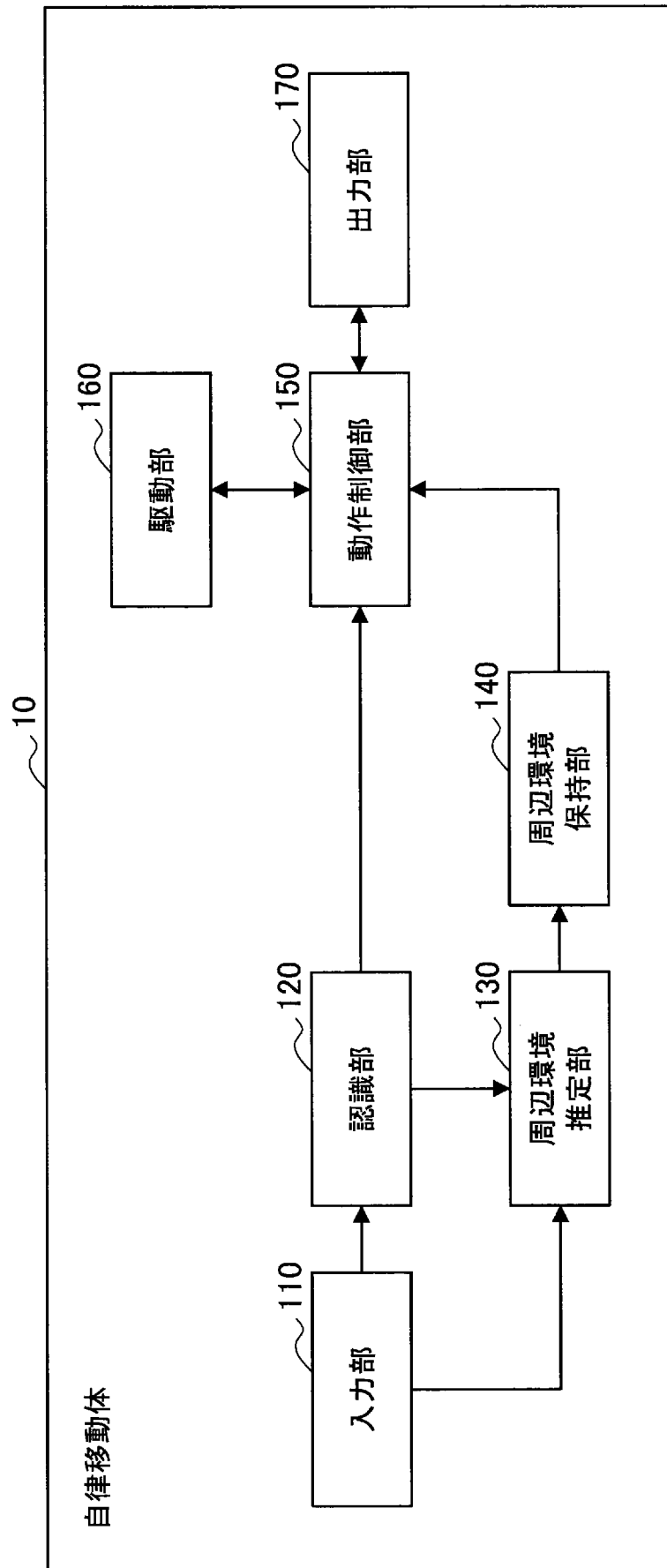
[図5]



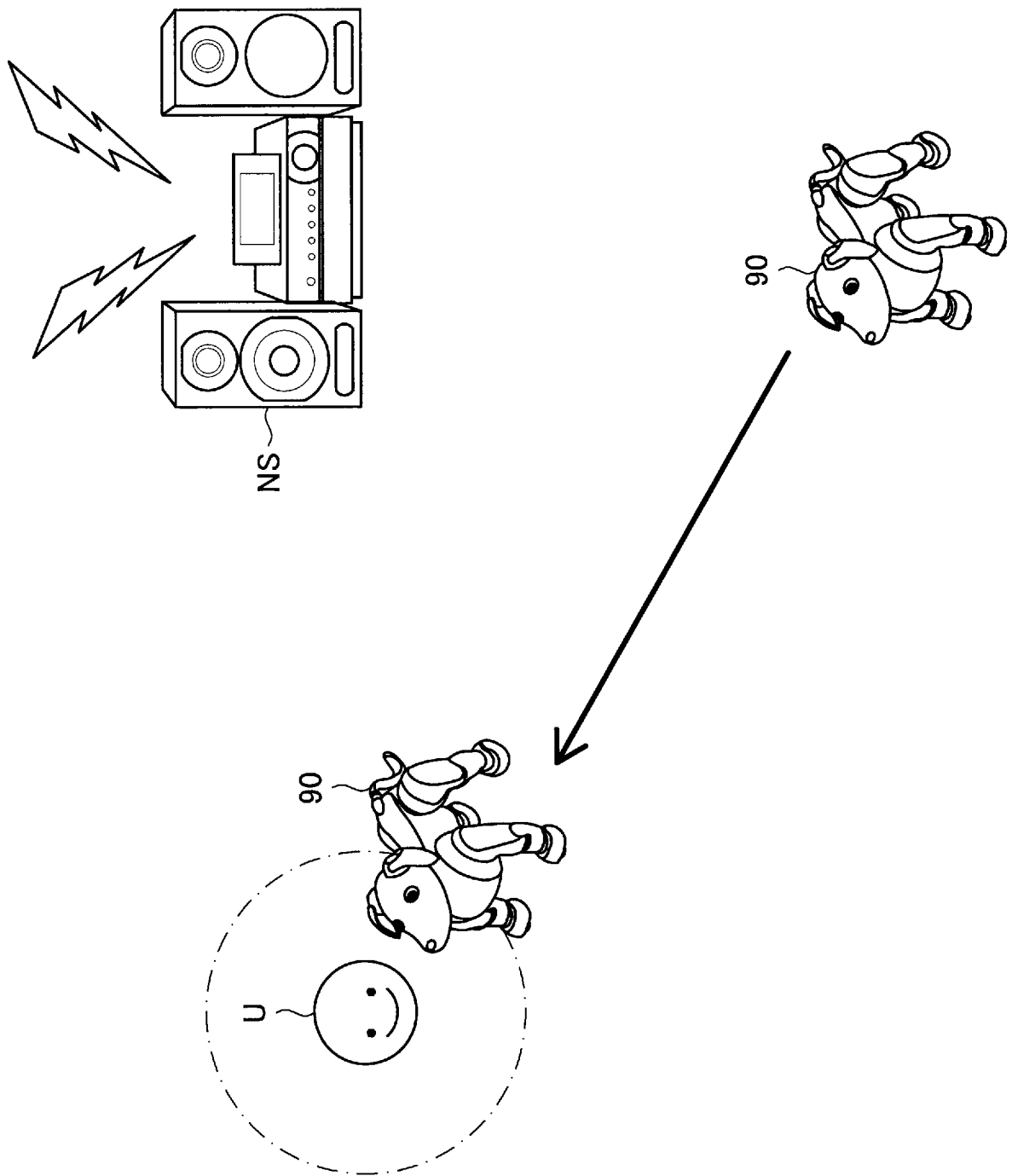
[図6]



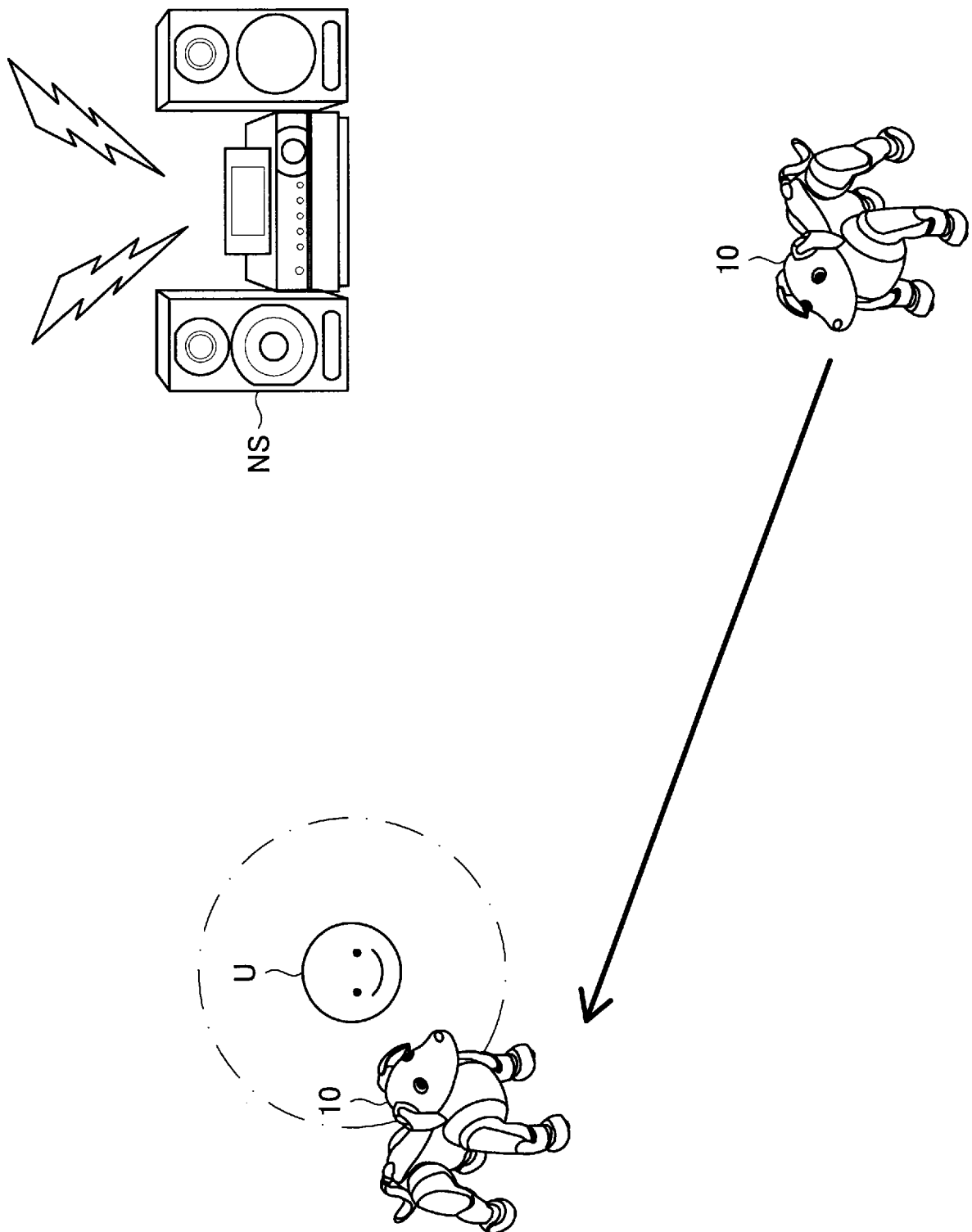
[図7]



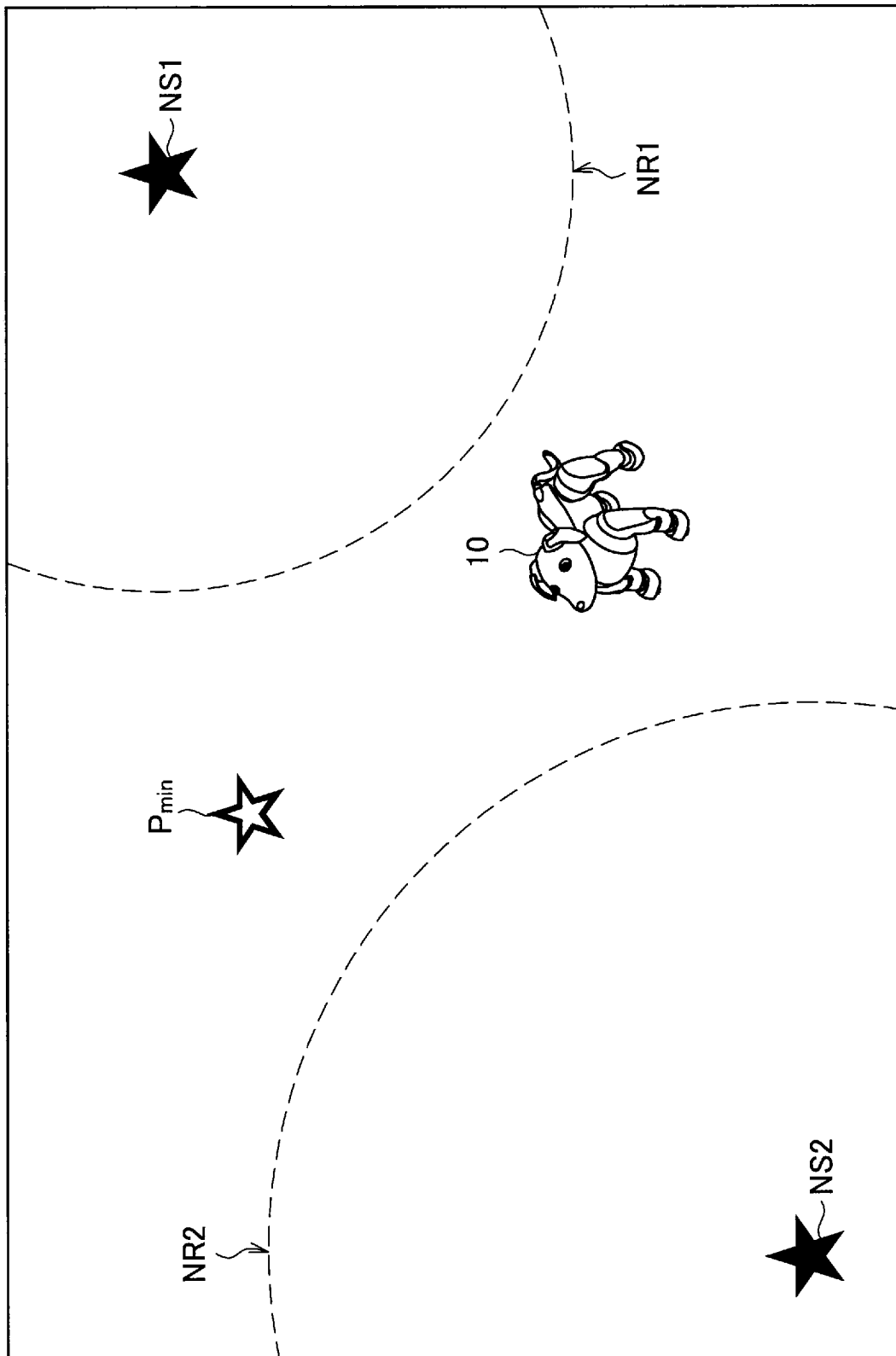
[図8A]



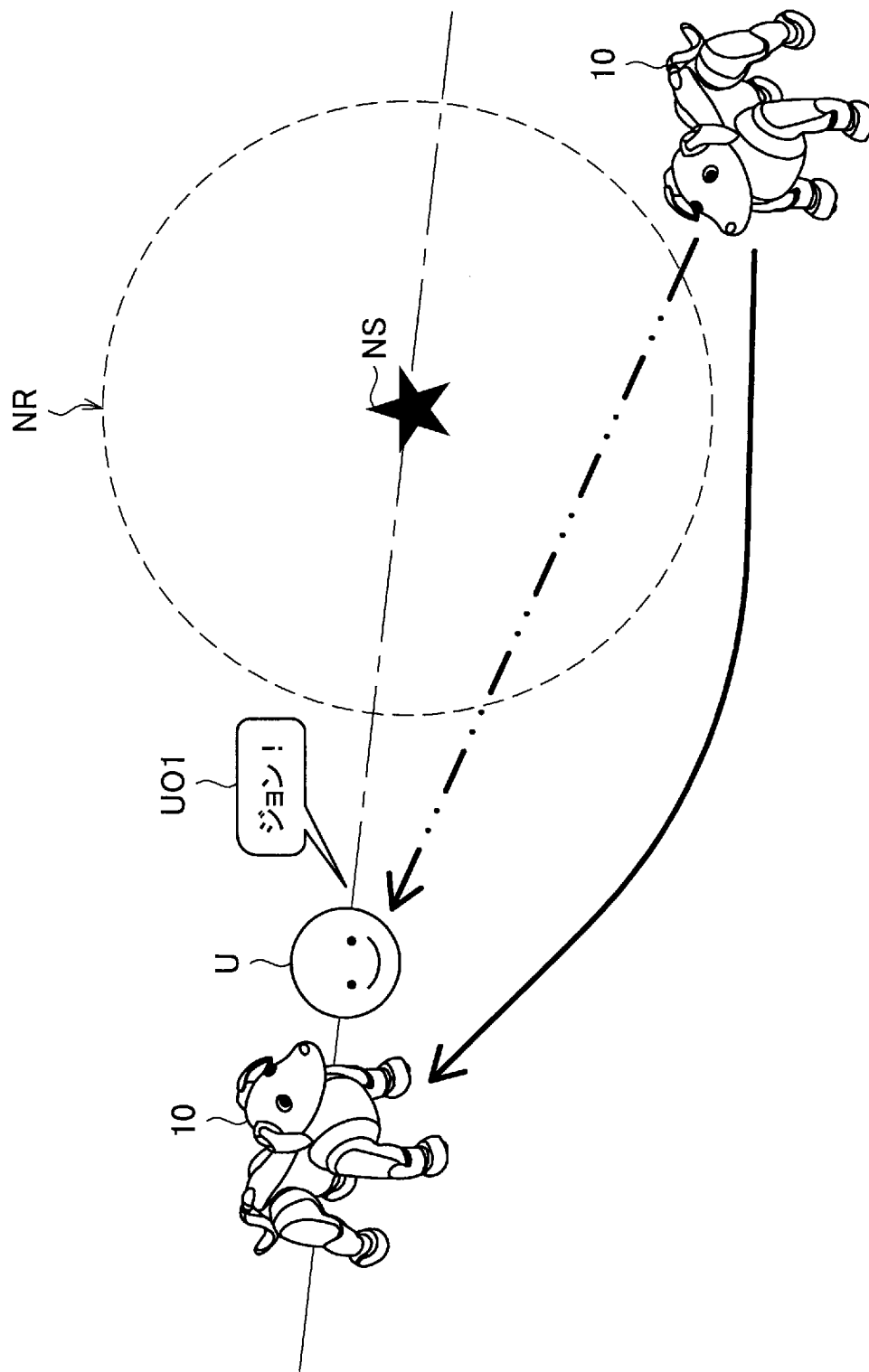
[図8B]



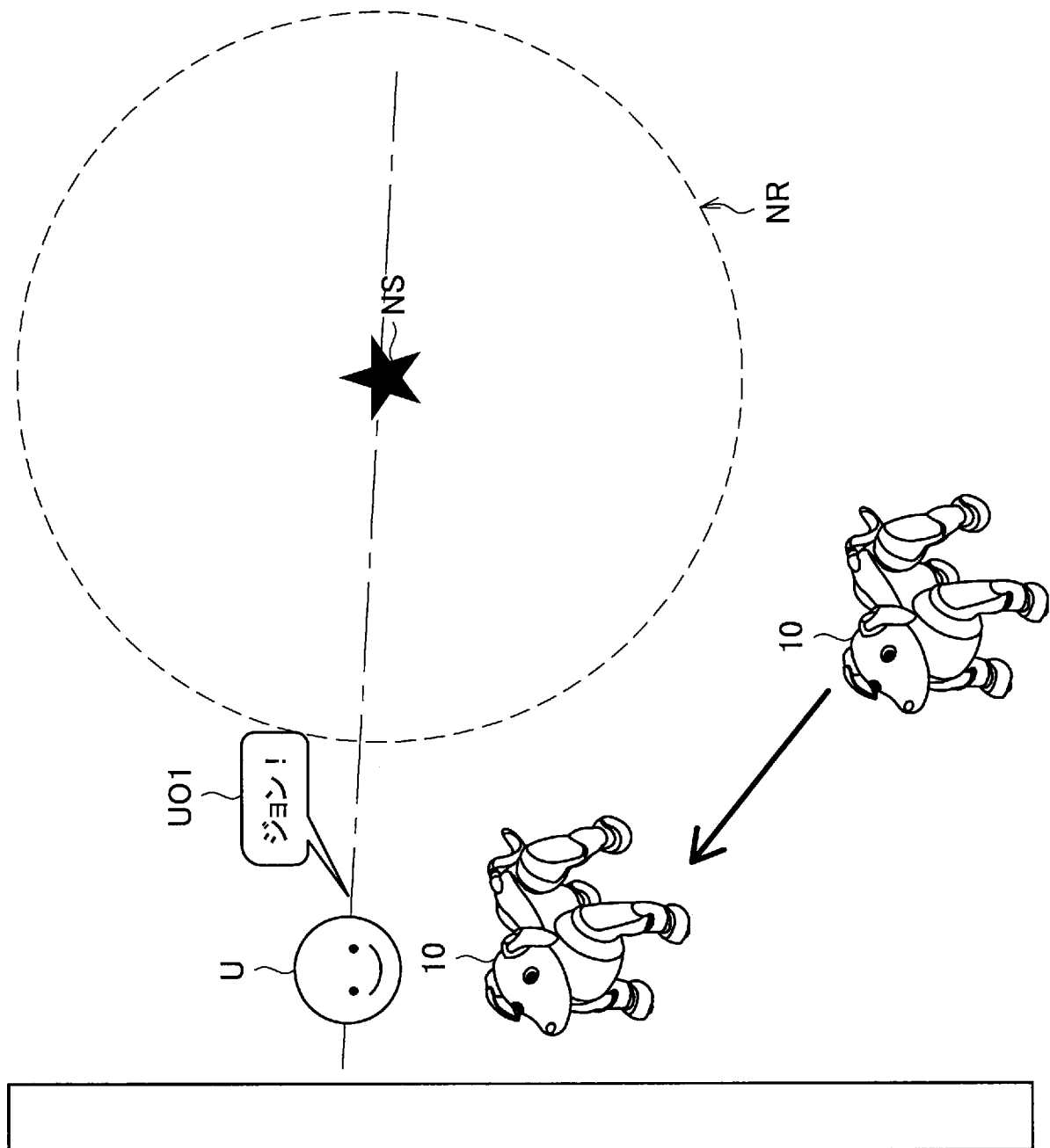
[図9]



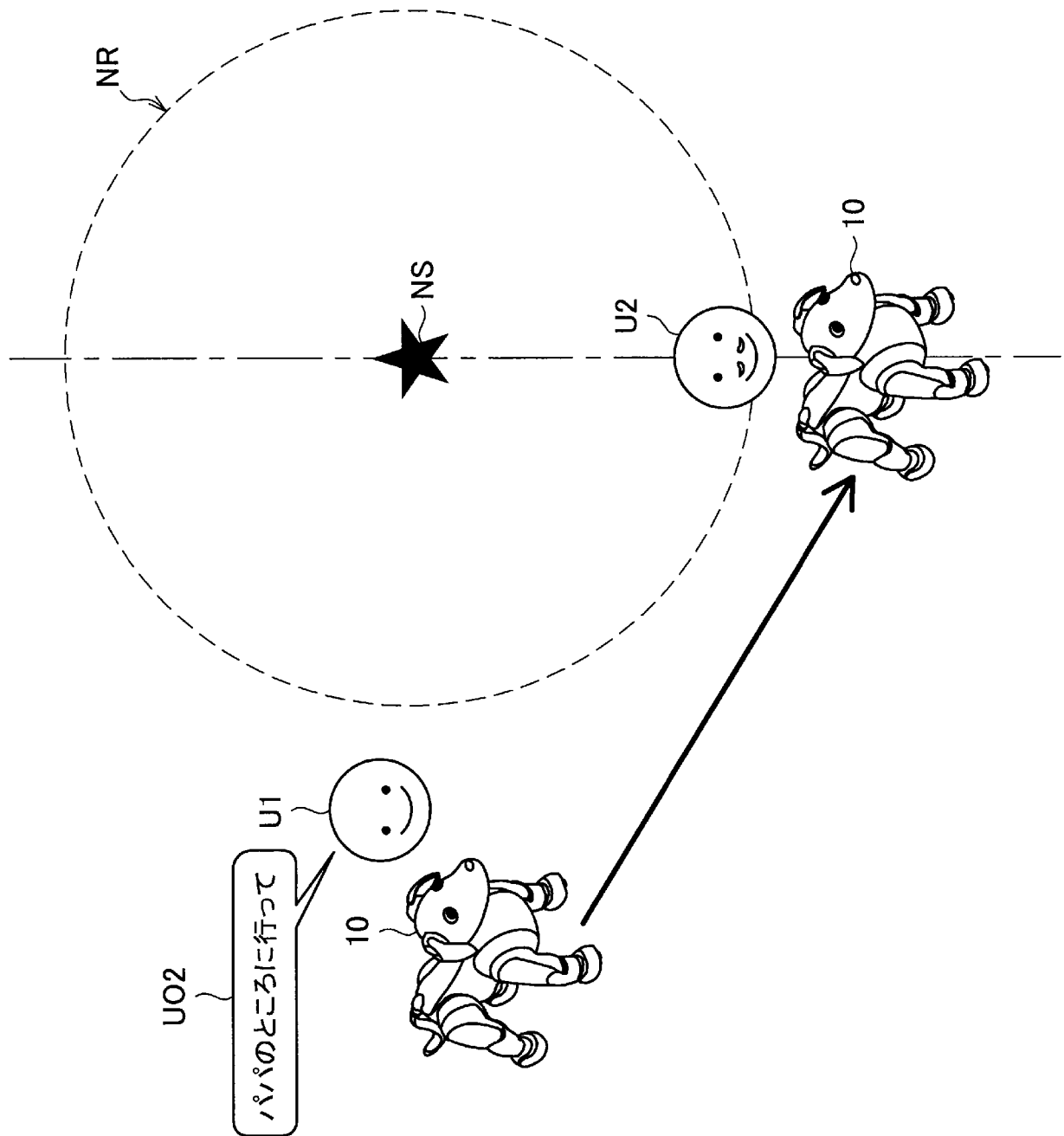
[図10]



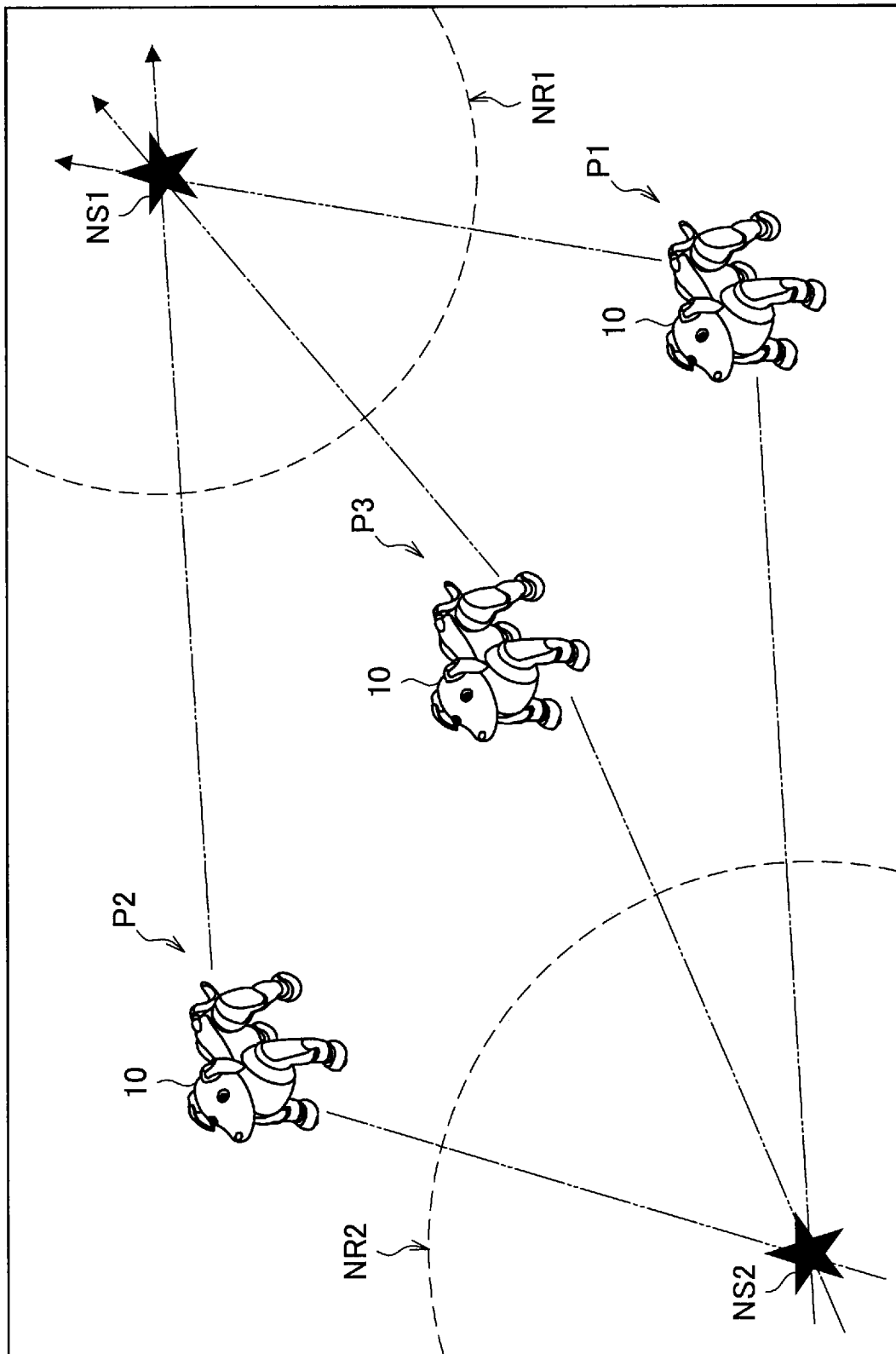
[図11]



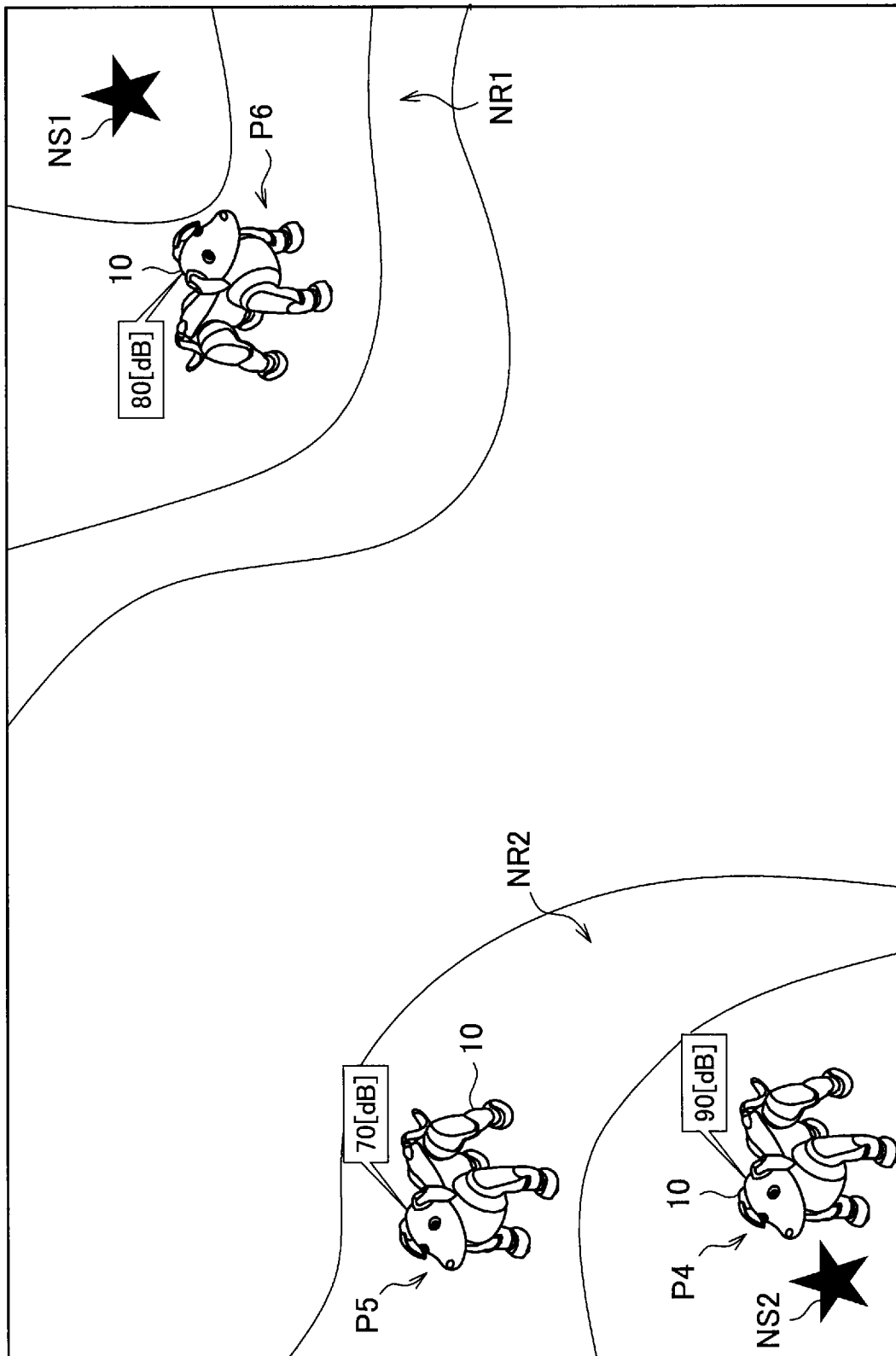
[図12]



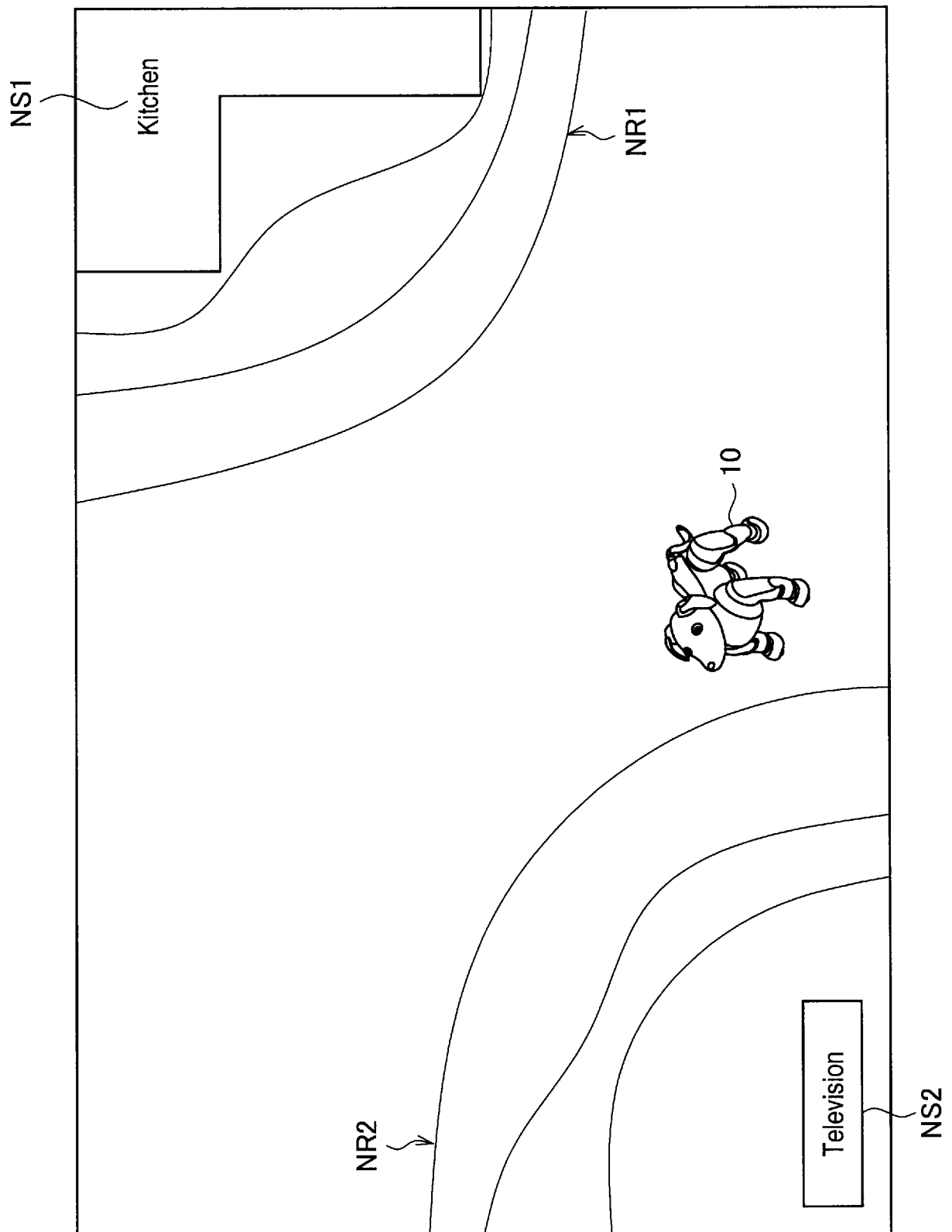
[図13]



[図14]



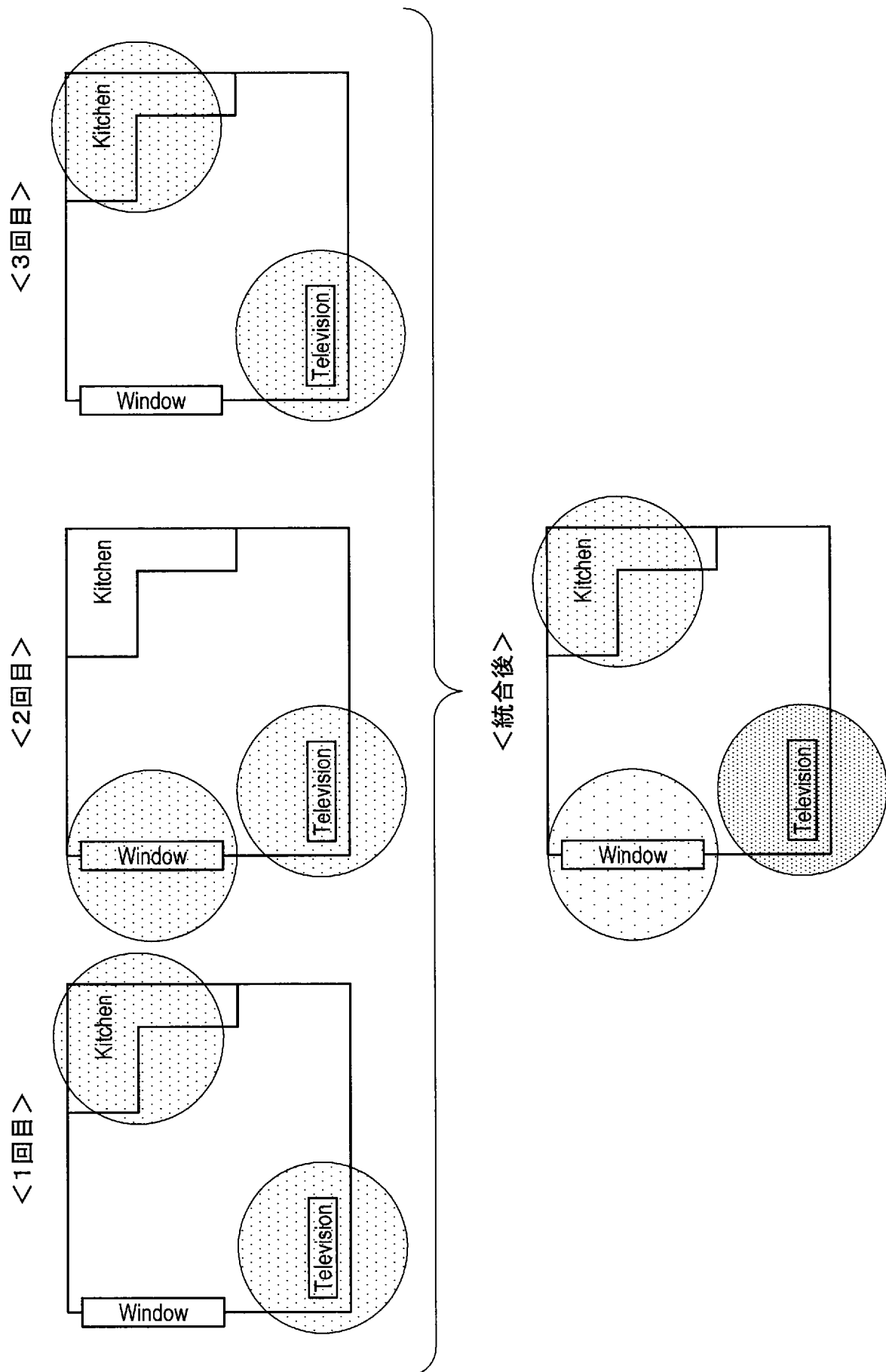
[図15]



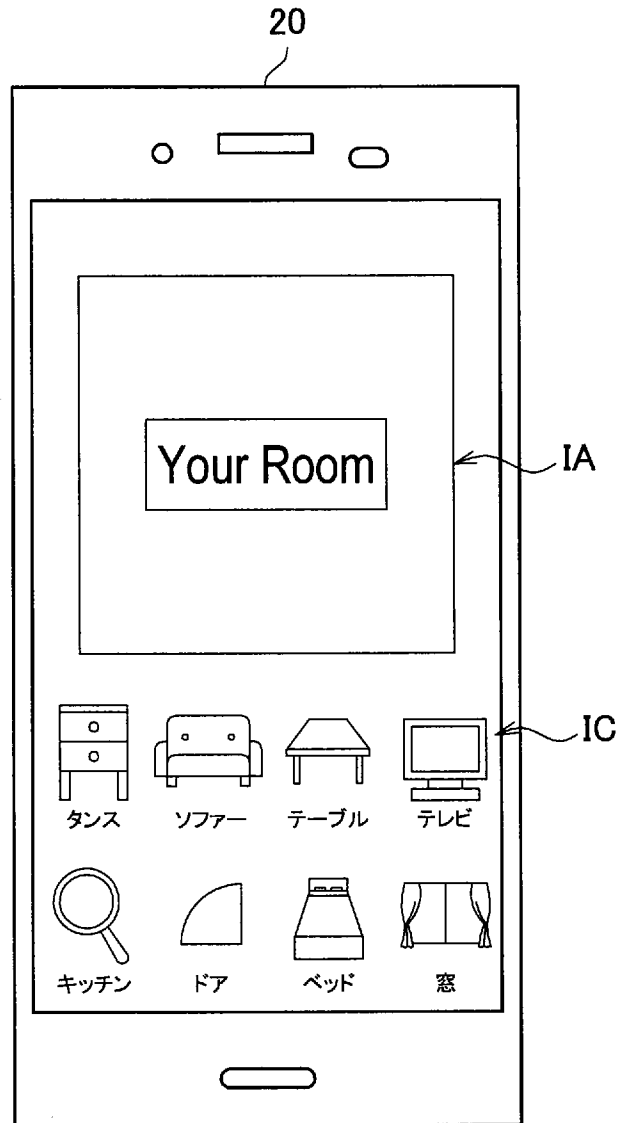
[図16]

予定把握	鍵音検出	ドア開閉音検出	「ただいま」などの 発話検出	人感センサ	...	実施
✓	✓	✓	✓	✓	...	YES
	✓	✓	✓	✓	...	NO
✓		✓	✓	✓	...	NO
✓	✓		✓	✓	...	NO
✓	✓	✓		✓	...	NO
✓	✓	✓	✓		...	NO
		✓	✓	✓	...	NO
✓			✓		...	NO
✓	✓		✓	✓	...	NO
✓	✓			✓	...	NO
✓	✓	✓			...	NO
	✓		✓	✓	...	NO
✓				✓	...	NO
✓	✓			✓	...	NO
	✓	✓			...	NO
	✓	✓	✓	✓	...	NO
✓		✓	✓		...	NO
		✓			...	NO
	✓	✓	✓		...	NO
...	...	...	...	...	...	...

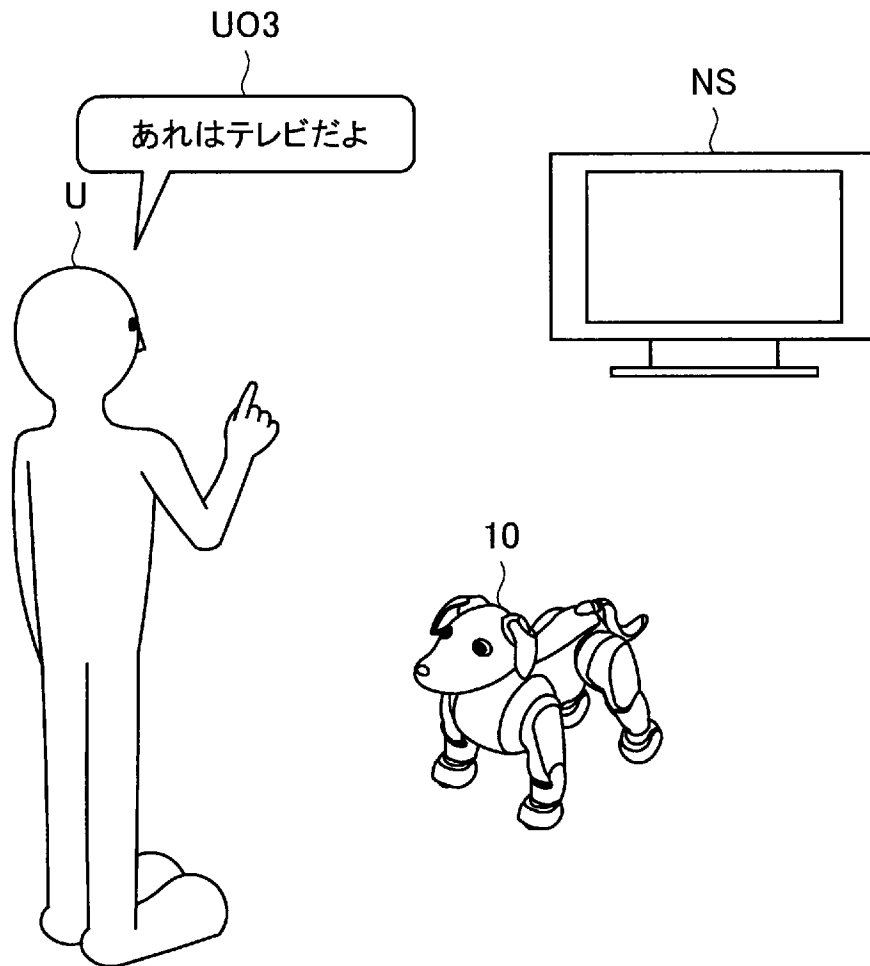
[図17]



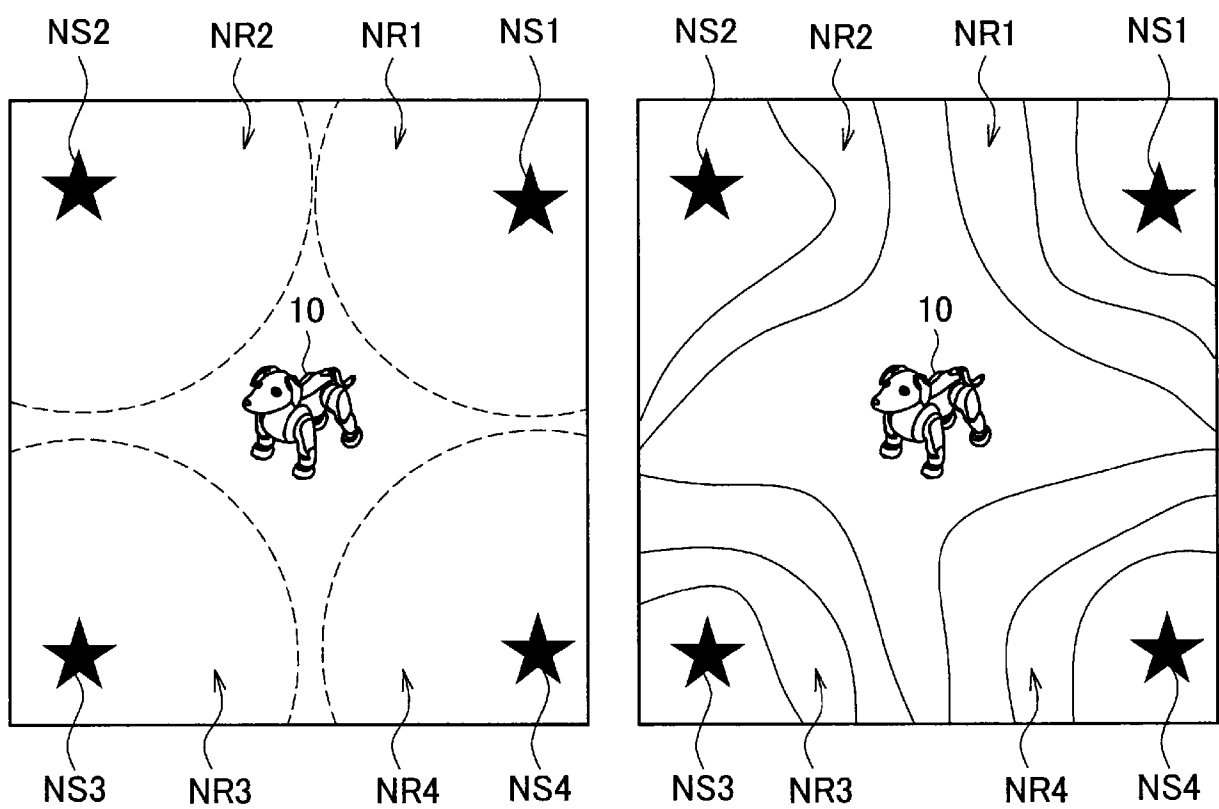
[図18]



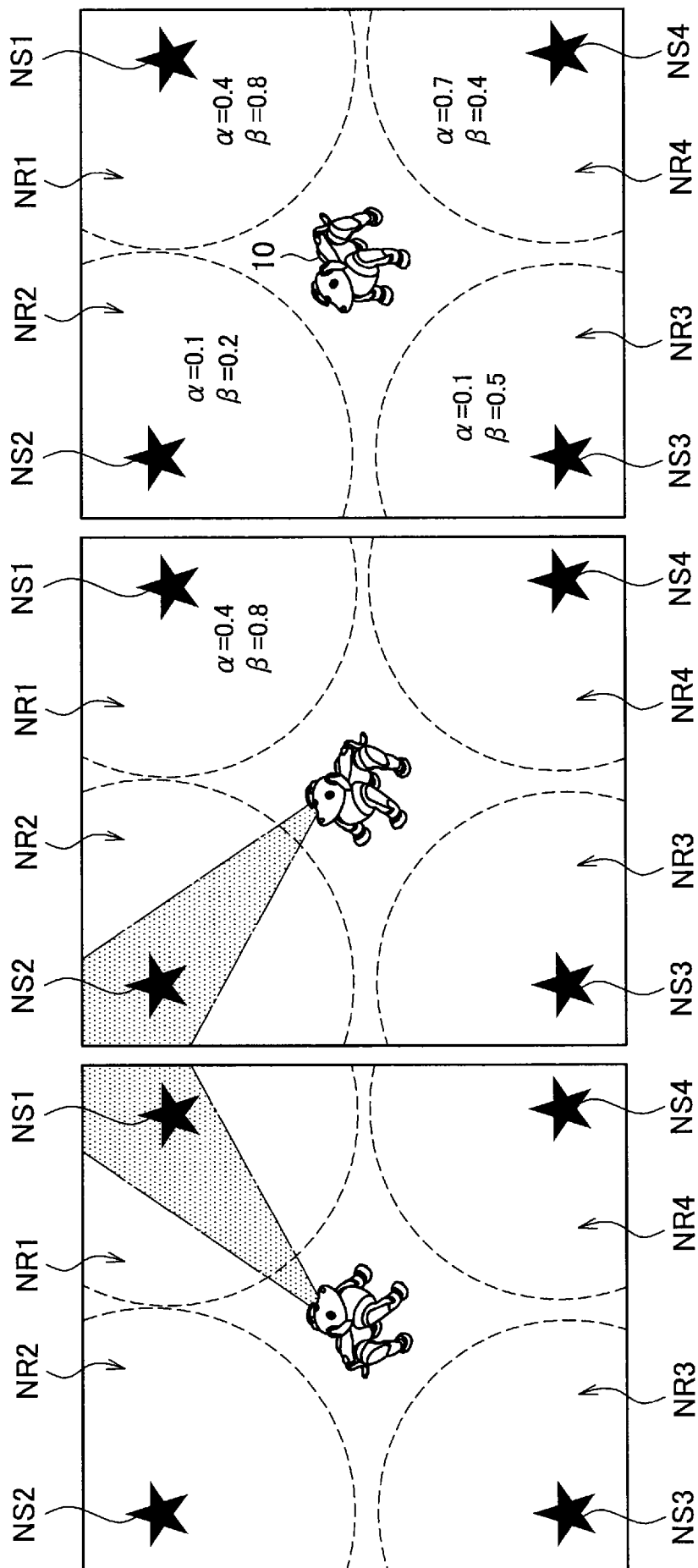
[図19]



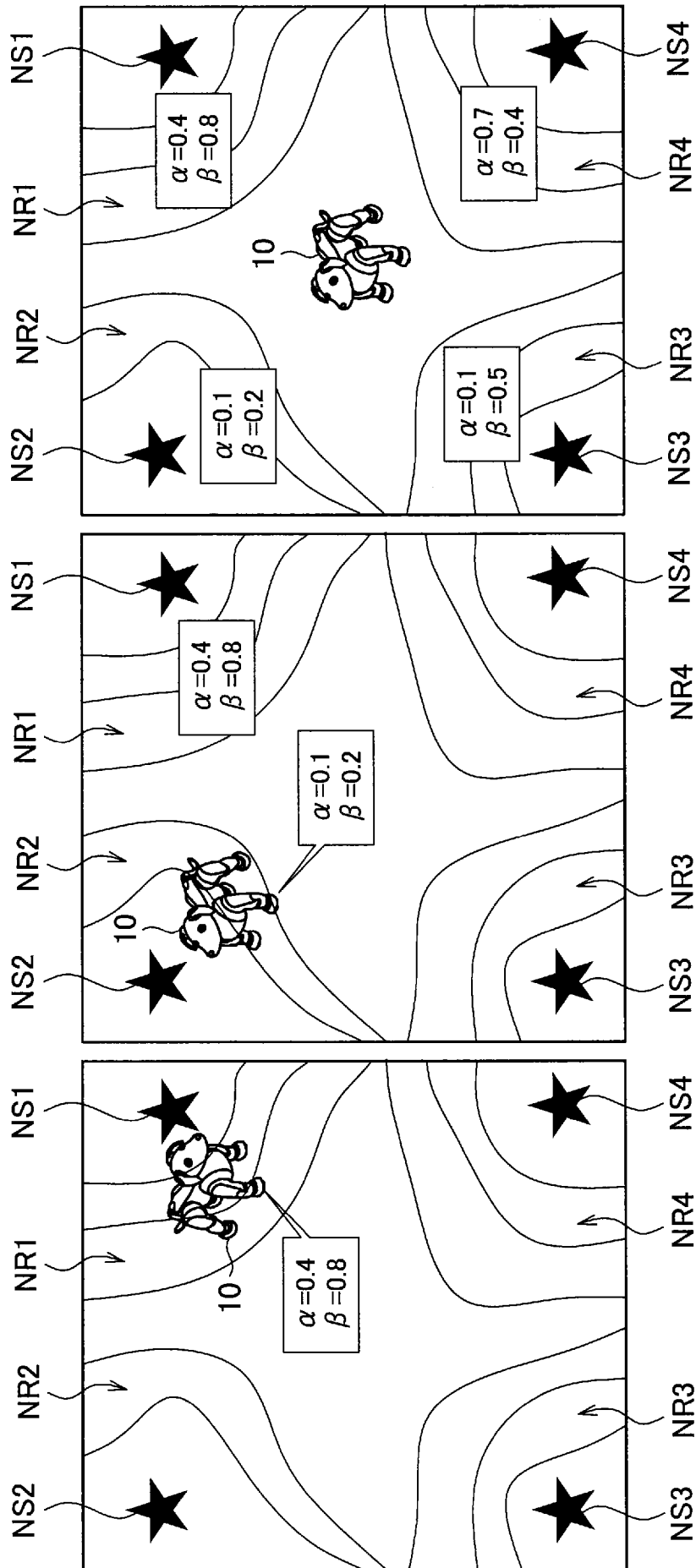
[図20]



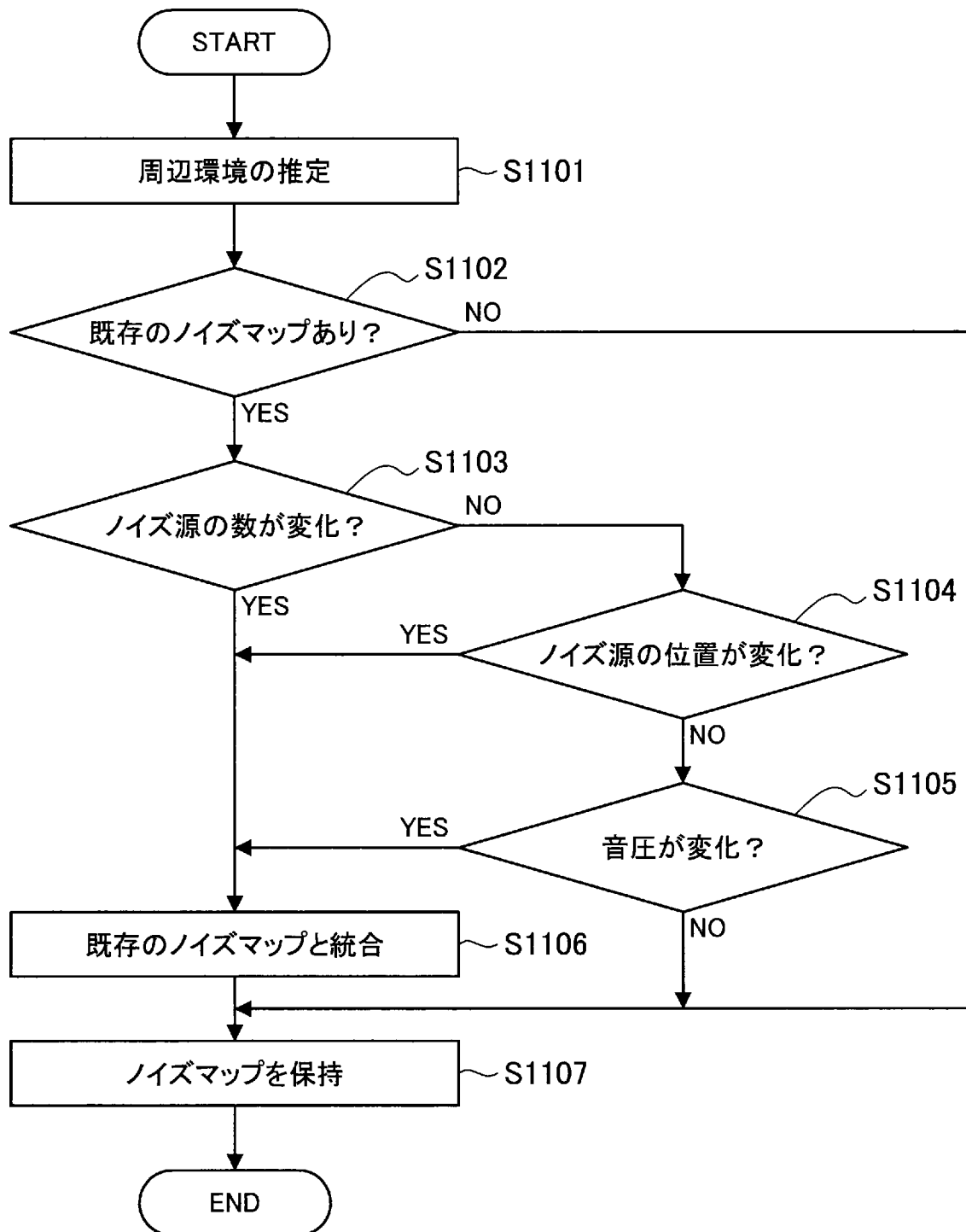
[図21]



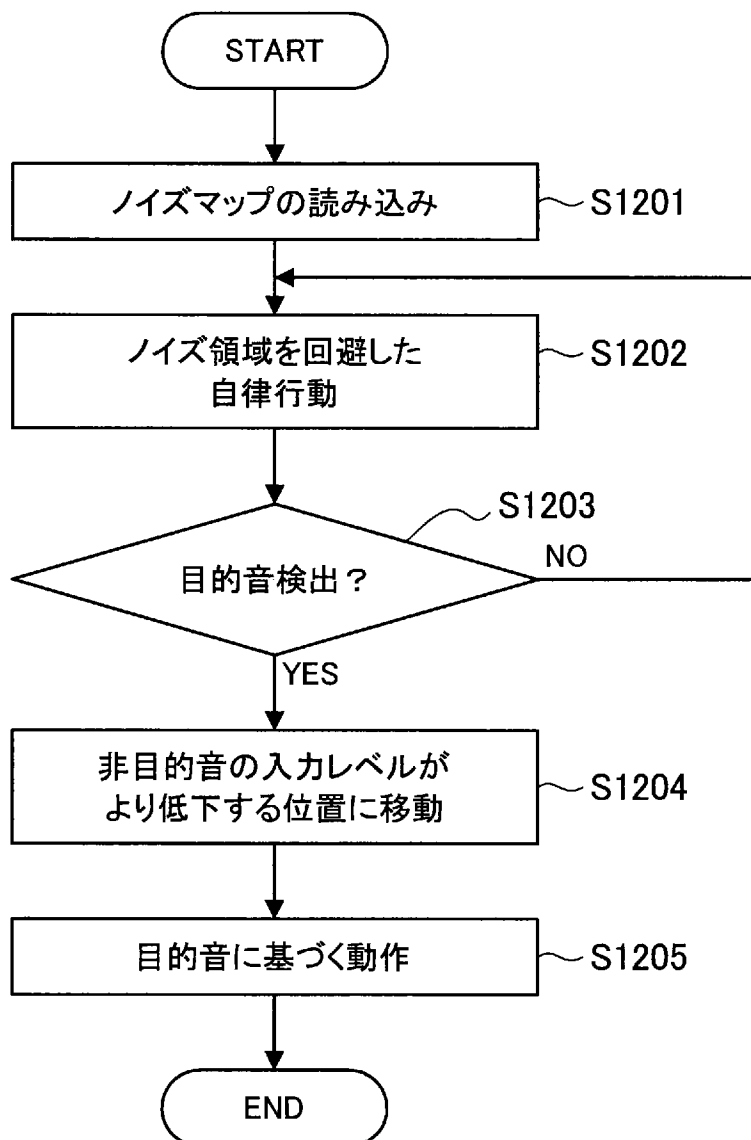
[図22]



[図23]



[図24]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/006580

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. B25J13/00 (2006.01) i, A63H3/33 (2006.01) i, A63H11/00 (2006.01) i, A63H11/18 (2006.01) i, A63H30/02 (2006.01) i, G10L15/00 (2013.01) i, G10L25/84 (2013.01) 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)

Int. Cl. B25J13/00, A63H3/33, A63H11/00, A63H11/18, A63H30/02, G10L15/00, G10L25/84

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2019
Registered utility model specifications of Japan 1996-2019
Published registered utility model applications of Japan 1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2006-181651 A (TOSHIBA CORP.) 13 July 2006, paragraphs [0014]-[0098], fig. 1-15 & US 2006/0143017 A1, paragraphs [0033]-[0117], fig. 1-15	1-20
Y	JP 2002-140092 A (NEC CORP.) 17 May 2002, paragraphs [0008]-[0018], fig. 1-9 (Family: none)	1-20
Y	JP 2007-164379 A (HONDA MOTOR CO., LTD.) 28 June 2007, paragraphs [0033]-[0133], fig. 1-11 & EP 1796079 A1 & US 2007/0135962 A1, paragraphs [0057]-[0295], fig. 1-11	6



Further documents are listed in the continuation of Box C.



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
09.05.2019

Date of mailing of the international search report
21.05.2019

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2019/006580

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2007-264472 A (TOSHIBA CORP.) 11 October 2007, paragraphs [0014]-[0269], fig. 1-13 & US 2007/0233321 A1, paragraphs [0029]-[0300], fig. 1-13	7-18
Y	WO 2017/169826 A1 (GROOVE X, INC.) 05 October 2017, paragraphs [0015]-[0131], fig. 1-10 & GB 2563785 A & DE 112017001573 T5 & CN 109153127 A & US 2018/0333862 A1, paragraphs [0021]-[0199], fig. 1-10	18
A	JP 2005-103679 A (TOSHIBA CORP.) 21 April 2005, paragraphs [0012]-[0065], fig. 1-11 & US 2005/0091684 A1, paragraphs [0024]-[0080], fig. 1-10	1-20
A	JP 2007-152443 A (MITSUBISHI HEAVY INDUSTRIES, LTD.) 21 June 2007, paragraphs [0025]-[0060], fig. 1-10 (Family: none)	1-20
A	US 2012/0130716 A1 (SAMSUNG ELECTRONICS CO., LTD.) 24 May 2012, paragraphs [0035]-[0097], fig. 1-7 & KR 10-2012-0054845 A	1-20

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. B25J13/00(2006.01)i, A63H3/33(2006.01)i, A63H11/00(2006.01)i, A63H11/18(2006.01)i, A63H30/02(2006.01)i, G10L15/00(2013.01)i, G10L25/84(2013.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. B25J13/00, A63H3/33, A63H11/00, A63H11/18, A63H30/02, G10L15/00, G10L25/84

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2006-181651 A (株式会社東芝) 2006.07.13, 段落[0014]-[0098], 図 1-15 & US 2006/0143017 A1, 段落[0033]-[0117], FIGS. 1-15	1-20
Y	JP 2002-140092 A (日本電気株式会社) 2002.05.17, 段落[0008]-[0018], 図 1-9 (ファミリーなし)	1-20

☑ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- | | |
|--|---|
| 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの | 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの |
| 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの | 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの |
| 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) | 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの |
| 「O」 口頭による開示、使用、展示等に言及する文献 | 「&」 同一パテントファミリー文献 |
| 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願 | |

国際調査を完了した日

09.05.2019

国際調査報告の発送日

21.05.2019

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

白井 卓巳

3 U

4 5 5 0

電話番号 03-3581-1101 内線 3364

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2007-164379 A (本田技研工業株式会社) 2007.06.28, 段落[0033]-[0133], 図 1-11 & EP 1796079 A1 & US 2007/0135962 A1, 段落[0057]-[0295], FIGS. 1-11	6
Y	JP 2007-264472 A (株式会社東芝) 2007.10.11, 段落[0014]-[0269], 図 1-13 & US 2007/0233321 A1, 段落[0029]-[0300], FIGS. 1-13	7-18
Y	WO 2017/169826 A1 (GROOVE X株式会社) 2017.10.05, 段落[0015]-[0131], 図 1-10 & GB 2563785 A & DE 112017001573 T5 & CN 109153127 A & US 2018/0333862 A1, 段落[0021]-[0199], FIGS. 1-10	18
A	JP 2005-103679 A (株式会社東芝) 2005.04.21, 段落[0012]-[0065], 図 1-11 & US 2005/0091684 A1, 段落[0024]-[0080], FIGS. 1-10	1-20
A	JP 2007-152443 A (三菱重工業株式会社) 2007.06.21, 段落[0025]-[0060], 図 1-10 (ファミリーなし)	1-20
A	US 2012/0130716 A1 (SAMSUNG ELECTRONICS CO., LTD.) 2012.05.24, 段落[0035]-[0097], FIGS. 1-7 & KR 10-2012-0054845 A	1-20