

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7339124号
(P7339124)

(45)発行日 令和5年9月5日(2023.9.5)

(24)登録日 令和5年8月28日(2023.8.28)

(51)国際特許分類	F I
B 2 5 J 19/06 (2006.01)	B 2 5 J 19/06
B 2 5 J 13/00 (2006.01)	B 2 5 J 13/00 Z
G 1 0 L 15/00 (2013.01)	G 1 0 L 15/00 2 0 0 R
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/10 5 0 0 Z

請求項の数 21 (全26頁)

(21)出願番号	特願2019-197505(P2019-197505)	(73)特許権者	515130201
(22)出願日	令和1年10月30日(2019.10.30)		株式会社 Preferred Networks
(65)公開番号	特開2020-138314(P2020-138314 A)		東京都千代田区大手町1丁目6番1号
(43)公開日	令和2年9月3日(2020.9.3)		大手町ビル
審査請求日	令和4年10月27日(2022.10.27)	(74)代理人	100107766
(31)優先権主張番号	特願2019-32699(P2019-32699)		弁理士 伊東 忠重
(32)優先日	平成31年2月26日(2019.2.26)	(74)代理人	100070150
(33)優先権主張国・地域又は機関	日本国(JP)		弁理士 伊東 忠彦
		(72)発明者	米倉 健太
			東京都千代田区大手町1丁目6番1号
			大手町ビル 株式会社 Preferred Networks内
		(72)発明者	浅井 大史
			東京都千代田区大手町1丁目6番1号
			最終頁に続く

(54)【発明の名称】 制御装置、システム及び制御方法

(57)【特許請求の範囲】

【請求項1】

制御対象の動作中に検出される音データから音声区間を判定する判定部と、
前記音声区間における音データに基づいて算出される話速が第1閾値以上であり、かつ
、前記音声区間の長さが第2閾値以下である場合に、前記制御対象を停止させる制御部と
を有する制御装置。

【請求項2】

前記制御部は、
前記話速が前記第1閾値以上であり、かつ、前記音声区間の長さが前記第2閾値以下で
あり、かつ、前記音声区間における音データの音圧の大きさが第3閾値以上である場合に
、前記制御対象を停止させる、請求項1に記載の制御装置。

10

【請求項3】

前記制御部は、
前記話速が前記第1閾値以上であり、かつ、前記音声区間の長さが前記第2閾値以下で
あり、かつ、前記音声区間における音データに基づいて認識された音素列が予め定められ
た音素列を含む場合に、前記制御対象を停止させる、請求項1に記載の制御装置。

【請求項4】

前記制御部は、
前記話速が前記第1閾値以上であり、かつ、前記音声区間の長さが前記第2閾値以下で
あり、かつ、前記音声区間における音データに基づいて認識された音素列が予め定められ

20

た音素列を含まない場合に、前記制御対象を停止させる、請求項 1 に記載の制御装置。

【請求項 5】

前記音声区間において撮影された画像データより顔領域を検出し、検出した顔領域に基づいて、顔の向きを認識する認識部を更に有し、

前記制御部は、

前記話速が前記第 1 閾値以上であり、かつ、前記音声区間の長さが前記第 2 閾値以下であり、かつ、前記顔の向きが予め設定された方向を向いている場合に、前記制御対象を停止させる、請求項 1 に記載の制御装置。

【請求項 6】

前記認識部は、前記音声区間において撮影された画像データのうち、前記音声区間における音の到来方向に位置する顔領域を検出する、請求項 5 に記載の制御装置。

10

【請求項 7】

制御対象の動作中に検出される音データから音声区間を判定する判定部と、

前記音声区間の長さが第 1 閾値以下であり、かつ、前記音声区間における音データの音圧の大きさが第 2 閾値以上である場合に、前記制御対象を停止させる制御部とを有する制御装置。

【請求項 8】

制御対象の動作中に検出される音データから音声区間を判定する判定部と、

前記音声区間の長さが第 1 閾値以下であり、かつ、前記音声区間における音データに基づいて認識された音素列が予め定められた音素列を含む場合に、前記制御対象を停止させる制御部とを有する制御装置。

20

【請求項 9】

制御対象の動作中に検出される音データから音声区間を判定する判定部と、

前記音声区間の長さが第 1 閾値以下であり、かつ、前記音声区間における音データに基づいて認識された音素列が予め定められた音素列を含まない場合に、前記制御対象を停止させる制御部とを有する制御装置。

【請求項 10】

制御対象の動作中に検出される音データから音声区間を判定する判定部と、

前記音声区間における音データに基づいて算出される話速が第 1 閾値以上であり、かつ、前記音声区間における音素列に繰り返し部分が含まれる場合に、前記制御対象を停止させる制御部とを有する制御装置。

30

【請求項 11】

前記制御部は、前記音素列に含まれる繰り返し部分の第 1 時刻における音データの音圧の大きさが、前記音素列に含まれる繰り返し部分の前記第 1 時刻より後の第 2 時刻における音データの音圧の大きさより小さい場合、前記制御対象を停止させる、請求項 10 に記載の制御装置。

【請求項 12】

前記音声区間における音データの音圧の大きさは、前記音声区間でない非音声区間における音データの音圧の大きさとの差分を用いて表される値である、

請求項 2 又は請求項 7 に記載の制御装置。

40

【請求項 13】

前記音声区間における音データの音圧の大きさは、前記音声区間に含まれる複数の区間それぞれにおいて算出された音圧の大きさの中央値、最大値、又は、平均値のいずれかを用いて表される値である、

請求項 2 又は請求項 7 に記載の制御装置。

【請求項 14】

前記音声区間における音データの音圧の大きさは、前記音声区間における音データの音

50

圧レベルである、

請求項 2、請求項 7、請求項 1 2 又は請求項 1 3 に記載の制御装置。

【請求項 1 5】

前記制御部は、複数種類の停止方法により前記制御対象を停止させることが可能である、
請求項 1 乃至 1 4 のいずれか 1 項に記載の制御装置。

【請求項 1 6】

前記複数種類の停止方法には、前記制御対象への電源供給の少なくとも一部を停止することで、前記制御対象を停止させる停止方法と、前記制御対象への電源供給を維持したまま、前記制御対象を停止させる停止方法とが含まれる、請求項 1 5 に記載の制御装置。

【請求項 1 7】

前記制御部は、前記制御対象の可動部を停止させる、
請求項 1 乃至 1 6 のいずれか 1 項に記載の制御装置。

【請求項 1 8】

前記制御対象はロボットである、
請求項 1 乃至 1 7 のいずれか 1 項に記載の制御装置。

【請求項 1 9】

請求項 1 乃至請求項 1 8 のいずれか 1 項に記載の制御装置と、
前記制御装置によって動作が制御される制御対象と
を有するシステム。

【請求項 2 0】

制御対象の動作中に検出される音データから音声区間を判定する判定工程と、
前記音声区間における音データに基づいて算出される話速が第 1 閾値以上であり、かつ
、前記音声区間の長さが第 2 閾値以下である場合に、前記制御対象を停止させる制御工程と
を有する制御方法。

【請求項 2 1】

制御対象の動作中に検出される音データから音声区間を判定する判定工程と、
前記音声区間の長さが第 1 閾値以下であり、かつ、前記音声区間における音データの音圧
の大きさが第 2 閾値以上である場合に、前記制御対象を停止させる制御工程と
を有する制御方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示は、制御装置、システム及び制御方法に関する。

【背景技術】

【0 0 0 2】

従来より、ユーザの音声指示に基づいて、制御対象（ロボット等）を緊急停止させる制御方法が知られている。例えば、下記特許文献 1 には、"止まれ - "等のように、母音の引きのばしを伴う発声を識別して、制御対象を緊急停止させる制御方法が開示されている。

【0 0 0 3】

一方で、制御対象に対するユーザの発声は、状況に応じて様々である。このため、母音の引きのばしを伴う発声であっても、制御対象を緊急停止させたくないケースや、反対に、母音の引きのばしを伴わない発声であっても、制御対象を緊急停止させたいケースもある。つまり、上記の制御方法では、ユーザの意図したとおりに制御対象を緊急停止させることができないケースが想定される。

【先行技術文献】

【特許文献】

【0 0 0 4】

【文献】特開 2 0 0 8 - 4 9 4 6 2 号公報

【発明の概要】

【発明が解決しようとする課題】

10

20

30

40

50

【 0 0 0 5 】

本開示は、音声指示に基づいて制御対象を緊急停止させる制御装置、システム及び制御方法を提供することを目的としている。

【課題を解決するための手段】

【 0 0 0 6 】

本開示の一態様による制御装置は、例えば、以下のような構成を有する。即ち、
制御対象の動作中に検出される音データから音声区間を判定する判定部と、
前記音声区間における音データに基づいて算出される話速が第1閾値以上であり、かつ、
前記音声区間の長さが第2閾値以下である場合に、前記制御対象を停止させる制御部と
を有する。

10

【図面の簡単な説明】

【 0 0 0 7 】

【図1】ロボットの外觀構成の一例を示す図である。

【図2】ロボットの制御装置のハードウェア構成の一例を示す図である。

【図3】制御装置の緊急停止部の機能構成の一例を示す第1の図である。

【図4】緊急停止部による緊急停止処理の流れを示す第1のフローチャートである。

【図5】制御装置の緊急停止部の機能構成の一例を示す第2の図である。

【図6】緊急停止部による緊急停止処理の流れを示す第2のフローチャートである。

【図7】制御装置の緊急停止部の機能構成の一例を示す第3の図である。

【図8】制御装置の緊急停止部の機能構成の一例を示す第4の図である。

20

【図9】緊急停止部による緊急停止処理の流れを示す第3のフローチャートである。

【図10】緊急停止部による緊急停止処理の流れを示す第4のフローチャートである。

【図11】制御装置の緊急停止部の機能構成の一例を示す第5の図である。

【図12】緊急停止部による緊急停止処理の流れを示す第5のフローチャートである。

【図13】制御装置の緊急停止部の機能構成の一例を示す第6の図である。

【図14】緊急停止部による緊急停止処理の流れを示す第6のフローチャートである。

【図15】制御装置の緊急停止部の機能構成の一例を示す第7の図である。

【図16】緊急停止部による緊急停止処理の流れを示す第7のフローチャートである。

【発明を実施するための形態】

【 0 0 0 8 】

30

以下、各実施形態について添付の図面を参照しながら説明する。なお、以下の各実施形態では、制御対象の動作中に、制御対象を緊急停止させるべき状況が生じた場合（例えば、ユーザに危険が差し迫った場合）、一般に、ユーザは短い言葉（場合によっては意味をなさない言葉）で早口に叫ぶことが多い点に着目している。具体的には、ユーザが発声した音声の音声区間長と、話速とに着目し、音声区間長が短く（所定の閾値以下であり）、話速が速い（所定の閾値以上であった）場合に、制御装置では、制御対象を緊急停止させるべき状況が生じたと判定する。

【 0 0 0 9 】

これにより、以下の各実施形態によれば、制御対象を緊急停止させるべき状況が生じた際にユーザが咄嗟に発声する音声（緊急発声の音声）を捉えて緊急停止させることが可能となり、ユーザの意図したとおりに制御対象を緊急停止させることが可能となる。

40

【 0 0 1 0 】

なお、以下の各実施形態において、“話速”とは、単位時間あたりの音素数であり、“音声区間長”とは、発声の開始から終了までの時間長さ（音声データの開始から終了までの時間長さ）を指す。また、“音声”とは人の声を指し、“音”とは、聴覚に対する任意の刺激であり、音声を含む概念である。

【 0 0 1 1 】

以下、各実施形態の詳細を説明する。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した説明を省略する。

50

【 0 0 1 2 】

[第 1 の実施形態]

< ロボットの外観構成 >

はじめに、本実施形態に係る制御装置の制御対象の一例であるロボットの外観構成について説明する。図 1 は、ロボットの外観構成の一例を示す図である。

【 0 0 1 3 】

図 1 に示すように、ロボット 1 0 は、可動部 4 0 (複数の関節 4 1 ~ 4 4 、エンドエフェクタ 4 5) と、マイクロフォン 2 0 と、カメラ 3 0 とを有する。

【 0 0 1 4 】

可動部 4 0 は、ロボット 1 0 が各種作業 (例えば、物体を把持し、把持した物体を移動先に移動する作業) を行うための動作を実行する。可動部 4 0 に含まれる複数の関節 4 1 ~ 4 4 及びエンドエフェクタ 4 5 は、それぞれ、アクチュエータを備える。

【 0 0 1 5 】

マイクロフォン 2 0 は集音装置の一例であり、空気などを伝播してくる音を検出し、音データを生成する。マイクロフォン 2 0 が検出する音には、ロボット 1 0 の周囲の環境音のほか、ロボット 1 0 のユーザが発声した音声等が含まれる。

【 0 0 1 6 】

なお、図 1 の例では、1つのマイクロフォン 2 0 のみを示したが、ロボット 1 0 は、複数のマイクロフォンを有していてもよい。複数のマイクロフォンを有することで、音の到来方向 (例えば、発声したユーザが位置する方向) を特定することができる。

【 0 0 1 7 】

また、図 1 の例では、マイクロフォン 2 0 がロボット 1 0 に搭載される場合について示したが、マイクロフォン 2 0 は、例えば、ロボット 1 0 が設置された空間の周囲の壁や天井等に配されてもよい。更に、集音装置はマイクロフォンに限定されず、マイクロフォン以外の集音装置がロボット 1 0 等に搭載されてもよい。

【 0 0 1 8 】

カメラ 3 0 は撮像装置の一例であり、ロボット 1 0 の周囲を撮影し、画像データを生成する。なお、図 1 の例では、カメラ 3 0 がロボット 1 0 に搭載される場合について示したが、カメラ 3 0 は、ロボット 1 0 とは別体に配されてもよい。また、撮像装置はカメラに限定されず、カメラ以外の撮像装置がロボット 1 0 等に搭載されてもよい。

【 0 0 1 9 】

また、図 1 に示すように、ロボット 1 0 は、制御装置 1 0 0 を内蔵する。制御装置 1 0 0 は、アクチュエータを制御することで、可動部 4 0 を動作させ、ロボット 1 0 に各種作業を実行させる実行部 1 1 0 を有する。また、制御装置 1 0 0 は、マイクロフォン 2 0 により生成された音データ (及びカメラ 3 0 により生成された画像データ) に基づいて、ロボット 1 0 を緊急停止させる緊急停止部 1 2 0 を有する。

【 0 0 2 0 】

なお、制御装置 1 0 0 は、ロボット 1 0 に内蔵されている必要はなく、ロボット 1 0 の外部に、ロボット 1 0 と別体に設けられてもよい。この場合、ロボット 1 0 と制御装置 1 0 0 とを有するロボット制御システムが形成される。

【 0 0 2 1 】

< 制御装置のハードウェア構成 >

次に、ロボット 1 0 に内蔵される制御装置 1 0 0 のハードウェア構成について説明する。図 2 は、制御装置のハードウェア構成の一例を示す図である。

【 0 0 2 2 】

図 2 は、制御装置のハードウェア構成の一例を示す図である。制御装置 1 0 0 は、プロセッサ 2 0 1 と、主記憶装置 2 0 2 と、補助記憶装置 2 0 3 と、デバイスインタフェース 2 0 4 と、を備え、これらの各構成要素がバス 2 1 0 を介して接続されたコンピュータとして実現される。

【 0 0 2 3 】

10

20

30

40

50

なお、図 2 の例では、制御装置 100 は、各構成要素を 1 個ずつ備えているが、同じ構成要素を複数個ずつ備えていてもよい。また、図 2 の例では、1 台の制御装置が示されているが、複数台の制御装置を配し、ソフトウェア（例えば、制御プログラム）が当該複数台の制御装置にインストールされて、各制御装置がソフトウェアの異なる一部の処理を実行するように構成してもよい。この場合、複数の制御装置それぞれがネットワークインタフェース等を介して、相互に通信してもよい。

【0024】

プロセッサ 201 は、演算装置を含む電子回路（処理回路、Processing circuit、Processing circuitry）である。プロセッサ 201 は、制御装置 100 内の各構成要素などから入力されたデータやプログラムに基づいて演算処理を行い、演算結果や制御信号を各構成要素などに出力する。具体的には、プロセッサ 201 は、OS（Operating System）やアプリケーションなどを実行することにより、制御装置 100 内の各構成要素を制御する。なお、プロセッサ 201 は、上記のような処理を行うことができれば特定の処理回路に限られるものではない。ここで、処理回路とは、1 チップ上に配置された 1 又は複数の電子回路を指してもよいし、2 つ以上のチップあるいはデバイス上に配置された 1 又は複数の電子回路を指してもよい。複数の電子回路を用いる場合、各電子回路は有線または無線により通信してもよい。

【0025】

主記憶装置 202 は、プロセッサ 201 が実行する命令およびデータなどの電子情報を記憶する記憶装置である。主記憶装置 202 に記憶された電子情報はプロセッサ 201 により直接読み出される。補助記憶装置 203 は、主記憶装置 202 以外の記憶装置である。なお、これらの記憶装置は、電子情報を記憶可能な任意の電子部品を意味するものとし、メモリでもストレージでもよい。また、メモリには、揮発性メモリと、不揮発性メモリがあるが、いずれであってもよい。制御装置 100 内において電子情報を保存するためのメモリは、主記憶装置 202 または補助記憶装置 203 により実現されてもよい。

【0026】

デバイスインタフェース 204 は、マイクロフォン 20、カメラ 30、可動部 40 と接続する USB（Universal Serial Bus）などのインタフェースである。

【0027】

< 制御装置の機能構成 >

次に、制御装置 100 の機能構成について説明する。上述したとおり、制御装置 100 は、実行部 110 及び緊急停止部 120 として機能するが、ここでは緊急停止部 120 の機能について説明する。図 3 は、制御装置の緊急停止部の機能構成の一例を示す第 1 の図である。制御装置 100 は、プロセッサ 201 が、主記憶装置 202 または補助記憶装置 203 に格納された制御プログラムの一例である緊急停止プログラムを実行することで緊急停止部 120 として機能する。

【0028】

図 3 に示すように、緊急停止部 120 は、音取得部 301、音声区間検出器 302、音素認識器 303、緊急発声判定器 304、話速算出器 305、制御部 306 を有する。

【0029】

音取得部 301 は、マイクロフォン 20 により生成された音データを取得し、音声区間検出器に出力する。

【0030】

音声区間検出器 302 は判定部の一例であり、音取得部 301 により出力された音データを受け取り、受け取った音データが音声データであるか否かを判定する。また、音声区間検出器 302 は、受け取った音データが音声データであると判定した場合、判定した音声データを音素認識器 303 に出力する。

【0031】

また、音声区間検出器 302 は、受け取った音データが音声データであると判定した場合、音声データの開始時刻を記録する。また、音声区間検出器 302 は、受け取った音デ

10

20

30

40

50

ータが音声データでなくなったと判定した場合、音声データの終了時刻を記録する。これにより、音声区間検出器 302 では、記録した音声データの開始時刻と終了時刻とに基づいて音声区間長を算出し、算出した音声区間長を緊急発声判定器 304 及び話速算出器 305 に出力する。

【0032】

音素認識器 303 は、音声区間検出器 302 により出力された音声データを受け取り、受け取った音声データについて音素認識処理を行うことで、音素列を話速算出器 305 に出力する。

【0033】

話速算出器 305 は、音声区間検出器 302 により出力された音声区間長と、音素認識器 303 により出力された音素列とを受け取り、音声区間長と音素列の長さとに基づいて、単位時間あたりの音素数を算出する。また、話速算出器 305 は、算出結果を話速として、緊急発声判定器 304 に出力する。

【0034】

緊急発声判定器 304 は、音声区間検出器 302 により出力された音声区間長と、話速算出器 305 により出力された話速とに基づいて、下記条件（緊急発声の条件）を満たすか否かを判定する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。

【0035】

緊急発声判定器 304 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 306 に出力する。なお、上記条件のいずれか一方または両方を満たさないと判定した場合、緊急発声判定器 304 は、制御部 306 に対して緊急停止信号を出力しない。

【0036】

制御部 306 は、緊急発声判定器 304 により出力された緊急停止信号を受け取ると、可動部 40 が緊急停止命令に基づいて動作するよう制御し、ロボット 10 を緊急停止させる。

【0037】

< 緊急停止処理の流れ >

次に、緊急停止部 120 による緊急停止処理の流れについて説明する。図 4 は、緊急停止部による緊急停止処理の流れを示す第 1 のフローチャートである。

【0038】

ステップ S101 において、緊急停止部 120 は、ロボット 10 が作業中であるか否かを判定する。ステップ S101 において、ロボット 10 が停止中であると判定した場合には、緊急停止処理を終了する。一方、ステップ S101 において、ロボット 10 が作業中であると判定した場合には、ステップ S102 に進む。

【0039】

ステップ S102 において、音取得部 301 は、マイクロフォン 20 において生成された音データを取得し、主記憶装置 202 に格納する。音取得部 301 は、一定時間（例えば、10ms あるいは 20ms）の音データを取得し、ステップ S103 に進む。

【0040】

ステップ S103 において、音声区間検出器 302 は、音取得部 301 において取得された音データが、音声データであるか否かを判定する。ステップ S103 において、音声データであると判定した場合には、ステップ S104 に進む。

【0041】

ステップ S104 において、音声区間検出器 302 は、音取得部 301 において前回取得された音データが、音声データであったか否かを判定する。ステップ S104 において、音声データであったと判定した場合には、ステップ S101 に戻る。一方、ステップ S104 において、非音声データであったと判定した場合には、ステップ S105 に進む。

10

20

30

40

50

【 0 0 4 2 】

ステップ S 1 0 5 において、音声区間検出器 3 0 2 は、現在時刻、または、現在処理している音声データ（ステップ S 1 0 2 において取得した音データ）が主記憶装置 2 0 2 に格納された時刻を記録し、ステップ S 1 0 1 に戻る。なお、前回取得された音データが非音声データで、今回取得された音データが音声データである場合とは、今回取得された音声データが、音声データの開始点（発声の開始点）であることを示している。つまり、ステップ S 1 0 5 において記録される時刻は、音声データ（発声）の開始時刻に他ならない。

【 0 0 4 3 】

一方、ステップ S 1 0 3 において、非音声データであったと判定した場合には、ステップ S 1 0 6 に進む。ステップ S 1 0 6 において、音声区間検出器 3 0 2 は、音取得部 3 0 1 において前回取得された音データが、音声データであったか否かを判定する。ステップ S 1 0 6 において非音声データであったと判定した場合には、ステップ S 1 0 1 に戻る。

10

【 0 0 4 4 】

一方、ステップ S 1 0 6 において、音声データであったと判定した場合には、ステップ S 1 0 7 に進む。なお、前回取得された音データが音声データで、今回取得された音データが非音声データである場合とは、今回取得された非音声データが、音声データの終了点（発声の終了点）であることを示している。

【 0 0 4 5 】

ステップ S 1 0 7 において、音素認識器 3 0 3 は、音声データ（発声）の開始時刻から、現在時刻（音声データ（発声）の終了時刻）までの間（音声区間）に、主記憶装置 2 0 2 に格納された音声データについて、音素列を認識する。

20

【 0 0 4 6 】

ステップ S 1 0 8 において、音声区間検出器 3 0 2 は、音声データ（発声）の開始時刻から、現在時刻（音声データ（発声）の終了時刻）までの時間長さを音声区間長として算出する。また、話速算出器 3 0 5 は、音素認識器 3 0 3 において認識された音素列の長さを、音声区間検出器 3 0 2 において算出された音声区間長で割ることで、話速を算出する。

【 0 0 4 7 】

ステップ S 1 0 9 において、緊急発声判定器 3 0 4 は、算出された話速及び音声区間長が下記条件を満たすか否かを判定する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。

30

【 0 0 4 8 】

ステップ S 1 0 9 において、緊急発声判定器 3 0 4 が、上記条件を満たさないと判定した場合には、ステップ S 1 0 1 に戻る。一方、ステップ S 1 0 9 において、緊急発声判定器 3 0 4 が、上記条件を満たすと判定した場合には、ステップ S 1 1 0 に進む。

【 0 0 4 9 】

ステップ S 1 1 0 において、制御部 3 0 6 は、可動部 4 0 が緊急停止命令に基づいて動作するように制御し、ロボット 1 0 を緊急停止させる。

【 0 0 5 0 】

< まとめ >

40

以上の説明から明らかなように、第 1 の実施形態に係る制御装置の緊急停止部は、作業中のロボットにおいて緊急停止させるべき状況が生じると、一般に、ユーザは短い言葉で早口に叫ぶことが多い点に着目し、

- ・話速が予め定められた閾値以上であり、かつ、
- ・音声区間長が予め定められた閾値以下である、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【 0 0 5 1 】

これにより、第 1 の実施形態に係る制御装置の緊急停止部によれば、作業中のロボットにおいて緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。この結果、ロボットをユーザの意図したとおり

50

に動作させることが可能となる。

【 0 0 5 2 】

[第 2 の実施形態]

上記第 1 の実施形態では、緊急停止させるべき状況が生じた場合にユーザが咄嗟に発声する音声の音声区間長と話速とに着目した。一方で、緊急停止させるべき状況が生じた場合、ユーザは、比較的大きな声で叫ぶことが想定される。そこで、第 2 の実施形態では、緊急停止させるべき状況が生じたか否かを判定するにあたり、音声区間長と話速とに加えて、音圧レベルに着目する。以下、第 2 の実施形態について、上記第 1 の実施形態との相違点を中心に説明する。

【 0 0 5 3 】

< 制御装置の機能構成 >

はじめに、第 2 の実施形態に係る制御装置 1 0 0 の緊急停止部の機能構成について説明する。図 5 は、制御装置の緊急停止部の機能構成の一例を示す第 2 の図である。図 3 に示した機能構成との相違点は、第 2 の実施形態に係る制御装置 1 0 0 の緊急停止部 5 0 0 の場合、音圧レベル算出器 5 0 1 を有する点、及び、緊急発声判定器 5 0 2 が緊急発声判定器 3 0 4 とは異なる機能を有する点である。

【 0 0 5 4 】

音圧レベル算出器 5 0 1 は、音声区間検出器 3 0 2 により出力された音声データを受け取り、受け取った音声データの音圧レベルを算出する。また、音圧レベル算出器 5 0 1 は、算出した音圧レベルを、緊急発声判定器 3 0 4 に出力する。

【 0 0 5 5 】

緊急発声判定器 5 0 2 は、音声区間検出器 3 0 2 により出力された音声区間長と、話速算出器 3 0 5 により出力された話速と、音圧レベル算出器 5 0 1 により出力された音圧レベルとに基づいて、下記条件を満たすか否かを判定する。

- ・ 話速が予め定められた閾値以上である。
- ・ 音声区間長が予め定められた閾値以下である。
- ・ 音圧レベルが予め定められた閾値以上である。

【 0 0 5 6 】

緊急発声判定器 5 0 2 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 5 0 2 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 0 5 7 】

このように、音声区間長及び話速に加えて、音圧レベルを条件にすることで、緊急停止部 5 0 0 によれば、緊急停止させるべき状況が生じたか否かを判定する際の判定精度を向上させることができる。

【 0 0 5 8 】

< 緊急停止処理の流れ >

次に、緊急停止部 5 0 0 による緊急停止処理の流れについて説明する。図 6 は、緊急停止部による緊急停止処理の流れを示す第 2 のフローチャートである。図 4 で示したフローチャートとの相違点は、ステップ S 2 1 1 である。

【 0 0 5 9 】

ステップ S 2 1 1 において、音圧レベル算出器 5 0 1 は、音声データ（発声）の開始時刻から、現在時刻（音声データ（発声）の終了時刻）までの間（音声区間）に、主記憶装置 2 0 2 に格納された音声データの音圧レベルを算出する。

【 0 0 6 0 】

なお、音圧レベル算出器 5 0 1 が算出する音圧レベルは、音声データ（発声）の開始時刻から、現在時刻（音声データ（発声）の終了時刻）までの間（音声区間）に、主記憶装置 2 0 2 に格納された音声データの音圧レベルの平均値を算出する。ただし、音圧レベル算出器 5 0 1 が算出する音圧レベルは、音声区間における音圧レベルの平均値に限定され

10

20

30

40

50

ない。

【 0 0 6 1 】

例えば、音声区間を一定区間ごとに分割し、それぞれの区間において算出された音圧レベルのうち、最大値または中央値を、音圧レベルとして算出してもよい。

【 0 0 6 2 】

あるいは、音圧レベル算出器 5 0 1 は、音声区間における音圧レベルの平均値（あるいは最大値または中央値）と、音声区間の直前の非音声区間における音データの音圧レベルの平均値（あるいは最大値または中央値）との差分を算出してもよい。ロボット 1 0 の周囲の環境音が大きく、うるさい場合と、ロボット 1 0 の周囲の環境音が小さく、静かな場合とで、緊急停止させるべき状況が生じた際にユーザが発声する音声の大きさも変わってくるからである。

【 0 0 6 3 】

< まとめ >

以上の説明から明らかなように、第 2 の実施形態に係る制御装置の緊急停止部は、作業中のロボットにおいて緊急停止させるべき状況が生じると、一般に、ユーザは短い言葉で早口に大声で叫ぶことが多い点に着目し、

- ・ 話速が予め定められた閾値以上であり、かつ、
- ・ 音声区間長が予め定められた閾値以下であり、かつ、
- ・ 音圧レベルが予め定められた閾値以上である、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【 0 0 6 4 】

これにより、第 2 の実施形態に係る制御装置の緊急停止部によれば、作業中のロボットにおいて緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。この結果、ロボットをユーザの意図したとおりに動作させることが可能となる。

【 0 0 6 5 】

[第 3 の実施形態]

上記第 1 の実施形態では、緊急停止させるべき状況が生じた場合にユーザが咄嗟に発声する音声の音声区間長と話速とに着目した。一方で、緊急停止させるべき状況が生じた場合に、ユーザが発声する可能性の高い音声内容は限られている。また、緊急停止させるべき状況が生じた場合に、ユーザが発声する可能性の低い音声内容もある。そこで、第 3 の実施形態では、緊急停止させるべき状況が生じたか否かを判定するにあたり、音声区間長と話速とに加えて、認識された音素列に着目する。以下、第 3 の実施形態について、上記第 1 の実施形態との相違点を中心に説明する。

【 0 0 6 6 】

< 制御装置の機能構成 >

はじめに、第 3 の実施形態に係る制御装置 1 0 0 の緊急停止部の機能構成について説明する。図 7 は、制御装置の緊急停止部の機能構成の一例を示す第 3 の図である。図 3 に示した機能構成との相違点は、緊急発声判定器 7 0 2 が緊急発声判定器 3 0 4 とは異なる機能を有する点である。

【 0 0 6 7 】

緊急発声判定器 7 0 2 は、音声区間検出器 3 0 2 により出力された音声区間長と、話速算出器 3 0 5 により出力された話速と、音素認識器 3 0 3 により出力された音素列とに基づいて、下記条件を満たすか否かを判定する。ただし、音素列格納部 7 0 1 には、緊急停止させるべき状況が生じた場合に、ユーザが発声する可能性の高い音声内容の音素列の集合が格納されているものとする。

- ・ 話速が予め定められた閾値以上である。
- ・ 音声区間長が予め定められた閾値以下である。
- ・ 音素列が、音素列格納部 7 0 1 に格納された、予め定められた音素列の集合に含まれる音素列を含む。

10

20

30

40

50

【 0 0 6 8 】

緊急発声判定器 7 0 2 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 7 0 2 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 0 6 9 】

なお、音素列格納部 7 0 1 には、緊急停止させるべき状況が生じた場合に、ユーザが発声する可能性が低い音声内容の音素列の集合を格納してもよい。ただし、その場合、緊急発声判定器 7 0 2 は、下記条件を満たすか否かを判定する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。
- ・音素列が、音素列格納部 7 0 1 に格納された、予め定められた音素列の集合に含まれる音素列を含まない。

【 0 0 7 0 】

緊急発声判定器 7 0 2 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 7 0 2 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 0 7 1 】

< まとめ >

以上の説明から明らかなように、第 3 の実施形態に係る制御装置の緊急停止部は、作業中のロボット 1 0 において緊急停止させるべき状況が生じると、一般に、ユーザは特定の短い言葉で早口に叫ぶことが多い点に着目し、

- ・話速が予め定められた閾値以上であり、かつ、
 - ・音声区間長が予め定められた閾値以下であり、かつ、
 - ・音素列が、予め定められた音素列を含む（あるいは含まない）、
- 場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【 0 0 7 2 】

これにより、第 3 の実施形態に係る制御装置の緊急停止部によれば、作業中のロボットにおいて緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。また、緊急停止させるべき状況が生じていないにも関わらず、緊急停止させてしまうといった事態を回避することができる。この結果、ロボットをユーザの意図したとおりに動作させることが可能となる。

【 0 0 7 3 】

[第 4 の実施形態]

上記第 1 の実施形態では、緊急停止させるべき状況が生じた場合にユーザが咄嗟に発声する音声の音声区間長と話速とに着目した。一方で、緊急停止させるべき状況が生じた場合、ユーザは、ロボットの方向に向かって発声する可能性が高い。そこで、第 4 の実施形態では、緊急停止させるべき状況が生じたか否かを判定するにあたり、音声区間長と話速とに加えて、ユーザが発声した方向に着目する。以下、第 4 の実施形態において、上記第 1 の実施形態との相違点を中心に説明する。

【 0 0 7 4 】

< 制御装置の機能構成 >

はじめに、第 4 の実施形態に係る制御装置 1 0 0 の緊急停止部の機能構成について説明する。図 8 は、制御装置の緊急停止部の機能構成の一例を示す第 4 の図である。図 3 に示した機能構成との相違点は、第 4 の実施形態に係る制御装置 1 0 0 の緊急停止部 8 0 0 の場合、画像取得部 8 0 1 及び顔認識器 8 0 2 を有する点と、緊急発声判定器 8 0 3 が緊急発声判定器 3 0 4 とは異なる機能を有する点である。

【 0 0 7 5 】

画像取得部 8 0 1 は、カメラ 3 0 により生成された画像データを取得し、顔認識器 8 0

10

20

30

40

50

2 に出力する。

【 0 0 7 6 】

顔認識器 8 0 2 は認識部の一例であり、画像取得部 8 0 1 により出力された画像データを受け取り、受け取った画像データから顔領域を検出し、検出した顔領域に基づいて、顔の向きを認識する。また、顔認識器 8 0 2 は、認識した顔の向きが、正面を向いているか否かを判定し、認識した顔の向き（または、正面を向いているか否か）の判定結果を、緊急発声判定器 8 0 3 に出力する。顔認識は公知の方法により実現できる。例えば、W. Zhao, R. Chellappa, A. Rosenfeld, "Face recognition: A literature survey," ACM Comput. Surv., Vol.35, No.4, pp.399-458, 2003 に示されている。顔の向きの検出についても、例えば、Michael D. Breitenstein, Daniel Kuettel, Thibaut Weise, Luc van Gool, and Hanspeter Pfister, "Real-time face pose estimation from single range images," CVPR, 2008 に示されている。

10

【 0 0 7 7 】

緊急発声判定器 8 0 3 は、音声区間検出器 3 0 2 により出力された音声区間長と、話速算出器 3 0 5 により出力された話速と、顔認識器 8 0 2 により出力された判定結果とに基づいて、下記条件を満たすか否かを判定する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。
- ・顔が正面を向いている。

【 0 0 7 8 】

20

緊急発声判定器 8 0 3 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 8 0 3 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 0 7 9 】

このように、音声区間長及び話速に加えて、顔の向きを条件にすることで、緊急停止部 8 0 0 によれば、緊急停止させるべき状況が生じたか否かを判定する際の判定精度を向上させることができる。

【 0 0 8 0 】

< 緊急停止処理の流れ >

30

次に、緊急停止部 8 0 0 による緊急停止処理の流れについて説明する。図 9 は、緊急停止部による緊急停止処理の流れを示す第 3 のフローチャートである。図 4 で示したフローチャートとの相違点は、ステップ S 4 1 1、S 4 1 2 である。

【 0 0 8 1 】

ステップ S 4 1 1 において、画像取得部 8 0 1 は、音声データ（発声）の開始時刻から、現在時刻（音声データ（発声）の終了時刻）までの間（音声区間）に、主記憶装置 2 0 2 に格納された画像データを取得する。

【 0 0 8 2 】

ステップ S 4 1 2 において、顔認識器 8 0 2 は、ステップ S 4 1 1 において取得された画像データから顔領域を検出し、検出した顔領域に基づいて、顔の向きを認識する。また、顔認識器 8 0 2 は、認識した顔の向きが、正面を向いているか否かを判定する。

40

【 0 0 8 3 】

< 変形例 1 >

上記説明では、画像取得部 8 0 1 が取得した画像データから、顔認識器 8 0 2 が顔領域を検出し、検出した顔領域に基づいて、顔の向きを認識するものとして説明した。しかしながら、画像取得部 8 0 1 が取得した画像データに、複数のユーザが含まれ、複数の顔領域が検出されることも考えられる。そこで、変形例 1 では、音取得部 3 0 1 が音の到来方向（発声したユーザが位置する方向）を特定し、顔認識器 8 0 2 では、画像データに含まれる複数の顔領域のうち、特定した到来方向に位置する顔領域について、顔の向きを認識する。

50

【 0 0 8 4 】

これにより、発声したユーザとは異なるユーザの顔の向きを判定してしまうといった事態を回避することができる。なお、変形例 1 の場合、マイクロフォンは複数配されているものとし、また、音取得部 3 0 1 は、複数のマイクロフォンにおいて生成された音データに基づいて、音の到来方向（発声したユーザが位置する方向）を特定することができるものとする。音の到来方向の特定には、公知の手法を用いればよい。例えば、

・ H. Pan, R. Scheibler, E. Bezzam, I. Dokmanic, and M. Vetterli, "FRIDA: FRI-based DOA estimation for arbitrary array layouts, " Proc. ICASSP, pp. 3186-3190, 2017

・ R. Schmidt, "Multiple emitter location and signal parameter estimation," IE 10
EE Trans. Antennas Propag., Vol. 34, Issue 3, pp. 276-280, 1986

が知られている。

【 0 0 8 5 】

図 1 0 は、緊急停止部による緊急停止処理の流れを示す第 4 のフローチャートである。図 9 に示す第 3 のフローチャートとの相違点は、ステップ S 5 1 1、S 5 1 2 である。

【 0 0 8 6 】

ステップ S 5 1 1 において、音取得部 3 0 1 は、音の到来方向（発声したユーザが位置する方向）を特定する。

【 0 0 8 7 】

ステップ S 5 1 2 において、顔認識器 8 0 2 は、ステップ S 4 1 1 において取得された 20
画像データにおいて、音取得部 3 0 1 により特定された到来方向（発声したユーザが位置する方向）の顔領域を検出し、検出した顔領域に基づいて、顔の向きを認識する。また、顔認識器 8 0 2 は、認識した顔の向きが、正面を向いているか否かを判定する。

【 0 0 8 8 】

このように、音声区間長及び話速に加えて、音の到来方向の顔領域から認識される顔の向きを条件にすることで、緊急停止部 8 0 0 によれば、緊急停止させるべき状況が生じたか否かを判定する際の判定精度を更に向上させることができる。

【 0 0 8 9 】

< 変形例 2 >

上記説明では、マイクロフォン 2 0、カメラ 3 0 がロボット 1 0 と一体的に構成されて 30
いることを前提としていた。しかしながら、マイクロフォン 2 0 またはカメラ 3 0 のいずれか一方または両方が、ロボット 1 0 と別体として構成されることも考えられる。

【 0 0 9 0 】

このような場合、制御装置 1 0 0 に対しては、ユーザが発声する際に向くべき方向を予め設定しておく。そして、顔認識器 8 0 2 では、画像データから、音の到来方向（発声したユーザが位置する方向）の顔領域を検出し、検出した顔領域に基づいて顔の向きを認識する際、認識した顔の向きが、予め設定した方向を向いているか否かを判定する。

【 0 0 9 1 】

そして、緊急発声判定器 8 0 3 では、音声区間検出器 3 0 2 により出力された音声区間長と、話速算出器 3 0 5 により出力された話速と、顔認識器 8 0 2 により出力された判定 40
結果とに基づいて、下記条件を満たすか否かを判定する。

- ・ 話速が予め定められた閾値以上である。
- ・ 音声区間長が予め定められた閾値以下である。
- ・ 顔が予め設定された方向を向いている。

【 0 0 9 2 】

緊急発声判定器 8 0 3 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 8 0 3 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 0 9 3 】

このように、音声区間長及び話速に加えて、音の到来方向の顔領域から認識される顔の向き（所定の顔の向き）を条件にすることで、緊急停止部 800 によれば、緊急停止させるべき状況が生じたか否かを判定する際の判定精度を向上させることができる。

【0094】

<まとめ>

以上の説明から明らかなように、第4の実施形態に係る制御装置の緊急停止部は、作業中のロボットにおいて緊急停止させるべき状況が生じると、一般に、ユーザは短い言葉で早口に所定の方向に向かって叫ぶことが多い点に着目し、

- ・話速が予め定められた閾値以上であり、かつ、
- ・音声区間長が予め定められた閾値以下であり、かつ、
- ・顔の向きが予め定められた方向を向いている、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【0095】

これにより、第4の実施形態に係る制御装置の緊急停止部によれば、所定の方向に向けられていない発声により、ロボットを緊急停止させてしまうといった事態を回避することができる。つまり、緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。この結果、ロボットをユーザの意図したとおりに緊急停止させることが可能となる。

【0096】

[第5の実施形態]

上記第1の実施形態では、緊急停止させるべき状況が生じたか否かを判定する場合について説明した。これに対して、第5の実施形態では、緊急停止させるべき状況が生じたか否かを判定することに加えて、通常停止させるべき状況が生じたか否かについても判定する。

【0097】

なお、緊急停止とは、緊急停止命令に基づいてロボット10への電源供給を停止することをいう。ただし、電源供給を停止する直前に、ロボット10の可動部に対してブレーキやロックをかけてもよい。また、電源供給を停止する前後に、ロボット10の可動部に対して、電源を用いない機構を利用してブレーキやロックをかけてもよい。

【0098】

一方、通常停止とは、通常停止命令に基づいて、ロボット10への電源供給を維持したまま、ロボット10の可動部を、その位置で外力が加わらない限り動かないように固定することをいう。ここでいう固定の方法には、

- a) 外部からの力によって動かされても、通常停止命令があった位置に戻るよう位置制御して固定する方法、
 - b) 外部からの力によって動かされた場合には、外部からの力がなくなった時点の位置を保持するように制御して固定する方法、
 - c) ブレーキを配し、当該ブレーキにより位置を固定する方法、
- 等が含まれる。

【0099】

なお、電源供給を停止するとは、ロボット10における全ての要素に対する電源供給を停止することのほか、ロボット10における動作に関する要素に対する電源供給を停止すること、が含まれるものとする。

【0100】

以下、第5の実施形態について、上記第1の実施形態との相違点を中心に説明する。

【0101】

<制御装置の機能構成>

はじめに、第5の実施形態に係る制御装置100の緊急停止部の機能構成について説明する。図11は、制御装置の緊急停止部の機能構成の一例を示す第5の図である。図3に示した機能構成との相違点は、緊急停止部1100が、停止発声判定器1101を有する

10

20

30

40

50

点である。

【 0 1 0 2 】

停止発声判定器 1 1 0 1 は、緊急停止させるべき状況が生じていないと判定されることで緊急発声判定器 3 0 4 より出力される、緊急発声でないことを示す信号（非緊急停止信号）を受け取る。また、停止発声判定器 1 1 0 1 は、音素認識器 3 0 3 より出力される音素列を受け取る。

【 0 1 0 3 】

更に、停止発声判定器 1 1 0 1 は、緊急発声判定器 3 0 4 より、緊急発声でないことを示す信号を受け取ると、音素認識器 3 0 3 より受け取った音素列に、予め定められた停止に関する音素列が含まれるか否かを判定する。また、停止発声判定器 1 1 0 1 は、予め定められた停止に関する音素列が含まれると判定した場合に、制御部 3 0 6 に対して、通常停止信号を出力し、ロボット 1 0 を通常停止させる。

10

【 0 1 0 4 】

< 緊急停止処理の流れ >

次に、緊急停止部 1 1 0 0 による緊急停止処理の流れについて説明する。図 1 2 は、緊急停止部による緊急停止処理の流れを示す第 5 のフローチャートである。図 4 で示したフローチャートとの相違点は、ステップ S 6 0 1、S 6 0 2 である。

【 0 1 0 5 】

ステップ S 6 0 1 において、停止発声判定器 1 1 0 1 は、音素認識器 3 0 3 より、音素列を受け取り、受け取った音素列に、停止に関する音素列が含まれるか否かを判定する。

20

【 0 1 0 6 】

ステップ S 6 0 1 において、受け取った音素列に、停止に関する音素列が含まれないと判定した場合には、ステップ S 1 0 1 に戻る。一方、ステップ S 6 0 1 において、受け取った音素列に、停止に関する音素列が含まれると判定した場合には、通常停止させるべき状況が生じたと判定し、ステップ S 6 0 2 に進む。

【 0 1 0 7 】

ステップ S 6 0 2 において、停止発声判定器 1 1 0 1 は、制御部 3 0 6 に対して通常停止信号を出力する。制御部 3 0 6 は、可動部 4 0 が通常停止命令に基づいて動作するように制御し、ロボット 1 0 を通常停止させる。

【 0 1 0 8 】

30

< まとめ >

以上の説明から明らかなように、第 5 の実施形態に係る制御装置の緊急停止部は、ロボットにおいて緊急停止させるべき状況が生じていないと判定した場合でも、音素列に通常停止すべき音素列が含まれていた場合には、ロボットを通常停止させる。

【 0 1 0 9 】

これにより、第 5 の実施形態に係る制御装置の緊急停止部によれば、緊急停止以外の停止方法を含む、複数種類の停止方法（緊急停止方法、通常停止方法）でロボットを停止させることができる。

【 0 1 1 0 】

[第 6 の実施形態]

40

上記第 1 の実施形態では、緊急停止させるべき状況が生じた場合にユーザが咄嗟に発声する音声の音声区間長と話速とに着目した。一方で、緊急停止させるべき状況が生じた場合、ユーザは、同じ単語、単語断片、又は言い回しを繰り返し発声することが想定される。例えば、「まってまって」（英語の場合には、"Wait, Wait, Wait"）、「ストップストップ」（英語の場合には、"Stop, Stop, Stop"）、「危ない危ない危ない」（英語の場合には、"Danger, Danger, Danger"）「ちょっちょっちょっ」（英語の場合には、"Hey, Hey, Hey"）などである。そこで、第 6 の実施形態では、緊急停止させるべき状況が生じたか否かを判定するにあたり、同一単語の繰り返しに着目する。以下、第 6 の実施形態について、上記第 1 の実施形態との相違点を中心に説明する。

【 0 1 1 1 】

50

< 制御装置の機能構成 >

はじめに、第 6 の実施形態に係る制御装置 1 0 0 の緊急停止部の機能構成について説明する。図 1 3 は、制御装置の緊急停止部の機能構成の一例を示す第 6 の図である。図 3 に示した機能構成との相違点は、第 6 の実施形態に係る制御装置 1 0 0 の緊急停止部 1 3 0 0 の場合、繰り返し検出器 1 3 0 1 を有する点、及び、緊急発声判定器 1 3 0 2 が緊急発声判定器 3 0 4 とは異なる機能を有する点である。

【 0 1 1 2 】

繰り返し検出器 1 3 0 1 は、音素認識器 3 0 3 により出力された音素列を受け取り、受け取った音素列における繰り返し部分を検出する。また、繰り返し検出器 1 3 0 1 は、検出結果を、緊急発声判定器 1 3 0 2 に出力する。

【 0 1 1 3 】

なお、繰り返し検出器の処理例は、例えば、プログラミング言語 Python を用いて以下のように記述することができる。

```
def detect(input):
```

```
    R = []
```

```
    for w in range(2, len(input)//2+1):
```

```
        for b in range(len(input)-w+1):
```

```
            s = input[b:b+w]
```

```
            R.append(s*2)
```

```
    for r in R:
```

```
        if r in input:
```

```
            return True
```

```
    return False
```

なお、上記記述において、戻り値が "True" とは、input に入力された音素列において繰り返し部分が検出されたことを表す。一方、戻り値が "False" とは、input に入力された音素列において繰り返し部分が検出されなかったことを表す。

【 0 1 1 4 】

なお、音素認識器が任意の音素列を認識するのではなく、文法により認識すべき音素列を指定できる機能を備える場合には、想定される音素列の繰り返しをすべて文法に記載しておくことで繰り返し部分を検出することができる。この場合、繰り返し検出器 1 3 0 1 は無くてもよい。文法を用いた音素認識器の実現方法の一例として Recognition Grammar Toolkit for Julius が挙げられる。認識の単位を単語ではなく音素とすることで、本実施の形態の音素認識器とすることができる。

【 0 1 1 5 】

緊急発声判定器 1 3 0 2 は、話速算出器 3 0 5 により出力された話速と、繰り返し検出器 1 3 0 1 により出力された検出結果とに基づいて、下記条件を満たすか否かを判定する。

- ・ 話速が予め定められた閾値以上である。
- ・ 音素列に繰り返し部分が含まれる。

【 0 1 1 6 】

緊急発声判定器 1 3 0 2 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 1 3 0 2 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 1 1 7 】

このように、話速が閾値以上であることに加えて、音素列に繰り返し部分が含まれることを条件にすることで、緊急停止部 1 3 0 0 によれば、緊急停止させるべき状況が生じたか否かを高精度に判定することができる。

【 0 1 1 8 】

< 緊急停止処理の流れ >

次に、緊急停止部 1 3 0 0 による緊急停止処理の流れについて説明する。図 1 4 は、緊

10

20

30

40

50

急停止部による緊急停止処理の流れを示す第 6 のフローチャートである。図 4 で示したフローチャートとの相違点は、ステップ S 1 4 0 1 である。

【 0 1 1 9 】

ステップ S 1 4 0 1 において、繰り返し検出器 1 3 0 1 は、ステップ S 1 0 7 で認識した音素列における繰り返し部分を検出する。

【 0 1 2 0 】

なお、繰り返し検出器 1 3 0 1 は、ステップ S 1 0 7 で認識された音素列において、予め定められた長さ以上の音素列の繰り返しが現れるか否かによって検出する。

【 0 1 2 1 】

< まとめ >

以上の説明から明らかなように、第 6 の実施形態に係る制御装置の緊急停止部は、作業中のロボットにおいて緊急停止させるべき状況が生じると、一般に、ユーザは短い言葉を繰り返し早口で発話することが多い点に着目し、

- ・話速が予め定められた閾値以上であり、かつ、
- ・ユーザの発声に対応する音素列に繰り返し部分が含まれる、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【 0 1 2 2 】

これにより、第 6 の実施形態に係る制御装置の緊急停止部によれば、作業中のロボットにおいて緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。この結果、ロボットをユーザの意図したとおり

【 0 1 2 3 】

なお、本実施形態において、緊急発声判定器 1 3 0 2 は、話速算出器 3 0 5 により出力された話速及び繰り返し検出器 1 3 0 1 により出力された検出結果に加えて、音声区間検出器 3 0 2 により出力された音声区間長を用いて緊急発声を検出してもよい。この場合、緊急発声判定器 1 3 0 2 は、下記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。
- ・音素列に繰り返し部分が含まれる。

【 0 1 2 4 】

[第 7 の実施形態]

上記第 6 の実施形態では、緊急停止させるべき状況が生じた場合にユーザが咄嗟に発声する音声の話速と発声に対応する音素列における繰り返し部分とに着目した。一方で、緊急停止させるべき状況が生じた場合かつユーザにとっての危険度が時間とともに上昇していく場合、同じ単語や言い回しを繰り返し発声しつつ、その音圧レベルが時間と共に上昇してくことが想定される。例えば、回転機械に服を巻き込まれたものの数秒程度は猶予があるような場合、徐々に大きな声になることが想定される。そこで、第 7 の実施形態では、緊急停止させるべき状況が生じたか否かを判定するにあたり、話速と同一単語の繰り返しとに加えて、音圧レベルの上昇に着目する。以下、第 7 の実施形態について、上記第 6 の実施形態との相違点を中心に説明する。なお、以下では、説明を簡単にするため、単語が繰り返されるものとして説明するが、単語の繰り返しに代えて単語断片や単語列が繰り返されてもよい。

【 0 1 2 5 】

< 制御装置の機能構成 >

はじめに、第 7 の実施形態に係る制御装置 1 0 0 の緊急停止部の機能構成について説明する。図 1 5 は、制御装置の緊急停止部の機能構成の一例を示す第 7 の図である。図 1 3 に示した機能構成との相違点は、第 7 の実施形態に係る制御装置 1 0 0 の緊急停止部 1 5 0 0 の場合、第 2 の実施形態に記載した音圧レベル算出器と同様の音圧レベル算出器 1 5 0 1 を有する点、及び、緊急発声判定器 1 5 0 2 が緊急発声判定器 1 3 0 2 とは異なる機

10

20

30

40

50

能を有する点である。

【 0 1 2 6 】

繰り返し検出器 1 3 0 1 は、単語の繰り返しの検出及び検出結果の出力に加え、繰り返された各単語に対応する音声区間（繰り返し区間）についても算出し、音圧レベル算出器 1 5 0 1 に出力する。音圧レベル算出器 1 5 0 1 は、繰り返された各単語に対応する音声区間（繰り返し区間）の音圧レベルを算出する。また、音圧レベル算出器 1 5 0 1 は算出した音圧レベルを、緊急発声判定器 1 5 0 2 に出力する。

【 0 1 2 7 】

緊急発声判定器 1 5 0 2 は、話速算出器 3 0 5 により出力された話速と、繰り返し検出器 1 3 0 1 により出力された検出結果と、音圧レベル算出器 1 5 0 1 により出力された音圧レベルとに基づいて、下記条件を満たすか否かを判定する。

- ・話速が予め定められた閾値以上である。
- ・音素列に繰り返し部分が含まれる。
- ・音圧レベルが増加している（繰り返し区間における第 1 時刻の音圧レベルが、第 1 時刻より後の第 2 時刻の音圧レベルより小さい）。

【 0 1 2 8 】

緊急発声判定器 1 5 0 2 は、上記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。なお、上記条件のいずれか 1 つまたは複数を満たさないと判定した場合、緊急発声判定器 1 5 0 2 は、制御部 3 0 6 に対して緊急停止信号を出力しない。

【 0 1 2 9 】

このように、話速が閾値以上であること及び音素列に繰り返し部分が含まれることに加え、音圧レベルが増加していることを条件にすることで、緊急停止部 1 5 0 0 によれば、緊急停止させるべき状況が生じたか否かを高精度に判定することができる。

【 0 1 3 0 】

< 緊急停止処理の流れ >

次に、緊急停止部 1 5 0 0 による緊急停止処理の流れについて説明する。図 1 6 は、緊急停止部による緊急停止処理の流れを示す第 7 のフローチャートである。図 6 で示したフローチャートとの相違点は、ステップ S 1 6 0 1、S 1 6 0 2 である。

【 0 1 3 1 】

ステップ S 1 6 0 1 では、第 6 の実施形態と同様、繰り返し検出器 1 3 0 1 が、ステップ S 1 0 7 で認識した音素列における繰り返し部分を検出する。加えて、繰り返し検出器 1 3 0 1 は、繰り返された各単語に対応する音声区間（繰り返し区間）を算出する。

【 0 1 3 2 】

ステップ S 1 6 0 2 において、音圧レベル算出器 1 5 0 1 は、ステップ S 1 6 0 1 において算出された繰り返し区間の音圧レベルを算出する。

【 0 1 3 3 】

< まとめ >

以上の説明から明らかなように、第 7 の実施形態に係る制御装置の緊急停止部は、作業中のロボットにおいて緊急停止させるべき状況が生じると、ユーザは短い言葉を繰り返し早口で発声し、その発声に対応する音声データの音圧レベルは徐々に大きくなることが多い点に着目し、

- ・話速が予め定められた閾値以上であり、かつ、
- ・ユーザの発声に対応する音素列に繰り返し部分が含まれ、かつ、
- ・繰り返し区間の音圧レベルが増加している、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力する。

【 0 1 3 4 】

これにより、第 7 の実施形態に係る制御装置の緊急停止部によれば、作業中のロボットにおいて緊急停止させるべき状況が生じた場合に、ユーザの音声指示に基づいて、ロボットを適切に緊急停止させることができる。この結果、ロボットをユーザの意図したとおり

10

20

30

40

50

に動作させることが可能となる。

【 0 1 3 5 】

なお、本実施形態において、緊急発声判定器 1 5 0 2 は、話速算出器 3 0 5 により出力された話速、繰り返し検出器 1 3 0 1 により出力された検出結果、及び音圧レベル算出器 1 5 0 1 により出力された繰り返し区間の音圧レベルに加えて、音声区間検出器 3 0 2 により出力された音声区間長を用いて緊急発声を検出してもよい。この場合、緊急発声判定器 1 5 0 2 は、下記条件のいずれも満たすと判定した場合、緊急発声を検出した旨を表す緊急停止信号を制御部 3 0 6 に出力する。

- ・話速が予め定められた閾値以上である。
- ・音声区間長が予め定められた閾値以下である。
- ・音素列に繰り返し部分が含まれる。
- ・音圧レベルが増加している。

10

【 0 1 3 6 】

[その他の実施形態]

上記各実施形態では、単位時間あたりの音素数を算出することで話速を算出するものとしたが、話速の算出方法はこれに限定されない。例えば、音素より小さい単位を用いて話速を算出してもよい。あるいは、音素より大きい単位（例えば、モーラ、音節、単語）を用いて話速を算出してもよい。

【 0 1 3 7 】

また、上記第 2 乃至第 5 の実施形態では、上記第 1 の実施形態において示した条件に、上記第 2 乃至第 5 の実施形態において新たに示した条件のうちのいずれかの条件を加えた場合について説明した。しかしながら、上記第 1 の実施形態において示した条件に、上記第 2 乃至第 5 の実施形態において新たに示した条件のうちの複数の条件を組み合わせる場合であってもよい。

20

【 0 1 3 8 】

あるいは、上記第 1 の実施形態において示した条件の一部に、上記第 2 乃至第 7 の実施形態において新たに示した条件のうちの一部の条件を組み合わせる場合であってもよい。例えば、

- ・音声区間長が予め定められた閾値以下であり、かつ、
- ・音圧レベルが予め定められた閾値以上である、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力してもよい。

30

【 0 1 3 9 】

あるいは、

- ・音声区間長が予め定められた閾値以下であり、かつ、
- ・音素列が、予め定められた音素列を含む（あるいは含まない）、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力してもよい。

【 0 1 4 0 】

あるいは、

- ・音声区間長が予め定められた閾値以下であり、かつ、
- ・音素列に繰り返し部分が含まれる、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力してもよい。

40

【 0 1 4 1 】

あるいは、

- ・音声区間長が予め定められた閾値以下であり、かつ、
- ・音素列に繰り返し部分が含まれ、かつ、
- ・音圧レベルが増加している、

場合に、緊急停止させるべき状況が生じたと判定し、緊急停止信号を出力してもよい。

【 0 1 4 2 】

また、上記各実施形態では、制御部 3 0 6 が、可動部 4 0 を制御するものとして説明したが、制御部 3 0 6 は、可動部 4 0 以外の要素を制御するように構成してもよい。例えば、ロボット 1 0 が音声出力部を有している場合にあっては、音声出力部が緊急停止命令に

50

基づいて動作するように制御してもよい（つまり、音声出力を停止し、スピーカをOFF状態にするように制御してもよい）。あるいは、ロボット10が表示部を有している場合にあっては、表示部が緊急停止命令に基づいて動作するように制御してもよい（つまり、表示を停止し、画面をOFF状態にするように制御してもよい）。ロボット10が不適切な音声出力や、不適切な表示出力を行った場合に有効である。

【0143】

また、上記各実施形態において、緊急停止部の機能は、プロセッサ201が、緊急停止プログラムを実行することで実現されるものとして説明した。しかしながら、緊急停止部の機能は、アナログ回路、デジタル回路又はアナログ・デジタル混合回路で構成された回路により実現されてもよい。また、緊急停止部の機能を実現する制御回路を備えていてもよい。各回路の実装は、ASIC（Application Specific Integrated Circuit）、FPGA（Field Programmable Gate Array）等によるものであってもよい。

10

【0144】

また、上記各実施形態において、緊急停止プログラムを実行するにあたっては、緊急停止プログラムをフレキシブルディスクやCD-ROM等の記憶媒体に収納し、コンピュータに読み込ませて実行させるものであってもよい。記憶媒体は、磁気ディスクや光ディスク等の着脱可能なものに限定されず、ハードディスク装置やメモリなどの固定型の記憶媒体であってもよい。また、ソフトウェアによる処理は、FPGA等の回路に実装され、ハードウェアが実行するものであってもよい。ジョブの実行は、例えば、GPU（Graphics Processing Unit）等のアクセラレータを使用して行ってもよい。

20

【0145】

なお、上記実施形態に挙げた構成等に、その他の要素との組み合わせ等、ここで示した構成に本発明が限定されるものではない。これらの点に関しては、本発明の趣旨を逸脱しない範囲で変更することが可能であり、その応用形態に応じて適切に定めることができる。

【符号の説明】

【0146】

10	: ロボット
20	: マイクロフォン
30	: カメラ
40	: 可動部
41 ~ 44	: 関節
45	: エンドエフェクタ
100	: 制御装置
300	: 緊急停止部
301	: 音取得部
302	: 音声区間検出器
303	: 音素認識器
304	: 緊急発声判定器
305	: 話速算出器
306	: 制御部
500	: 緊急停止部
501	: 音圧レベル算出器
502	: 緊急発声判定器
700	: 緊急停止部
701	: 音素列格納部
702	: 緊急発声判定器
800	: 緊急停止部
801	: 画像取得部
802	: 顔認識器
803	: 緊急発声判定器

30

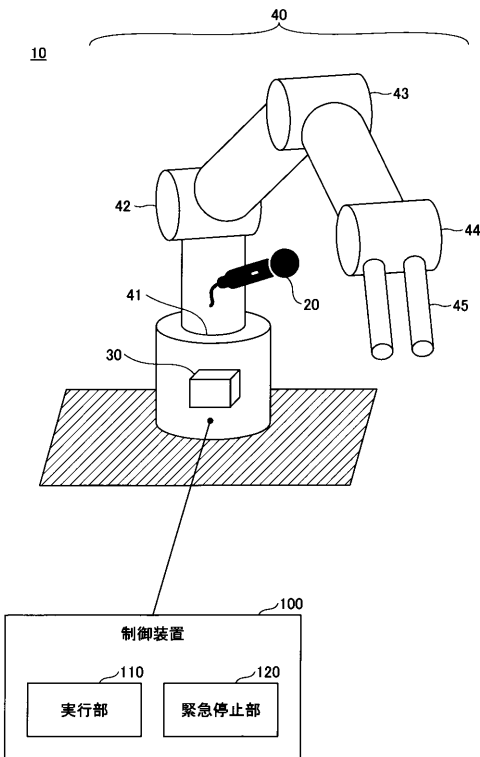
40

50

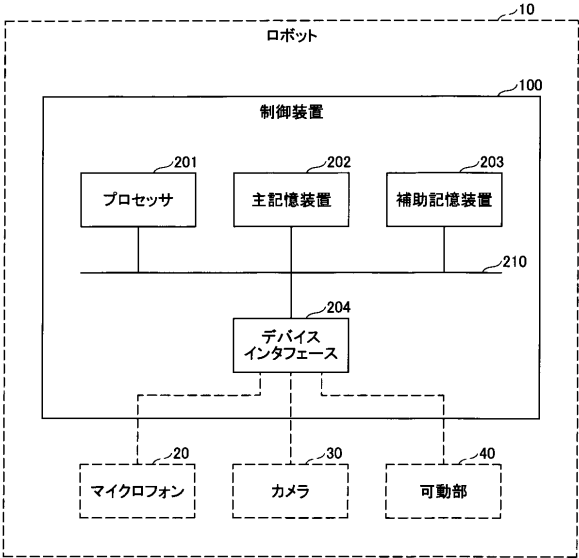
- 1 1 0 0 : 緊急停止部
- 1 1 0 1 : 停止発声判定器
- 1 3 0 0 : 緊急停止部
- 1 3 0 1 : 繰り返し検出器
- 1 3 0 2 : 緊急発声判定器
- 1 5 0 0 : 緊急停止部
- 1 5 0 1 : 音圧レベル算出器
- 1 5 0 2 : 緊急発声判定器

【図面】

【図 1】



【図 2】



10

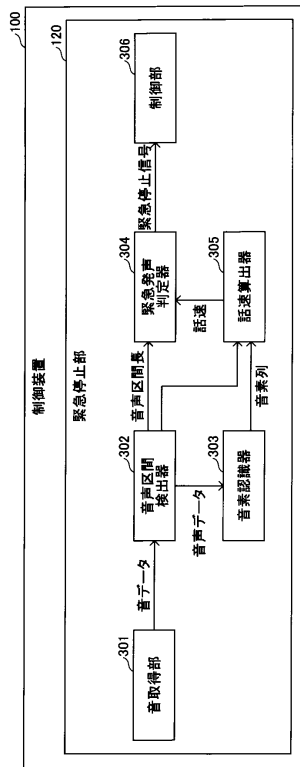
20

30

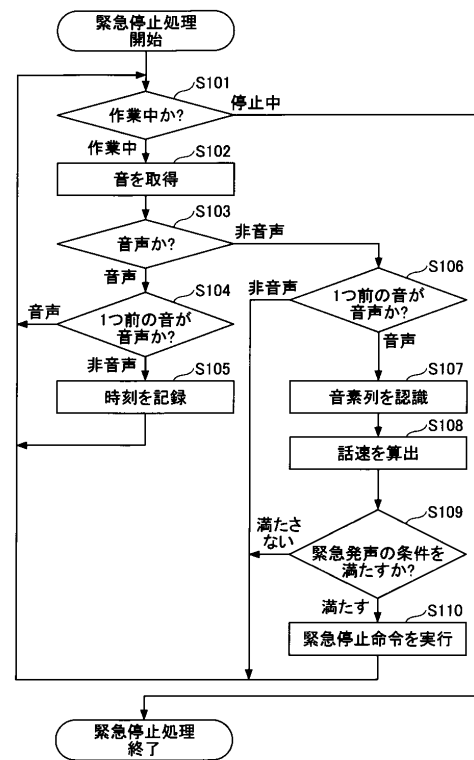
40

50

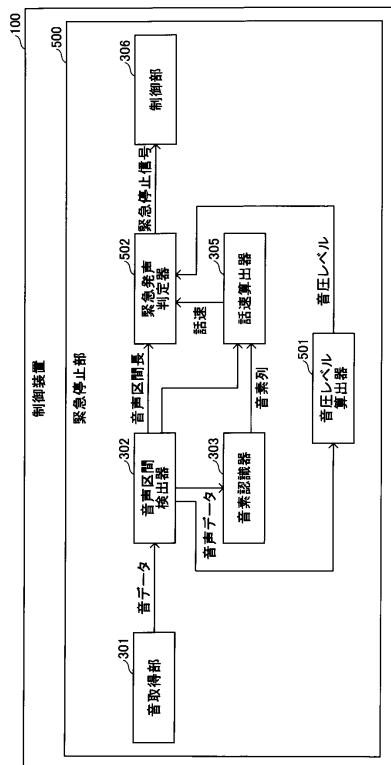
【図 3】



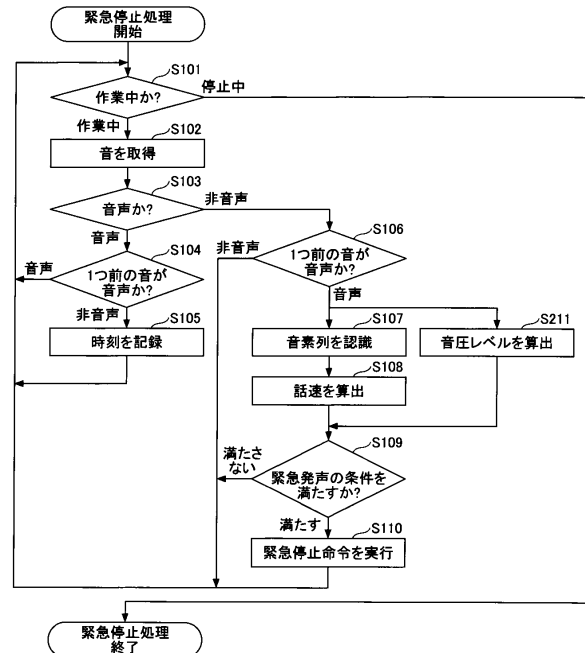
【図 4】



【図 5】



【図 6】



10

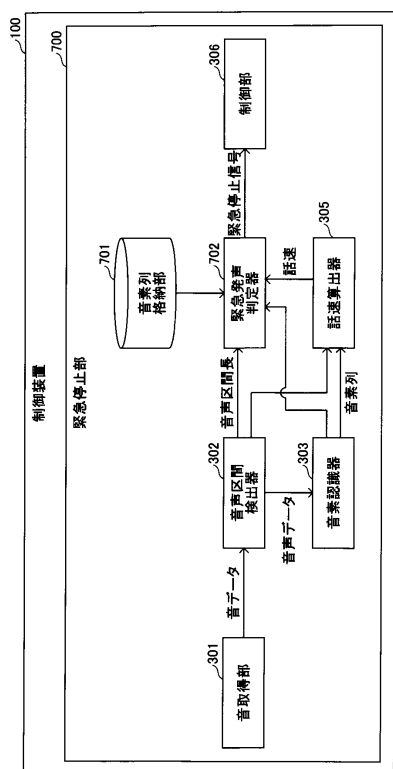
20

30

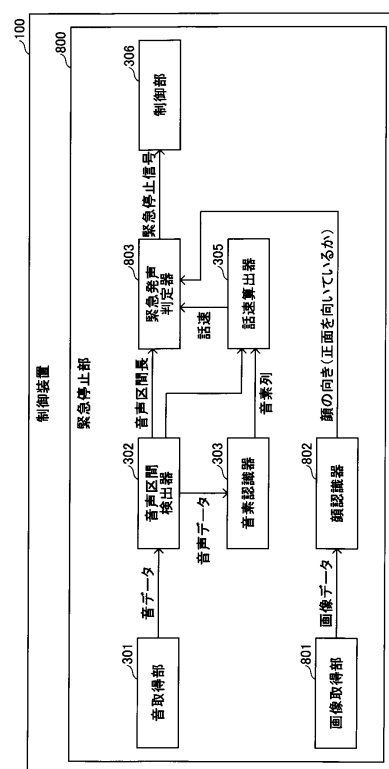
40

50

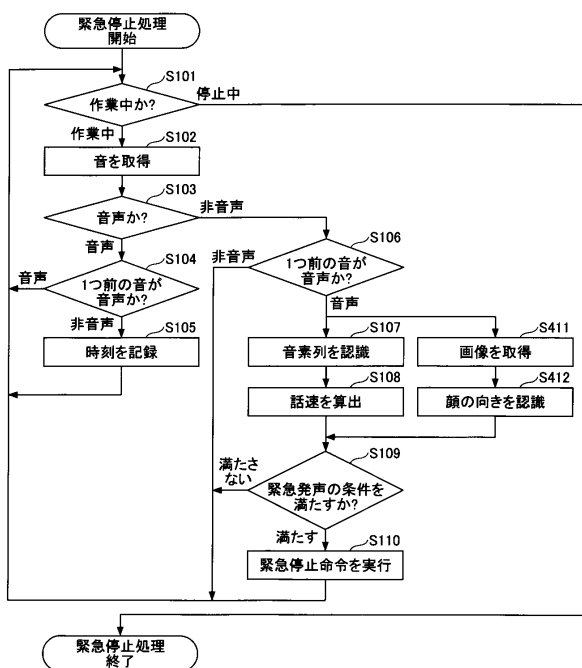
【圖 7】



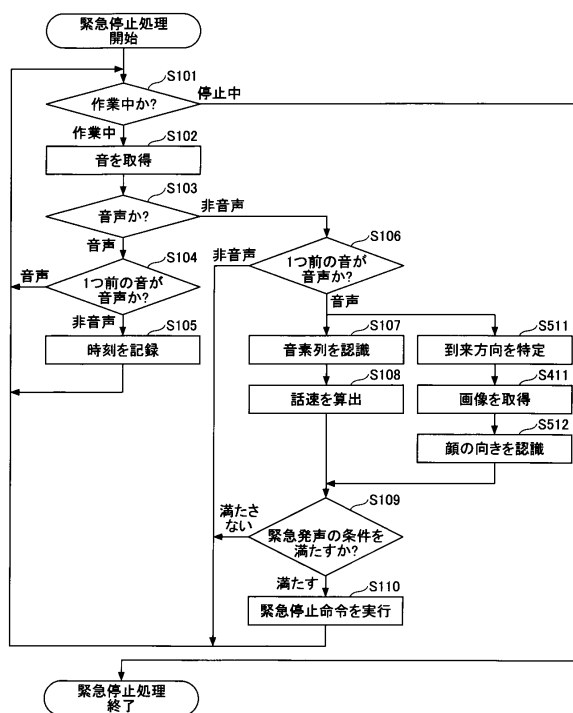
【圖 8】



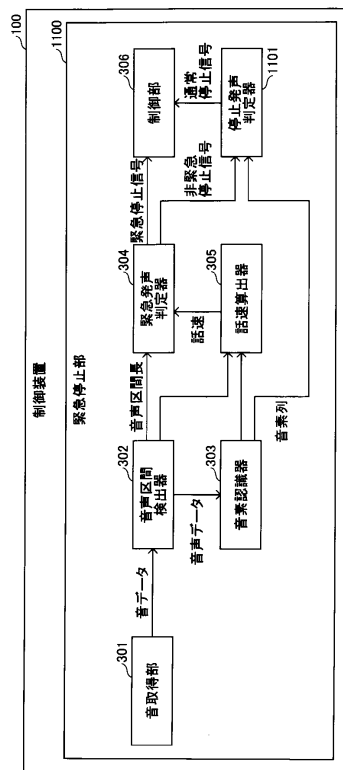
【 図 9 】



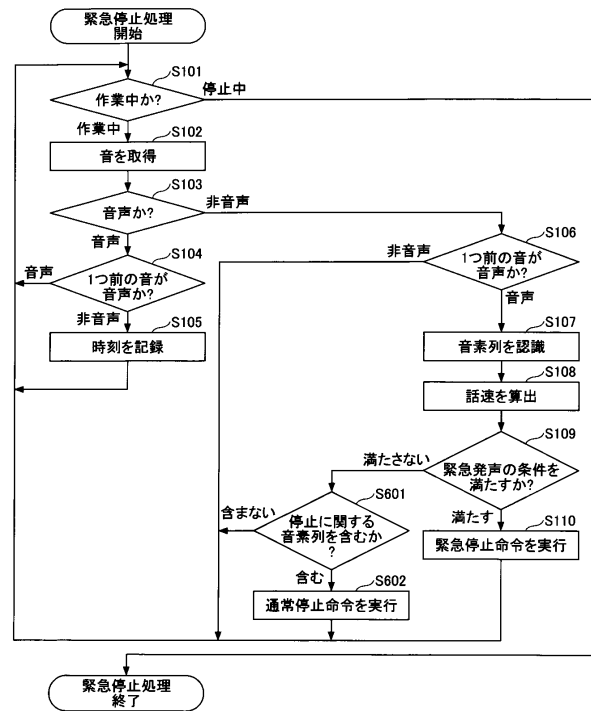
【圖 10】



【 図 1 1 】



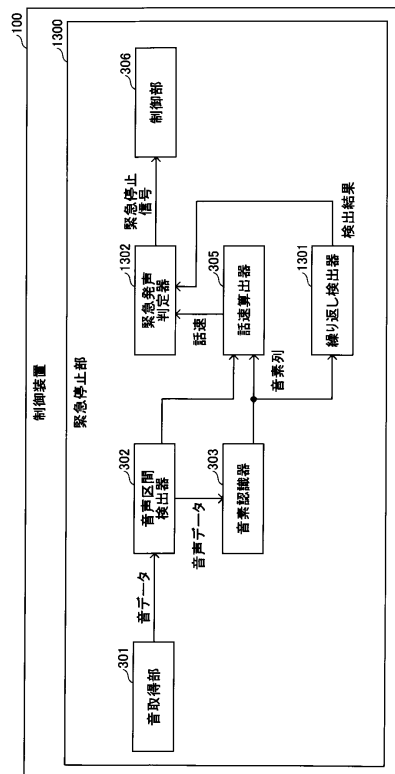
【 図 1 2 】



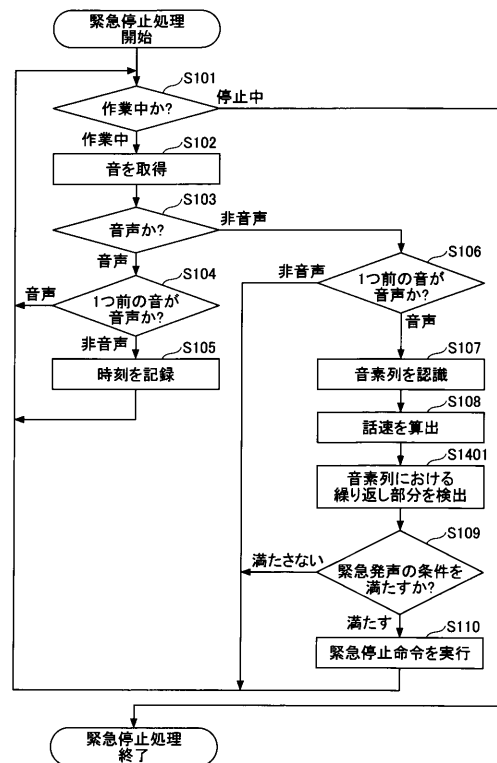
10

20

【 图 1 3 】



【 図 1 4 】



30

40

50

フロントページの続き

大手町ビル 株式会社Preferred Networks内
(72)発明者 鍋島 厚太
東京都千代田区大手町1丁目6番1号 大手町ビル 株式会社Preferred Networks内
(72)発明者 永尾 学
東京都千代田区大手町1丁目6番1号 大手町ビル 株式会社Preferred Networks内
審査官 石川 薫
(56)参考文献 特開平05-119792(JP,A)
特開2007-027167(JP,A)
特開2017-145066(JP,A)
特開2014-228644(JP,A)
特開2014-235345(JP,A)
(58)調査した分野 (Int.Cl., DB名)
B25J 1/00 - 21/02
G10L 15/00
G10L 15/10