

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5393544号
(P5393544)

(45) 発行日 平成26年1月22日(2014.1.22)

(24) 登録日 平成25年10月25日(2013.10.25)

(51) Int. Cl.	F I	
B 2 5 J 13/00 (2006.01)	B 2 5 J 13/00	Z
B 2 5 J 5/00 (2006.01)	B 2 5 J 5/00	F
G 1 0 L 13/00 (2006.01)	G 1 0 L 13/00	1 0 0 A
G 1 0 L 13/08 (2013.01)	G 1 0 L 13/08	1 3 2
A 6 3 H 3/33 (2006.01)	A 6 3 H 3/33	C
請求項の数 6 (全 12 頁) 最終頁に続く		

(21) 出願番号	特願2010-56265 (P2010-56265)	(73) 特許権者	000005326
(22) 出願日	平成22年3月12日 (2010.3.12)		本田技研工業株式会社
(65) 公開番号	特開2011-189439 (P2011-189439A)		東京都港区南青山二丁目1番1号
(43) 公開日	平成23年9月29日 (2011.9.29)	(74) 代理人	100064908
審査請求日	平成24年11月27日 (2012.11.27)		弁理士 志賀 正武
特許法第30条第1項適用 社団法人 日本ロボット学会、第27回 日本ロボット学会学術講演会CD-ROM、平成21年9月15日		(74) 代理人	100108578
			弁理士 高橋 詔男
		(74) 代理人	100146835
			弁理士 佐伯 義文
		(74) 代理人	100094400
			弁理士 鈴木 三義
		(74) 代理人	100107836
	弁理士 西 和哉		
(74) 代理人	100108453		
	弁理士 村山 靖彦		
			最終頁に続く

(54) 【発明の名称】 ロボット、ロボット制御方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

基体に可動連結した可動部の駆動を制御する駆動制御手段と、
音声を生成する音声生成手段と、
前記音声生成手段において生成された音声を出力する音声出力手段と
を有するロボットにおいて、

前記音声生成手段は、
前記駆動制御手段によって制御される前記可動部の前記基体に対する姿勢に基づいて、
生成する音声を補正することを特徴とするロボット。

【請求項2】

前記可動部の1つは頭部であって、
前記音声生成手段は、
前記頭部のピッチ軸方向の前記姿勢に基づいて、声道フィルタを用いて生成する音声を
補正することを特徴とする請求項1に記載のロボット。

【請求項3】

前記音声生成手段は、
音圧レベルが前記頭部の前記姿勢を示すピッチ角度に依存する周波数帯域の音声信号を
増幅する声道フィルタを用いて、生成する音声信号に含まれる前記ピッチ角度に依存する
周波数帯域を増幅する
ことを特徴とする請求項2に記載のロボット。

【請求項 4】

前記可動部の 1 つは頭部であって、
 前記音声生成手段は、
 前記頭部のヨー軸方向の前記姿勢に基づいて、生成する音声の左右の出力レベル比を補正することを特徴とする請求項 1 から請求項 3 の何れかに記載のロボット。

【請求項 5】

基体に可動連結した可動部の駆動を制御する駆動制御手段と、
 音声を生成する音声生成手段と、
 前記音声生成手段において生成された音声を出力する音声出力手段と
 を備えるロボットを制御するロボット制御方法であって、
 前記音声生成手段は、
 前記駆動制御手段によって制御される前記可動部の前記基体に対する姿勢に基づいて、生成する音声を補正することを特徴とするロボット制御方法。

10

【請求項 6】

基体に可動連結した可動部を備え、音声を生成して出力するロボットのコンピュータに、
 前記可動部の駆動を制御する駆動制御ステップと、
 音声を生成する音声生成ステップと、
 前記音声生成ステップにおいて生成された音声を出力する音声出力ステップと
 を実行させるプログラムであって、
 前記音声生成ステップは、
 前記駆動制御ステップによって制御される前記可動部の前記基体に対する姿勢に基づいて、生成する音声を補正することを特徴とするプログラム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、姿勢に基づいて出力音声を補正するロボット、ロボット制御方法およびプログラムに関する。

【背景技術】

【0002】

従来、コミュニケーションロボットにおいて、言語以外の情報である韻律や感情情報を発話音声に付加する技術がある。韻律や感情情報の付加によって、発話内容は、よりの確に伝達される。また、音声合成において、表現力を向上するために人の声道をモデル化する手法が開示されている（例えば、特許文献 1 参照）。

30

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2002 - 23775 号公報

【発明の概要】

【発明が解決しようとする課題】

40

【0004】

しかしながら、韻律や感情情報を付加する技術では、コミュニケーションロボットの姿勢（特に顔の向き）が変化しても発話音声に変化しないため違和感がある。本発明はかかる事情に鑑みてなされたものであり、姿勢に応じて、違和感のない自然な音声を出力する技術を提供することを目的とする。

【課題を解決するための手段】

【0005】

上記問題を解決するために、請求項 1 記載のロボット（例えば、実施形態におけるロボット 4）は、基体に可動連結した可動部（例えば、実施形態における頭部 4）の駆動を制御する駆動制御手段（例えば、実施形態における制御部 100）と、音声を生成する音声

50

生成手段（例えば、実施形態における音声生成部 110）と、前記音声生成手段において生成された音声出力する音声出力手段（例えば、実施形態における音声出力部 140L、140R）とを有し、前記音声生成手段は、前記駆動制御手段によって制御される前記可動部の前記基体に対する姿勢に基づいて、生成する音声を補正することを特徴とする。

請求項 2 記載のロボットは、前記可動部の 1 つが頭部であって、前記音声生成部が、前記頭部のピッチ軸方向の前記姿勢に基づいて、声道フィルタを用いて生成する音声を補正することを特徴とする。

請求項 3 記載のロボットは、前記音声生成手段が、音圧レベルが前記頭部の前記姿勢を示すピッチ角度に依存する周波数帯域の音声信号を増幅する声道フィルタを用いて、生成する音声信号に含まれる前記ピッチ角度に依存する周波数帯域を増幅することを特徴とする。

10

請求項 4 記載のロボットは、前記可動部の 1 つが頭部であって、前記音声生成部は、前記頭部のヨー軸方向の前記姿勢に基づいて、生成する音声の左右の出力レベル比を補正することを特徴とする。

請求項 5 記載のロボット制御方法は、基体に可動連結した可動部の駆動を制御する駆動制御手段と、音声を生成する音声生成手段と、前記音声生成手段において生成された音声出力する音声出力手段とを備えるロボットを制御するロボット制御方法であって、前記音声生成手段は、前記駆動制御手段によって制御される前記可動部の前記基体に対する姿勢に基づいて、生成する音声を補正することを特徴とする。

請求項 6 記載のプログラムは、基体に可動連結した可動部を備え、音声を生成して出力するロボットのコンピュータに、前記可動部の駆動を制御する駆動制御ステップと、音声を生成する音声生成ステップと、前記音声生成ステップにおいて生成された音声出力する音声出力ステップとを実行させるプログラムであって、前記音声生成ステップは、前記駆動制御ステップによって制御される前記可動部の前記基体に対する姿勢に基づいて、生成する音声を補正することを特徴とする。

20

【発明の効果】

【0006】

請求項 1 記載のロボットによれば、音声生成手段によって生成された音声はそのまま音声出力手段から出力されるのではなく、音声生成手段によって生成された音声は可能部の姿勢に基づいて補正されてから音声出力手段から出力されるため、例えば、可動部の姿勢に応じてスピーカの位置、角度を動かすことなく、可動部の姿勢に応じた違和感のない自然な音声を出力することができる。

30

請求項 2 記載のロボットによれば、人間における音声の変化を模して、頭部の垂直方向の姿勢に応じて音声を変化させるため、より自然な音声を出力することができる。

請求項 3 記載のロボットによれば、ピッチ角度の変化に対し依存度が高い周波数帯域を変化させるため、効果的に音声を出力することができる。

請求項 4 記載のロボットによれば、頭部の水平方向の姿勢に応じて音声を変化させるため、より自然な音声を出力することができる。

請求項 5 記載の方法および請求項 6 記載のプログラムによれば、請求項 1 と同様の効果を得ることができる。

40

【図面の簡単な説明】

【0007】

【図 1】本実施形態によるロボットの概略概観図である。

【図 2】本実施形態によるロボットのブロック図である。

【図 3】音声信号補正部による補正に関連する内容を説明するための説明図である。

【図 4】音声信号補正部による補正を説明するための説明図である。

【発明を実施するための形態】

【0008】

以下、本発明の実施形態について図面を参照して詳細に説明する。図 1 (a) は本実施形態によるロボット 4 の概略正面図、図 1 (b) はロボット 4 の概略側面図である。ロボ

50

ット4は、図1(a)(b)に示すように、基体部41と、基体部41にそれぞれ可動連結される頭部42(可動部)と、脚部43L(可動部)、脚部43R(可動部)と、腕部44L(可動部)、腕部44R(可動部)とを備えている。また、ロボット4は、図1(b)に示すように、背負う格好で基体部41に収納部45を装着している。なお、基体部41には、スピーカ140L(音声出力部140L)およびスピーカ140R(音声出力部R)が収納されている。

【0009】

図2は、ロボット4のブロック図である。ロボット4は、図2に示すように、制御部100(駆動制御手段)、音声生成部110(音声生成手段)、頭機能部120、腕機能部130L、腕機能部130R、音声出力部140L(音声出力手段)、音声出力部140R(音声出力手段)、脚機能部150L、脚機能部150Rを備えている。

10

【0010】

頭機能部120は、頭部42または基体41上部に収納され、制御部100からの制御信号に基づいて、頭部42のピッチ軸方向(垂直方向)およびヨー軸方向(水平方向)の動作を制御する。これにより、ロボット4の頭部42の基体41に対する姿勢は、ピッチ軸方向およびヨー軸方向に変化する。つまり、ロボット4は、首を上下左右に振ることができる。

【0011】

脚機能部150L、Rは、それぞれ脚部43L、Rに収納され、制御部100からの制御信号に基づいて二足歩行などの動作を制御する。腕機能部130L、Rは、腕部44L、Rに収納され、制御部100からの制御信号に基づいて、二足歩行時の腕の動作などを制御する。

20

【0012】

制御部100は、収納部45に収納され、頭機能部120、脚機能部150および腕機能部140に制御信号を出力し、頭部42、脚部43L、R、腕部44L、Rの駆動を制御する。また、制御部100は、頭部42の基体41に対する姿勢を示す情報(頭部姿勢情報)を音声生成部110に出力する。

【0013】

例えば、制御部100は、頭部42の基体41に対する変化後(制御後)の姿勢を示す情報を制御情報として頭機能部120に出力する場合、当該制御情報を頭部姿勢情報として音声生成部110に出力する。また例えば、制御部100は、頭部42の基体41に対する現在の姿勢からの変化量を示す情報を制御情報として頭機能部120に出力する場合、頭部42の基体41に対する変化後(制御後)の姿勢を算出し、算出結果を頭部姿勢情報として音声生成部110に出力する。

30

【0014】

音声生成部110は、収納部45に収納され、音声信号生成部112および音声信号補正部114を備える。音声信号生成部112は、所定の音声に係る音声信号を生成する。音声信号生成部112は、生成した音声信号を音声信号補正部114に出力する。なお、所定の音声は、例えば、外部(例えば、ロボット4の操作者)から指定された言語に係る音声、外部からの命令に応じて選択した所定の判定処理の結果に対応付けられている音声、ロボット4が具備する種々のセンサ(例えば、障害物を検出するセンサ)の検出結果に対応付けられている音声などである。なお、ロボット4は、頭部42を動かしている間においても音声を出力することができる。

40

【0015】

音声信号補正部114は、音声信号生成部112から音声信号を取得する。また、音声信号補正部114は、制御部100から頭部姿勢情報を取得する。音声信号補正部114は、制御部100から取得した頭部姿勢情報に基づいて、音声信号生成部112から取得した音声信号を補正する。

【0016】

具体的には、音声信号補正部114は、頭部42のピッチ軸方向の姿勢を示す角度と、

50

頭部 4 2 のヨー軸方向の姿勢を示す角度とを含む頭部姿勢情報を取得する。音声信号補正部 1 1 4 は、頭部 4 2 の姿勢のうちピッチ軸方向の姿勢に関しては、声道モデルによる声道フィルタを用いて、音声信号生成部 1 1 2 から取得した音声信号を補正する。ピッチ軸方向に姿勢が変化した場合、声道形状は変化し、音声信号のスペクトル包絡が変化するためである。音声信号補正部 1 1 4 は、頭部 4 2 の姿勢のうちヨー軸方向の姿勢に関しては、音声出力部 1 4 0 に出力する左右の出力レベル比を補正する。ヨー軸方向に姿勢が変化した場合、声道形状は変化なく、左右の音量を調整すれば足りるからである。なお、声道モデルによる声道フィルタ、左右の出力レベル比、ピッチ軸方向の姿勢に応じた補正、および、ヨー軸方向の姿勢に応じた補正に関する詳細は後述する。

【 0 0 1 7 】

10

音声信号補正部 1 1 4 は、補正後の音声信号（または頭部姿勢情報によっては、音声信号生成部 1 1 2 からの音声信号）を音声出力部 1 4 0 L、R に出力する。

【 0 0 1 8 】

音声出力部 1 4 0 L、R は、上述の如く、基体部 4 1 に収納され、音声生成部 1 1 0 において生成された音声信号（音声信号生成部 1 1 2 によって生成された音声信号または音声信号補正部 1 1 4 によって補正された音声信号）を取得し、音声として外部に出力する。

【 0 0 1 9 】

以下、音声信号補正部 1 1 4 について詳細に説明する。図 3 は、音声信号補正部 1 1 4 による補正に関連する内容を説明するための説明図である。図 4 は、音声信号補正部 1 1 4 による補正を説明するための説明図である。

20

【 0 0 2 0 】

（声道モデルによる声道フィルタに関する説明）

声道フィルタは、ピッチ軸方向の角度 θ_p を用いて、 $H(\theta_p)$ とモデル化される。本実施形態においては、人の声を検査して声道フィルタ $H(\theta_p)$ をモデル化する。具体的には、男性被験者の頭部を下向き 50° から上向き 50° （ 10° 単位）まで変化させた種々のピッチ角度 θ_p における同じ音量の音声を無響室内で接話マイクロホンを用いて録音し、声道フィルタ $H(\theta_p)$ をモデル化する。より詳細には、基本周波数が $201 \sim 523$ (Hz) である 10 秒間のスイープ音、母音「a」の発声を録音する。次いで、録音された音声の音声信号を 3 分の 1 オクターブバンドで解析し、各周波数帯域のピッチ角度 0° を基準とする音圧レベル比を計算する。

30

【 0 0 2 1 】

ピッチ角度 θ_p に応じて補正する音声信号の周波数帯域には、ピッチ角度 θ_p の変化に応じて音圧レベルが顕著に変化する周波数帯域を用いる。具体的には、補正する周波数帯域として、図 3 (a) に示すように、ピッチ角度 θ_p の変化に応じて音圧レベルが顕著に変化し、かつ、大抵の母音のフォルマントを有する、 500 (Hz)、 $800 \sim 1000$ (Hz)、 $2500 \sim 4000$ (Hz) の 3 つの周波数帯域を用いる。

【 0 0 2 2 】

なお、図 3 (a) の横軸は、ピッチ角度 θ_p (度) であって、(Up) は上向き、(Down) は下向きである。図 3 (a) の縦軸は、ピッチ角度 0° を基準とした各ピッチ角度 θ_p の音圧レベル比、即ちゲイン g^p_i (dB) である。図 3 (a) に示すように、ゲイン g^p_i は、被験者が上を向いているときは大きく、被験者が下を向いているときは小さい。なお、女性被験者の場合も同様である。

40

【 0 0 2 3 】

（左右の出力レベル比に関する説明）

ヨー角度 θ_y に応じた左右の出力レベル比の補正に関しては、図 3 (b) に示すモデルを利用する。図 3 (b) は、ヨー角度 θ_y の変化による音圧レベル比を示す図である。図 3 (b) の横軸は、ヨー角度 θ_y (度) であって、マイナスは右向き、プラスは左向きである。図 3 (b) の縦軸は、ヨー角度 0° のときの左右のチャンネルの音圧レベルの差（左チャンネル - 右チャンネル）を基準とした、各ヨー角度 θ_y のときの左右のチャンネルの音圧レベ

50

ルの差（左チャンネル - 右チャンネル）の比、即ちゲイン g_{yL} (dB) である。図 3 (b) には、理論値と実測値とを示したが、本実施形態においては、顔の向きの変化がより明確に表現できるように、ヨー角度 y の変化の影響がより大きい実測値を使用する。

【0024】

なお、図 3 (b) に示す実測値は、例えば、無響室内において、図 3 (c) に示すように、各マイク (mic.) のインパルス応答を測定して得ることができる。図 3 (c) のヨー角度 y は、図 3 (b) との整合から、中央位置（破線方向）を 0° 、左方向をプラスとしている。

【0025】

(ピッチ軸方向およびヨー軸方向の姿勢に応じた補正)

10

音声信号補正部 114 は、図 4 に示すように、音声信号生成部 112 から取得した音声信号（入力音声信号 $x(t)$ ）を、頭部 42 のピッチ角度 p に基づいて補正し、続いて頭部 42 のヨー角度 y に基づいて補正する。

【0026】

音声信号補正部 114 は、ピッチ角度 p に基づいて補正するために、まず、入力音声信号 $x(t)$ を、ピッチ角度 p に依存する p 依存性成分 $x_i(t)$ と、ピッチ角度 p に依存しない p 非依存性成分 $x_{NULL}(t)$ とに分解する。なお、入力音声信号 $x_i(t)$ における i は、補正する周波数帯域のインデックスであって、入力音声信号 $x_1(t)$ は 500 (Hz)、入力音声信号 $x_2(t)$ は $800 \sim 1000$ (Hz)、入力音声信号 $x_3(t)$ は $2500 \sim 4000$ (Hz) である。

20

【0027】

具体的には、音声信号補正部 114 は、下記式 (1) に従って、入力音声信号 $x(t)$ から p 依存性成分 $x_i(t)$ を算出し、下記式 (2) に従って、 p 非依存性成分 $x_{NULL}(t)$ を算出する。

【0028】

【数 1】

$$x_i(t) = (x * h_i)(t) \quad \dots (1)$$

$$x_{NULL}(t) = (x * h_{NULL})(t) \quad \dots (2)$$

但し、* は畳み込み演算である。

30

【0029】

上記式 (1) において、 h_i は、 $x_i(t)$ 、即ち、各周波数帯域 (500 、 $800 \sim 1000$ 、 $2500 \sim 4000$ (Hz)) を透過させるバンドパスフィルタである。また、 $x_{NULL}(t)$ は、入力音声信号 $x(t)$ から $x_i(t)$ を引いたものである。つまり、 h_{NULL} に関する周波数応答のゲインは、 500 、 $800 \sim 1000$ 、 $2500 \sim 4000$ Hz で 0 であり、他の 3 分の 1 オクターブバンド中心周波数で 1 である。

【0030】

次いで、音声信号補正部 114 は、 p 依存性成分 $x_i(t)$ について、下記式 (3) に従って、角度 $p(t)$ のときの i 番目の周波数帯域のゲイン $g_i^p(t)$ を得る。

40

【0031】

【数 2】

$$g_i^p(t) = \frac{g_{i,\theta_m}^p(\theta_{m+1} - \theta_p(t)) + g_{i,\theta_{m+1}}^p(\theta_p(t) - \theta_m)}{10} \quad \dots (3)$$

【0032】

上記式 (3) において、角度 θ_m および角度 θ_{m+1} は、制御部 100 から取得した頭

50

部姿勢情報によって示される頭部 4 2 のピッチ軸方向の姿勢を示す角度 $\theta_p(t)$ に基づく角度である。具体的には、角度 θ_m および角度 θ_{m+1} は、下記式 (4) (5) によって、 $\theta_p(t)$ から得る。

【0033】

【数3】

$$\theta_{m+1} = (\lfloor \theta_p(t)/10 \rfloor + 1) \times 10 \quad \dots (4)$$

$$\theta_m = \lfloor \theta_p(t)/10 \rfloor \times 10 \quad \dots (5)$$

但し、 $\lfloor \cdot \rfloor$ は、床関数である。例えば、 $\theta_p(t) = 35^\circ$ のとき、 $\theta_{m+1} = 40^\circ$ および $\theta_m = 30^\circ$ である。 10

【0034】

つまり、音声信号補正部 114 は、 θ_p 依存性成分 $x_i(t)$ について、図 3 (a) に示す $\theta_p(t)$ 毎のゲインに係るスペクトル包絡モデル (図 4 中 A、B によって示されるモデル) を、上記式 (4) (5) に示すように 10 度ずつ dB に対して線形補間し、増幅する。

【0035】

次いで、音声信号補正部 114 は、下記式 (6) に従って、 $x_i(t)$ を増幅する。 20

【0036】

【数4】

$$x_{i,g}(t) = x_i(t) \times 10^{\frac{g_i^p(t)}{10}} \quad \dots (6)$$

【0037】

次いで、音声信号補正部 114 は、下記式 (7) に従って、 $x_i(t)$ と $x_{NULL}(t)$ とを合成し、ピッチ角度 θ_p に基づいて補正した補正後のモノラル信号 x_p を得る。

【0038】

30

【数5】

$$x_p(t) = x_{NULL}(t) + \sum_{i=1}^3 x_{i,g}(t) \quad \dots (7)$$

【0039】

以上のように、音声信号補正部 114 は、上記式 (1) ~ (7) によって表される声道フィルタ (周波数領域で表されている声道フィルタ $H(\theta_p)$ を時間領域で表した声道フィルタ) を用いて、音声信号生成部 112 によって生成された音声信号 $x(t)$ を、頭部 4 2 のピッチ軸方向の姿勢 (ピッチ角度 θ_p) に応じて補正する。 40

【0040】

続いて、音声信号補正部 114 は、ヨー角度 θ_y に基づいて補正するために、モノラル信号 x_p からステレオ信号 x_{stereo} (左チャンネル $x_L(t)$ および右チャンネル $x_R(t)$) を生成する。

【0041】

次いで、音声信号補正部 114 は、下記式 (8) に従って、角度 $\theta_y(t)$ のときの左チャンネル $x_L(t)$ のゲイン $g_L^p(t; \theta_y(t))$ を得る。即ち、音声信号補正部 114 は、頭部 4 2 のヨー軸方向の姿勢に基づいて、左チャンネル $x_L(t)$ および右チャンネル $x_R(t)$ の出力レベル比を補正する。

【0042】

50

【数6】

$$g_L^y(t; \theta_y(t)) = \frac{g_{\theta_n}^y(\theta_{n+1} - \theta_y(t)) + g_{\theta_{n+1}}^y(\theta_y(t) - \theta_n)}{10} \dots (8)$$

【0043】

上記式(8)において、角度 θ_n および角度 θ_{n+1} は、制御部100から取得した頭部姿勢情報によって示される頭部42のヨー軸方向の姿勢を示す角度 $\theta_y(t)$ に基づく角度である。具体的には、角度 θ_n および角度 θ_{n+1} は、下記式(9)(10)によって、 $\theta_y(t)$ から得る。

【0044】

【数7】

$$\theta_{n+1} = (\lfloor \theta_y(t)/10 \rfloor + 1) \times 10 \dots (9)$$

$$\theta_n = (\lfloor \theta_y(t)/10 \rfloor) \times 10 \dots (10)$$

但し、 $\lfloor \cdot \rfloor$ は、床関数である。例えば、 $\theta_y(t) = 35^\circ$ のとき、 $\theta_{n+1} = 40^\circ$ および $\theta_n = 30^\circ$ である。

20

【0045】

なお、左右チャンネルのゲインは、下記式(11)に示すように、対称的なものであるから、音声信号補正部114は、角度 $\theta_y(t)$ のときの右チャンネル $x_R(t)$ のゲイン $g_{P_R}(t; \theta_y(t))$ を得る。

【0046】

【数8】

$$g_R^y(t; \theta_y(t)) = g_L^y(t; -\theta_y(t)) \dots (11)$$

30

【0047】

つまり、音声信号補正部114は、図3(c)に示す $\theta_y(t)$ 毎のゲインに係るスペクトル包絡モデル(図4中Cによって示されるモデル)を、上記式(9)(10)に示すように10度ずつ線形補間し、増幅する。

【0048】

最後に、音声信号補正部114は、左チャンネル信号 $x_L(t)$ を下記式(12)に従って増幅した出力音声信号 $x'_L(t)$ を音声出力部140Lに出力し、右チャンネル信号 $x_R(t)$ を下記式(13)に従って増幅した出力音声信号 $x'_R(t)$ を音声出力部140Rに出力する

40

【0049】

【数9】

$$x'_L(t) = x_L(t) \times 10 \frac{g_L^P(t)}{10} \dots (12)$$

$$x'_R(t) = x_R(t) \times 10 \frac{g_R^P(t)}{10} \dots (13)$$

【0050】

50

以上のように、音声信号補正部 114 は、頭部 42 のヨー軸方向の姿勢（ヨー角度 θ_y ）に基づいて、左チャンネル $x_L(t)$ および右チャンネル $x_R(t)$ を、左チャンネル $x'_L(t)$ および右チャンネル $x'_R(t)$ に補正する。即ち、音声信号補正部 114 は、頭部 42 のヨー軸方向の姿勢に基づいて、左チャンネルおよび右チャンネルの出力レベル比を補正する。

【0051】

以上、上記実施形態によれば、音声生成手段によって生成された音声はそのまま音声出力手段から出力されるのではなく、音声生成手段によって生成された音声は可能部の姿勢に基づいて補正されてから音声出力手段から出力されるため、可動部の姿勢に応じた違和感のない自然な音声を出力することができる。人間における音声の変化を模して、頭部 42 の垂直方向の姿勢に応じて音声を変化させるとともに、頭部 42 の水平方向の姿勢に応じて音声を変化させるため、より自然な音声を出力することができる。

10

また、補正対象の周波数は実験結果に基づく所定の周波数に限定して、ピッチ軸方向の姿勢に応じて音声信号を補正しているため、効果的に計算量を低減させることができる。

【0052】

なお、上記実施形態においては、制御部 100 が頭部姿勢情報（音声信号補正部 114 が音声信号を補正する際に参照する、頭部 42 の基体 41 に対する姿勢を示す情報）を音声信号補正部 114 に出力する態様、即ち、音声信号補正部 114 が頭部姿勢情報を制御部 100 から取得する態様を説明したが、音声信号補正部 114 が頭部姿勢情報を取得する態様はこれに限定されない。例えば、ロボット 1 は、頭部 42 の基体 41 に対する姿勢を測定する測定部（非図示）を備え、音声信号補正部 114 は、測定部による測定結果を頭部姿勢情報として取得してもよい。

20

【0053】

また、上記実施形態においては、音声信号補正部 114 は頭部姿勢情報に基づいて音声信号を補正する態様を説明したが、音声信号補正部 114 が音声信号を補正する態様はこれに限定されない。例えば、音声信号補正部 114 は、脚部 43、腕部 44L、腕部 44R の基体 41 に対する姿勢を示す情報に基づいて音声信号を補正してもよい。

【0054】

なお、本発明の一実施形態によるロボット 4 の各処理を実行するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、当該記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより、本発明の一実施形態によるロボット 4 に係る上述した種々の処理を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OS や周辺機器等のハードウェアを含むものであってもよい。また、「コンピュータシステム」は、WWW システムを利用している場合であれば、ホームページ提供環境（あるいは表示環境）も含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、フラッシュメモリ等の書き込み可能な揮発性メモリ、CD-ROM 等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。

30

【0055】

さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ（例えば DRAM (Dynamic Random Access Memory)）のように、一定時間プログラムを保持しているものも含むものとする。また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステムから、伝送媒体を介して、あるいは、伝送媒体中の伝送波により他のコンピュータシステムに伝送されてもよい。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク（通信網）や電話回線等の通信回線（通信線）のように情報を伝送する機能を有する媒体のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであってもよい。さらに、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プ

40

50

ログラム)であっても良い。

【0056】

以上、この発明の実施形態について図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計等も含まれる。

【符号の説明】

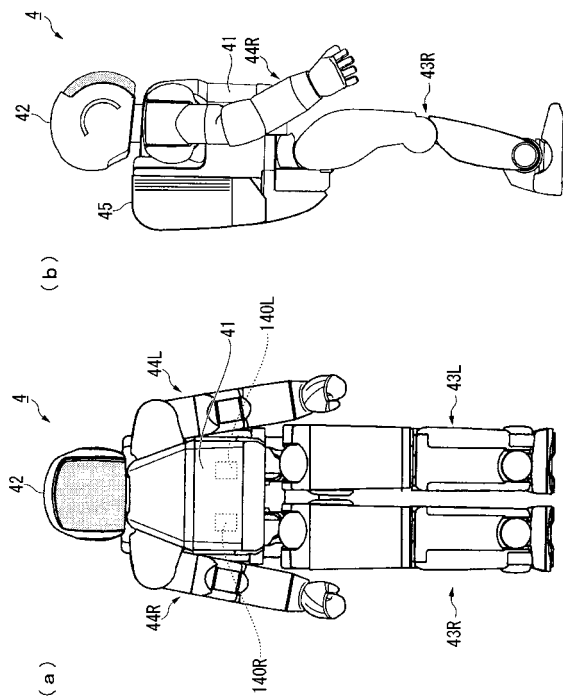
【0057】

- 4 ロボット
- 4 1 基体部
- 4 2 頭部(可動部)
- 4 3 L, 4 3 R 脚部(可動部)
- 4 4 L, 4 4 R 腕部(可動部)
- 4 5 収納部
- 1 0 0 制御部(駆動制御手段)
- 1 1 0 音声生成部(音声生成手段)
- 1 1 2 音声信号生成部
- 1 1 4 音声信号補正部
- 1 2 0 頭機能部
- 1 3 0 腕機能部
- 1 4 0 音声出力部(音声出力手段)
- 1 5 0 脚機能部

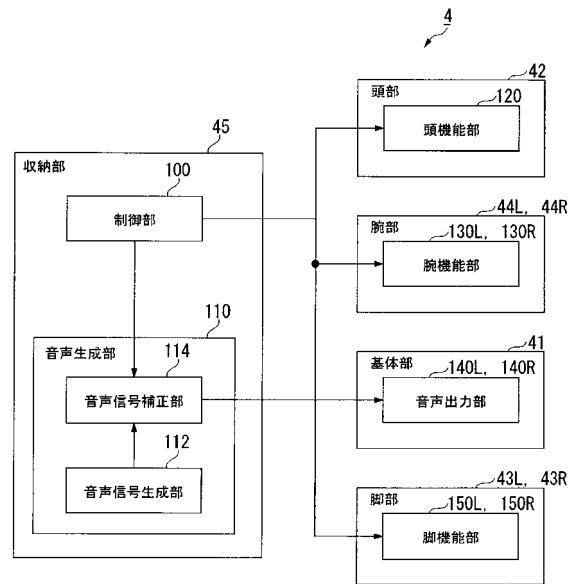
10

20

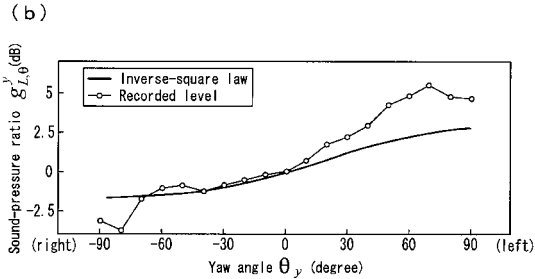
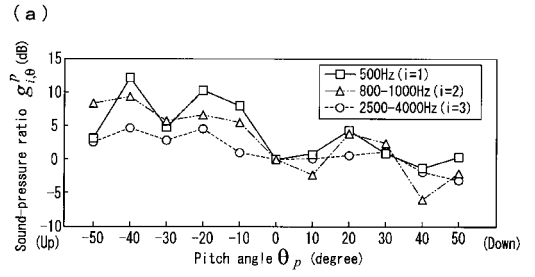
【図1】



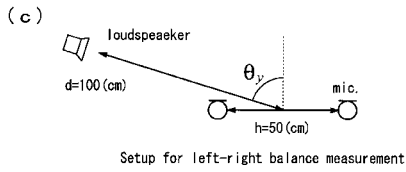
【図2】



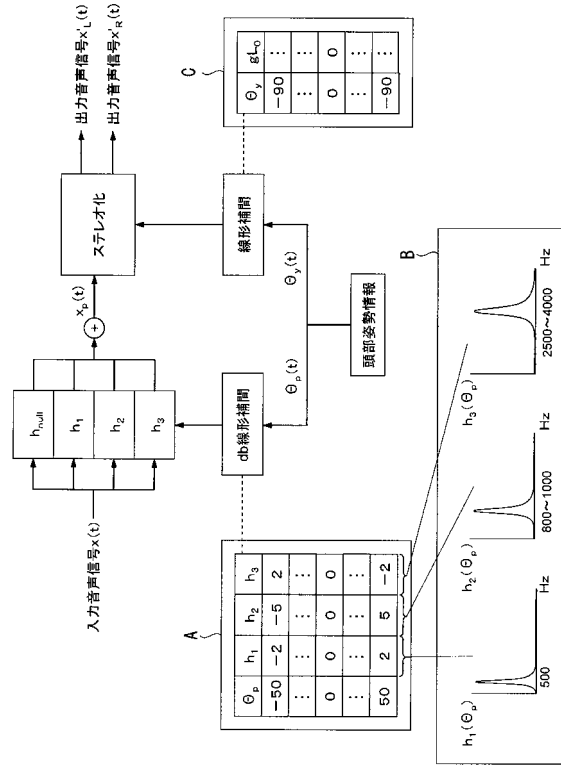
【 図 3 】



Empirical and theoretical θ_y -gain model for the left channel



【 図 4 】



フロントページの続き

(51)Int.Cl. F I
 A 6 3 H 11/00 (2006.01) A 6 3 H 11/00 Z
 G 1 0 L 21/007 (2013.01) G 1 0 L 21/04 1 2 0 E

- (72)発明者 中臺 一博
 埼玉県和光市本町 8 - 1 株式会社ホンダ・リサーチ・インスティテュート・ジャパン内
- (72)発明者 大塚 琢馬
 京都府京都市左京区吉田本町 国立大学法人京都大学 大学院情報学研究科内
- (72)発明者 奥乃 博
 京都府京都市左京区吉田本町 国立大学法人京都大学 大学院情報学研究科内

審査官 佐藤 彰洋

- (56)参考文献 特開 2 0 0 5 - 2 6 6 6 7 1 (J P , A)
 特開 2 0 0 6 - 1 3 6 9 6 3 (J P , A)
 国際公開第 2 0 0 5 / 0 7 6 6 6 0 (W O , A 1)

(58)調査した分野(Int.Cl. , D B名)

B 2 5 J 1 / 0 0 - 2 1 / 0 2
 A 6 3 H 3 / 3 3
 A 6 3 H 1 1 / 0 0
 G 1 0 L 1 3 / 0 0
 G 1 0 L 1 3 / 0 8
 G 1 0 L 2 1 / 0 0 7