

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5382780号
(P5382780)

(45) 発行日 平成26年1月8日(2014.1.8)

(24) 登録日 平成25年10月11日(2013.10.11)

(51) Int.Cl.	F I
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/10 5 0 0 N
G 1 0 L 25/63 (2013.01)	G 1 0 L 11/00 4 0 2 H

請求項の数 5 (全 22 頁)

(21) 出願番号	特願2009-64131 (P2009-64131)	(73) 特許権者	393031586
(22) 出願日	平成21年3月17日(2009.3.17)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2010-217502 (P2010-217502A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成22年9月30日(2010.9.30)	(74) 代理人	100099933
審査請求日	平成24年2月29日(2012.2.29)		弁理士 清水 敏
		(72) 発明者	石井カルロス寿憲
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	石黒 浩
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	萩田 紀博
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内

(出願人による申告) 平成20年度独立行政法人新エネルギー・産業技術総合開発機構「次世代ロボット知能化技術開発プロジェクト コミュニケーション知能(社会・生活分野)の開発 公共空間における情報支援知能モジュール群の開発」委託研究、産業技術力強化法第19条の適用を受ける特許出願

最終頁に続く

(54) 【発明の名称】 発話意図情報検出装置及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

人間の発話音声信号から、発話内容に依存しないパラ言語情報である発話意図を検出するための発話意図情報検出装置であって、

前記発話音声信号の韻律に関する情報を処理するための第1の音声処理手段と、

前記発話音声信号の声質に関する情報を処理するための第2の音声処理手段と、

感動詞の種類毎に、発話スタイルと発話意図との関係を予め記憶するための記憶手段と

、
前記発話音声信号の音声認識の結果を用いて前記記憶手段に記憶されている感動詞の種類を特定し、特定された感動詞について前記韻律に関する情報及び前記声質に関する情報の少なくともいずれかを用いて発話スタイルを特定するための特定手段と、

特定された感動詞の種類及びその発話スタイルに基づいて発話意図を抽出するための抽出手段とを含む、発話意図情報検出装置。

【請求項2】

前記第2の音声処理手段は、前記発話音声信号の発話区間中にボーカル・フライ区間が占める割合を算出するための手段と、前記発話音声信号の低周波帯域に含まれる最大周波数成分と高周波帯域に含まれる最大周波数成分との差であるスペクトル傾斜を算出するための手段とを含む、

前記特定手段は、ボーカル・フライ区間が占める割合とスペクトル傾斜とに基づくりきみ度合いについての情報を用いて発話スタイルを特定するための手段を含む、請求項1に

10

20

記載の発話意図情報検出装置。

【請求項 3】

前記第 2 の音声処理手段は、前記発話音声信号の発話区間中に非周期性／ダブル周期性区間が占める割合を算出するための手段を含み、

前記特定手段は、非周期性／ダブル周期性区間が占める割合についての情報を用いて発話スタイルを特定するための手段を含む、請求項 1 に記載の発話意図情報検出装置。

【請求項 4】

前記第 2 の音声処理手段は、前記発話音声信号の発話区間中に気息性区間が占める割合を算出するための手段を含み、

前記特定手段は、気息性区間が占める割合についての情報を用いて発話スタイルを特定するための手段を含む、請求項 1 に記載の発話意図情報検出装置。

【請求項 5】

コンピュータにより実行されると、当該コンピュータを請求項 1～請求項 4 のいずれかに記載の発話意図情報検出装置として動作させる、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、人間の発話音声から、発話内容に依存しないパラ言語に関する発話者の情報を検出するための装置に関し、特に、人間の発話音声に含まれる韻律に関する情報と声質に関する情報とから、パラ言語についての発話意図に関する情報を検出するための技術に関する。

【背景技術】

【0002】

近年の技術進歩により、人語を発する様々な装置が生産されるようになってきている。この様な装置としては、一例として、カーナビゲーションシステムが挙げられる。カーナビゲーションシステムは、機械が人間に対して一方通行の発話を行なうものであるが、人間との対話が必要とされる装置も存在する。例えば、ロボット等がこれにあたる。

【0003】

ロボットのような装置は、カーナビゲーションシステムよりもさらに人間の生活に密着する可能性が高い。従ってそうした装置で円滑に人間と対話を行なうためには、人間の発話内容だけでなく、感情まで考慮する必要性がある。

【0004】

発話に伴う発話者の感情を推定する場合、発話内容だけでなく、発話内容に依存しない情報である発話意図、態度及び感情等のパラ言語情報をさらに考慮する事が合理的である。つまり、予想されるすべての発話内容に対応する人間の感情を予め学習させるよりは、発話内容と、発話内容に付随するパラ言語情報とを用いて人間の感情を推定する方が合理的でかつ正確であると言える。

【0005】

このようなパラ言語情報の抽出に関する従来技術として、特開 2007-079363 号公報（特許文献 1）に開示されたパラ言語情報検出装置がある。

【0006】

このパラ言語情報検出装置は、パラ言語情報を、韻律情報だけでなく声質情報も用いて抽出する事により、精度高くパラ言語情報を検出することを目的としている。このパラ言語情報検出装置は、人間の発話音声信号から、発話内容に依存しないパラ言語情報を検出するためのパラ言語情報検出装置であって、発話音声信号の韻律に関する情報を処理するための第 1 の音声処理手段と、発話音声信号の声質に関する情報を処理するための第 2 の音声処理手段と、韻律に関する情報と声質に関する情報とから発話音声に関するパラ言語情報を抽出するためのパラ言語情報抽出手段とを含む。

【0007】

このパラ言語情報検出装置によると、情報検出の際に韻律に関する情報のみならず、声

10

20

30

40

50

質に関する情報も使用できる。それゆえ、パラ言語情報検出の精度を上げる事ができる。従って、より精度の高いパラ言語情報検出装置を提供する事ができる。

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特開2007-079363号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

上述したパラ言語情報検出装置は、発話内容に依存しない情報である発話意図、態度及び感情等のパラ言語情報を的確に考慮できる点で好ましい。

10

【0010】

ところで、人間同士が会話する際において発する言葉として、発話内容に依存しない感動詞（または応答詞、感嘆詞、間投詞）がある。この感動詞は、感動、応答、呼びかけを表わし、活用がなく、特に、発話スタイル（韻律・声質）によって、伝達されるパラ言語機能が変化するものである。このような感動詞は、言葉のみでは意味または意図があいまいで発話スタイルにより意図が伝達する非語彙である。

【0011】

本出願人は、自然発話に出現するさまざまな感動詞について分析したところ、感動詞の種類及び発話意図を、パラ言語情報を用いた音声会話システムにおいて考慮する必要があることを見出した。すなわち、感動詞の発話スタイル（韻律・声質）に基づいて、発話意図を検出して、その発話意図に基づいて発話者の感情を分析して、その感情に対応するようにロボットに発話させることにより、人間との会話を円滑にさらに行なわせることができる。

20

【0012】

そこで、本発明の目的は、感動詞について、パラ言語情報としての発話意図情報を発話スタイル（韻律・声質）に基づいて、精度高く検出できる装置を提供する事である。

【課題を解決するための手段】

【0013】

本発明の第1の局面に係る発話意図情報検出装置は、人間の発話音声信号から、発話内容に依存しないパラ言語情報である発話意図を検出するための発話意図情報検出装置であって、発話音声信号の韻律に関する情報を処理するための第1の音声処理手段と、発話音声信号の声質に関する情報を処理するための第2の音声処理手段と、発話スタイルと発話意図との関係を予め記憶するための記憶手段と、韻律に関する情報及び声質に関する情報の少なくともいずれかを用いて発話スタイルを特定するための特定手段と、特定された発話スタイルに関係がある発話意図を抽出するための抽出手段とを含む。

30

【0014】

好ましくは、第2の音声処理手段は、発話音声信号の発話区間中にボーカル・フライ区間が占める割合を算出するための手段と、発話音声信号の低周波帯域に含まれる最大周波数成分と高周波帯域に含まれる最大周波数成分との差であるスペクトル傾斜を算出するための手段とを含み、特定手段は、ボーカル・フライ区間が占める割合とスペクトル傾斜とに基づくりきみ度合いについての情報を用いて発話スタイルを特定するための手段を含む。

40

【0015】

また、好ましくは、第2の音声処理手段は、発話音声信号の発話区間中に非周期性／ダブル周期性区間が占める割合を算出するための手段を含み、特定手段は、非周期性／ダブル周期性区間が占める割合についての情報を用いて発話スタイルを特定するための手段を含む。

【0016】

また、好ましくは、第2の音声処理手段は、発話音声信号の発話区間中に気息性区間が

50

占める割合を算出するための手段を含み、特定手段は、氣息性区間が占める割合についての情報を用いて発話スタイルを特定するための手段を含む。

【0017】

より好ましくは、記憶手段は、感動詞について、発話スタイルと発話意図との関係を予め記憶するための手段を含み、抽出手段は、感動詞についての発話意図を抽出するための手段を含む。

【0018】

本発明の第2の局面に係るコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを上記したいずれかの発話意図情報検出装置として動作させる。

【発明の効果】

10

【0019】

この発話意図情報検出装置によると、パラ言語情報としての発話意図情報を検出する際に、韻律に関する情報及び声質に関する情報の少なくともいずれかを用いて発話スタイルを特定して、記憶された発話スタイルと発話意図との関係に基づいて、発話意図を抽出できる。それゆえ、感動詞について、パラ言語情報としての発話意図情報を発話スタイル（韻律・声質）に基づいて、精度高く検出できる。

【図面の簡単な説明】

【0020】

【図1】本実施の形態に係るパラ言語情報検出装置60についての機能ブロック図である。

20

【図2】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その1）である。

【図3】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その2）である。

【図4】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その3）である。

【図5】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その4）である。

【図6】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その5）である。

30

【図7】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その6）である。

【図8】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その7）である。

【図9】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その8）である。

【図10】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その9）である。

【図11】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その10）である。

40

【図12】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その11）である。

【図13】図1の非語彙辞書50に記憶される発話スタイルと発話意図との関係を説明するための図（その12）である。

【図14】韻律による音声処理部70の処理の詳細を示す機能ブロック図である。

【図15】韻律特徴処理部80の詳細を示す機能ブロック図である。

【図16】声質による音声処理部72の詳細を示す機能ブロック図である。

【図17】ボーカル・フライ検出部120の詳細を示す機能ブロック図である。

【図18】非周期性／ダブル周期性検出部124の詳細を示す機能ブロック図である。

【図19】正規化自己相関関数算出部160の詳細を示す機能ブロック図である。

50

【図20】氣息性検出部128の詳細を示す機能ブロック図である。

【図21】本発明の一実施の形態に係るパラ言語情報抽出装置60を実現するコンピュータシステムの外觀図である。

【図22】図21に示すコンピュータのブロック図である。

【発明を実施するための形態】

【0021】

以下、図面を参照し、本発明の一実施の形態を説明する。本実施の形態は、発話音声信号から韻律による音声処理と声質による音声処理とを行ない、発話意図情報を抽出するためのパラ言語情報検出装置に関するものである。

【0022】

＜構成＞

図1に、本実施の形態に係るパラ言語情報検出装置60の機能ブロック図を示す。図1を参照して、このパラ言語情報検出装置60は、韻律に基づいて発話音声信号を処理してパラ言語情報（発話意図）の抽出に使用するパラメータを出力するための韻律による音声処理部70と、声質に基づいて発話音声信号を処理してパラ言語情報（発話意図）の抽出に使用するパラメータを出力するための声質による音声処理部72と、韻律による音声処理部70と声質による音声処理部72とから得られたパラメータから、予め学習用データを用いて学習した、パラメータとパラ言語情報（発話意図）との関係を規定した非語彙辞書50を用いて、パラ言語情報を抽出して出力するためのパラ言語情報抽出部74とを含む。

【0023】

このパラ言語情報抽出部74においては、特許文献1に開示されたパラ言語情報に加えて、感動詞を発話スタイル（声質に基づくパラメータ及び韻律に基づくパラメータ）により分析して、発話意図を抽出して出力する。このため、パラ言語情報検出装置60は、パラ言語情報抽出部74に接続され、このパラ言語情報抽出部74において使用される非語彙辞書50を備える。この非語彙辞書50は、上述したように、感動詞についての発話スタイルと発話意図との関係を規定したものである。

【0024】

図2～図13を参照して、この非語彙辞書50について説明する。この非語彙辞書には、代表的な感動詞について、以下の手順で分析した結果に基づく、発話スタイルと発話意図との関係が規定されている。

【0025】

この分析にあたり、3種類の自然発話音声データベースにおける、音声データと書き起こしデータとを使用した。すべてのデータベースの書き起こしを1行1発話となる一つのファイルにまとめ、特定の感動詞で始まる発話をテキスト検索により識別した。検索された発話を、被験者により、感動詞であったか否かを識別した。

【0026】

このようなテキスト検索作業の結果、図2～図13に示すような感動詞のグループが得られた。例えば、「え・えー」（図2）、「うん・ふん」（図3）、「あ・あー」（図4）、「へ・へー」（図5）、「はい・はい」（図6）等である。これらが、対話相手への反応として最も頻繁に出現する感動詞のグループである。なお、「なるほど」、「なんか」、「えっと」、「でー」などの感動詞もテキスト検索で識別されたが、これらがもたらすパラ言語機能の変化（バリエーション）は比較的少ない。

【0027】

感動詞「え・えー」グループについては、以下のように分析され、図2に示すように、発話スタイルと発話意図との関係が規定されている。

【0028】

短い「え」は、下降調の場合には肯定、同意、承諾、理解などを表現し、上昇調の場合には聞き返しまたは意外を表現する。氣息音発声または非周期性を含むh a r s h発声が伴う場合、驚きを表現する。長い「えー」については、上昇調の場合には否定的な反応（

10

20

30

40

50

不満、非難、疑い、嫌悪など）を表現し、弱く平坦な場合にはフィラー（考え中）を表現する。文脈により、感心または同情を表現する場合もある。また、りきみ発声が伴う場合には深い驚きまたは深い同情を表わす場合がある。

【0029】

感動詞「うん・ふん」グループについては、以下のように分析され、図3に示すように、発話スタイルと発話意図との関係が規定されている。

【0030】

この感動詞「うん・ふん」グループは、「え・えー」グループとほとんど同様の機能を持つ。違いは、下降上昇調の「ううん」は否定（打ち消し）を表現することである。また、りきみ発声の「えー」に対し、りきみ発声の「うーん」は、深い躊躇、深い困惑または深く考えている状態を表現するという違いがある。

10

【0031】

感動詞「あ・あー」グループについては、以下のように分析され、図4に示すように、発話スタイルと発話意図との関係が規定されている。

【0032】

短い「あ」は、何かに気付いた時、思い出した時または驚いた時に発せられる感動詞である。驚きの場合、breathy/whispery（氣息音）発声が伴うことが多い傾向があり、直後に強い吐息を発することが多い傾向がある。下降調の「あー」は、相槌として理解、同意または肯定の表現に用いられる。連続する「あー」については、全体的に下降するイントネーションを伴う場合には強い理解または強い同意を表現し、最後の「あー」で強いピッチの立て直しがある場合には同情、残念またはがっかりなどを表現する傾向がある。

20

【0033】

感動詞「へ・へー」グループについては、以下のように分析され、図5に示すように、発話スタイルと発話意図との関係が規定されている。

【0034】

短い「へ」については、下降調の場合には肯定または相槌、上昇調の場合には聞き返し、breathy/whispery（氣息音）発声の場合には驚きまたは意外が表現される。長い「へー」は、対話相手に感心または興味を示す場合に使用されるが、単独で単調の場合には逆に聞き流しまたは無関心と捉えられることもある。また、長くてりきみ発声が伴う場合には深みのある感心または驚きが表現される。

30

【0035】

感動詞「はい・はい」グループについては、以下のように分析され、図6に示すように、発話スタイルと発話意図との関係が規定されている。

【0036】

この感動詞「はい・はい」グループは、肯定、承諾または相槌として、フォーマルまたはカジュアルな場面で用いられる。「あ」または「あー」が先行することが多く、相槌と共に気付きまたは理解を示す。上昇調の短い「はい」は聞き返しを示す。連続する「はい」については、イントネーションが下降する場合は理解または共感を示すが、最後の「はい」のアクセント成分にピッチの立て直しがある場合は「邪魔くさい、もう分かった」の意味が表現される。

40

【0037】

感動詞「は・はーん」グループについては、図7に示すように、発話スタイルと発話意図との関係が規定されている。

【0038】

平坦調または下降調の「はー・はーん」は「はい」と「うん・ふん」との間の会話音声であって相槌または理解を示す。（柔らかい声質で）短いまたは長い「はー・はーん」であって連続性のある場合は対話者との間に距離があり謙遜を示す。長い「はー」は感心、驚きまたは理解を示し、長い「はー」であってりきみ発声の場合には感心または驚きの度合いが増す。長い「はー」については、下降調かつ氣息音発声の場合にはため息を示し、

50

短い上昇調であって柔らかい声質の場合には聞き返しを示し、長い上昇調（で硬い声質）の場合には否定的反応（非難）を示す。

【0039】

感動詞「お・おーん」グループについては、図8に示すように、発話スタイルと発話意図との関係が規定されている。

【0040】

短い「お」の場合には驚きまたは気付きを示し、長い「おー」または「おん」の場合には「あー」「うん」のカジュアル形であって理解を示し、「おお」または「おん」の連続の場合には強い理解を示す。

【0041】

感動詞「ほん・ほー」グループについては、図9に示すように、発話スタイルと発話意図との関係が規定されている。

【0042】

平坦調または下降調の「ほん」は、「うん・ふん」のカジュアル形であって相槌または理解を示し、長い「ほー」でりきみ発生の場合には深い感心または深い驚きを示し、長い「ほー」で下降調かつ氣息音発声の場合にはため息を示す。

【0043】

感動詞「いや・やー」グループについては、以下のように分析され、図10に示すように、発話スタイルと発話意図との関係が規定されている。

【0044】

この感動詞「いや・やー」グループは、「いいえ」のように、否定（打消し）としても用いられ、否定的な反応として使用されることが多い。h a r s h / w h i s p e r y（氣息音）発声または高ピッチが伴うと、強い驚きが表現され、りきみ発声が伴うと、深い驚き、深い困惑または深い躊躇が表現される。

【0045】

感動詞「あら・ありゃ」グループについては、以下のように分析され、図11に示すように、発話スタイルと発話意図との関係が規定されている。

【0046】

この感動詞「あら・ありゃ」グループは、驚き、意外または感動の表現に主に女性が用いる。しかし、疑問または疑いを示す際にも用いられる。この場合、上昇調に伴うことがある。また、「ら」が伸長して柔らかい氣息性を含んだ発声に伴う場合には同情、残念または共感などを示すことがある。「あらあら」は、望ましくない出来事に対する反応として使用される。

【0047】

感動詞「あれ・あれー」グループについては、以下のように分析され、図12に示すように、発話スタイルと発話意図との関係が規定されている。

【0048】

この感動詞「あれ・あれー」グループは、「あら」グループのように、驚きまたは意外を示すが、疑問または疑いの気持ちを多く含む。「れ」が伸長する場合、疑問または疑いの度合いが増す傾向がある。また、「あれー」が平坦で伸長する場合には、「えーっと」のように、フィラー（考え中）として発話されることもある。

【0049】

感動詞「わ・うわー」グループについては、以下のように分析され、図13に示すように、発話スタイルと発話意図との関係が規定されている。

【0050】

この感動詞「わ・うわー」グループは、驚きまたは感動などを示す。h a r s h / w h i s p e r y（氣息音）が伴う場合には驚きの度合いが増し、りきみ発声が伴う場合には驚き、感心または嫌悪などの度合いが増す傾向がある。

【0051】

このように非語彙辞書50には、発話スタイル（韻律パラメータ及び声質パラメータ）

10

20

30

40

50

と発話意図との関係を規定している。韻律パラメータとして発話持続時間及び音程の変化を採用し、声質パラメータとして「りきみ」の有無、「非周期性」の有無、「周期性（連続性）」の有無、「気息性」の有無を採用している。このように、自然対話音声に出現する感動詞がもたらすパラ言語情報（発話意図）と発話スタイルとの関係を規定しておいて、この非語彙辞書50を用いて、感動詞の種類及びその発話スタイルに基づいて発話意図を検出する。以下に、発話意図を検出する具体的なパラ言語情報検出装置60の構成を説明する。

【0052】

図14に、韻律による音声処理部70の詳細を機能ブロック図で示す。図14を参照して、韻律による音声処理部70は、発話音声信号をピッチの動き、つまり音程の変化を表わすパラメータであるF0moveに変換する処理を行なうための韻律特徴処理部80と、発話持続時間に関する情報を抽出するための発話持続時間抽出部84とを含む。なお、韻律による音声処理部70は、韻律特徴処理部80で得られたF0moveと発話持続時間抽出部84で得られた発話時間情報とからトーンパラメータを抽出するためのトーンパラメータ抽出部を含むように構成することもできる。ここで、トーンパラメータとは、言葉の中に含まれる音程の上下をパラメータ化したものである。このような場合、トーンパラメータを用いてパラ言語情報を抽出することができる（特許文献1）。

【0053】

図15に、韻律特徴処理部80の詳細を機能ブロック図で示す。図15を参照して、韻律特徴処理部80は、発話音声信号から音程に関する情報であるパラメータF0を得るためのF0抽出部90と、パラメータF0を用いてある音節内のピッチの動き（方向と度合い）つまり音程の変化を半音単位で表わすパラメータであるF0moveを抽出するためのF0move抽出部92とを含む。F0抽出部90は、発話音声信号から音の高さに関する情報であるF0のみを抽出し、音階で表わす様に変換する。

【0054】

図16に、声質による音声処理部72の詳細を機能ブロック図で示す。図16を参照して、声質による音声処理部72は、発話音声信号からボーカル・フライを検出するためのボーカル・フライ検出部120と、全発話区間に占めるボーカル・フライ区間の割合を算出するためのボーカル・フライ割合算出部122とを含む。ここで、ボーカル・フライとは、声道の励振がほとんど減衰した事により生じる7Hz～78Hzくらいの非常に低い周波数のパルス音声の事である。

【0055】

声質による音声処理部72はさらに、与えられた発話音声信号のうちで、ボーカル・フライ区間以外でかつ音声波形が非周期である区間及びダブル周期性である区間の情報である非周期性区間情報及びダブル周期性区間情報を検出するための非周期性／ダブル周期性検出部124と、非周期性／ダブル周期性検出部124で検出された非周期性及びダブル周期性区間情報からボーカル・フライ検出部120で検出されたボーカル・フライ区間情報132を除き、これらの非周期性区間情報及びダブル周期性区間情報が全発話区間中で占める割合を算出するための非周期性／ダブル周期性割合算出部126とを含む。ここで、非周期性とは、音声波形が非周期的である事である。また、ダブル周期性とは、音声波形が、ピーク長及びピーク幅の異なる二つの波形からなる波形のセットが周期的に繰返された形状をもつ事をいう。

【0056】

声質による音声処理部72はさらに、与えられた発話音声信号から、気息性区間情報を検出するための気息性検出部128と、気息性区間が全発話区間中で占める割合を算出するための気息性割合算出部130とを含む。ここで、気息性とは、音声に含まれる息漏れの度合いの事である。気息性のある声としては例えば、ささやき声等が挙げられる。

【0057】

声質による音声処理部72はさらに、気息性検出部128から出力されるスペクトル傾斜値と、ボーカル・フライ割合算出部122から出力されるボーカル・フライ区間の割合

10

20

30

40

50

(ボーカル・フライ区間割合情報) とから、声質におけるりきみの割合を算出するりきみ割合算出部 40 を含む。りきみ割合算出部 40 は、スペクトル傾斜値が予め定められたしきい値よりも小さくかつボーカルフライであると(ボーカルフライ区間割合が高いと)、りきみ割合が高いと算出する。なお、スペクトルの傾斜等の詳細については後述する。また、多くの感動詞において、りきみ割合が高まると、驚き、感心、嫌悪等の感情または態度表現を強めることになる。

【0058】

図 17 に、ボーカル・フライ検出部 120 の詳細を機能ブロック図で示す。図 17 を参照して、ボーカル・フライ検出部 120 は、発話信号のうち 100 Hz ~ 1500 Hz の周波数成分のみを通過させるためのバンドパスフィルタ 140 と、バンドパスフィルタ 140 を通過した発話信号 154 を超短期フレーム長でフレーム化し、各フレームについて、その前後 2 フレームと比較してパワーが大きく、かつその差が予め定められたパワーしきい値よりも大きいフレームのフレーム位置を示す情報 150 をパワーピーク候補の位置情報として出力するための超短期ピーク検出処理部 142 と、発話信号 154 を短期フレーム長でフレーム化したものについてフレーム内周期性 (Intra-frame periodicity: IFP 値) に関する値を算出し、フレーム内周期性が所定個数以上存在するフレーム以外のフレームの IFP 値をヌルに設定するための短期周期性検出部 144 と、超短期ピーク検出処理部 142 から与えられたピーク位置情報 150 のうち、短期周期性検出部 144 から与えられた短期周期性情報 152 により、フレーム値がヌルとなっている部分の情報 156 のみを類似性検査部 148 に与えるための周期性検査部 146 と、情報 156 によって特定されるパワーピーク候補の付近の波形とその前のパワーピーク付近の波形との間のパルス間類似性 (inter-pulse similarity: IPS 値) に関する値が所定のしきい値以上であるもののピーク位置情報を検出し、このピーク位置情報に基づき、隣接するパルス間で IPS 値の高いものの間のフレームからボーカル・フライ区間情報を検出し、ボーカル・フライ割合算出部 122 と非周期性/ダブル周期性割合算出部 126 とに与えるための類似性検査部 148 とを含む。

【0059】

図 18 に、非周期性/ダブル周期性検出部 124 の詳細を機能ブロック図で示す。図 18 を参照して、非周期性/ダブル周期性検出部 124 は、発話音声信号をフィルタリング処理して音声波形のピークを検出する事によって、正規化自己相関関数を算出するための正規化自己相関関数算出部 160 と、正規化自己相関関数算出部 160 で算出された正規化自己相関関数に基づいた正規化自己相関関数の波形から、ピーク値またはピーク位置の関係等で表わされる正規化自己相関関数パラメータを算出するための正規化自己相関関数パラメータ算出部 162 と、算出された正規化自己相関関数パラメータの値から、非周期性及びダブル周期性区間情報を検出するための非周期性/ダブル周期性区間情報検出部 164 とを含む。

【0060】

正規化自己相関関数パラメータ算出部 162 では、正規化自己相関関数算出部 160 で得られた正規化自己相関関数より最初の 2 ピーク (P1 及び P2) を検出する。ただし、ピーク値は 0.2 を超えるもののみピークとみなす。

【0061】

これらのピークの正規化自己相関値を NAC (P1)、NAC (P2) 及び、正規化自己相関位置を TL (P1)、TL (P2) と呼び、正規化自己相関関数パラメータとして扱う。

【0062】

図 19 に、正規化自己相関関数算出部 160 の詳細を機能ブロック図で示す。図 19 を参照して、正規化自己相関関数算出部 160 は、発話信号のうち 60 Hz 以上の周波数成分のみを通すためのハイパスフィルタ 170 と、ハイパスフィルタ 170 の出力する音声信号の高域部分を強調する処理を行なうための高域強調部 172 と、高域強調部 172 の出力する音声信号に線型予測分析を行ない、声道パラメータ抽出部 174 で声道パラメー

タを抽出し、逆フィルタ176で、ハイパスフィルタ170の出力する音声信号に声道パラメータ抽出部174で抽出された声道パラメータを使用して、逆フィルタを行ない、声帯音源波形に対応する残差信号が得られると、後の処理に必要なピーク検出を容易にするために2kHz以下の音声信号のみを通すためのローパスフィルタ178と、ローパスフィルタ178を通った音声信号が与えられるとウィンドウの大きさを80msにし、そのウィンドウに含まれる音声信号から自己相関関数を算出するための自己相関関数算出部180と、自己相関関数算出部180で算出された自己相関関数の波形から、各々のフレームに含まれた最大のピークを検出するためのピーク検出部182と、ピーク検出部182で検出された最大ピークとその直前もしくは直後の最大ピークとの間の時間のずれを抽出し、ずれた時間の4倍の時間を1フレームとする様にフレーム長を再調節し、再調節されたフレームに含まれる自己相関関数の算出を行なうための自己相関関数再算出部184と、得られた自己相関関数を正規化する処理を行なうための正規化部186とを含む。

10

【0063】

図20に、気息性検出部128の詳細を機能ブロック図で示す。図20を参照して、気息性検出部128は、発話音声信号のうちで、100Hz～1500Hzの周波数成分のみを通過させるためのF1パスフィルタ202と、このF1パスフィルタ202を通過した波形全体から、振幅の変化を抽出するための振幅包絡抽出部204と、発話信号のうち、1800Hz～4000Hzの周波数成分のみを通過させるためのF3パスフィルタ200と、F3パスフィルタ200を通過した波形全体から、振幅の変化を抽出するための振幅包絡抽出部210と、振幅包絡抽出部204から得られた振幅の変化と振幅包絡抽出部210から得られた振幅の変化との間の相互相関を計算するための相互相関計算部214とを含む。ここで、F1パスフィルタ202を通過した周波数をF1波と呼び、F3パスフィルタ200を通過した周波数をF3波と呼ぶ。また、振幅包絡抽出部204で抽出された振幅の変化をF1振幅包絡と呼び、振幅包絡抽出部210で抽出された振幅の変化をF3振幅包絡と呼ぶ。

20

【0064】

気息性検出部128はさらに、F1パスフィルタ202を通過した成分からなるF1波から、最大周波数成分を抽出するための第1の最大周波数成分抽出部206と、F3パスフィルタ200を通過した成分からなるF3波から、最大周波数成分を抽出するための第2の最大周波数成分抽出部212と、F1波中に含まれる最大周波数成分とF3波中に含まれる最大周波数成分との差であるスペクトル傾斜A1-A3値を算出するためのスペクトル傾斜算出部216とを含む。なお、スペクトル傾斜A1-A3値は、スペクトル傾斜算出部216からりきみ割合算出部40へ出力される。

30

【0065】

気息性検出部128はさらに、相互相関計算部214から得られたF1F3相関値があるしきい値未満であり、かつ、スペクトル傾斜算出部216から得られたスペクトル傾斜A1-A3値があるしきい値未満であるか否かにより気息性区間か否かを判定し、気息性区間情報を出力するための気息性判定部218を含む。

【0066】

<動作>

40

図1を参照して、まず、ユーザが発話をする、その発話音声信号が図示しないマイクにより発話音声信号に変換される。マイクによって変換された発話音声信号は、韻律による音声処理部70と声質による音声処理部72とに与えられる。この韻律による音声処理部70での処理によって、発話持続時間についての情報及び音程の変化についての情報が得られる。声質による音声処理部72での処理によって、りきみの度合いに関する情報、発話全体に占めるボーカル・フライの割合に関する情報、非周期性及びダブル周期性の割合に関する情報、及び気息性の割合に関する情報が得られる。韻律による音声処理部70及び声質による音声処理部72での処理の詳細については後述する。

【0067】

図14を参照して、韻律による音声処理部70の動作の詳細について述べる。発話音声

50

信号を受信すると、韻律特徴処理部80では、まず、その発話音声信号をピッチの動きつまり音程の変化を表わすパラメータであるF0moveに変換する処理が行なわれる。F0moveは、音程に関する情報であるF0から得られる。

【0068】

図15を参照して、韻律特徴処理部80での動作の詳細について述べる。発話音声信号を受信すると、F0抽出部90では発話音声信号から音の高さに関する情報のみを抽出し、音階情報に変換してパラメータF0を得る。

【0069】

パラメータF0を用いてF0move抽出部92で、ある音節内のピッチの動き（方向と度合い）つまり音程の変化を半音単位で表わすパラメータであるF0moveが抽出される。F0moveは、複数のF0の差から求める事が可能である。

10

【0070】

図14を参照して、発話持続時間抽出部84で、発話音声信号から発話持続時間に関する情報が抽出される。

【0071】

韻律特徴処理部80で抽出されたF0moveと発話持続時間抽出部84で抽出された発話持続時間に関する情報とが、非語彙辞書50における韻律情報の発話持続時間と音程の変化とに対応している。

【0072】

図16を参照して、声質による音声処理部72は以下の様に動作する。まず、発話音声信号から、ボーカル・フライ検出部120でボーカル・フライ区間情報が検出される。

20

【0073】

図17を参照して、ボーカル・フライ検出部120は以下の様に動作する。バンドパスフィルタ140は、発話信号のうち100Hz～1500Hzの周波数成分のみを通過させる。バンドパスフィルタ140を通過した発話信号154は、超短期ピーク検出処理部142、短期周期性検出部144及び類似性検査部148に与えられる。超短期ピーク検出処理部142は、発話信号154を超短期フレーム化し、各フレームに対し超短期パワーを算出する。そして、各フレームについて、その前後2フレームと比較してパワーの差がパワーしきい値よりも大きいフレームをパワーピーク候補とし、そのフレーム位置を示す情報150を出力する。

30

【0074】

短期周期性検出部144は、発話信号154をフレーム化し、その各フレームについてIFP値を算出する。算出されたIFP値としきい値とを比較し、しきい値未満であれば、そのフレームのIFP値をヌルに設定する。ヌルではないフレームが少なくとも3フレームだけ連続していなければ、それらのフレームのIFP値をヌルに補正する。そして補正されたIFP値が周期性検査部146に与えられる。

【0075】

周期性検査部146は、超短期ピーク検出処理部142から与えられたピーク位置情報150のうち、短期周期性検出部144から与えられた短期周期性情報152により、フレームIFP値がヌルとなっている部分の情報156のみを類似性検査部148に与える。

40

【0076】

類似性検査部148は、情報156によって特定される区間に存在するパワーピーク候補の各パワーピーク付近の波形とその前のパワーピーク付近の波形との間のIPS値を算出する。そしてそのIPS値としきい値とを比較し、しきい値以上のパワーピークのピーク位置情報を検出する。このピーク位置情報に基づき、隣接するパルス間でIPS値の高いものの間のフレームをボーカル・フライ区間として検出し、それらを示す情報（ボーカル・フライ区間情報）を出力する。

【0077】

図16を参照して、検出されたボーカル・フライ区間情報はボーカル・フライ割合算出

50

部122に与えられる。ボーカル・フライ区間情報から、全発話区間中でボーカル・フライ区間の占める割合がボーカル・フライ割合算出部122で算出される。この算出はボーカル・フライ区間を全発話区間で割る事によって得られる。算出されたボーカル・フライ区間割合情報は、後の処理のために、りきみ割合算出部40及びパラ言語情報抽出部74に与えられる。

【0078】

非周期性／ダブル周期性検出部124により、発話音声信号のうちで、音声波形が非周期である区間及びダブル周期である区間の情報である非周期性区間及びダブル周期性区間が検出され、それらを示す非周期性区間情報及びダブル周期性区間情報が出力される。

【0079】

図18を参照して、非周期性／ダブル周期性検出部124は以下の様に動作する。発話音声信号が与えられると、正規化自己相関関数算出部160は、その音声信号をフィルタリング処理した音声波形を解析する事によって自己相関関数を算出する。そしてその自己相関関数を正規化し正規化自己相関関数を算出する。この正規化自己相関関数算出部160での処理の詳細については以下に述べる。

【0080】

図19を参照して、発話信号が与えられると、ハイパスフィルタ170によって、60Hz以上の周波数成分のみが通過させられる。60Hz以上の音声信号は、高域強調部172と逆フィルタ176とに与えられる。高域強調部172は与えられた音声信号の高域部分を強調する処理を行なう。そして、声道パラメータ抽出部174で、声道を特徴付けるフィルタパラメータを推測する。その後、ハイパスフィルタ170の出力音声信号に声道パラメータ抽出部174で与えられた声道パラメータを用いて、声帯音源信号を求めるために逆フィルタ176を行なう。

【0081】

逆フィルタ176で処理された残差信号は、次にローパスフィルタ178に与えられる。このローパスフィルタ178は、後の処理に必要となるピーク検出を容易にするために2kHz以下の周波数成分のみを通過させる。ローパスフィルタ178を通過した周波数成分は、自己相関関数算出部180と自己相関関数再算出部184とに与えられる。自己相関関数算出部180では、検出処理の際に使用するフレームの大きさを80msとし、フレーム中の音声信号波形から自己相関関数を得る。そしてこの自己相関関数を出力する。

【0082】

ピーク検出部182では、自己相関関数算出部180で得られた自己相関関数に含まれた最大のピークを検出する処理が行なわれる。

【0083】

自己相関関数再算出部184では、まず、ピーク検出部182で検出された最大ピークの位置の4倍の時間を新しいフレーム長とする。この様なフレームの再調節が行なわれるのは、自己相関関数の適切な算出を行なうためである。つまり、固定のフレーム長の場合、フレームが大きすぎても小さすぎても自己相関関数の適切な算出をする事が難しいからである。そして、そのフレームから再度自己相関関数を得る。

【0084】

次に、正規化部186で得られた自己相関関数を正規化する処理を行なう。図18を参照して、正規化自己相関関数算出部160で算出された正規化自己相関関数に基づいて、正規化自己相関関数パラメータ算出部162での算出処理が行なわれる。そして、音波の非周期性及びダブル周期性を抽出するために正規化自己相関関数の波形から、ピーク値及びピーク位置を検出する。そしてその後それらピーク値の比率とピーク位置の比率とを算出する。ピーク値の比率は、 $1000 * NAC(P2) / NAC(P1)$ で求められる。また、ピーク位置の比率は $2000 * TL(P2) / TL(P1)$ で求められる。

【0085】

さらに、算出された正規化自己相関関数パラメータを使用して、非周期性／ダブル周期

10

20

30

40

50

性区間情報検出部 164 で当該音声信号が非周期性もしくはダブル周期性を持つ区間が検出される。この検出処理の詳細は以下の通りである。

【0086】

つまり、上述した自己相関関数パラメータがいずれも 1000 に近似した値であれば、その自己相関関数の波形で表わされる区間の発話音声波形は周期性を持つと言える。そこで、それ以外の値を取る発話区間を非周期性及びダブル周期性区間として抽出する事ができる。

【0087】

非周期性／ダブル周期性区間情報検出部 164 で検出された非周期性／ダブル周期性区間情報が非周期性／ダブル周期性割合算出部 126 に与えられる。

10

【0088】

図 16 を参照して、全発話区間中で非周期性区間及びダブル周期性区間の占める割合が非周期性／ダブル周期性割合算出部 126 で算出される。この算出は、非周期性区間及びダブル周期性区間を全発話区間で割る事によって行なわれる。

【0089】

この算出処理の前にまず、ボーカル・フライ検出部 120 で、ボーカル・フライ区間として検出された区間情報を非周期性／ダブル周期性区間情報から除去する処理が行なわれる。ボーカル・フライも非周期性特徴を持つが、ここでは、ボーカル・フライ以外の非周期性／ダブル周期性を対象としているからである。

【0090】

20

図 20 を参照して、気息性検出部 128 は以下の様に動作する。発話音声信号が与えられると、F1 パスフィルタ 202 は、まず、その発話音声信号のうち、100 Hz ～ 1500 Hz の周波数成分のみを通過させる。振幅包絡抽出部 204 では、F1 パスフィルタ 202 を通った F1 波の波形から、振幅包絡を抽出する。

【0091】

F3 パスフィルタ 200 でも同様に、発話音声信号のうち、1800 Hz ～ 4000 Hz の周波数成分のみを通過させる。そして振幅包絡抽出部 210 では、F3 パスフィルタ 200 を通った F3 波の波形から、振幅包絡を抽出する。

【0092】

振幅包絡抽出部 204 から得られた F1 振幅包絡と振幅包絡抽出部 210 から得られた F3 振幅包絡との相互相関を相互相関計算部 214 で計算する。この処理により、F1 振幅包絡と F3 振幅包絡の相互の関係を示す F1 F3 相関値が得られる。

30

【0093】

F1 パスフィルタ 202 を通過した F1 波からはまた、最大周波数成分抽出部 206 でこの周波数帯域中に含まれるもののうち最大の周波数成分が抽出される。そして、F3 パスフィルタ 200 を通過した F3 波にも、最大周波数成分抽出部 212 で同様の処理が行なわれる。F1 波中に含まれる最大周波数成分と F3 波中に含まれる最大周波数成分との差、つまりスペクトル傾斜を算出する処理がスペクトル傾斜算出部 216 で行なわれる。このスペクトル傾斜を A1 - A3 とする。なお、算出されたスペクトル傾斜 A1 - A3 値は、りきみ割合算出部 40 に出力される。

40

【0094】

気息性判定部 218 では、F1 F3 相関値とスペクトル傾斜 A1 - A3 値とを用いて気息性であるか否かを判定して、気息性区間情報を出力する。ここでの処理では、F1 F3 相関値があるしきい値未満で、かつ A1 - A3 値があるしきい値未満であれば、気息性区間であると判定する。これらのしきい値は予め学習によって得られる。このしきい値と実際に得られた F1 F3 相関値と A1 - A3 とを比較参照する事により、気息性の有無が判定できる。

【0095】

気息性区間情報は、気息性割合算出部 130 に与えられる。図 16 を参照して、気息性割合算出部 130 は、全発話区間中で気息性区間の占める割合を、気息性区間を全発話区

50

間で割る事によって算出する。算出された気息性区間割合は、後の処理のためにパラ言語情報抽出部74に与えられる。

【0096】

図20に示す気息性検出部128から出力されたスペクトル傾斜A1-A3値及びボーカル・フライ割合算出部122から出力されたボーカル・フライ区間割合情報とに基づいて、図16に示すりきみ割合算出部40が、りきみの割合（りきみの有無）を判定する。このとき、りきみ割合算出部40は、スペクトル傾斜値が予め定められたしきい値よりも小さくかつボーカルフライであると（ボーカルフライ区間割合が高いと）、りきみ割合が高く、りきんでいると（りきみがある）と算出する。

【0097】

このような動作により、発話持続時間抽出部84により韻律パラメータとしての発話持続時間が算出され、韻律特徴処理部80により韻律パラメータとしての音程の変化（パラメータであるF0move）が算出され、りきみ割合算出部40により声質パラメータとしての「りきみ」の有無が判定され、非周期性／ダブル周期性割合算出部126により声質パラメータとしての「非周期性」の有無及び「周期性（連続性）」の有無が判定され、気息性割合算出部130により声質パラメータとしての「気息性」の有無が判定される。

【0098】

このような韻律パラメータ（韻律情報）及び声質パラメータ（声質情報）と、非語彙辞書50（図2～図13）とを用いて、パラ言語情報抽出部74でパラ言語情報として発話意図が抽出される。

【0099】

たとえば、図7に示すように、感動詞「は・はーん」グループについては、韻律情報（発話持続時間及び音程変化）に加えて、声質情報の「周期性」を加えて判断することにより発話意図として「謙遜」を抽出できたり、「りきみ」を加えて判断することにより発話意図として「強い感心・強い驚き」を抽出できたり、「気息音」を加えて判断することにより発話意図として「ため息」を抽出できたりする。

【0100】

ここでの処理においては、発話持続時間、音程変化、りきみ割合、非周期性及び（ダブル）周期性区間の割合及び、気息性の割合に関する情報とパラ言語情報（発話意図）との関係に関するデータを集積する必要がある。この集積されたデータによってさらに、どの様なパラメータが入力されれば、どの様なパラ言語情報が検出できるかというモデルを学習によって作成する事ができる。

【0101】

このモデルに使用されるものとしては、決定木（分類木）、ニューラルネットワーク、SVM（Support Vector Machine）、Gaussian Model、Bayesian Model等が考えられる。

【0102】

[コンピュータによる実現]

この実施の形態のシステムは、コンピュータハードウェアと、そのコンピュータハードウェアにより実行されるプログラムと、コンピュータハードウェアに格納されるデータとにより実現される。図21はこのコンピュータシステム330の外観を示し、図22はコンピュータシステム330の内部構成を示す。

【0103】

図21を参照して、このコンピュータシステム330は、CD-ROM（コンパクトディスク読出専用メモリ）ドライブ350を有するコンピュータ340と、キーボード346と、マウス348と、モニタ342とを含む。

【0104】

図22を参照して、コンピュータ340は、CD-ROMドライブ350に加えて、CPU（中央処理装置）356と、CPU356及びCD-ROMドライブ350に接続されたバス366と、ブートアッププログラム等を記憶する読出専用メモリ（ROM）35

10

20

30

40

50

8と、バス366に接続され、プログラム命令、システムプログラム、及び作業データ等を記憶するランダムアクセスメモリ（RAM）360とを含む。コンピュータシステム330はさらに、プリンタ344を含んでいる。

【0105】

ここでは示さないが、コンピュータ340はさらにローカルエリアネットワーク（LAN）への接続を提供するネットワークアダプタボードを含んでもよい。

【0106】

コンピュータシステム330にパラ言語情報抽出装置60としての動作を行なわせるためのコンピュータプログラムは、CD-ROMドライブ350に挿入されるCD-ROM362に記憶され、さらにハードディスク354に転送される。または、プログラムは図示しないネットワークを通じてコンピュータ340に送信されハードディスク354に記憶されてもよい。プログラムは実行の際にRAM360にロードされる。CD-ROM362から、またはネットワークを介して、直接にRAM360にプログラムをロードしてもよい。

【0107】

このプログラムは、コンピュータ340にこの実施の形態のパラ言語情報抽出装置60として動作を行なわせる複数の命令を含む。この動作を行なわせるのに必要な基本的機能のいくつかはコンピュータ340上で動作するオペレーティングシステム（OS）もしくはサードパーティのプログラム、またはコンピュータ340にインストールされる各種ツールキットのモジュールにより提供される。従って、このプログラムはこの実施の形態のシステム及び方法を実現するのに必要な機能全てを必ずしも含まなくてよい。このプログラムは、命令のうち、所望の結果が得られるように制御されたやり方で適切な機能または「ツール」を呼出すことにより、上記したパラ言語情報抽出装置60としての動作を実行する命令のみを含んでいればよい。コンピュータシステム330の動作は周知であるので、ここでは繰返さない。

【0108】

以上の様に、感動詞についてのパラ言語情報を検出する際に、韻律に関する情報及び声質に関する情報を使用する事により、パラ言語情報の検出精度が高くなる。特に、この声質情報における、りきみ割合、非周期性及び（ダブル）周期性区間の割合、及び、氣息性の割合に関する情報とパラ言語情報（感動詞における発話意図）との関係が規定された非語彙辞書を用いて、感動詞のパラ言語情報の検出精度を向上させることができる。

【0109】

今回開示された実施の形態に使用された具体的な数字は例示である。

【0110】

また、今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味及び範囲内でのすべての変更を含む。

【符号の説明】

【0111】

- 40 りきみ割合算出部
- 50 非語彙辞書
- 60 パラ言語情報検出装置
- 70 韻律による音声処理部
- 72 声質による音声処理部
- 74 パラ言語情報抽出部
- 122 ボーカル・フライ割合算出部
- 126 非周期性／ダブル周期性割合算出部
- 130 氣息性割合算出部

10

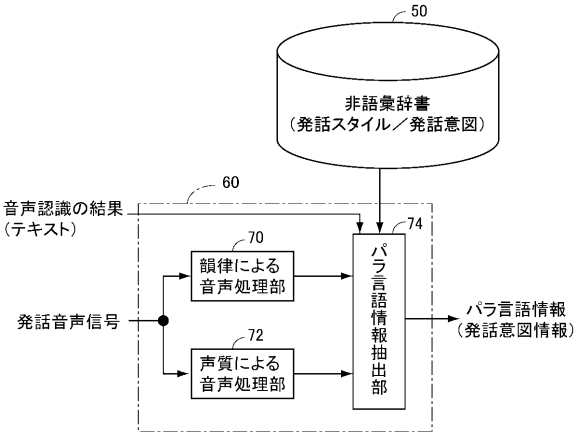
20

30

40

50

【図 1】



【図 2】

え・えー	発話スタイル					発話意図
	韻律情報		声質情報			
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	
え	短い	下降調				肯定的反応(肯定、同意、承諾、理解、相槌等)
	短い	上昇調				聞き返し、意外
	長い	上昇調	(あり)			意外、驚き
えー	弱く長い	平坦調				否定的反応(不満、嫌悪、非難、疑い等)
	長い	平坦調				考え中、発話準備
	長い	平坦調／上昇調				感心、同情
え	短い	下降調	あり		あり	深い驚き、深い同情 同意、理解

【図 3】

うん・ふん	発話スタイル					発話意図
	韻律情報		声質情報			
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	
うん・ふん	短い	下降調				肯定的反応(肯定、同意、承諾、理解、相槌等)
	短い	上昇調				聞き返し、意外
	短い		(あり)			意外、驚き
うーん・ふーん	長い	上昇調				否定的反応(不満、嫌悪、非難、疑い等)
	弱く長い	平坦調				考え中、発話準備
	長い	平坦調／上昇調				感心、同情
うん・ふん	短い	下降調			あり	同意、理解
うーん		上昇下降調				否定(打ち消し)
うーん・ふーん	長い		あり			深く考えている、深く悩んでいる、深い躊躇、深い困惑

【図 4】

あ、あー	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	気息性	
あ	短い	平坦調					気づき、思い出した時
	短い	平坦調				あり	気づき＋驚き
	長い	下降調					理解、同意、肯定
あー	長い	下降調 (ピッチの立て直しが小)			あり		理解、同意(強い肯定)
	長い	下降調 (ピッチの立て直しが明確)			あり		同情、残念、落胆(がっかり)
	長い	下降上昇調					同情、残念、落胆(がっかり)

【図 5】

	発話スタイル						発話意図
	韻律情報			声質情報			
	発話持続時間 (発話状態)	音程の変化	きみ	非周期性 (連続性)	周期性 (連続性)	気息性	
へ、へー	短い	下降調					肯定的反応(肯定、同意、承諾、理解、相槌等)
へ	短い	上昇調					聞き返し
	短い	上昇調				あり	意外、驚き
	長い	上昇調					相槌(対話相手に興味又は関心)
へー	長い	単独で単調					聞き流し、無関心
	長い		あり				深い関心、深い驚き

【図 7】

	発話スタイル					発話意図	
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)		気息性
はー・はーん	短い／長い	平坦調／下降調					相槌、理解
はー	長い				あり		謙遜
	長い						感心、驚き、理解
	長い	下降調	あり				強い感心、強い驚き
	短い (柔らかい)	上昇調				あり	ため息
はー	長い (硬い)	上昇調					聞き返し
		上昇調					否定的反応(非難)

【図 6】

	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性	周期性 (連続性)	気息性	
はい、はい	短い	上昇調					聞き返し
はいはい…		下降調			あり		理解、共感
		最後の「ない」でピッチの 立て直しが明確			あり		邪魔くさい、もう分かった

【図 8】

発話スタイル	発話スタイル					発話意図
	韻律情報		声質情報			
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	
お・おーん	短い					驚き、気づき
おー・おーん	長い					理解(「あー」、「うん」のカジュアル形)
おお・おん				あり		強い理解

【図 9】

ま・ほ	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	氣息性	
ほん		平坦下降調					相植、理解(「うん」「ふん」のカジュアル形)
ほー	長い		あり				深い感心、深い驚き
ほー	長い	下降調				あり	ため息

【図 10】

いや・やー	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	氣息性	
いや	短い						否定的応答
いや・やー			あり				深い驚き、深い困惑、深い躊躇
いや・やー						あり	強い驚き

【図 11】

	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性	周期性 (連続性)	氣息性	
あら・ありや		上昇調					疑問、疑い
あらー	長い					あり	同情、残念、落胆(がっかり)
あらあら						あり	望ましくない出来事に対する反応

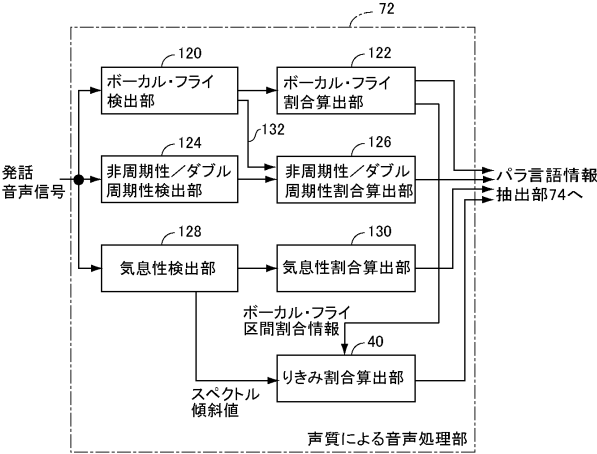
【図 12】

	発話スタイル						発話意図
	韻律情報		声質情報				
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	気息性	
あれ							
あれー	長い	上昇調					疑問、疑い
あれー							強い疑問、強い疑い

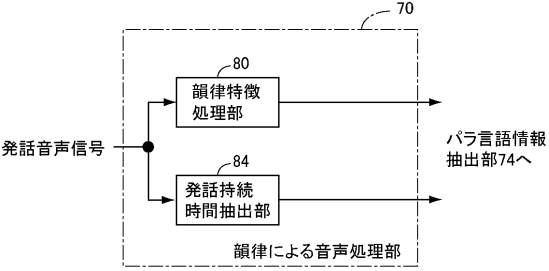
【図 1 3】

	発話スタイル					発話意図
	韻律情報		声質情報			
	発話持続時間 (発話状態)	音程の変化	りきみ	非周期性 (連続性)	周期性 (連続性)	
わ・うわー						驚き、感心、嫌悪 強い驚き、強い感心、強い嫌悪 強い驚き、強い感心、強い嫌悪
わ・うわー						
わ・うわー			あり			
わ・うわー					あり	
わ・うわー						

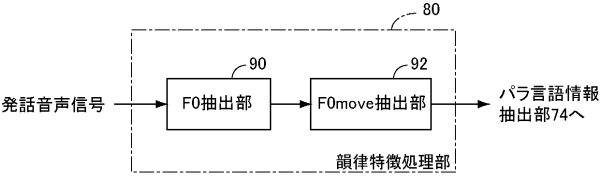
【図 1 6】



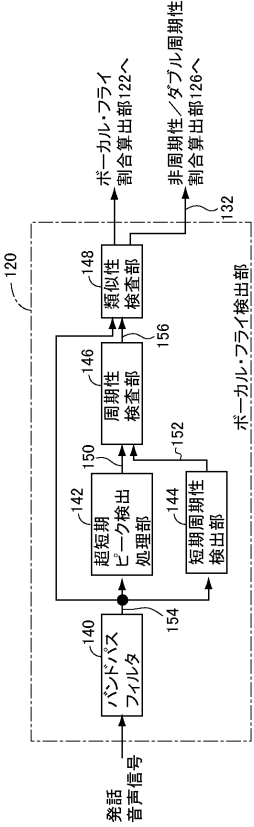
【図 1 4】



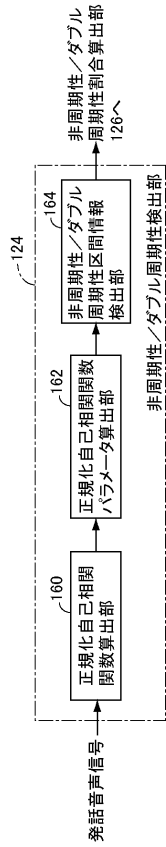
【図 1 5】



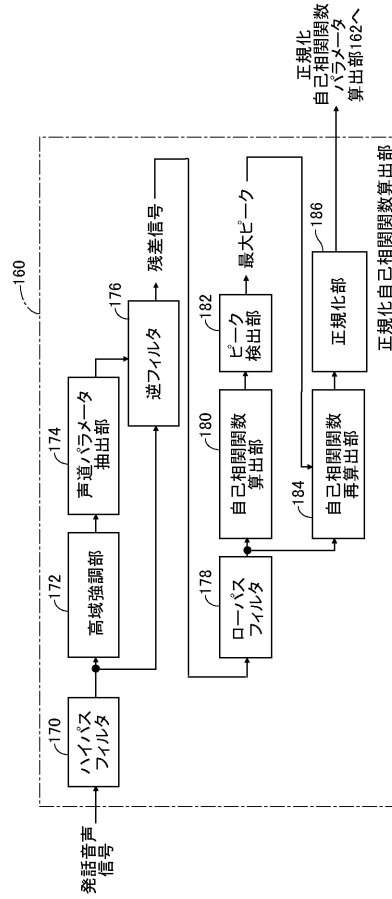
【図 1 7】



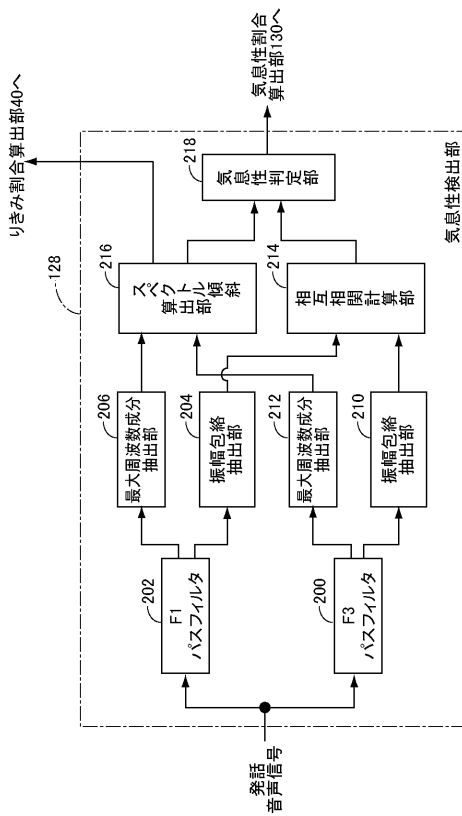
【図 18】



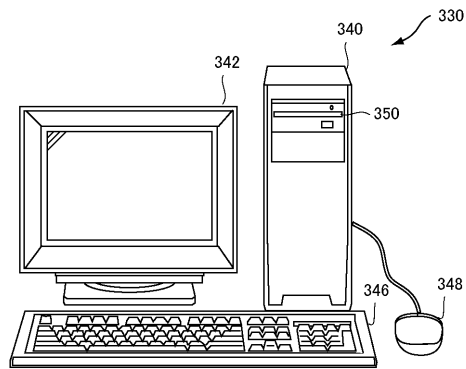
【図 19】



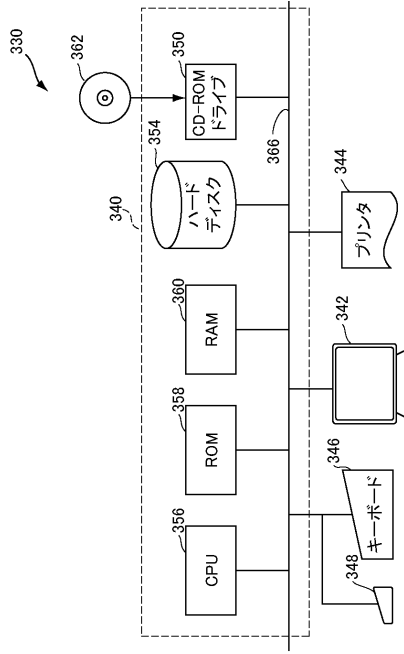
【図 20】



【図 21】



【図 22】



フロントページの続き

審査官 毛利 太郎

- (56)参考文献 特開2007-079363 (JP, A)
特開2003-330478 (JP, A)
特開2009-003162 (JP, A)
石井 カルロス 寿憲 Carlos T. ISHI, 自然発話に現れる感動詞の発話スタイルと機能の分析, 日本音響学会 2008年 秋季研究発表会講演論文集CD-ROM [CD-ROM], 日本, 社団法人日本音響学会, 2008年 9月10日, p.269-270
石井 カルロス寿憲 Carlos T. ISHII, 喉頭を力んだ発声の音響特徴の分析, 日本音響学会 2006年 春季研究発表会講演論文集CD-ROM [CD-ROM], 日本, 社団法人日本音響学会, 2006年 3月16日, p.227-228
石井 カルロス 寿憲 Carlos Toshinori ISHI, りきみの自動検出のための音響分析 Acoustic analysis for automatic detection of pressed voice, 電子情報通信学会技術研究報告 Vol. 106 No. 178 IEICE Technical Report, 日本, 社団法人電子情報通信学会 The Institute of Electronics, Information and Communication Engineers, 2006年 7月14日, p.1-6
石井 カルロス寿憲 Carlos Toshinori ISHI, Vocal Fry発声区間の自動検出法 A Method for Automatic Detection of Vocal Fry Segments, 電子情報通信学会論文誌 (J89-D) 第12号 THE IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS (JAPANESE EDITION), 日本, 社団法人電子情報通信学会 THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 2006年12月 1日, p.2679-2687
石井 カルロス寿憲 Carlos T. ISHI, EGGを用いた「りきみ」発声の音響分析, 日本音響学会 2007年 春季研究発表会講演論文集CD-ROM [CD-ROM], 日本, 社団法人日本音響学会, 2007年 3月13日, p.221-222
- (58)調査した分野(Int.Cl., DB名)
G10L 15/00-15/34
G10L 25/00-25/93