

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-212641  
(P2004-212641A)

(43) 公開日 平成16年7月29日(2004.7.29)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
G10L 15/06	G10L 3/00 521T	5D015
G06F 3/16	G06F 3/16 320Z	
G10L 15/00	G10L 3/00 551P	
G10L 15/20	G10L 3/00 571J	
G10L 15/28	G10L 3/02 301D	

審査請求 未請求 請求項の数 14 O L (全 23 頁) 最終頁に続く

(21) 出願番号	特願2002-382028 (P2002-382028)	(71) 出願人	000003078 株式会社東芝 東京都港区芝浦一丁目1番1号
(22) 出願日	平成14年12月27日 (2002.12.27)	(74) 代理人	100076233 弁理士 伊藤 進
		(72) 発明者	蟻生 政秀 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内
		Fターム(参考)	5D015 EE05 LL10

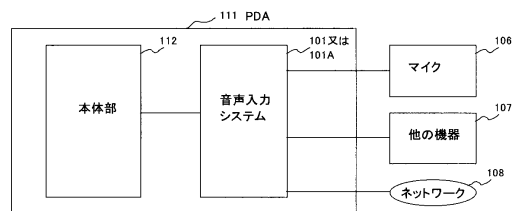
(54) 【発明の名称】 音声入力システム及び音声入力システムを備えた端末装置

(57) 【要約】

【課題】 周囲状況に合わせた音声処理を効率よく行えるようにする。

【解決手段】 時計機能による時間情報を取得する(301)。その時間情報から環境情報を参照し(302)、信号処理を行う内容を決定する(303)。入力した音声信号(304)について、環境情報を用いることで入力音声信号の処理(305)を効率よく行い、その処理結果を出力する(306)。

【選択図】 図4



## 【特許請求の範囲】

## 【請求項 1】

音声信号を受信する受信手段と、  
前記音声信号に対して信号処理を施す信号処理手段と、  
時間に関連づけられた環境情報を記憶する記憶手段と、  
時間を計測する時間計測手段と、  
前記計測された時間から関連する環境情報を前記記憶手段から取り出し、当該環境情報に基づいて前記信号処理手段を制御する制御手段と、  
を備えたことを特徴とする音声入力システム。

## 【請求項 2】

さらに、前記信号処理に用いるパラメータを時間に関連づけて記憶する手段を具備し、  
前記制御手段は、前記環境情報だけでなく、前記パラメータに基づいて、入力音声の前記信号処理を制御することを特徴とする請求項 1 に記載の音声入力システム。

## 【請求項 3】

さらに、前記信号処理の結果を反映して前記環境情報や前記パラメータの内容を変更する手段を具備したことを特徴とする請求項 2 に記載の音声入力システム。

## 【請求項 4】

前記環境情報が場所の情報であることを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 5】

前記環境情報が人間の情報であることを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 6】

前記信号処理の内容が、入力対象である音声以外の雑音の影響を抑圧する処理であることを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 7】

さらに、音声認識を行う手段を具備し、  
前記環境情報や前記パラメータを基に前記信号処理を行うと共に前記音声認識を行うことを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 8】

さらに、他の音声入力システムと情報の授受を行う手段を具備し、  
他の音声入力システムの前記環境情報や前記パラメータを利用して、自身の音声入力システムの前記環境情報や前記パラメータを更新可能としたことを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 9】

さらに、音声信号以外のセンサ情報を入力する手段を具備し、  
前記センサ情報を利用して、自身の音声入力システムの前記環境情報や前記パラメータを更新可能としたことを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 10】

さらに、記憶手段若しくは信号処理手段の一部を他の音声入力システムと共有化することを特徴とする請求項 1 乃至 3 のいずれか 1 つに記載の音声入力システム。

## 【請求項 11】

音声信号を受信する受信手段と、  
時間に関連づけられた環境情報を記憶する記憶手段と、  
時間を計測する時間計測手段と、  
前記音声信号に対して信号処理を施す信号処理手段と、  
時間に関連づけられた前記環境情報を記憶する記憶領域がシステム外部にある場合に、その環境情報を読み出すためのアドレス情報を時間と関連づけて記憶する記憶部を有し、前記計測された時間から関連する環境情報を前記外部の記憶領域から読み出し、当該環境情

10

20

30

40

50

報に基づいて前記信号処理手段を制御する制御手段と、  
を備えたことを特徴とする音声入力システム。

【請求項 1 2】

音声信号を受信する受信手段と、  
前記音声信号に対して信号処理を施す信号処理手段と、  
時間に関連づけられた環境情報を記憶する記憶手段と、  
時間を計測する時間計測手段と、  
前記計測された時間から関連する環境情報を前記記憶手段から取り出し、当該環境情報に基づいて前記信号処理手段を制御する制御手段と、  
を具備したことを特徴とする、音声入力システムを備えた端末装置。

10

【請求項 1 3】

さらに、前記信号処理に用いるパラメータを記憶する手段を具備し、  
前記制御手段は、前記環境情報だけでなく、前記パラメータに基づいて、入力音声の前記信号処理を制御することを特徴とする請求項 1 2 に記載の音声入力システムを備えた端末装置。

【請求項 1 4】

さらに、前記信号処理の結果を反映して前記環境情報や前記パラメータの内容を変更する手段を具備したことを特徴とする請求項 1 3 に記載の音声入力システムを備えた端末装置。

【発明の詳細な説明】

20

【0001】

【発明の属する技術分野】

本発明は、使用状況に合わせて、適切な信号処理を行い、常に好適な音声信号を取得できるようにする音声入力システム及び該システムを備えた端末装置に関する。

【0002】

【従来の技術】

近年、電子機器回路技術の向上により、ウェアラブル・コンピュータや、個人用携帯情報端末（以下、PDA、Personal Digital Assistantの略）、ハンドヘルドコンピュータなどの情報処理機器が身近なものとなっている。このような機器では、機器とユーザ間のインタフェースとして音声は重要な役割を果たしている。

30

【0003】

以下、音声を電子機器で扱う装置、方式、プログラムの総称を「音声入力システム」とする。ユーザが電子機器を使用する様々な状況において、適切に音声を処理し、明確な音声信号を取得することが音声入力システムに求められる。

【0004】

例えば雑踏の中で話された音声を聞き取ることは、現在のコンピュータ技術にとっては難しく（人間同士なら必ずしも難しいことではないが）、様々な状況の中で適切な音声処理（信号処理）を行う必要がある。

【0005】

例えば、PDAの操作を音声で行う場合には、静かなオフィス環境で入力される音声と雑踏において入力される音声とでは、入力される音声の特性が異なるため、音声操作を行った場合、静かな環境と雑踏の環境とで同一の音声処理を行ったのでは十分な期待通りの操作性能が得られないことが予想される。

40

【0006】

何故なら、周囲が静かな環境とそうでない環境では音声の信号対雑音比（以下、SN比）が異なることや、ユーザの話し方（囁き声になったり、大声になったり等）が変化するためである。そのため、入力された音声のSN比に合わせて雑音を抑圧したり、音声をフィルタリングして変動を吸収したりするといった、周囲の環境変化に合わせた音声処理が必要になってくる。

【0007】

50

このような場合に、従来の解決策としては、一般的には、適応的信号処理によってどんな周囲状況でもある程度対応できるような信号処理を行っているものがある（例えば、非特許文献1参照）。具体的な例としては、音声から随時に周囲雑音を推定して影響を入力音声から除去することで、周囲状況が変化しても雑音を抑圧できるようなものが挙げられる。

【0008】

しかし、このような適応的信号処理はどのような周囲状況にもある程度対応できる反面、適応に時間がかかったり、周囲状況の変化が大きい場合には漸次的な適応処理では対応できないという問題点がある。

【0009】

それに対して、適応的処理に用いるパラメータの初期値を周囲状況に合わせてユーザの操作や音声入力システムの上位システムによって与えてしまえば、適応までの時間も少なく済むし、処理の誤差も軽減されると考えられる。

【0010】

また、適応的信号処理でなくとも、周囲状況に合わせた信号処理用パラメータを用いることが音声入力システムにとって有益であることは勿論である。ただし、この場合、従来では、音声入力システムの操作者が周囲状況を判断して、状況に合わせた信号処理の設定を行っていたため、操作の煩雑さや、余分で複雑な処理が必要となる場合があった。

【0011】

一方、使用状況に応じて音声を処理するという目的なら、例えば状況を判断するのに時間のみを用い、音声入力時の時刻を含む時間帯に応じて装置の機能を切り替え、その機能状態に応じて音声認識可能な音声（すなわち、受理可能な音声）を決めているものもある（例えば、特許文献1参照）。

【0012】

また、スケジュールに従って通話の音声以外の音声を付け加えるものもある。すなわち、プライバシー保護の観点から予めスケジュールされた生活環境音を携帯電話における発声音に重畳して送信するというものである（例えば、特許文献2参照）。

【0013】

【特許文献1】

特開平8-190470号公報（第1-5頁、図1）

【0014】

【特許文献2】

特開2002-27136号公報（第8-10頁、図10）

【0015】

【非特許文献1】

サイド・バイ・ヴァセッジ (Saeed V. Vaseghi) 著, 「アドバンスド デジタル シグナル プロセッシング アンド ノイズ リダクション (Advanced Digital Signal Processing and Noise Reduction)」, (英国), 第二版, ワイリー (WILEY), 2000年9月

【0016】

【非特許文献2】

古井貞熙著 音響・音声工学 近代科学社 1992年刊のpp176-219

【0017】

【発明が解決しようとする課題】

前述のように、信号処理においては、ユーザの操作や音声入力システムの上位システムによってパラメータを与えるか、あるいは適応的信号処理によって周囲状況に合わせてある程度の処理を行うことができた。その手法については非特許文献1を始めとして既存の手法がある。ただし、ユーザ操作や上位システムでパラメータを与えるには煩雑であるし、適応的信号処理では適応時間が掛かる等の問題があった。

10

20

30

40

50

## 【0018】

また、システムの置かれた状況は時間だけでは判断できない場合があるが、先に挙げた特許文献1では時間以外の情報に応じた信号処理を行うことができなかった。

## 【0019】

さらに、特許文献2では、携帯電話の使用においてプライバシーを保護するのが主眼であって、例えば1日のスケジュールに合わせて生活環境の音を発声音に重畳することによって、通話時にユーザの音声と共にその時の周囲の音声をリアルな状態で送信してしまうことのないようにしている。

## 【0020】

従って、特許文献2では、通話の音声に対してスケジュールに合わせた生活環境音（例えば、雑踏の音、駅構内の音、空港の音など）を重畳しているため、次のような不具合を生じていた。すなわち、スケジュール上での環境がオフィス、実際の環境が雑踏であった場合に、通話している相手側へ出力される音声は（発声＋オフィスの雑音＋雑踏の雑音）、あるいは実際の環境が駅のホームであった場合には（発声＋オフィスの雑音＋駅構内の雑音）となる。また、実際の環境の背景音がスケジュールされた重畳音よりも特異的、もしくは音量が大きい場合には生活環境音を重畳しても実環境の背景音の方が支配的になりやすいという問題がある。

## 【0021】

そこで、本発明は以上の問題に鑑みてなされたもので、時間に関連づけられた環境に応じて入力する音声を適切に信号処理し、常に明確な音声信号を得ることが可能な音声入力システム及び音声入力システムを備えた端末装置を提供することを目的とするものである。

## 【0022】

## 【課題を解決するための手段】

本発明に係る音声入力システムは、音声信号を受信する受信手段と、前記音声信号に対して信号処理を施す信号処理手段と、時間に関連づけられた環境情報を記憶する記憶手段と、時間を計測する時間計測手段と、前記計測された時間から関連する環境情報を前記記憶手段から取り出し、当該環境情報に基づいて前記信号処理手段を制御する制御手段と、を具備したものである。

## 【0023】

本発明において、受信手段は、マイクのような音声を計測する機器からの音声信号や、直接に接続された機器からの音声信号、あるいはネットワークを介して接続された機器からの音声信号を音声入力システムに取り込むものである。

## 【0024】

時間計測手段は、システム内で時間の進行をカウントする手段によって時間情報を得る、または電波時計のように音声入力システム外の時間信号を取り込むことによって時間情報を得るものとする。ここで言う時間とは、計測開始時点からの経過時間のような相対的な形の時間情報でも、年月日及び時刻のような絶対的な形の時間情報でも構わない。

## 【0025】

時間に関連づけられた環境情報とは、時間と共に推移する情報、所定の時間帯に対応付けられる情報、時間を変数とする関数的情報などであり、例えばスケジュール情報である。従って、予め時間の経過に対する状況変化が分かっている場合は、スケジュール情報と同様に扱うことができる。なお、予め時間と環境情報の対応関係がない状況（不意に起こる状況の変化や予定の範囲外の位置的な変化など）は、後述するセンサ情報を用いてそれに合わせて環境情報を更新することによって対処する。スケジュール情報は、その属性として、時間と、これに関連した場所や人間（より具体的には、時間に関連して行く予定の場所や会う予定の人間など）の情報を含んだものである。

## 【0026】

信号処理手段は、雑音処理，音声強調，音声認識などの各種の信号処理を行うものである。時間に関連づけられた環境情報に基づいて適切な信号処理を行える。

## 【0027】

制御手段は、計測された時間から関連する環境情報を前記記憶手段から取り出し、当該環境情報に基づいて前記信号処理手段を制御することで、使用状況に合わせて信号処理手段で入力音声に対し適切な信号処理を行わせることができる。

【0028】

以上のように、時間に関連づけられた環境情報をもとに、入力した音声を信号処理する。環境情報が時間に関連づけられた、“場所”であるときは、スケジュール内容である場所に応じて入力音声に対する雑音除去などの信号処理を変えることにより、音声入力システムの使用環境が、時間経過に対応して、駅構内であるときと、オフィス内であるときと、工場内であるときとで、雑音除去処理の仕方を変えることにより、従来と比べてより確実な雑音除去を行うことができるようになる。

10

【0029】

また、本発明による音声入力システムは、さらに、前記信号処理に用いるパラメータを時間と関連づけて記憶する手段を具備し、前記制御手段は、前記環境情報だけでなく、前記パラメータに基づいて、入力音声の前記信号処理を制御することを特徴とする。

【0030】

従来は適応的な信号処理に基づいて、例えば雑音除去であれば、最初は周りの雑音を収集及び学習してそれに応じた雑音除去を行っているが、本発明では、環境情報のほかに信号処理用のパラメータを記憶しておき、環境情報が時間に関連づけられた場所であるときは、駅構内であればこのパラメータ、会議室であれば別のこのパラメータ、といったように、自動的（或いは人為的）に信号処理パラメータを呼び出して（選択して）使用することができる。

20

【0031】

また、本発明による音声入力システムは、さらに、前記信号処理の結果を反映して前記環境情報や前記パラメータの内容を変更する手段を具備したことを特徴とする。

【0032】

本発明では、環境情報や信号処理用パラメータを後から追加したり修正する機能を付加したものである。これによって、例えば、環境情報としての場所を追加したり、雑音処理するパラメータを更新することができる。

【0033】

また、本発明による音声入力システムは、前記環境情報が場所の情報であることを特徴とする。

30

【0034】

本発明では、環境情報として、場所の情報に限定したものである。どこに居るかという場所の情報に対応して、入力音声の信号処理の仕方を変えることができる。すなわち、どういう場所に居るということが分れば、その場所（例えば、オフィスの静かな所、街中の雑踏）に合わせた音声処理を行うことができる。雑音抑圧処理であれば、雑音の種類に対応した雑音抑圧のパラメータを変えることで雑音除去率を上げることができる。

【0035】

また、本発明による音声入力システムは、前記環境情報が人間の情報であることを特徴とする。

40

【0036】

本発明では、環境情報として、人間の情報に限定したものである。誰と会う（或いは喋っている）かという、“人間”の情報に対応して、入力音声の信号処理の仕方を変えることができる。すなわち、誰と会うということが分れば、その人に合わせた音声処理を行うことができる。例えば音声認識処理において、相手が男性か女性か、大人か子供かについて話者を特定できるので認識率を上げることができる。

【0037】

また、本発明による音声入力システムは、前記信号処理の内容が、入力対象である音声以外の雑音の影響を抑圧する処理であることを特徴とする。

【0038】

50

本発明では、信号処理の内容が、雑音を抑圧する処理である。環境情報の内容に従って、例えば、オフィスならオフィスに合った雑音抑圧処理、また街中の雑踏では雑踏に合った雑音抑圧処理を行うことができる。

【0039】

また、本発明による音声入力システムは、さらに、音声認識を行う手段を具備し、前記環境情報や前記パラメータを基に前記信号処理を行うと共に前記音声認識を行うことを特徴とする。

【0040】

本発明では、音声認識するときは、環境情報の内容に合った雑音等の処理を行いながら、環境情報の内容に合った認識語彙等を用いて認識処理を行う。これにより、場所や人などの環境情報に合った認識処理が行え、認識率を上げることができる。

10

【0041】

また、本発明による音声入力システムは、さらに、他の音声入力システムと情報の授受を行う手段を具備し、他の音声入力システムの前記環境情報や前記パラメータを利用して、自身の音声入力システムの前記環境情報や前記パラメータを更新可能としたことを特徴とする。

【0042】

本発明では、他の音声入力システムとの情報のやり取りができる。例えば、ある会社の工場を初めて訪問する人は、その会社や工場に合った認識語彙や雑音処理の情報を持っていないが、訪問時にその工場の音声入力システムで使っている環境情報や信号処理用パラメータを利用することができれば、その場に合った適切な音声処理を行うことが可能となる。

20

【0043】

また、本発明による音声入力システムは、さらに、音声信号以外のセンサ情報を入力する手段を具備し、前記センサ情報を利用して、自身の音声入力システムの前記環境情報や前記パラメータを更新可能としたことを特徴とする。

【0044】

本発明では、センサ情報を入力可能としたものであり、これまで述べた環境情報だけでは捕捉しえない予定外の環境変化にも対処し得るようにしている。例えば、GPS（全地球測位システム、Global Positioning Systemの略）とかカメラ

30

【0045】

また、本発明による音声入力システムは、さらに、構成要素 他にも通信部でもいいのだから、“構成要素の一部を他の音声入力...”でもいいのでは？]の一部を他の音声入力システムと共有化することを特徴とする。

【0046】

本発明では、例えば記憶領域の実体の一部が、ネットワークを介して別の場所にあり、その別の場所に置かれた記憶領域の一部を他の音声入力システムと共有化して使用する。記憶領域の一部が置かれる場所は、例えばサーバである。このようにすれば、環境情報や信号処理用パラメータを複数の他の音声入力システムと共有化できるので、どこに居ても時間に対応した最適な条件で音声入力処理を行えたり、同じデータを使うことにより簡便に共通サービスを得られたりという利点が見られる。

40

【0047】

また、本発明による音声入力システムは、音声信号を受信する受信手段と、時間に関連づけられた環境情報を記憶する記憶手段、時間を計測する時間計測手段と、前記音声信号に対して信号処理を施す信号処理手段と、時間に関連づけられた前記環境情報を記憶する記憶領域がシステム外部にある場合に、その環境情報を読み出すためのアドレス情報を時間と関連づけて記憶する記憶部を有し、前記計測された時間から関連する環境情報を前記外部の記憶領域から読み出し、当該環境情報に基づいて前記信号処理手段を制御する制御手

50

段と、を備えたことを特徴とする。

【0048】

本発明では、記憶する手段は、自身の音声入力システム内に記憶領域を備えるのではなく、記憶領域自体は自身の音声入力システム外にあってその記憶領域のアドレスは自身の音声入力システムにあり、必要な場合に外部にある記憶領域と情報の授受を行う場合を指している。

【0049】

本発明に係る音声入力システムを備えた端末装置は、音声信号を受信する受信手段と、前記音声信号に対して信号処理を施す信号処理手段と、時間に関連づけられた環境情報を記憶する記憶手段と、時間を計測する時間計測手段と、前記計測された時間から関連する環境情報を前記記憶手段から取り出し、当該環境情報に基づいて前記信号処理手段を制御する制御手段と、を具備したものである。

10

【0050】

本発明の端末装置では、時間に関連づけられた環境情報をもとに、入力した音声を信号処理する。環境情報が時間に関連づけられた場所であるときは、スケジュール内容である場所に依りて入力音声に対する雑音除去などの信号処理を変えることにより、音声入力システムの使用環境が、時間経過に対応して、駅構内であるときと、オフィス内であるときと、工場内であるときとで、雑音除去処理の仕方を変えることにより、従来と比べてより確実な雑音除去を行うことができるようになる。

【0051】

本発明による音声入力システムを備えた端末装置は、さらに、前記信号処理に用いるパラメータを時間に関連づけて記憶する手段を具備し、前記制御手段は、前記環境情報だけでなく、前記パラメータに基づいて、入力音声の前記信号処理を制御することを特徴とする。

20

【0052】

従来は適応的な信号処理に基づいて、例えば雑音除去であれば、最初は周りの雑音を収集及び学習してそれに依りた雑音除去を行っているが、本発明では、環境情報のほかに信号処理用のパラメータを記憶しておき、環境情報が時間に関連づけられた場所であるときは、駅構内であればこのパラメータ、会議室であれば別のこのパラメータ、といったように、自動的に（或いは人為的）に信号処理パラメータを呼び出して（選択して）使用することができる。

30

【0053】

本発明による音声入力システムを備えた端末装置は、さらに、前記信号処理の結果を反映して前記環境情報や前記パラメータの内容を変更する手段を具備したことを特徴とする。

【0054】

本発明の端末装置では、環境情報や信号処理用パラメータを後から追加したり修正する機能を付加したものである。これによって、例えば、環境情報としての場所を追加したり、雑音処理するパラメータを更新することができる。

【0055】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態について詳細に説明する。

図1は本発明に係る音声入力システムを示すブロック図である。

40

【0056】

図1において、音声入力システム101は、音声入力を受信する受信手段としての通信部102と、複数の環境情報を時間に対応付けて記憶する記憶手段である記憶領域部103と、雑音処理、音声認識処理などの各種の信号処理を行う信号処理部104と、中央演算装置（以下、CPU）などで構成され、前記記憶領域部103に記憶してある環境情報に基づいて前記信号処理部104における入力音声の信号処理を制御する制御部105と、を有している。制御部105には、時間を計測する時間計測部105-1（実時間を計測する時計手段や、経過時間をカウントするタイムカウンタなど）が含まれている。但し、

50



時間計測部は外部にある計測部で計測してもよい。

【0057】

通信部102は、マイク106、情報機器や記録再生機器や他音声システムなどの他の機器107、及びネットワーク108と有線或いは無線で接続して、外部からシステム内に音声入力を受け取るだけでなくシステムの外部へ音声出力を送出したりすることが可能となっている。

【0058】

通信部102は、信号処理部104での信号処理に適した形式に適宜、データを変換する機能も含んでいてもよい。

なお、図1における各部の構成要素は、必ずしも物理的な構成によって区分されたものではない。各部の構成要素は電子部品によって構成することもできるし、処理内容の働きや処理を実行するプログラムによっても機能的に区分できるものとして説明する。つまり、図1における各構成要素については、他の働きをする機器とその構成の一部を共有化したり、構成の要素は電子部品でなくとも同種の働きをするプログラムやプロセスによって同様の機能を実現することもできる。

【0059】

そして、音声入力システム101で信号処理した結果を音声入力システム101の外部回路等で使用する場合には、信号処理部104は制御部105の制御によってその使用目的に応じた信号処理結果を外部回路等へ出力することが可能である。

【0060】

マイク106は音声を信号に変換して伝達する。このマイク106は一般にある任意のマイクで実現可能である。このマイクが複数個あってもよく、通信部102からの信号を受けてマイク106の制御を行うことも可能である。例えば、通信部102からの信号によってマイク入力のスイッチや、マイクの向きを変えることが可能であってもよい。

【0061】

他の機器107とは、音声入力システム101が処理できる形式の情報を持っている機器であり、音声入力システム101以外の機器を表す。例えば、他の機器107がPDAであって、他の機器107にはユーザの詳細なスケジュール情報が記憶されているとする。音声入力システム101の制御部105は、通信部102を介してそのスケジュール情報のうち、音声入力システム101が任意のタイミングで自身が処理出来る形式のデータだけ他の機器107から取り込んだり、或いは他の機器107側から任意のタイミングで音声入力システム101に送信してもらうよう要求したりすることが可能となっている。これによって、音声入力システム101ではユーザが直接的に入力しなくても、時間に関連づけられた環境情報、例えば上記のスケジュール情報の例では場所の情報やそのときに会う予定の人の情報を得ることができる。この他の機器107は、複数あっても構わないし、他の音声入力システムであってもよい。

【0062】

ネットワーク108は、ブルートゥース(Bluetooth)や無線ローカル・エリア・ネットワーク(無線LAN)に代表される無線による通信ネットワーク、もしくはインターネットに代表される大規模通信網ネットワークであってもよい。そのような通信ネットワークを介して先のマイク106や他の機器107とも情報の授受を行うこともできる。この通信技術の詳細については省略する。

【0063】

記憶領域部103は、各種の環境情報を時間に対応させて記憶している。環境情報と音声システムが使われる周囲の環境状況や音声システム内部機器の動作環境状況などを含む。記憶領域部103は、それぞれの環境下での処理の手順や必要となる処理パラメータ、処理に必要な一時記憶領域、音声信号や出力結果等を記憶する領域等、の各種記憶領域を有している。記憶領域部103は、半導体メモリや磁気ディスク等の電子部品や、同様の機能を有する電子部品によって構成することができる。

【0064】

10

20

30

40

50

信号処理部 104 は、音声入力システム 101 の目的に応じて、制御部 105 の制御によって通信部 102 からの音声入力信号を処理することができる。時間に関連づけられた環境情報に基づいて信号処理を行う。例えば、雑音処理、音声強調、音声認識などの信号処理を行う。また、その信号処理に必要なパラメータを記憶領域部 103 から取り出して処理を行うこともできる。ここでの具体的な処理手順や内容については後述する。信号処理部 104 は、ソフトウェアでも、また信号処理チップ等の電子部品でも構成することができる。

【0065】

制御部 105 は、CPU などによって構成され、記憶領域部 103 に記憶してある環境情報や信号処理用パラメータに基づいて信号処理部 104 における入力音声の信号処理を制御する。また音声入力システムの動作を制御する。

10

【0066】

次に、図 1 の動作を図 2 を参照して説明する。

図 2 は図 1 の音声入力システムの処理動作を示している。まず、制御部 105 は時間情報として現在の時刻を時間計測部 105 - 1 から取得する（ステップ 301）。この時間情報は他の機器 107 やネットワーク 108 上の他の機器（図示せず）から取得してもよい。次に、制御部 105 は取得した時間情報に関連する環境情報を記憶領域部 103 の所定領域から取得し（ステップ 302）、当該環境情報に基づいて入力音声の信号処理の内容を決定する（ステップ 303）。そして、その決定された処理内容に従って、入力される音声信号に対して信号処理を行い、その処理結果を記憶領域部 103 の所定領域へ出力する（ステップ 304 ~ 306）。

20

【0067】

尚、記憶する手段は、自身の音声入力システム 101 内に記憶領域を備えるのではなく、記憶領域自体は自身の音声入力システム外にあってその記憶領域のアドレスは自身の音声入力システム 101 にあり、必要な場合に外部にある記憶領域と情報の授受を行うように構成してもよい。

【0068】

図 3 は本発明に係る音声入力システムの他の構成例を示すブロック図である。図 1 と同一部分には同一符号を付してある。

【0069】

図 3 において、音声入力システム 101 A は、入力音声信号を受信する受信手段としての通信部 102 と、複数の環境情報を時間に対応付けて記憶する記憶手段である記憶領域部 103 と、雑音処理、音声強調、音声認識処理などの各種の信号処理を行う信号処理部 104 と、CPU などによって構成され、システム外部の記憶領域部（図示せず）に記憶してある環境情報に基づいて前記信号処理部 104 における入力音声信号の信号処理を制御する制御部 105 A と、を有している。制御部 105 には、時間を計測する時間計測部 105 - 1（実時間を計測する時計手段や、経過時間をカウントするタイムカウンタなど）と、システム外部の記憶領域に記憶してある環境情報を読み出すためのアドレス情報を時間と関連づけて記憶する記憶部 105 - 2 とが含まれている。

30

【0070】

図 3 の構成では、制御部 105 A は、時間に関連づけられた前記環境情報を記憶する記憶領域がシステム外部にある場合に、その環境情報を読み出すためのアドレス情報を時間と関連づけて記憶部 105 - 2 に記憶してあり、時間計測部 105 - 1 で計測された時間から関連する環境情報を前記システム外部の記憶領域から読み出し、当該環境情報に基づいて信号処理部 104 を制御する。音声入力システムの処理動作は図 2 と同様であるので、説明を省略する。

40

【0071】

上記のような音声入力システム 101（又は 101 A）は端末装置例えば PDA に適用することができる。

図 4 は本発明に係る、音声入力システムを備えた PDA を示すブロック図である。

50

## 【0072】

図4に示すPDA111は、マイクなどから音声を入力し、環境情報の内容に基づいて信号処理を行う、図1のように構成された音声入力システム101と、ユーザ指示入力部、表示部、データ記録部及び制御部(CPU)(いずれも図示せず)を備え、スケジュール表やメールを作成したり、インターネット情報を送受信したり、前記音声入力システム101で信号処理された音声データを記録再生することが可能なPDA本体部112と、を具備して構成される。なお、PDA本体部112内のデータ記録部は音声入力システム101の記憶領域部103に比べて大きな容量を有しており、PDA本体部112で扱う画像データ、音声データ、文字データなどの多量のデータを保存することが可能となっている。

10

## 【0073】

図5は図4のPDA111の使用例を説明するものである。

## 【0074】

図5において、時計201は単に時間情報としての時刻を示すためのものであり、必ずしも物理的に存在しているものではない。図5(a)は16時の状態を、図5(b)は18時の状態を示している。なお、以下の説明では、時刻を24時制で表現している。

## 【0075】

16時のとき、図5(a)のようにユーザ202は外出中であり、雑踏の中において、音声入力システム101を備えたPDA111を所持している。そして、ユーザ202は自分の発声する命令語の音声でPDA111を音声操作しているものとする。また、この時間に外出中であることは、PDA本体部112のデータ記録部にスケジュール表として記録してあるものとする。

20

## 【0076】

この場合、音声入力システム101の制御部105に対して予め環境情報としてPDA本体部112のデータ記録部内のスケジュール表を使用する旨の設定(操作)をユーザが行うことで、記憶領域部103には前記スケジュール表に基づく時間に関連づけられた環境情報が記憶されている。

## 【0077】

PDA111内における音声入力システム101の制御部105は、システム101内の時間計測部105-1による時間に基づいて記憶領域部103から環境情報を取得することで、この時間にはユーザ202が外出中であるという情報を得る。

30

## 【0078】

そこで、音声入力システム101の制御部105は、そのときPDA111に対してユーザ202が入力する音声について、ユーザが外出中であることから雑踏の中での音響処理パラメータ、処理手順を記憶領域部103から呼び出して信号処理部104で音声認識の信号処理を行う。それによって、雑踏の中で発声された音声に対して適切な音声処理を行って正確な音声認識を行うことができる。そして、制御部105はその信号処理結果に基づいてPDA本体部112の制御部に対して命令内容に応じた動作をさせる。例えば、インターネット受信動作を開始させて、所望の情報を得ることが可能となる。或いは、音声メモとして本体部112に記録することもできる。

40

## 【0079】

また、時間が経過して18時になったとき、図5(b)のようにユーザ202はオフィス内に居て、自分の発声する命令語の音声でPDA111を音声操作しているものとする。音声入力システム101の制御部105は、先ほどと同様に現在時刻と記憶領域部103内の環境情報から、この時間にはユーザ202がオフィスにいるという情報を得る。

## 【0080】

そこで、音声入力システム101の制御部105は、その情報からオフィスで得られる音声について音響処理で必要となるパラメータと処理手順を記憶領域部103から呼び出して、オフィスデスク203にいるユーザ202の音声を、信号処理部104において適切に信号処理して音声認識することができる。

50

## 【0081】

以上のようにして、雑音抑圧，音声強調，音声認識といった信号処理において、環境情報を利用して使用環境に応じた適切な音声処理ができる。

## 【0082】

さらに、適応的な信号処理を用いたときに、適応後のパラメータを記憶しておくこともできる。これによって、翌日になったとしたときに例えば時間に対応する環境情報から、18時から同じオフィスにいるという情報が得られた場合には、前日の前記の適応後パラメータを呼び出して音声処理に用いることにより、より簡単にかつ精度よく音声処理することが可能となる。

## 【0083】

尚、本発明の音声入力システムはPDAのみに対して適用されるものではなく例えば携帯電話，録音機器，パーソナルコンピュータなどの他の端末装置に対しても応用できる。また、環境情報はスケジュール情報のみに限定されるものではない。

## 【0084】

## 〔第1の実施の形態〕

次に、本発明の第1の実施の形態の音声入力システムについて説明する。ここでの音声入力システム101は、PDA本体部112での音声入力に使われるものとする。また、PDA本体部112では、音声入力システム101の処理結果である音声信号を本体部112内のデータ記録部に音声メモとして記録しておくことが可能である。処理動作の流れは図2と同様である。

## 【0085】

すなわち、処理動作の流れは、まず時間情報として現在の時刻を時間計測部105-1で取得し、次に取得した時間情報に関連する環境情報を記憶領域部103から取得して、当該環境情報に基づいて入力音声の信号処理の処理内容を決定する。そして、その決定された処理内容に従って、入力される音声信号に対して信号処理を行う。

## 【0086】

次に、上記処理フローにおける信号処理内容の決定について図6を参照して説明する。図6は、第1の実施の形態における環境情報と音声処理内容の対応関係を示している。

## 【0087】

図6では、音声入力システム101を含めたPDA111に節電のために通常モードと省電力モードを設けて、これらの処理モードを環境情報とし、該情報に合わせて音声処理内容も変える例を示している。

## 【0088】

図6では、時間とそれに関連付けられた環境情報として「処理モード」が規定されており、この環境情報にはさらに音声の「処理内容」が対応付けられている。

## 【0089】

すなわち、時刻が10時で通常モードに設定されていた場合は、時間帯としてユーザが活動時の音声入力である可能性が高いし、節電する必要もないので、音声入力に対して高性能な音声検出を行い、かつ高品質な音声取り込みの結果を音声入力システム101の処理結果としてPDA本体部112に送る。これによってユーザが一般的に使う状況で、それに見合った適切な音声処理を行うことができる。なお、ここでの音声検出方法は非特許文献2の177ページに見られるような形で実現可能であり、ここでは処理を節約しない高性能な手法を用いたものとして詳細は省略する。音声の取り込みについても、コンパクト・ディスク(CD)なみの高音質な音声で信号を取り込む技術は一般的に存在し、それらの技術で実現可能である。ここではその詳細は省略する。

## 【0090】

次に24時で通常モードに設定されていたり、10時で省電力モードに設定されていたりした場合は、夜間で滅多に音声入力はない、或いは省電力モードであるという理由で、先ほどよりは簡易化した音声検出や音質を落とした処理(例えばサンプリング周波数を銅線電話品質(8kHz)とする)を行って、状況に見合った音声処理を行う。

10

20

30

40

50

## 【0091】

そして24時で省電力モードに設定されていた場合は、処理に必要な電力がない、或いはユーザからの音声が入ってくることは殆んどないという理由で音声処理を行わないとする。これによって音声の処理をすべきでない場合や必要が余りない場合に見合った処理ができる。また、取得した時刻に対して環境情報に対応するものがない場合には、そういう場合用の信号処理内容を前もって決めておいてそれを用いたり、最も近い関係の条件の内容を参照するといったやり方で対処することができる。

## 【0092】

〔第2の実施の形態〕

次に、本発明の第2の実施の形態の音声入力システムについて説明する。処理動作の流れは図2と同様である。 10

## 【0093】

図7は、第2の実施の形態における環境情報と音声処理内容の対応関係を示している。時間と関連付けた環境情報である処理モードとして、通常モードと通勤モードが設けられている。通勤モードとは、電車内や雑踏のように喧しいところでも、聴き取り易い音声入力を行うためのモードである。

## 【0094】

時間が1時～6時、10時～15時のようなラッシュのない状況では、通常モードとし、低い精度の音声検出及び取り込みを行い、且つ周りが喧しくないから音声取り込みのボリュームは普通にする。一方、6時～10時、15時～1時のようなラッシュの状況では、通勤モードとし、高精度な音声検出及び音声取り込みを行い、且つ周りが喧しく、本人の声も大きくなるので音声取り込みのボリュームを若干抑える（即ち音声信号レベルを若干低下させる）音声処理を行う。 20

## 【0095】

〔第3の実施の形態〕

次に、本発明の第3の実施の形態の音声入力システムについて説明する。処理動作の流れは図2と同様である。

## 【0096】

図8は、第3の実施の形態における環境情報と信号処理パラメータとの対応関係を示している。時間に関連づけられた環境情報である処理モードとしては、通常モードと省電力モードが設けられている。音声に対する信号処理の内容は、入力音声信号に対するサンプリング周波数の変更のみとなっている。 30

## 【0097】

本実施形態では、「信号処理内容の決定」の内容が、信号処理に用いるパラメータの設定となる。そして、サンプリング周波数が信号処理に用いるパラメータとなっている。また本実施形態ではサンプリング周波数というパラメータの性質上離散値のパラメータだが、本発明においては環境情報とパラメータとの関係が連続関数になっていても構わない。

## 【0098】

例えば、10時で通常モードに設定されていた場合は、高音質で音声を取り込むためにサンプリング周波数は44.1kHz（CD品質）とし、24時での通常モードや10時での省電力モードの場合はサンプリング周波数は22.05kHz、24時の省電力モードならサンプリング周波数は8kHz（電話品質）とする。このようにして決まったサンプリング周波数で音声をサンプリングしてデジタル信号にする手法は、既存の手法で十分実現可能であり、ここではその詳細は省略する。 40

## 【0099】

以上の第1、第3の実施の形態では、時間と関連付けられた環境情報を用いることで、日常の一般的な状況では高音質で音声を取り込み、音声処理する電力が余りない場合や、夜間などのそれほど高音質で取り込む必要がない場合には、音声入力システムに負担が掛からない低い精度の処理を行うというように、使用状況に合わせた音声処理を行うことが可能である。

## 【0100】

また、第2の実施の形態では、周囲が喧しい状況では高音質で音声を取り込み、比較的静かな状況では低い精度での処理を行う。

## 【0101】

〔第4の実施の形態〕

次に、本発明の第4の実施の形態を、図9と図10を用いて説明する。本実施形態は、会社などで用いられるノート型コンピュータ（以下、NPC）と共に用いられる音声入力システムの例である。この場合は、音声入力システムは、音声処理用のアプリケーションプログラムとして実現することが可能である。

## 【0102】

環境情報は、使用時の時刻と関連してNPCが使用される場所、例えば会議室A、B、Cの情報である。この環境情報が音声入力システム101の記憶領域部103に記憶される。音声入力システム101の音声処理の内容としては、ユーザの発声に対して雑音抑圧処理を行うもので、雑音抑圧処理した音声信号を前記NPCに出力し、該NPCがその音声を議事録用に記録する。環境情報である会議室に対して、雑音抑圧処理に用いられる信号処理用パラメータが対応付けられる。

10

## 【0103】

雑音抑圧の信号処理はスペクトル・サブトラクション法（以下、SS）を用いるとする。SSの実現については、前述の非特許文献1を始めとして、多くの公知例文献に示されているので、ここでは省略する。本実施形態では、例えばSSで用いる推定雑音特徴ベクトルを信号処理で用いるパラメータとする。また、この推定雑音特徴ベクトルは使用会議室における非音声区間から随時更新されるものとする。

20

## 【0104】

図10はこれらの環境情報とパラメータとの対応関係を示している。この対応関係は、予め記憶領域部103に記憶されている。そして、音声入力信号を雑音抑圧処理するに際して、音声入力システム用のアプリケーションプログラムを起動した状態で、表示される設定用画面上の所定箇所に使用する時刻と会議室名を入力することにより、雑音抑圧処理動作が実行可能となる。

## 【0105】

図9に第4の実施の形態の音声入力システムに係る処理フローを示す。処理の手順は図2とほぼ同様である。まず、制御部105は、時間情報として現在の時刻を時間計測部105-1から取得する（ステップ401）。次に、制御部105は取得した時間情報に関連する環境情報を取得し（ステップ402）、当該環境情報に基づいて入力音声の信号処理のパラメータすなわち推定雑音特徴ベクトルを記憶領域部103から取り出して信号処理部104に設定する（ステップ403）。

30

## 【0106】

このとき図10のような対応関係を参照し、対応する環境情報があればそれに対応した推定雑音特徴ベクトルを用いて信号処理が行われる。

## 【0107】

一方、対応する環境情報がなかった場合には、記憶領域部103に記憶可能な領域があるか確認後、新しく環境情報を作成する。つまりこの例では、ある時間にある会議室で使われるのが初めてだったなら、記憶領域部103に新たな環境情報とパラメータを記憶する領域が割り当てられる。このときのパラメータの初期値は、全推定雑音の平均から求めたり、予め設定した初期値用の値にするなどして決定される。また、対応する環境情報がなかった場合に、新しく作成せずに、所定に決まった処理を割り当ててもよい。

40

## 【0108】

こうして信号処理に用いるパラメータを信号処理部104に設定した後、入力音声（ステップ404）に対して雑音抑圧処理及び非音声区間においては雑音の推定を行う（ステップ405）。そして信号処理後の信号を結果としてNPCへ出力する一方（ステップ406）、（更新された推定雑音のパラメータを使ってさらに信号処理を行ったり、）処理終

50

了後に環境情報に対応付けられたパラメータの記憶領域を更新された推定雑音に基づいて書き換えたりする（符号407）。

【0109】

本実施の形態では、環境情報やパラメータを書き換える際に、上述のように新しい条件ができる度に新たな記憶領域を割り当てたり、処理のたびに情報を更新したりする例を挙げた。

【0110】

新しい条件を判断するのに、環境情報としては時間（或いは時間帯）の別にのみ着目したり、会議室の情報にのみ着目したり、また、逆にパラメータ側に着目することで判断することもできる。具体的には、新しい時間、新しい会議室で処理した後、推定雑音に注目して、これまでに記憶されたパラメータの中で許容範囲内の近さのものについてはパラメータを共用化してしまうことが考えられる。すなわち図10で時刻は異なるが同じ会議室Aでの推定雑音特徴ベクトルA1とA2が十分近ければまとめてA1で代表させることもできる。

10

【0111】

〔第5の実施の形態〕

次に、本発明の第5の実施の形態について説明する。本実施の形態の説明に用いる例は、第4の実施の形態の説明に用いた例と同様とする。すなわち、NPCと共に用いられる音声入力システムの例である。

【0112】

本実施形態で、第4の実施の形態と異なる点は、NPC内にはスケジュール表が記録されており、環境情報はスケジュール表から取り出されるものである。スケジュール表には使用される時刻と会議室がその他の情報（例えばパラメータ）と共に記載されている。

20

【0113】

スケジュール情報を用いるので、使用される時刻に対応して使用される会議室が決まり、それに対応付けられて記憶領域部103に記憶されているパラメータを用いて雑音抑圧の信号処理を適切にすることが可能となる。

【0114】

このようにすることで、今日はA会議室を使用して、明日の別の時刻にA会議室を使用するスケジュールになっていた場合、明日の当該時刻になると自動的にA会議室の雑音処理パラメータを用いて信号処理を行うことができることとなる。

30

【0115】

〔第6の実施の形態〕

次に、本発明の第6の実施の形態について説明する。本実施の形態に用いる例は、第5の実施の形態に用いた例と同様とする。

【0116】

本実施形態で、第5の実施の形態と異なる点は、スケジュール情報には、時刻とそれに対応して誰と会うかという人間の情報が入っている場合である。すなわち、時刻と共に人間の情報を含んでいる。

【0117】

このようにすることにより、人と会う時刻になると自動的に相手にあった音声入力ができるようになる。音声認識処理であれば、誰と会うという情報によって話者を特定できるので認識率を上げることができる。スケジュール情報に対応するイベント（人の出会い）がなかった場合には代表ユーザの情報で代用することで、信号処理してもよい。ここで用いられる信号処理については、話者に合わせた雑音抑圧や、音声強調などが挙げられる。その実現方法については、一般的に用いられている周知の手法で実現可能であるので、その詳細は省略する。

40

【0118】

〔第7の実施の形態〕

次に、本発明の第7の実施の形態について、図11を参照して説明する。本実施の形態の

50

説明に用いる例は、第5の実施の形態の説明に用いた例と同様とする。

【0119】

本実施形態で、第5の実施の形態と異なる点は、信号処理において音声認識も含める点である。音声認識の方法については、前述の非特許文献2に代表される多くの公知例文献があり、ここではその詳細は省略する。ここでは前記公知例文献でも述べられているHMM (Hidden Markov Model)を用いた音声認識を用いるものとする。音声認識の対象となる語彙は、予め設定された一般用語を対象とし、さらに場所によって追加語彙があるものとしてこれを信号処理のパラメータとする。この追加語彙の登録については、ここでは前もって場所に応じて登録してあったものとするが、ユーザや音声入力システムより上位のシステムが適宜語彙を登録してもよい。図11は環境情報の場所とパラメータとしての追加語彙との対応関係を示している。

10

【0120】

本実施形態における処理の手順は図2と同様である。すなわち、得られた時刻から環境情報を取得し、図11の対応関係のようにして音声処理の内容となる音声認識で用いる追加語彙を設定する。それを受けて、一般の認識語彙と処理パラメータの追加語彙とで音声認識を行い、認識結果を音声入力システムの出力とする。

【0121】

〔第8の実施の形態〕

次に、本発明の第8の実施の形態について説明する。本実施の形態の説明に用いる例は、第7の実施の形態の説明に用いた例(音声認識を含める)と同様とする。

20

【0122】

本実施の形態で、第7の実施の形態と異なる点は、通信部102を介して音声入力システムが情報の授受をできるものとし、通信可能な範囲内に別の音声入力システムがあるものとした場合である。

【0123】

音声入力システム間の通信路はローカル・エリア・ネットワーク(LAN)やBluetoothに代表される、実現可能な既存の機器間通信技術で実現するものとする。ここではその詳細は省略する。その場合の機器間通信路において他の通信可能な機器の検出、及び通信路の確保、実際の通信の手順に関しても、先に例として挙げた実現可能な技術で用いられる代表的な手順に従うものとする。

30

【0124】

図12は本発明の第8の実施の形態に係る、通信部102を介した音声入力システム間の情報授受の概念図を示している。前述のように通信路を介して情報の授受が可能な音声入力システムが2つあるとする。片方がユーザ1の音声入力システム、他方がユーザ2の音声入力システムとする。各音声入力システムには前述のような環境情報と、それと対応付けられたパラメータである追加語彙の情報がある。すなわち、ユーザ1の音声入力システムの環境情報と追加語彙間の対応関係501と、ユーザ2の音声入力システムの環境情報と追加語彙間の対応関係502がある。

【0125】

なお、音声認識の信号処理を行う信号処理部104で行われるパラメータとしての追加語彙は、各音声入力システムの記憶領域部103に記憶されている。

40

【0126】

ユーザ1の音声入力システムは時刻を取得して、環境情報を参照するとき、通信路上でアクセスできる他の音声入力システムの環境情報についても問い合わせるメッセージを送る(符号503にて示す)。ユーザ2の音声入力システムは、その問い合わせが受理できる場合に自身のシステムの環境情報と追加語彙間の関係をユーザ1の音声入力システムに送信する(符号504にて示す)。こうすることでユーザ1の音声入力システムは、ユーザ1のシステムの対応関係501に対してユーザ2のシステムの対応関係502を加えた対応関係505を得ることで、ユーザ1の音声入力システムはこれまで自身のシステムで持っていなかった環境情報とパラメータの追加語彙との対応を利用することができるよう

50



になる。

【0127】

これによって、通常とは異なる新しい状況に入ったユーザでも、これまでにその場面を経験したり、設定したりしたことのあるユーザの音声入力システムの情報を利用することで、新しい状況にすぐ対応した音声処理を行うことができる。ここではユーザ1の音声入力システムの立場から記述したが、通信部を介した情報の問い合わせ（符号503）とその返答（符号504）を交互に行うことで、2つの音声入力システムがそれぞれ和集合の環境情報と追加語彙間の対応関係の情報を得るようにしてもよい。これによって2つの音声入力システムが、環境情報とパラメータの追加語彙間の対応関係の情報を共有化することができる。

10

【0128】

また、上記では処理が始まって時刻を取得した後に情報を授受した例を挙げたが、処理全体が始まる前で、時刻を取得する前に前もって情報の授受を音声入力システム間で済ましていても構わない。また、上記では環境情報とパラメータの追加語彙間の対応の情報全体を授受するように説明したが、取得された時間に係る対応関係のみを授受しても構わない。

【0129】

また、ユーザや音声入力システムの上位のシステムからの設定によって、他の音声入力システムに与えない情報があったり、他の音声入力システムと自身の音声入力システムの情報に差分があった場合の変更方法（上書き、あるいは変更しないなど）を制御したりしても構わない。

20

【0130】

〔第9の実施の形態〕

次に、本発明の第9の実施の形態について図13及び図14を参照して説明する。

【0131】

図13は本実施形態の音声入力システムを示している。通信部102にセンサ109からの情報を入力できるようにした点以外は、図1の構成と同様である。

【0132】

本実施形態では、図13に示すように音声入力システムは音声信号以外のセンサ情報をセンサ109より入力できるようにするものである。そのセンサは音声入力システム内であってもよいし、音声入力システム外にあって、その出力が通信路を介して得られるものでもよい。

30

【0133】

音声信号以外のセンサ109の情報は、例えばグローバル・ポジショニング・システム（GPS）と地図情報とから得られた現在地情報とする。なお、GPSからは、その原理上、正確な時間情報も同時に得ることが可能である。つまりGPSで得られた現在地と地図情報から、制御部105はユーザがどのようなカテゴリーの場所に現在いるかを判断した結果を音声信号以外のセンサの情報とする。その判断の仕方は、例えば現在地と最寄りのランドマークや地図情報から得られる建造物の情報から判断することで現実化できる。その判断手法は本発明の主旨とは外れるので詳細は省略する。またここでの信号処理は雑音抑圧とし、パラメータは使用状況における推定雑音特徴ベクトルとする。

40

【0134】

図14は、記憶領域部103に記憶される時間に関連づけられた環境情報、例えば場所と、信号処理用パラメータとしての推定雑音特徴ベクトルとの対応関係を示している。この対応関係は予めユーザ操作もしくは上位システムにより記憶領域部103に記憶してある。しかしながら、時間に関連する環境情報に必要な対応がなければ、後述するように、センサ109の情報を利用して、自身の音声入力システムの環境情報や信号処理用パラメータを更新可能とされる。

【0135】

処理の手順はこれまでに図2で説明したものと同様である。ただし、ここでは、時刻情報

50

と共に、音声信号以外のセンサ情報、例えば現在地情報が得られるとする。時間情報とGPSなどのセンサ109から得られる現在地情報との対応関係が図13の対応関係に該当すれば、記憶領域部103内の推定雑音ベクトルを読み出して適切な雑音抑圧処理を信号処理部104で行うことが可能となる。

【0136】

例えば、ユーザが11時に駅構内にいれば、図14の参照結果から繁華街用の推定雑音特徴ベクトルが得られる。これをパラメータとしてスペクトル・サブトラクション法(SS)等の雑音抑圧法を用いれば、速やかに状況に合わせた信号処理を行うことができる。

【0137】

環境情報に対応するものがなかった場合、前述のように新しく条件を新設してもよいし、どこか満たしている条件があればそれで代用してもよい。例えば9時に駅構内にいた場合に図14には対応する条件がないが、駅構内ということで駅周辺のある条件で代用しても構わない。このような代用の仕組みは上記の例だけでなく、応用用途に合った形で選ぶことができる。また、この代用に関しては本実施の形態だけでなく、これまでの実施の形態に関しても適用することができる。

10

【0138】

〔第10の実施の形態〕

次に、本発明の第10の実施の形態について説明する。

【0139】

本実施形態では、自身の音声入力システムの記憶機能の一部を他の音声入力システムと共有化するものである。

20

【0140】

図15は本実施形態の音声入力システムの概略構成のブロック図を示している。インターネット等のネットワークにデータを共有化するためのサーバ110が接続された構成となっている。その他の構成は図1と同様である。

【0141】

例えば、ある会社の社内で複数の音声入力システムを備えた機器(例えばPDA)を使う場合には、時間に関連付けられた環境情報を、サーバ110内に記憶することでその会社の社員用情報として共有化できる。

【0142】

このように環境情報を共有化すると、いちいち他の社員から環境情報を貰わなくても、社員なら会社施設内のどこでも時間に対応した最適な使用環境で音声入力を行うことが可能となる。

30

【0143】

〔第11の実施の形態〕

第11の実施形態では、自身の音声入力システムの信号処理機能の一部を他の音声入力システムと共有化するものである。

【0144】

前述の環境情報の共有化のほかに、音声の信号処理を適切なサーバで共有の信号処理パラメータを用いて一括処理する形の共有化も可能である。信号処理のパラメータを共有化することで、複数人が使う状況、すなわち同じ場所(部屋など)に同じ時間にいる人は使用環境が同じとなり使用パラメータ値も複数の音声入力システムで同じものとなることによって、同じ信号処理がなされる結果、音声を入力し処理する際に、簡便に共通サービスを受けられるというメリットが得られる。

40

【0145】

図16は本発明の第11の実施形態の音声入力システムの概略構成のブロック図を示している。インターネット等のネットワークに信号処理動作を行う部分を共有化するためのサーバ110Aが接続され、音声入力システム101Bには信号処理部が無い構成となっている。

【0146】

50

このような構成では、マイク 106 などから音声が入力システム 101B に入力されると、通信部 102 を介して記憶領域部 103 に音声データとして一旦取り込まれ、制御部 105 の制御によって、ネットワーク 108 を経由してサーバ 110A に転送される。そして、サーバ 110A に格納されている時間に関連づけられた信号処理パラメータを用いて信号処理した後、その処理結果のデータをネットワーク経由で音声入力システム 101B に返送し、記憶領域部 103 の所定の領域に処理後のデータとして格納するか、音声入力システム 101B を備えた端末装置の本体部（図示せず）のデータ記録部に格納する。

#### 【0147】

尚、本発明における音声入力システムを備えた端末装置は、音声認識を用いた個人認証装置に応用できることは勿論である。より具体的には、本発明の音声入力システムを携帯端末における個人認証に用いても有用である。

#### 【0148】

##### 【発明の効果】

以上述べたように本発明によれば、時間の情報から、環境情報を参照し、その環境情報に基づいて入力音声の信号処理を制御することができる。これによって、ユーザや音声入力システムの上位システムの制御を必要とすることなく、周囲の状況に合わせた信号処理を行うことができるという効果を有する。

##### 【図面の簡単な説明】

【図1】本発明に係る音声入力システムの概略構成を示すブロック図。

【図2】本発明の第1の実施の形態の処理フローを示す図。

【図3】本発明に係る音声入力システムの他の構成例を示すブロック図。

【図4】本発明に係る音声入力システムを備えた端末の概略構成を示すブロック図。

【図5】音声入力システムの使用例を説明する図。

【図6】本発明の第1の実施の形態での環境情報と音声処理内容の対応関係を示す図。

【図7】本発明の第2の実施の形態での環境情報と音声処理内容の対応関係を示す図。

【図8】本発明の第3の実施の形態での環境情報とパラメータの対応関係を示す図。

【図9】本発明の第4の実施の形態の処理フローを示す図。

【図10】本発明の第4の実施の形態での環境情報とパラメータの対応関係を示す図。

【図11】本発明の第7の実施の形態での環境情報とパラメータの対応の概念図

【図12】本発明の第8の実施の形態に係る、通信部を介した音声入力システム間の情報授受を示す図。

【図13】本発明の第9の実施の形態に係る音声入力システムの概略構成を示すブロック図。

【図14】本発明の第9の実施の形態での環境情報とパラメータの対応関係を示す図。

【図15】本発明の第10の実施の形態に係る音声入力システムの概略構成を示すブロック図。

【図16】本発明の第11の実施の形態に係る音声入力システムの概略構成を示すブロック図。

##### 【符号の説明】

101 ... 音声入力システム、102 ... 通信部、103 ... 記憶領域部、104 ... 信号処理部、105 ... 制御部、106 ... マイク、107 ... 他の機器（PDA など）、108 ... ネットワーク、111 ... PDA（端末）、112 ... PDA 本体部（本体部）。

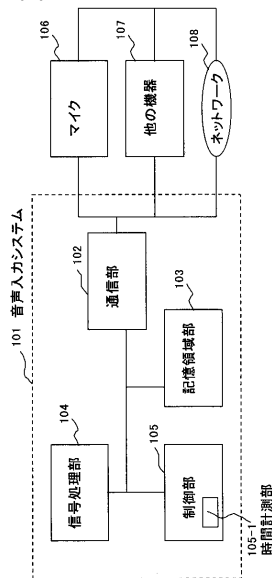
10

20

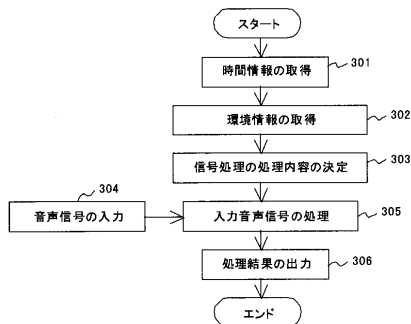
30

40

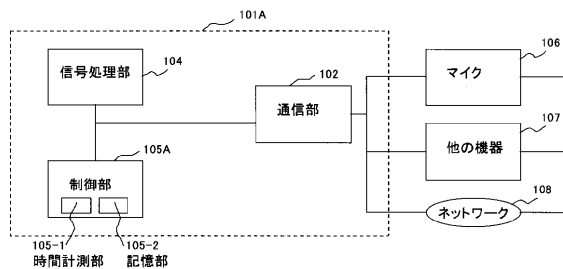
【図 1】



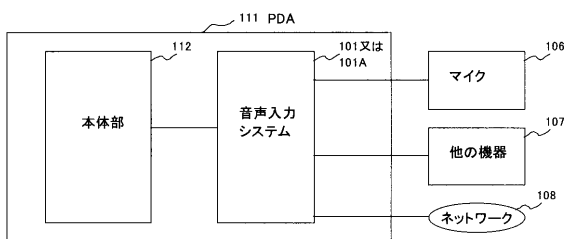
【図 2】



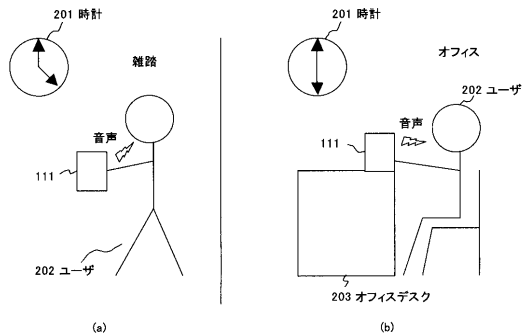
【図 3】



【図 4】



【図 5】



【図 6】

時間	処理モード	処理内容
23:00-6:00	通常	低精度な音声検出&音声取り込み
23:00-6:00	省電力	音声処理を行わない
6:00-23:00	通常	高精度な音声検出&音声の取り込み
6:00-23:00	省電力	低精度な音声検出&音声取り込み

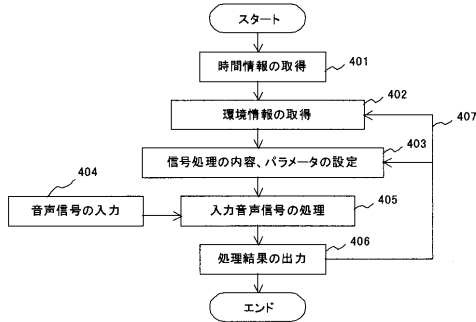
【図 7】

時間	処理モード	処理内容
1:00-6:00	通常	低精度な音声検出&音声取り込み、及び音声取り込みのボリューム普通
6:00-10:00	通勤	高精度な音声検出&音声取り込み、及び音声取り込みのボリューム若干低下
10:00-15:00	通常	低精度な音声検出&音声取り込み、及び音声取り込みのボリューム普通
15:00-1:00	通勤	高精度な音声検出&音声取り込み、及び音声取り込みのボリューム若干低下

【図 8】

時間	処理モード	サンプリング周波数
23:00-6:00	通常	22.05kHz
23:00-6:00	省電力	8kHz
6:00-23:00	通常	44.1kHz
6:00-23:00	省電力	22.05kHz

【図9】



【図10】

時間	会議室	推定雑音特徴ベクトル
10:00-12:00	A	A1
13:00-15:00	B	B1
15:00-17:00	A	A2
10:00-12:00	C	C1

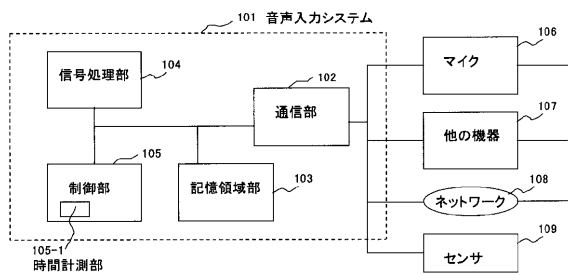
【図11】

時間	場所	追加語彙
10:00-12:00	A 会議室	A 会議室用
13:00-15:00	B 応接室	B 応接室用
15:00-17:00	A 会議室	A 会議室用
10:00-12:00	C 居室	C 居室用

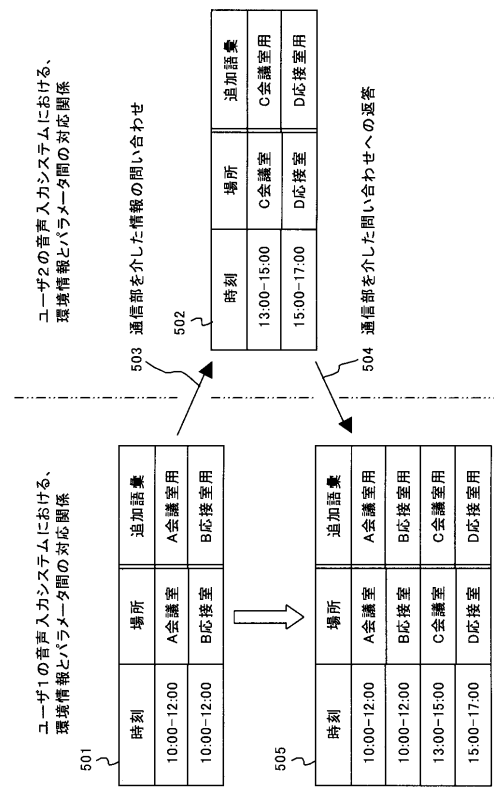
【図13】

時間	場所	推定雑音特徴ベクトル
10:00-12:00	駅周辺	繁華街用 (昼間)
13:00-15:00	住宅街	住宅街用
19:00-21:00	商店街	繁華街用 (夜間)
10:00-12:00	公園	自然用

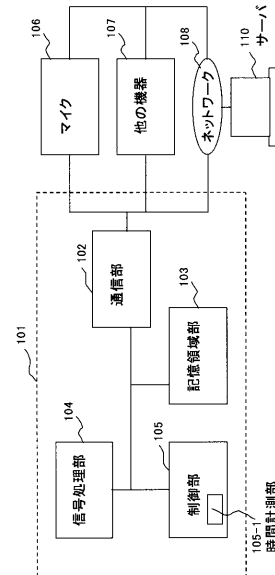
【図14】



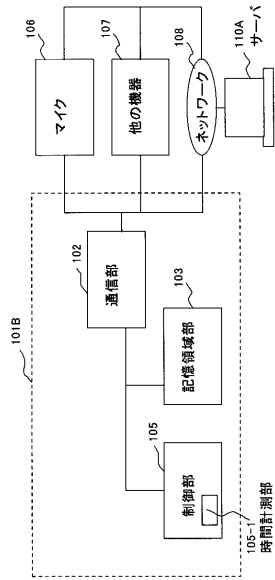
【図12】



【図15】



【図16】



フロントページの続き

(51)Int.Cl.<sup>7</sup>

G 1 0 L 21/02

F I

G 1 0 L 3/02 3 0 1 F

テーマコード(参考)