

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4904691号
(P4904691)

(45) 発行日 平成24年3月28日(2012.3.28)

(24) 登録日 平成24年1月20日(2012.1.20)

(51) Int.Cl.	F 1	
GO3B 15/00 (2006.01)	GO3B 15/00	R
GO3B 17/38 (2006.01)	GO3B 15/00	F
HO4N 5/225 (2006.01)	GO3B 17/38	B
HO4N 5/232 (2006.01)	HO4N 5/225	F
HO4N 101/00 (2006.01)	HO4N 5/232	A

請求項の数 6 (全 11 頁) 最終頁に続く

<p>(21) 出願番号 特願2004-378386 (P2004-378386)</p> <p>(22) 出願日 平成16年12月28日(2004.12.28)</p> <p>(65) 公開番号 特開2006-184589 (P2006-184589A)</p> <p>(43) 公開日 平成18年7月13日(2006.7.13)</p> <p>審査請求日 平成19年10月17日(2007.10.17)</p> <p>前置審査</p>	<p>(73) 特許権者 000001443 カシオ計算機株式会社 東京都渋谷区本町1丁目6番2号</p> <p>(72) 発明者 加福 滋 東京都羽村市栄町3丁目2番1号 カシオ 計算機株式会社羽村技術センター内</p> <p>審査官 菊岡 智代</p> <p>(56) 参考文献 特開2001-305642 (JP, A)) 特開2003-066990 (JP, A)) 特開2000-259198 (JP, A))</p> <p style="text-align: right;">最終頁に続く</p>
---	--

(54) 【発明の名称】 カメラ装置、及び撮影方法

(57) 【特許請求の範囲】

【請求項1】

所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置であって、

入力する音声を認識する音声認識手段と、

前記所定の命令語を記憶する命令語記憶手段と、

この命令語記憶手段に記憶されている命令語に設定されている複数の発声段階と、前記一連の撮影動作における各段階の動作との対応関係を示す対応情報を記憶する対応情報記憶手段と、

前記音声認識手段により認識された音声の認識段階が前記命令語の各発声段階に達する毎に、前記対応情報記憶手段に記憶されている段階情報により示される、当該発声段階に対応する前記一連の撮影動作における各段階の動作を順に開始させる制御手段とを備え、

前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、前記音声認識手段は、撮影動作における自動焦点調整の動作中に生じるノイズ成分を含む音響モデルと含まない音響モデルとを有し、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識することを特徴とするカメラ装置。

【請求項2】

前記一連の撮影動作におけるいずれかの段階の動作は、連続して行われる複数の動作か

らなることを特徴とする請求項 1 記載のカメラ装置。

【請求項 3】

前記命令語記憶手段に複数の命令語が記憶されるとともに、前記対応情報記憶手段に、前記命令語記憶手段に記憶されている複数の命令語の各々に設定されている各発声段階と、複数の一連の撮影動作の各々における各段階の動作との対応関係を示す複数の対応情報が記憶されたことを特徴とする請求項 1 又は 2 記載のカメラ装置。

【請求項 4】

前記音声認識手段により認識された音声からなる語句を新たな命令語として前記命令語記憶手段に記憶させる登録手段と、

この登録手段により前記命令語記憶手段に記憶された新たな命令語における複数の発声段階と、前記一連の撮影動作における各段階の動作との対応関係を示す新たな対応情報を生成し、前記対応情報記憶手段に記憶させる生成手段と

を備えたことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載のカメラ装置。

【請求項 5】

所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置における撮影方法であって、

入力する音声を逐次認識する工程と、

認識した音声の認識段階が前記命令語に設定されている複数の発声段階の各発声段階に達する毎に、各発声段階に対応付けられている前記一連の撮影動作における各段階の動作を順に開始する工程とを含み、

前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、音声を逐次認識する工程は、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識することを特徴とする撮影方法。

【請求項 6】

所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置が有するコンピュータに、

入力する音声を音声認識手段に逐次認識させる処理と、

前記音声認識手段により認識した音声の認識段階が前記命令語に設定されている複数の発声段階の各発声段階に達する毎に、各発声段階に対応付けられている前記一連の撮影動作における各段階の動作を装置各部に順に開始させる処理とを実行させ、

前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、音声を音声認識手段に逐次認識する処理は、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識することを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声シャッター機能を有するカメラ装置、及び撮影方法に関するものである。

【背景技術】

【0002】

従来、カメラ装置においては、登録されている命令語（認識対象語）の音声を認識したことをトリガとしてオートフォーカスによるピント合わせ、及び露光からなる一連の撮影動作を行う音声シャッター機能を備えたものが下記の特許文献 1 等に記載されている。

【特許文献 1】特開平 11 - 194392

【発明の開示】

【発明が解決しようとする課題】

【 0 0 0 3 】

しかしながら、上記の技術においては、例えば「ハイ、チーズ」という命令語の認識を完了した後で合焦動作と露光動作とを行うため、ユーザーが命令語を発してから、実質的な撮影動作が行われるまでの間に若干のタイムラグが発生するという問題があった。

【 0 0 0 4 】

本発明は、かかる従来の課題に鑑みてなされたものであり、撮影のための命令語を発してから、実際の撮影動作が行われるまでの間のタイムラグを殆どなくすることが可能となるカメラ装置、及び撮影方法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】

【 0 0 0 5 】

前記課題を解決するため請求項1の発明にあっては、所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置であって、入力する音声認識する音声認識手段と、前記所定の命令語を記憶する命令語記憶手段と、この命令語記憶手段に記憶されている命令語に設定されている複数の発声段階と、前記一連の撮影動作における各段階の動作との対応関係を示す対応情報を記憶する対応情報記憶手段と、前記音声認識手段により認識された音声の認識段階が前記命令語の各発声段階に達する毎に、前記対応情報記憶手段に記憶されている段階情報により示される、当該発声段階に対応する前記一連の撮影動作における各段階の動作を順に開始させる制御手段とを備え、前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、前記音声認識手段は、撮影動作における自動焦点調整の動作中に生じるノイズ成分を含む音響モデルと含まない音響モデルとを有し、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識するものとした。

【 0 0 0 6 】

かかる構成においては、所定の命令語が音声入力されたとき、命令語の音声認識が完了する以前の段階から、一連の撮影動作における各段階の動作が順に開始される。

【 0 0 0 7 】

また、請求項2の発明にあっては、前記一連の撮影動作におけるいずれかの段階の動作は、連続して行われる複数の動作からなるものとした。

【 0 0 0 9 】

また、請求項4の発明にあっては、前記命令語記憶手段に複数の命令語が記憶されるとともに、前記対応情報記憶手段に、前記命令語記憶手段に記憶されている複数の命令語の各々に設定されている各発声段階と、複数の一連の撮影動作の各々における各段階の動作との対応関係を示す複数の対応情報が記憶されたものとした。

【 0 0 1 0 】

また、請求項5の発明にあっては、前記音声認識手段により認識された音声からなる語句を新たな命令語として前記命令語記憶手段に記憶させる登録手段と、この登録手段により前記命令語記憶手段に記憶された新たな命令語における複数の発声段階と、前記一連の撮影動作における各段階の動作との対応関係を示す新たな対応情報を生成し、前記対応情報記憶手段に記憶させる生成手段とを備えたものとした。

【 0 0 1 2 】

また、請求項5の発明にあっては、所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置における撮影方法であって、入力する音声を逐次認識する工程と、認識した音声の認識段階が前記命令語に設定されている複数の発声段階の各発声段階に達する毎に、各発声段階に対応付けられている前記一連の撮影動作における各段階の動作を順に開始する工程とを含み、前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、音声を逐次認識する工程は、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変

10

20

30

40

50

更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識する方法とした。

【0013】

かかる方法によれば、所定の命令語が音声入力されたとき、命令語の音声認識が完了する以前の段階から、一連の撮影動作における各段階の動作が順に開始される。

【0014】

また、請求項6の発明にあっては、所定の命令語の音声入力をトリガとして複数段階の動作からなる一連の撮影動作を行う自動焦点調整機能を備えるカメラ装置が有するコンピュータに、入力する音声を音声認識手段に逐次認識させる処理と、前記音声認識手段により認識した音声の認識段階が前記命令語に設定されている複数の発声段階の各発声段階に達する毎に、各発声段階に対応付けられている前記一連の撮影動作における各段階の動作を装置各部に順に開始させる処理とを実行させ、前記一連の撮影動作における複数段階の動作には自動焦点調整機能による合焦動作及び露光動作を含み、音声を音声認識手段に逐次認識する処理は、前記一連の撮影動作における自動焦点調整を指示した命令語の認識に 10
応答して前記ノイズ成分を含む音響モデルに、含まない音響モデルから変更して、自動焦点調整を指示した前記命令語の発声段階の次の発声段階の音声を、認識するプログラムと
した。

【発明の効果】

【0015】

以上のように本発明においては、所定の命令語が音声入力されたとき、命令語の音声認識が完了する以前の段階から、一連の撮影動作における各段階の動作が順に開始されるようにした。よって、撮影のための命令語を発してから、実際の撮影動作が行われるまでの間のタイムラグを殆どなくすることが可能となる。

【発明を実施するための最良の形態】

【0016】

以下、本発明の一実施の形態を図にしたがって説明する。図1は本発明に係るデジタルカメラの電氣的な概略構成を示すブロック図である。

【0017】

このデジタルカメラは音声シャッター機能を備えたものであり、キー入力部1と、合焦部2、露光部3、画像入力部4、画像圧縮部5、画像記憶部6、画像表示部7、マイク8、音声入力部9、A/D変換部(A/D)10、ワークメモリ11、音声認識部12、プログラムメモリ13を含み、キー入力部1及びマイク8を除く上記各部2~7、8~13が制御部14によって駆動及び制御される構成である。

【0018】

前記キー入力部1は、電源キーやシャッターキー、撮影/再生モードのモード切替キーや、各種機能の設定に使用される操作キー等、ユーザーがデジタルカメラの操作に使用する各種のキーからなり、いずれかのキーが操作されるとそれが制御部14により検出される。前記合焦部2には、撮影モードにおいて、図示しない光学系におけるフォーカスレンズを被写体距離に応じた位置に駆動するフォーカス・モータや、その駆動回路が含まれる。

【0019】

前記露光部3には、前記光学系により結像された被写体像を撮像するCCD等の撮像素子と、その駆動回路、撮像素子から出力されるアナログの撮像信号をデジタル信号に変換するA/D変換器等を含みデジタルの撮像信号を出力する。前記画像入力部4は、デジタルに変換された撮像信号に種々の信号処理を施すための各種の信号処理回路により構成される。前記画像圧縮部5は、画像入力部4で処理された後の画像データを圧縮し、また圧縮状態の画像データを伸張する回路により構成される。前記画像記憶部6は圧縮された画像データを記憶する各種メモリカード等により構成される。前記画像表示部7は、液晶表示器や、その駆動回路等から構成され、画像入力部4で信号処理された被写体画像や、画像記憶部6から読み出された記録画像を液晶表示器に表示させる。

10

20

30

40

50

【 0 0 2 0 】

前記音声入力部 9 は、マイク 8 から入力した音声を増幅するアンプや他の音声処理回路から構成され、処理後の音声信号を出力する。A / D 変換部 1 0 は音声入力部 9 から出力されるアナログの音声信号をデジタル信号に変換する。ワークメモリ 1 1 は、変換後の音声信号（音声データ）を逐次記憶したり、制御部 1 4 が各部の制御に際して生成したり使用する各種のデータ等を随時記憶する R A M である。

【 0 0 2 1 】

前記音声認識部 1 2 は、音声シャッター機能がオン設定されているときの撮影待機状態において、ワークメモリ 1 1 に逐次記憶される入力音声に対し、そのデータの蓄積速度と同じ又はそれ以上（遅れた場合にすぐ追いつけるという意味）の速度で、前記プログラムメモリ 1 3 に記録されている音響モデルを使用し、特徴抽出とピタビアルゴリズムによる認識処理を行う。認識結果を制御部 1 4 に逐次送る。

10

【 0 0 2 2 】

制御部 1 4 は、主として C P U や入出力インターフェイスを含む周辺回路から構成されている。前記プログラムメモリ 1 3 は E E P R O M 等の不揮発性のメモリであり、音声認識部 1 2 が音声認識に際して使用する前述した音響モデル等のデータや、制御部 1 4 に上記各部を制御させるための各種のプログラム、例えば A E（自動露出）、A F（自動焦点調整）等の制御プログラム、さらには制御部 1 4 を本発明の制御手段、登録手段、生成手段として機能させるためのプログラムが格納されている。

【 0 0 2 3 】

また、プログラムメモリ 1 3 は本発明の命令語記憶手段、対応情報記憶手段であって、プログラムメモリ 1 3 には、前述したプログラムやデータと共に図 2 に示した登録テーブル T を構成するデータとが格納されている。登録テーブル T は、複数種の予約語（音声パターン）1 0 1、及び各々の予約語 1 0 1 に設定されている複数の発声段階と、一連の撮影動作における各段階の動作との関係を示す対応情報 1 0 2 から構成される。予約語 1 0 1 は音声シャッター機能を用いた撮影時に使用可能な命令語であり、予約語 1 0 1 には、その内容に応じた複数の発声段階が設定されている。

20

【 0 0 2 4 】

発声段階数は、予約語 1 0 1 「ハイ、チーズ」が「ハイ」まで（第 1 の発声段階）と、「・・・チーズ」まで（第 2 の発声段階）の 2 段階である以外は 3 段階である。また、本実施の形態において各々の発声段階に対応して設定可能な動作は、自動焦点調整、露光、記録の 3 種類であり、動作の数は予約語 1 0 1 「ハイ、トリマス、モウイチマイ」の第 3 の発声段階（「・・・モウイチマイ」）に対応する記録、露光、記録の 3 動作が最大である。また、前述した音響モデルのうち、上記複数の予約語 1 0 1 に含まれる「チーズ」と「トリマス」部分の音響モデルには、自動焦点調整に伴い生じるノイズ（光学系やモータの駆動音）を重畳した P C M で学習済みの H M M（H i d d e n M a r k o v M o d e l、隠れマルコフ・モデル）が用意されている。

30

【 0 0 2 5 】

次に、以上の構成からなるデジタルカメラの本発明に係る動作を説明する。図 3 は、音声シャッター機能がオン設定されているときの音声シャッターモードによる動作を示したフローチャートである。

40

【 0 0 2 6 】

音声シャッターモードでは、例えばユーザーによるシャッターキーの押下に応じてマイク 8 からの音声入力処理を開始する（ステップ S A 1）。そして、入力される新たな音節部分を音声認識部 1 2 により逐次認識し、認識中の音声パターンを逐次蓄積する（ステップ S A 2）。音声認識部 1 2 における音声認識の手法は以下の通りである。

【 0 0 2 7 】

プログラムメモリ 1 3 には前述した予約語 1 0 1 が図 4 に示したような文法でセットされている。同図は「ハイ、チーズ」である例である。「無」は無音を表すサイレントモデルであり、「ハイ、チーズ」の発話の前後と「ハイ」の後に「無音」があっても無くても

50

対応できる文法である。ここでは、説明を簡単にするため音節を単位とした音響モデル（HMM）で説明する。なお、音響モデルは、「はいチーズ」という単語HMMであっても良いし、音素や半音素、トライフォンなどもっと細かい単位のHMMでもよい。

【0028】

ここで、無音モデルは1状態モデル、音節モデルは3状態モデルであるとする。

例として「は」で説明すると、

「は」の第1状態とは

「は」の子音「h」の部分の音響的特徴を出力する確率が高いHMMモデル

「は」の第2状態とは

「は」の子音と母音のつなぎ部分の音響的特徴を出力する確率が高いHMMモデル

10

「は」の第3状態とは

「は」の母音「a」の部分の音響的特徴を出力する確率が高いHMMモデル

というように学習されたモデルである。

【0029】

累積尤度計算は図5のように行われる。

例えば時刻Nの「い」の第1状態の累積尤度は、

$A = (\text{時刻} N - 1 \text{ の「は」の第3状態の累積尤度}) \times (\text{時刻} N \text{ の特徴値が「い」の第1状態である出力確率}) \times (\text{「は」の第3状態から「は」を終了する状態遷移確率})$

$B = (\text{時刻} N - 1 \text{ の「い」の第1状態の累積尤度}) \times (\text{時刻何の特徴値が「い」の第1状態である出力確率}) \times (\text{「い」の第1状態から「い」の第1状態に遷移する(ループ遷移)状態遷移確率})$

20

として求めたA, Bを比較し、確率の大きい方を時刻Nの「い」の第1状態の累積尤度とする。

【0030】

このようにして、全ての状態についての累積尤度を時刻毎に更新していく。例えば時刻Nが、「はいチーズ」と発話する際の「い」の発話後、「ち」を発話する直前の瞬間であったとすれば、累積尤度では「い」の第3状態か、「い」の第3状態と「ち」の第1状態の間の「無」モデルの尤度が最も高くなっている筈である。

【0031】

しかし、ここで問題となるのが、無関係な発話をした場合やノイズを拾った場合であって、その際にも累積尤度が最も高い点は存在するので、偶然「い」の最終状態の尤度が高くなってしまふことが考えられる。そこで、発話の信頼度を計算する必要がある。「はいチーズ」に使われない音節（たとえば「あ」とか「う」「ん」など）についても毎フレーム出力確率を求め、各フレームで最も出力確率の高かったモデルの値を乗算（ピタビならばlogを使うので加算しておく）する。もし発話が本当に「はい」であったならば、各フレームで最も尤度の高いモデルは「は」又は「い」のどこかの状態である可能性が高いので、「はい」の累積尤度との差は少なくなる。これを基に信頼度を求める。もちろん信頼度の計算方法は、必ずしもこの方法である必要はない。

30

【0032】

以上の手法により音声認識部12は、「い」の最終状態、または「い」と「ち」の間の「無」モデルの累積尤度が最も高く、かつ信頼度がしきい値以上である時、制御部14に、現時点で「はい」の発話が終わったことを示す信号を送る。同様に「ず」の最終状態の尤度、またはそれに続く「無」モデルの累積尤度が最も高く、かつ信頼度がしきい値以上となった時、「はいチーズ」の発話が終わったことを示す信号を送る。

40

【0033】

一方、音声認識部12が音声認識を行う間に、制御部14は逐次認識された音声パターンを内部メモリに蓄積しており、音声認識部12から上記信号が送られる毎に、認識中の音声パターンと登録テーブルTの内容とを比較して、候補となる予約語101とその発声段階とを特定し（ステップSA3）、それらを示す認識状況情報を更新する（ステップSA4）。図6は、ある時点の認識状況情報103の内容を示したものであり、同図(a)

50

は「ハイ」までが認識された時点の内容である。

【0034】

引き続き、制御部14は、予約語101の候補が残されていれば(ステップSA5でYES)、さらに次の発声段階に進んだか否かを判別する(ステップSA6)。そして、未だ次の発声段階に進んでいなければ(ステップSA6でNO)、ステップS2以降の処理を繰り返す。やがて次の発声段階に進んだら、つまり前の発声段階が終了したことが確定したら(ステップSA6でYES)、その時点での認識状況情報103に基づき、残されている予約語候補に対応する現段階の動作を行うための処理を実施する(ステップSA7)。すなわち認識状況情報103の内容が図6(a)に示した内容であり、ユーザーが「はい」までを発話した段階では自動焦点調整を開始する。以後、現在の発声段階が、認識中の予約語(候補)の最終段階となるまで(ステップSA8でNO)、ステップAS2へ戻って前述した処理を繰り返す。

10

【0035】

その後、認識状況情報103の内容が例えば図6(b)に示したものとなり、ユーザーが「はい、撮ります」までを発話した段階では、ステップSA7において露光を行い、さらに認識状況情報103の内容が例えば図6(c)に示したものとなり、ユーザーが「ハイ、トリマス、オーケー」までを発話した段階では、ステップSA7において記録を行う。やがて上記処理を繰り返す間に予約語候補がなくなった場合や(ステップSA5でNO)、現在の発声段階が認識中の予約語101の最終段階となったら(ステップSA8でYES)、認識状況情報103の内容をクリアする(ステップSA9)。しかる後、ステップAS2へ戻り、以降の処理を初めから繰り返す。

20

【0036】

以上のように音声シャッターモードにおいては、予め登録されている予約語(命令語)を音声認識して一連の撮影動作を行うが、予約語全体の認識が完了する以前の認識途中において、その認識段階が予約語に設定されている発声段階に達する毎に、撮影に要する動作を開始する。したがって、ユーザーが予約語を発してから、実質的な撮影動作である露光が行われるまでの間のタイムラグを殆どなくすることができる。

【0037】

また、一連の撮影動作の内容が異なる複数の予約語が登録されているため、ユーザーは、複数の予約語を使い分けることにより、全体の動作内容が異なる撮影動作を指示することができる。しかも「チーズ」と「トリマス」部分、つまり自動焦点調整直後の露光を指示する部分の認識が、自動焦点調整に伴い生じるノイズの成分を含む音響モデルを用いて行われるため、上記部分が自動焦点調整中に発声されたとしても、その部分にも高い認識率を確保することができる。

30

【0038】

一方、図7は、前記デジタルカメラにおいて予め用意されているコマンド登録モードによる制御部14の処理手順を示すフローチャートである。

【0039】

コマンド登録モードが設定されると制御部14は、コマンド入力方法の選択画面を画像表示部7のLCD画面に表示させ、ユーザーに入力方法を選択させる(ステップSB1)。本実施の形態では、入力方法として「選択」と「自由」の2種類が用意されている。ここでユーザーが「選択」を選ぶと(ステップSB2でYES)、既に登録されている予約語(図2参照)の各発声段階部分からいずれかの言葉をコマンド要素として選択させる(ステップSB3)。また、ユーザーが「自由を」を選ぶと(ステップSB2でNO)、所定のキー操作によって任意の言葉をコマンド要素として入力させる(ステップSB4)。なお、任意の言葉の入力方法は、例えば画像表示部7のLCD画面に50音等の文字選択画面を表示し、キー操作で複数の文字を選択させることにより行う。

40

【0040】

引き続き、指定可能な撮影時における動作、すなわち本実施の形態では「自動焦点調整」、「露光」、「記録」の3種類を表示して、その中から選択又は入力されたコマンド要

50

素に対応する1又は複数の動作を選択させる(ステップS B 5)。そして、発声段階数をインクリメントした後(ステップS B 1)、コマンド入力の終了指示がなければ(ステップS B 7でNO)、ステップS B 1へ戻り、前述と同様の処理によってユーザーに次の発生段階のコマンド要素を選択又は入力させ、かつそれと対応する動作を選択させる。

【0041】

そして、ステップS B 6で1又は複数の動作を選択させた後、コマンド入力の終了指示があったら(ステップS B 7でYES)、選択又は入力された複数のコマンド要素の間に「、」(発声段階の区切り)を自動的に挿入し、新たな予約語を生成する(ステップS B 8)。なお、図示しないが係る処理はユーザーによって選択又は入力されたコマンド要素が1つであった場合には、そのコマンド要素をそのまま予約語とする。そして、生成した新たな予約語を追加して登録テーブルTを更新する(ステップS B 9)。

10

【0042】

したがって、ユーザーにおいてはコマンド登録モードを選択することにより、前述した音声シャッターモードでの撮影時に使用可能な命令語を、自由に追加することができる。

【0043】

なお、以上の説明においては本発明を専用機としてのデジタルカメラに採用した場合について説明したが、これに限らず本発明は、カメラ付き携帯電話、さらには銀塩カメラ等の他のカメラ装置にも採用することができる。

【図面の簡単な説明】

【0044】

20

【図1】本発明に係るデジタルカメラの概略構成を示すブロック図である。

【図2】プログラムメモリに記憶されているデータを示す概念図である。

【図3】音声シャッターモードでの動作を示すフローチャートである。

【図4】記録されている予約語の文法を示す図である。

【図5】音声認識中の累積尤度計算方法を示す図である。

【図6】音声認識中における認識状況情報の変化例を示した図である。

【図7】コマンド登録モードでの制御部の処理手順を示すフローチャートである。

【符号の説明】

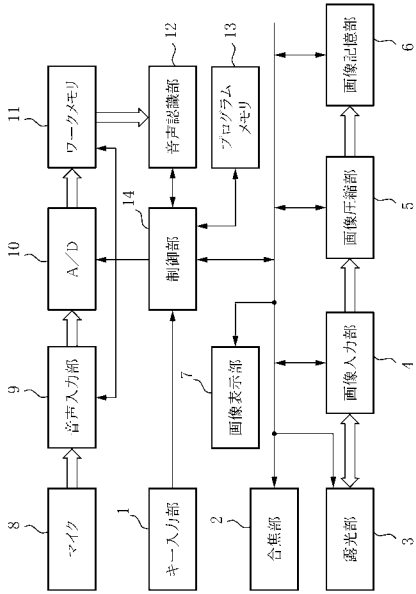
【0045】

- 1 キー入力部
- 2 合焦部
- 3 露光部
- 7 画像表示部
- 8 マイク
- 9 音声入力部
- 10 A/D変換部
- 11 ワークメモリ
- 12 音声認識部
- 13 プログラムメモリ
- 14 制御部
- 101 予約語
- 102 対応情報
- T 登録テーブル

30

40

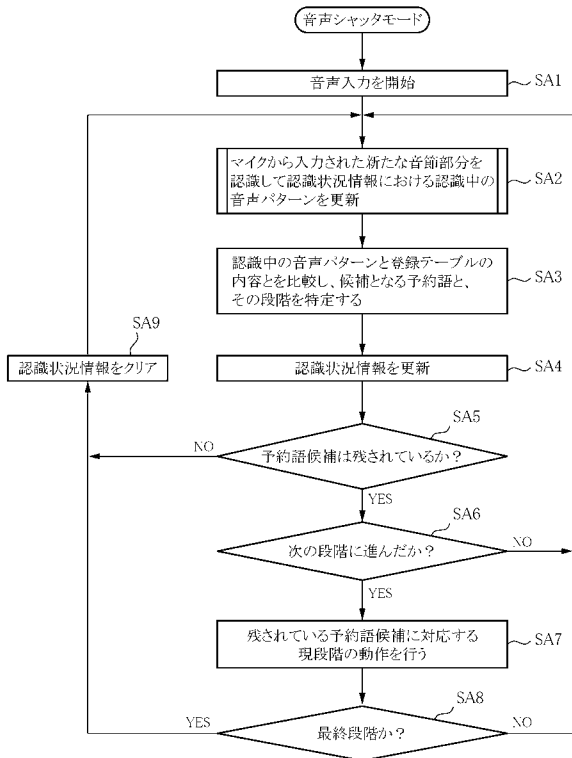
【図1】



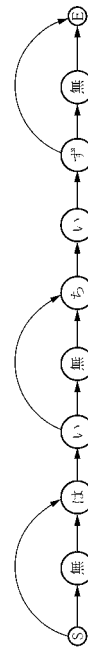
【図2】

NO.	予約語(音声パターン)	第1の発声段階	第1の動作段階	第2の発声段階	第2の動作段階	第3の発声段階	第3の動作段階
1	「ハイ、チーズ」	「ハイ」	自動焦点	「...チーズ」	露光&記録	—	—
2	「ハイ、トリマズ、オーケー」	「ハイ」	自動焦点	「...トリマズ」	露光	「...オーケー」	記録
3	「ハイ、トリマズ、トリマズ」	「ハイ」	自動焦点	「...トリマズ」	露光	「...トリマズ」	露光&記録
4	「ハイ、トリマズ、セウイチャマイ」	「ハイ」	自動焦点	「...セウイチャマイ」	露光	「...セウイチャマイ」	記録&露光&記録
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

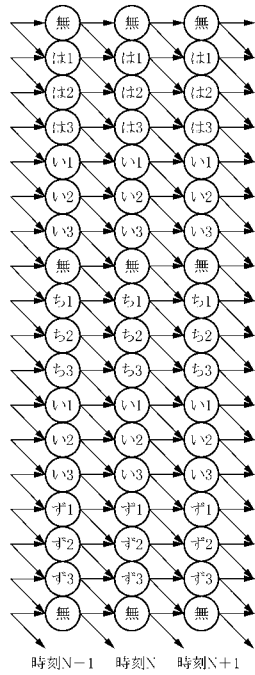
【図3】



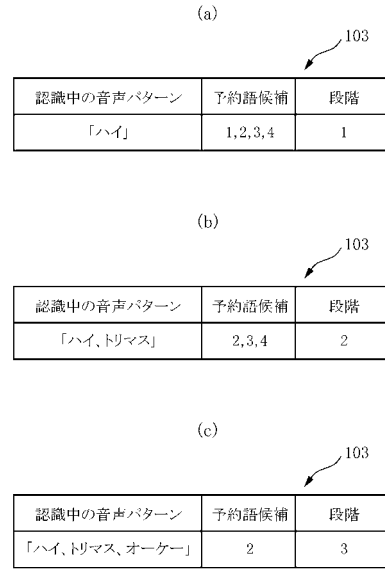
【図4】



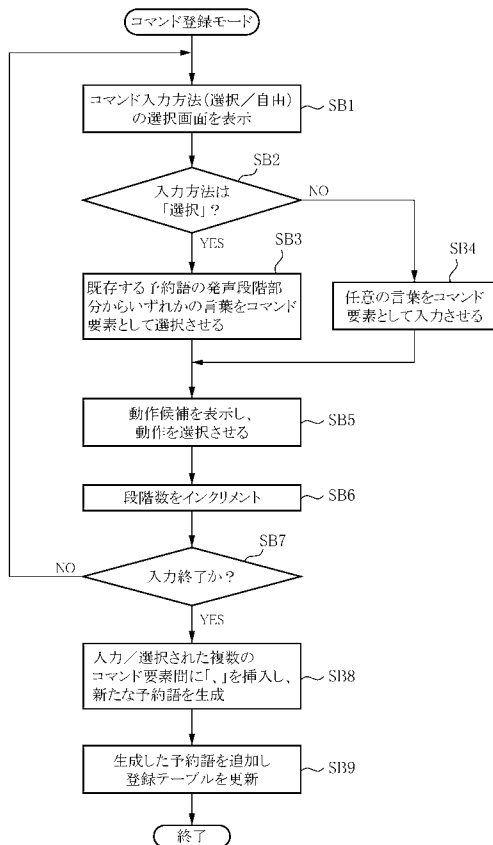
【 図 5 】



【 図 6 】



【 図 7 】



フロントページの続き

(51)Int.Cl.

F I

H 0 4 N 5/232

Z

H 0 4 N 101:00

(58)調査した分野(Int.Cl. , DB名)

G 0 3 B 1 5 / 0 0

G 0 3 B 1 7 / 3 8

H 0 4 N 5 / 2 2 2 - 5 / 2 5 7