(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第5146429号
(P5146429)

(45)発行日　平成25年2月20日(2013.2.20)　(24)登録日　平成24年12月7日(2012.12.7)

(51)Int.Cl.		F I		
G10L 15/06	(2013.01)	G10L	15/06	300C
G10L 15/00	(2013.01)	G10L	15/00	200F
G06F 3/16	(2006.01)	G10L	15/06	300J
		G06F	3/16	320H

請求項の数 9　(全 28 頁)

(21)出願番号　特願2009-218132(P2009-218132)
(22)出願日　平成21年9月18日(2009.9.18)
(65)公開番号　特開2011-65108(P2011-65108A)
(43)公開日　平成23年3月31日(2011.3.31)
審査請求日　平成22年3月24日(2010.3.24)

(73)特許権者　303000372
コニカミノルタビジネステクノロジーズ株式会社
東京都千代田区丸の内二丁目７番２号
(74)代理人　100086933
弁理士　久保　幸雄
(72)発明者　伊藤　歩
東京都千代田区丸の内一丁目６番１号　コニカミノルタビジネステクノロジーズ株式会社内

審査官　安田　勇太

最終頁に続く

(54)【発明の名称】画像処理装置、音声認識処理装置、音声認識処理装置の制御方法、およびコンピュータプログラム

(57)【特許請求の範囲】
【請求項１】
原稿の画像を読み取る画像読取装置を備えた画像処理装置であって、
ユーザに音声を入力させるための音声入力手段と、
前記音声を認識するための語句情報によって構成される辞書を記憶する辞書記憶手段と、
複数の所定の語句についての順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成する複合語句生成手段と、
生成された前記複数の複合語句を前記語句情報として前記辞書に登録する複合語句登録手段と、
前記複数の所定の語句を任意の順序で発声した発声語句に係る音声が入力されたときに、前記辞書の中から当該発声語句と一致する前記複合語句を検索することによって当該音声を認識する、音声認識手段と、
前記音声認識手段による前記音声の認識の結果に基づいて前記画像に対する処理を実行する画像処理手段と、
を有することを特徴とする画像処理装置。
【請求項２】
前記複合語句に関連付けられた関連情報を記憶する関連情報記憶手段を有し、
前記画像処理手段は、前記音声認識手段による前記音声の認識の結果に基づいて前記関連情報記憶手段から前記関連情報を抽出し、当該関連情報を用いて前記画像に対する処理

を実行する、
請求項１記載の画像処理装置。
【請求項３】
当該画像処理装置が有する機能に関する機能語句を記憶する機能語句記憶手段と、
ユーザにより指定されたユーザ語句を前記関連情報記憶手段に書き込むユーザ語句書込み手段と、を有し、
前記複合語句生成手段は、前記機能語句および前記ユーザ語句についての前記複合語句を生成する、
請求項２記載の画像処理装置。
【請求項４】
表示装置に表示するために準備された複数の操作画面のそれぞれに対応して前記辞書が設けられており、
前記音声認識手段は、音声が入力されたときに前記表示装置に表示されている前記操作画面に対応する前記辞書を用いて当該音声を認識する、
請求項１ないし請求項３のいずれかに記載の画像処理装置。
【請求項５】
ユーザにより指定されたユーザ語句が前記関連情報記憶手段に既に書き込まれている場合に、前記表示装置に当該ユーザに対する警告のための表示を行う、
請求項４に記載の画像処理装置。
【請求項６】
音声を認識するための音声認識処理装置であって、
ユーザに音声を入力させるための音声入力手段と、
前記音声を認識するための語句情報によって構成される辞書を記憶する辞書記憶手段と、
複数の所定の語句についての順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成する複合語句生成手段と、
生成された前記複数の複合語句を前記語句情報として前記辞書に登録する複合語句登録手段と、
前記複数の所定の語句を任意の順序で発声した発声語句に係る音声が入力されたときに、前記辞書の中から当該発声語句と一致する前記複合語句を検索することによって当該音声を認識する、音声認識手段と、
を有することを特徴とする音声認識処理装置。
【請求項７】
前記複合語句に関連付けられた関連情報を記憶する関連情報記憶手段を有し、
前記音声認識手段による前記音声の認識の結果に基づいて前記関連情報記憶手段から前記関連情報を抽出し、当該関連情報を用いて画像に対する処理を実行する、
請求項６記載の音声認識処理装置。
【請求項８】
音声認識処理装置において音声を認識するための制御方法であって、
音声を認識するための語句情報によって構成される辞書を前記音声認識処理装置に記憶しておき、
複数の所定の語句についての順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成し、
生成された前記複数の複合語句を前記語句情報として前記辞書に登録し、
前記複数の所定の語句を任意の順序で発声した発声語句に係る音声が入力されたときに、前記辞書の中から当該発声語句と一致する前記複合語句を検索することによって当該音声を認識する、
ことを特徴とする音声認識処理装置の制御方法。
【請求項９】
音声を認識するための音声認識処理装置を制御するコンピュータのためのコンピュータ

プログラムであって、
前記コンピュータによって実行されたときに、前記音声認識処理装置に、
複数の所定の語句についての順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成する手段と、
生成された前記複数の複合語句を、音声を認識するための語句情報として辞書に登録する手段と、
前記複数の所定の語句を任意の順序で発声した発声語句に係る音声が入力されたときに、前記辞書の中から当該発声語句と一致する前記複合語句を検索することによって当該音声を認識させる手段と、
を実現させることを特徴とす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】
本発明は、原稿の画像を読み取る画像読取装置を備えた画像処理装置、音声認識処理装置、音声認識処理装置の制御方法、およびコンピュータプログラムに関する。
【背景技術】
【０００２】
近年、画像処理装置に音声を認識するための機能（音声認識機能）が設けられるようになった。
【０００３】
例えば、画像処理装置に備えられたスキャナ機能などによって生成された画像データを、他の画像処理装置またはパーソナルコンピュータなどに送信する場合において、ユーザは音声によって画像データの送信先（宛先）を検索することができる。
【０００４】
音声認識には、音声を認識するための語句が登録された認識辞書テーブルが必要である。通常、認識辞書テーブルには単語単位に語句が登録されるので、単語単位でしか音声を認識することができない。そのため、複数の単語の音声を認識するには、複数の単語を一つの語句として認識するための辞書テーブルとなるグラマーに登録しておく必要がある。
【０００５】
例えば、ユーザによって既に登録されている「鈴木」という人のファックス番号を音声によって検索するために、「ファックス」、「スズキ」という２つの単語を連続して「ファックススズキ」と発声したとする。この場合は、図３３に示す従来の認識辞書テーブルＴＢ１０のように「ファックススズキ」という語句が登録されていれば、ユーザの音声は認識される。
【０００６】
ところが、ユーザは同じ組合せの語句をいつも特定の順序で発声するとは限らない。上記の例の場合に、ユーザは単語の順序を入れ替えて「スズキファックス」と発声することもある。
【０００７】
画像処理装置は、認識辞書テーブルに登録されている語句しか音声を認識することはできない。そのため、ユーザが、認識辞書テーブルに登録されている「ファックススズキ」とは異なる順序で「スズキファックス」と発声した場合は、ユーザの音声は認識されない。
【０００８】
よって、意味は同じでも順序を変えて発声した複数の語句の音声が認識されないことがあり、ユーザに戸惑いが生じることがある。また、ユーザにとっての画像処理装置の操作性も低下する。
【０００９】
従来において、音声によって個人名称を検索する場合において、音声認識用の名字辞書および名前辞書を作成し、音声により順番に入力された名字と名前についてそれぞれ音声

認識を実行する方法が提案されている（特許文献１）。
【先行技術文献】
【特許文献】
【００１０】
【特許文献１】特開２００２－１０８３８９号公報
【発明の概要】
【発明が解決しようとする課題】
【００１１】
しかし、特許文献１に記載された方法でも、音声により名字と名前とが正しい順番で入力されなければ、入力された名字および名前について音声を正しく認識することができない。つまり、特許文献１に記載される方法では、意味は同じで順序を変えて発声された複数の単語の音声を認識することができず、ユーザは不便を感じることがある。
【００１２】
本発明は、このような問題点に鑑み、発声する複数の単語または語句の順序をユーザが意識しなくても、音声認識を正しく行うことを目的とする。
【課題を解決するための手段】
【００１３】
本発明の実施形態に係る画像処理装置は、原稿の画像を読み取る画像読取装置を備えた画像処理装置であって、ユーザに音声を入力させるための音声入力手段と、前記音声を認識するための語句情報によって構成される辞書を記憶する辞書記憶手段と、複数の所定の語句についての順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成する複合語句生成手段と、生成された前記複数の複合語句を前記語句情報として前記辞書に登録する複合語句登録手段と、前記複数の所定の語句を任意の順序で発声した発声語句に係る音声が入力されたときに、前記辞書の中から当該発声語句と一致する前記複合語句を検索することによって当該音声を認識する、音声認識手段と、前記音声認識手段による前記音声の認識の結果に基づいて前記画像に対する処理を実行する画像処理手段と、を有する。
【００１４】
好ましくは、前記複合語句に関連付けられた関連情報を記憶する関連情報記憶手段を有し、前記画像処理手段は、前記音声認識手段による前記音声の認識の結果に基づいて前記関連情報記憶手段から前記関連情報を抽出し、当該関連情報を用いて前記画像に対する処理を実行する。
【００１５】
ユーザにより指定された語句、または予め登録された語句など、複数の所定の語句について、それらの順序の異なる全ての組合せからなるとともに、一文となるように所定の助詞および動詞を付加した複数の複合語句を生成して辞書に登録しておくことにより、音声で入力された語句の順序が異なってもそれを辞書の中から検索することができる。
【発明の効果】
【００１６】
本発明によると、発声する複数の単語または語句の順序をユーザが意識しなくても、音声認識を正しく行うことができる。
【図面の簡単な説明】
【００１７】
【図１】本発明の一実施形態に係る画像処理装置を含むネットワークシステムの例を示す図である。
【図２】一実施形態の画像形成装置のハードウェア構成の例を示す図である。
【図３】操作パネルの例を示す図である。
【図４】画像形成装置の機能的構成の例を示す図である。
【図５】第１実施形態における登録情報テーブルの例を示す図である。
【図６】機能語句情報テーブルの例を示す図である。

【図７】宛先登録画面の例を示す図である。
【図８】読み方登録画面の例を示す図である。
【図９】検索文字入力画面の例を示す図である。
【図１０】認識辞書テーブルの例を示す図である。
【図１１】辞書管理部の構成の例を示す図である。
【図１２】音声認識処理部の構成の例を示す図である。
【図１３】画像形成装置の音声認識処理の流れの例を示すフローチャートである。
【図１４】宛先選択画面の例を示す図である。
【図１５】宛先選択画面の例を示す図である。
【図１６】第２実施形態における登録情報テーブルの例を示す図である。
【図１７】機能語句情報テーブルの例を示す図である。
【図１８】ボックス登録画面の例を示す図である。
【図１９】読み方登録画面の例を示す図である。
【図２０】認識辞書テーブルの例を示す図である。
【図２１】ボックス選択画面の例を示す図である。
【図２２】ボックス選択画面の例を示す図である。
【図２３】第３実施形態における登録情報テーブルの例を示す図である。
【図２４】文書保存画面の例を示す図である。
【図２５】読み方入力画面の例を示す図である。
【図２６】警告表示画面の例を示す図である。
【図２７】認識辞書テーブルの例を示す図である。
【図２８】文書利用画面の例を示す図である。
【図２９】複合語句生成テーブルの例を示す図である。
【図３０】読み方登録画面の例を示す図である。
【図３１】第１～第３実施形態における画像形成装置全体の処理の流れの例を説明するためのフローチャートである。
【図３２】実施形態４における画像形成装置全体の処理の流れの例を説明するためのフローチャートである。
【図３３】従来の認識辞書テーブルの例を示す図である。
【発明を実施するための形態】
【００１８】
図１に示すネットワークシステムＳＹＳは、画像形成装置１を含むネットワークシステムである。画像形成装置１は、本発明に係る画像処理装置の一実施形態である。
【００１９】
図１に示すように、ネットワークシステムＳＹＳは、画像形成装置１、ファイルサーバ２ａ、端末装置３ａ、および通信回線ＮＴ１などによって構成される。
【００２０】
画像形成装置１、ファイルサーバ２ａ、および端末装置３ａは、それぞれ、通信回線ＮＴ１を介して互いに接続されている。通信回線ＮＴ１として、ＬＡＮ、ＷＡＮ、イントラネット、または専用線などが用いられる。
【００２１】
ファイルサーバ２ａは、電子文書（電子データ）を管理するためのファイルサーバである。
【００２２】
端末装置３は、ネットワークシステムＳＹＳのユーザが用いる端末装置である。各ユーザは、この端末装置３を用いて、後に説明する、画像形成装置１の「ＰＣプリント機能」を利用することができる。
【００２３】
ファックス端末４ａは、通信回線ＮＴ２を介して画像形成装置１または他のファックス端末との間でファクシミリによるデータの送受信を行う。なお、通信回線ＮＴ２として、

公衆回線などが用いられる。

【００２４】

画像形成装置１は、コピー機能、ＰＣプリント機能、ファックス送信機能、電子メール送信機能、ＳＭＢ送信機能、ＦＴＰ送信機能、およびボックス（Ｂｏｘ）機能などの様々なアプリケーション機能を集約した画像形成装置である。複合機またはＭＦＰ（Multi Function Peripherals）などと呼ばれることもある。

【００２５】

本実施形態の画像形成装置１には、上記のコピー機能などの基本的な機能のほかに、音声実行機能が備えられている。音声実行機能は、音声認識処理を実行し、その音声認識処理の結果に基づいて所定の処理を実行する機能である。音声認識処理は、図１に示す画像形成装置１の操作パネル１０ｆに供えられているマイクから入力される音声を、記憶媒体に記憶されている認識辞書テーブル（グラマー）に基づいて認識するための処理である。本実施形態では、ユーザが２つの単語（語句）を連続して発声することによって画像形成装置１を操作する場合を例に説明する。

【００２６】

図２に示すように、画像形成装置１は、ＣＰＵ１０ａ、ＲＡＭ１０ｂ、ＲＯＭ１０ｃ、ハードディスク１０ｄ、制御用回路１０ｅ、操作パネル１０ｆ、スキャナ部１０ｇ、プリンタ部１０ｈ、ファックス部１０ｊ、およびネットワークインタフェース１０ｋなどによって構成される。これらの全部または一部によって画像処理装置が形成される。

【００２７】

スキャナ部１０ｇは、原稿の用紙（以下、単に「原稿」と記載する。）に描かれている写真、文字、絵、図表などの画像を光学的に読み取って電子データ化する装置である。原稿として、紙以外の種々のシート、物などを用いることが可能である。

【００２８】

プリンタ部１０ｈは、スキャナ部１０ｇが読み取った画像データまたは他の装置から送信されてきた画像データを、イエロー、マゼンタ、シアン、およびブラックの４色のトナーを使用して用紙に印刷する装置である。

【００２９】

ファックス部１０ｊは、スキャナ部１０ｇが読み取った画像データを、通信回線を介してファックス端末に送信しまたはファックス端末から送信されてくる画像データを受信するための装置である。

【００３０】

ネットワークインタフェース１０ｋは、ＮＩＣ（Network Interface Card）であって、ＬＡＮまたはインターネットなどを介して他の装置への接続を行い、電子データの送受信などを行う。

【００３１】

制御用回路１０ｅは、ハードディスク１０ｄ、操作パネル１０ｆ、スキャナ部１０ｇ、プリンタ部１０ｈ、ファックス部１０ｊ、およびネットワークインタフェース１０ｋなどの装置を制御するための回路である。

【００３２】

操作パネル１０ｆは、ユーザが画像形成装置１に指示を与えまたは画像形成装置１からユーザにメッセージを通知するための装置である。

【００３３】

これらの各装置のうちの一部が連携することによって、前述のアプリケーション機能が実現される。

【００３４】

「コピー機能」は、スキャナ部１０ｇによって原稿の画像を読み取り、それによって得た画像データをプリンタ部１０ｈによって用紙に印刷する機能である。

【００３５】

「ＰＣプリント機能」は、ネットワークインタフェース１０ｋによって通信回線を介し

て画像形成装置１と互いに接続されている端末装置（以下、単に「端末装置」と記載する）などから画像データなどを受信し、プリンタ部１０ｈによってその画像データなどを用紙に印刷する機能である。この機能は、「ネットワークプリンティング」と呼ばれることもある。

【００３６】

「ファックス送信機能」は、スキャナ部１０ｇで読み取った画像データなどをファックス部１０ｊからファックス端末などに送信する機能である。

【００３７】

「ファックス受信機能」は、ファックス部１０ｊによってファックス端末などから画像データを受信し、プリンタ部１０ｈによってその画像データを用紙に印刷する機能である。

【００３８】

「電子メール送信機能」は、スキャナ部１０ｇで読み取った画像データなどを電子メールに添付してネットワークインタフェース１０ｋから送信する機能である。

【００３９】

「ＳＭＢ送信機能」は、スキャナ部１０ｇで読み取った画像データなどを、ＳＭＢ（Server Message Block）に基づいて、ＩＰアドレスなどで指定される送信先に直接送信する機能である。

【００４０】

「ＦＴＰ送信機能」は、スキャナ部１０ｇで読み取った画像データなどをＦＴＰ（File Transfer Protocol）に基づいて送信する機能である。

【００４１】

「ボックス機能」は、ユーザごとに「ボックス」または「パーソナルボックス」などと呼ばれる、パーソナルコンピュータにおけるフォルダまたはディレクトリなどに相当する記憶領域をハードディスク１０ｄに設けておき、ユーザがスキャナ部１０ｇを操作するなどして得た画像データなどを自分の記憶領域に保存できるようにする機能である。「ドキュメントサーバ機能」と呼ばれることもある。

【００４２】

ＰＣプリント機能を利用する際に端末装置において予めボックスを指定しておくことによって、プリントを行いつつプリントする画像データをそのボックス内に保存させることができる。ボックスの指定は、例えば、端末装置にインストールされている画像形成装置１のドライバの機能を用いて行うことができる。

【００４３】

図３に示すように操作パネル１０ｆには、タッチパネルＴＰ、マイクＭＫ、操作部ＴＫ、および音声認識ボタンＫＢ、コピーボタンＭＢ、スキャンＦＡＸボタンＳＢ、およびＢＯＸボタンＴＢなどが設けられている。

【００４４】

タッチパネルＴＰは、画像形成装置１からユーザに通知するメッセージまたは各種機能を利用するための画面（操作画面）などを表示する表示装置である。タッチパネルの機能によって入力装置としての役割も果たす。

【００４５】

マイク（マイクロフォン）ＭＫは、ユーザが音声（音声コマンド）を入力するための装置であり、入力された音声をアナログの音声データ（アナログ音声データＡＮＤＴ）に変換する。

【００４６】

操作部ＴＫは、プリントの枚数またはファックス番号などを入力するためのボタンのボタン群である。

【００４７】

音声認識ボタンＫＢは、音声認識モード（音声認識を行うためのモード）への切替えを行うためのボタンである。このボタンが押されると、画像形成装置１は、音声認識処理を

実行し、音声認識の結果に基づいて所定の処理を実行する。

【００４８】

コピーボタンＭＢは、印刷モード（画像データなどを用紙に印刷するためのモード）への切り替えを行うためのボタンである。このボタンが押されると、コピーについての設定を行う画面（コピー設定画面）がタッチパネルＴＰに表示される。

【００４９】

スキャンＦＡＸボタンＳＢは、送信モード（画像データなどを他の装置に送信するためのモード）への切り替えを行うためのボタンである。このボタンが押されると、画像データなどの送信先の設定を行うための画面がタッチパネルＴＰに表示される。

【００５０】

ＢＯＸボタンＴＢは、ボックスモード（ボックスへの画像データの保存またはボックスに保存されている画像データなどの参照を行うためのモード）への切り替えを行うためのボタンである。このボタンが押されると、画像データなどの保存先とするボックスまたは参照したいボックスの指定を行うための画面がタッチパネルＴＰに表示される。

【００５１】

図４において、画像形成装置１は、総合情報記憶部３０、機能語句情報記憶部３１、表示制御部３２、情報書込み部３３、辞書管理部３４、および音声認識処理部３５を有する。これらの機能を実現するためのプログラムは、当該プログラムが記録されたＣＤ－ＲＯＭ、ＤＶＤ－ＲＯＭ等の記録媒体ＭＳ１や半導体メモリ等の記録媒体ＭＳ２を含む可搬型の記録媒体ＭＳ（図２参照）からインストールすることが可能である。また、ネットワークを介してサーバからプログラムをダウンロードすることも可能である。

【００５２】

各部を実現するためのプログラムおよびデータは、ハードディスク１０ｄにインストールされ、これらのプログラムおよびデータは必要に応じてＲＡＭ１０ｂにロードされ、ＣＰＵ１０ａによってプログラムが実行される。なお、ＣＰＵ１０ａはコンピュータとして動作することが可能である。また、外部のコンピュータとＬＡＮなどを介して接続するようにしてもよい。また、図４の一部または全部の機能をプロセッサ（制御用回路）によって実現するように構成してもよい。

〔第１実施形態〕

次に、画像形成装置１の第１実施形態について、図４～図１５を参照して説明する。

【００５３】

以下、ユーザが、「電子メール送信機能」を利用して、スキャナ部１０ｇで読み取った画像データを電子メールで「上杉」という人に宛てに送信する場合を例に説明する。また、ユーザは、電子メールの宛先を検索する際に、検索対象の宛先名と宛先種別とを連続して発声するものとする。

【００５４】

図５は登録情報テーブルＴＢ１の例を示す図、図６は機能語句情報テーブルＴＢ２の例を示す図、図７は宛先登録画面ＧＭ１の例を示す図、図８は読み方登録画面ＧＭ２の例を示す図、図９は検索文字入力画面ＧＭ３の例を示す図、図１０は認識辞書テーブルＴＢ３の例を示す図、図１１は辞書管理部３４の構成の例を示す図、図１２は音声認識処理部３５の構成の例を示す図、図１３は画像形成装置１の音声認識処理の流れの例を示すフローチャート、図１４は宛先選択画面ＧＭ４の例を示す図、図１５は宛先選択画面ＧＭ５の例を示す図である。

【００５５】

これら、宛先登録画面ＧＭ１、読み方登録画面ＧＭ２、検索文字入力画面ＧＭ３、宛先選択画面ＧＭ４、宛先選択画面ＧＭ５は、操作画面の例である。

【００５６】

図４の各部の機能および処理内容などについて詳細に説明する。

【００５７】

総合情報記憶部３０は、登録情報テーブルＴＢ１を記憶している。登録情報テーブルＴ

Ｂ１には、図５に示すように、それぞれのユーザの宛先についての情報である宛先情報ＤＴ１が記憶され管理されている。宛先情報ＤＴ１の「登録番号」フィールドには、他の宛先情報ＤＴ１と区別するための登録番号（識別番号）ＢＴが示される。

【００５８】

「登録名」フィールドには、例えばユーザの氏名などの宛先名ＷＵ１が示される。「読み（登録名）」フィールドには、「登録名」フィールドに示される氏名などの読み方ＷＵ２が示される。「検索文字」フィールドには、宛先を登録名で検索するための文字（検索文字）ＷＵ３が示される。「宛先種別」フィールドには、「宛先」フィールドに示される宛先へのデータの送信方法が示される。送信方法としては、Ｅ－ｍａｉｌ、ＦＴＰ、ＳＭＢ、およびＦＡＸの他に、ＷｅｂＤＡＶやネットワークＢＯＸなどが挙げられる。「宛先」フィールドには、「宛先種別」フィールドに示される送信方法でデータを送信する際の宛先ＷＵ４が示される。

【００５９】

これら、宛先名ＷＵ１、読み方ＷＵ２、検索文字ＷＵ３、宛先ＷＵ４などは、ユーザによって指定されまたは入力される「ユーザ語句ＷＵ」である。第２～第４実施形態においても、同様に、符号ＷＵを付した要素は「ユーザ語句ＷＵ」である。なお、「ユーザ語句ＷＵ」は「設定語句ＷＵ」ということもある。

【００６０】

機能語句情報記憶部３１は、機能語句情報テーブルＴＢ２を記憶している。機能語句情報テーブルＴＢ２には、図６に示すように、画像形成装置１に備えられている送信機能を示す語句（機能語句）についての情報である機能語句情報ＤＴ２が記憶され管理されている。機能語句情報ＤＴ２の「機能名」フィールドには、画像形成装置１に備えられている送信機能を示す語句が示され、「読み（機能名）」フィールドには、「機能名」フィールドに示される語句の読み方が示される。

【００６１】

なお、「読み（機能名）」フィールドに書き込まれる語句は、機能語句であるが、これは同時に、画像形成装置１において予め準備されたシステム語句である。したがって、これらの機能語句およびシステム語句について、以降においては「システム語句ＷＳ」または「機能語句ＷＳ」と記載する。第２～第４実施形態においても、同様に、符号ＷＳを付した要素は「システム語句ＷＳ」または「機能語句ＷＳ」である。

【００６２】

表示制御部３２は、音声が入力されたタイミングまたは操作パネル１０ｆに備えられているボタンが押されたタイミングで、所定の画面を操作パネル１０ｆに表示させるための処理を行う。画像形成装置１にログインしているユーザがいないときは、ユーザＩＤおよびパスワードを入力するためのログイン用画面を表示させる。表示する画面にはそれぞれ他の画面と識別するための識別番号（画面番号）が付されており、表示制御部３２は総合制御部３６からの指令に係る画面番号の画面を操作パネル１０ｆに表示させる。

【００６３】

情報書込み部３３は、タッチパネルＴＰまたは操作部ＴＫの操作によってユーザにより指定されまたは入力された登録事項を、総合情報記憶部３０に書き込んで記憶させる（登録させる）ための処理を行う。

【００６４】

ここで、ユーザが宛先を登録するときの手順について、図７および図８の各画面を参照しながら説明する。

【００６５】

ユーザは、操作パネル１０ｆのスキャンＦＡＸボタンＳＢを押下することによって、宛先登録画面ＧＭ１（図７参照）を表示させる。ここで、ユーザは、登録番号、登録名、および送信先のアドレス（第１実施形態では、Ｅ－ｍａｉｌアドレス）を登録する。それぞれの項目は、登録番号ボタンＢＴＮ１、登録名ボタンＢＴＮ２、およびＥ－ｍａｉｌ宛先ボタンＢＴＮ３をそれぞれ押下することによって表示される、入力キーが表示された入力

キー画面（図示せず）から入力することができる。入力された項目は、登録番号ボタンＢＴＮ１、登録名ボタンＢＴＮ２、およびＥ－ｍａｉｌ宛先ボタンＢＴＮ３にそれぞれ対応付けられた表示欄ＲＮ１、ＲＮ２、ＲＮ３に表示される。

【００６６】

ユーザが、登録番号、登録名、および送信先のアドレスを入力しＯＫボタンを押下すると、読み方登録画面ＧＭ２（図８参照）が表示される。ここで、ユーザは、宛先登録画面ＧＭ１で登録した登録名を発声するときの読み方を登録する。ユーザは、登録名（読み方）ボタンＢＴＮ４を押下して入力キー画面（図示せず）を表示させ、読み方を入力する。読み方は、ひらがなまたはカタカナのどちらを入力してもよい。登録名が英語の場合は、英語の発音などを入力する。入力された読み方は、登録名（読み方）ボタンＢＴＮ４に対応付けられた表示欄ＲＮ４に表示される。

【００６７】

読み方登録画面ＧＭ２では、読み方以外に検索文字を入力し登録することもできる。この場合は、検索文字ボタンＢＴＮ５を押下することによって表示される検索文字入力画面ＧＭ３（図９参照）から、検索文字を入力することができる。ここで、常用ボタンＢＴＮ６を押下すると、検索対象の登録名を検索文字以外の「常用」という区分に振り分けることができる。

【００６８】

図８に戻って、ユーザが読み方を入力してＯＫボタンを押下すると、宛先登録画面ＧＭ１および読み方登録画面ＧＭ２での入力内容が宛先情報ＤＴ１として、総合情報記憶部３０に記憶されている登録情報テーブルＴＢ１に登録される。

【００６９】

図４に戻って、辞書管理部３４は、音声認識処理に用いる認識辞書テーブルＴＢ３（図１０参照）を保存し管理する。図１１に示すように、辞書管理部３４は、辞書データベース３４ａ、複合語句生成部３４ｂ、語句バッファ３４ｃ、および複合語句書込み部３４ｄを有する。

【００７０】

次に、図１１の各部の機能および処理内容などについて詳細に説明する。

【００７１】

図１１において、認識辞書データベース３４ａは、認識辞書テーブルＴＢ３を記憶し管理している。認識辞書テーブルＴＢ３は、操作パネル１０ｆに表示されるそれぞれの画面に対応付けて用意される。つまり、音声によって画像形成装置１を操作する場合、ユーザは、操作パネル１０ｆに表示されている画面（正確には、画面に表示されるボタンに示されるコマンド）などを見ながら語句（コマンド）を発声する。したがって、ユーザが発声した音声を認識するためには、表示されている画面に対応したコマンドを示す語句が、認識辞書テーブルＴＢ３に記憶されている必要がある。

【００７２】

図１０に示すように、認識辞書テーブルＴＢ３には、認識語句情報ＤＴ３が記憶され管理されている。認識語句情報ＤＴ３には、登録情報テーブルＴＢ１（図５参照）に記憶されている宛先情報ＤＴ１の登録番号に対応する語句番号ＢＴと、音声認識処理により音声データを変換して得られる文字列と比較するための語句（認識語句）ＷＮとが対応付けて示される。

【００７３】

認識語句ＷＮには、複合語句生成部３４ｂによって生成されて書き込まれた複合語句ＷＦと、画像形成装置１において各画面に対応してコマンドとして予め準備されたコマンド語句ＷＣとが含まれる。コマンド語句ＷＣは、画像データの送受信、画像データによる印刷などの画像処理、または画面の制御など、種々の処理または制御を実行させるコマンドとして用いられる。コマンド語句ＷＣには、上に述べたシステム語句ＷＳが含まれる。

【００７４】

語句番号ＢＴが１～９９９の認識語句情報ＤＴ３は、複合語句ＷＦについての情報であ

り、後に説明する複合語句書込み部３４ｄにより書き込まれて追加される。語句番号ＢＴが１０００～１００７の認識語句情報ＤＴ３は、コマンド語句ＷＣについての情報であり、画面番号欄ＧＲＮに示される画面番号の画面に表示されるコマンドに対応して予め登録されている。

【００７５】

複合語句生成部３４ｂは、複合語句ＷＦを生成するための処理を行う。すなわち、第１実施形態における複合語句ＷＦは、登録情報テーブルＴＢ１（図５参照）の「読み（登録名）」フィールドに示される語句ＷＵ２と、機能語句情報テーブルＴＢ２（図６参照）の「読み（機能名）」フィールドに示される語句ＷＳとの、２つの語句ＷＵ２，ＷＳについて、それらの順序の異なる全ての組合せとして生成される。なお、対象となる語句が２つの場合には、前後の配置を逆にして組み合わせることにより複合語句ＷＦが生成される。

【００７６】

なお、複合語句ＷＦを生成するための語句の組合せは、ユーザが適宜変更することができる。その場合に、ユーザは、コマンド語句ＷＣまたはシステム語句ＷＳの中から、その機能などに応じて適宜選択することができる。

【００７７】

次に複合語句ＷＦを生成するための処理について説明する。

【００７８】

複合語句生成部３４ｂは、総合情報記憶部３０で管理されている登録情報テーブルＴＢ１に宛先情報ＤＴ１が追加（登録）されると、その追加された宛先情報ＤＴ１を登録情報テーブルＴＢ１から抽出し語句バッファ３４ｃに一時的に記憶させる。それに伴い、機能語句情報記憶部３１で管理されている機能語句情報テーブルＴＢ２から、抽出した宛先情報ＤＴ１の「宛先種別」フィールドに示される送信方法と同じ機能名を示す機能語句情報ＤＴ２を抽出し、語句バッファ３４ｃに一時的に記憶させる。そして、語句バッファ３４ｃに記憶させた宛先情報ＤＴ１および機能語句情報ＤＴ２のそれぞれから「読み（登録名）」フィールドおよび「読み（機能名）」フィールドに示される文字の文字データを生成する。

【００７９】

例えば、登録情報テーブルＴＢ１に登録番号００６の宛先情報ＤＴ１が追加されると、その登録番号００６の宛先情報ＤＴ１を登録情報テーブルＴＢ１から抽出し、語句バッファ３４ｃに一時的に記憶させる。そして、登録番号００６の宛先情報ＤＴ１の「読み（登録名）」フィールドに示される語句ＷＵ２である「ウエスギ」の文字データを生成する。

【００８０】

また、登録番号００６の宛先情報ＤＴ１の「宛先種別」フィールドには、「Ｅ－ｍａｉｌ」が表示されているので、「機能名」フィールドに「Ｅ－ｍａｉｌ」が示される機能語句情報ＤＴ２を機能語句情報テーブルＴＢ２から抽出し、語句バッファ３４ｃに一時的に記憶させる。そして、その機能語句情報ＤＴ２の「読み（機能名）」フィールドに示される語句ＷＳである「イーメール」の文字データを生成する。

【００８１】

複合語句生成部３４ｂは、生成した２つの文字データを、前後の配置を逆にして組合せて、新たな文字データである複合語句ＷＦを生成する。例えば、上記の例の場合、「ウエスギ」および「イーメール」を前後の配置を逆にして組み合わせた、「ウエスギイーメール」および「イーメールウエスギ」の２つの複合語句ＷＦを生成する。

【００８２】

複合語句生成部３４ｂは、生成した複合語句ＷＦを、その複合語句ＷＦを生成するのに用いた宛先情報ＤＴ１の登録番号ＢＴとともに複合語句書込み部３４ｄに送る。

【００８３】

複合語句書込み部３４ｄは、複合語句生成部３４ｂから取得した複合語句ＷＦを、登録番号ＢＴと同じ語句番号ＢＴを用いて、認識語句ＷＮとして認識辞書テーブルＴＢ３に書き込む。

【００８４】
図４に戻って、音声認識処理部３５は、操作パネル１０ｆに備えられたマイクＭＫに入力される音声についての音声認識処理を行う。図１２に示すように、音声認識処理部３５は、Ａ／Ｄ変換部３５ａ、音声データバッファ３５ｂ、および音声解析部３５ｃを有する。
【００８５】
次に、音声認識処理について、図１２の各部の機能および処理内容などとともに説明する。
【００８６】
Ａ／Ｄ変換部３５ａは、マイクＭＫから出力されたアナログ音声データＡＮＤＴをデジタルの音声データ（デジタル音声データＤＧＤＴ）に変換する。
【００８７】
音声データバッファ３５ｂは、Ａ／Ｄ変換部３５ａから出力されたデジタル音声データＤＧＤＴを一時的に記憶する。
【００８８】
音声解析部３５ｃは、音声データバッファ３５ｂに記憶されているデジタル音声データＤＧＤＴを解析するための処理を行う。つまり、マイクＭＫから入力された音声を認識するための処理を行う。音声データバッファ３５ｂにデジタル音声データＤＧＤＴが記憶されると、そのデジタル音声データＤＧＤＴを文字列（認識対象語句）に変換する。そして、認識辞書テーブルＴＢ３の中から、認識対象語句と一致する認識語句ＷＮを検索する。認識対象語句と一致する認識語句ＷＮが検索された場合、つまり音声が認識された場合は、その認識語句ＷＮの語句番号ＢＴを、検索結果信号ＳＮとして総合制御部３６に送る。
【００８９】
総合制御部３６は、登録情報テーブルＴＢ１の中から、検索結果信号ＳＮに示される語句番号ＢＴと同じ登録番号ＢＴの宛先情報ＤＴ１を抽出し、その情報に基づいて処理を行う。
【００９０】
認識対象語句と一致する認識語句ＷＮが検索されなかった場合、つまり音声が認識されなかった場合は、音声が認識されなかった旨の検索結果信号ＳＮを総合制御部３６に送る。
【００９１】
図４に戻って、総合制御部３６は、画像処理部３７を制御するとともに、画像形成装置１の全体をも制御する。また、音声認識処理部３５からの検索結果信号ＳＮに基づいて、認識辞書テーブルＴＢ３に登録されている認識語句ＷＮに対応する処理を行う。
【００９２】
画像処理部３７は、総合制御部３６からの制御信号に応じて、スキャナ部１０ｇが読み取った画像データまたは他の装置から送信されてきた画像データなどに対し、種々の画像処理を行う。
【００９３】
次に、ユーザが音声によって電子メールの送信先の検索を行う場合における画像形成装置１の音声認識処理について、図１４、図１５の各画面、および図１３のフローチャートを参照しながら参照する。
【００９４】
ユーザが操作パネル１０ｆのスキャンＦＡＸボタンＳＢを押すと、図１４に示すような宛先選択画面ＧＭ４が表示される。宛先選択画面ＧＭ４には、登録情報テーブルＴＢ１に記憶されているそれぞれの宛先情報ＤＴ１に対応付けられている送信対象アイコンＳＭ（ＳＭ１～６）が表示される。それぞれの送信対象アイコンＳＭには、対応付けられている宛先情報ＤＴ１に示される登録名および宛先種別などが示される。ここで、ユーザが音声認識ボタンＫＢを押すと（図１３の＃４０１でＹｅｓ）、操作パネル１０ｆのマイクＭＫがオンになり（＃４０２）、音声の入力が可能となる。ユーザは、宛先種別（ここでは「

Ｅ－ｍａｉｌ」）および送信対象となる宛先の登録名（ここでは「上杉」）を示す語句（「イーメール、ウエスギ」または「イーメールウエスギ」）を発声することによって、音声を入力する。

【００９５】

画像形成装置１は、入力された音声について（＃４０３）、宛先選択画面ＧＭ４に対応付けられた認識辞書テーブルＴＢ３を用いて音声認識処理を行う（＃４０４）。つまり、入力された音声を認識対象語句、例えば「イーメールウエスギ」に変換し、認識辞書テーブルＴＢ３の中から「イーメールウエスギ」と一致する認識語句ＷＮを検索する。認識対象語句と一致する認識語句ＷＮが検索されると、これに基づいて音声が認識される。

【００９６】

第１実施形態では、語句番号ＢＴが００６の認識語句情報ＤＴ３に示される認識語句ＷＮが「イーメールウエスギ」と一致する。

【００９７】

音声が認識されると（＃４０５でＹｅｓ）、認識対象語句と一致する認識語句ＷＮの語句番号ＢＴと同じ登録番号ＢＴの宛先情報ＤＴ１に対応付けられている送信対象アイコンＳＭのみが表示された宛先選択画面ＧＭ５（図１５参照）を表示する（＃４０６）。

【００９８】

例えば、送信対象アイコンＳＭ６は、登録番号００６の宛先情報ＤＴ１に対応付けられている。画像形成装置１は、宛先選択画面ＧＭ５においてユーザが「オーケー」と発声しその音声が認識されると、登録情報テーブルＴＢ１に記憶されている登録番号００６の宛先情報ＤＴ１を参照し、「上杉」に対応した電子メールアドレスを示した確認画面（図示せず）を表示する（＃４０７）。ユーザは、電子メールアドレスを確認した後、送信ボタンを押すなどして画像データを電子メールで送信することができる。

【００９９】

認識対象語句と一致する認識語句ＷＮが検索されなかった場合、つまり音声が認識されなかった場合は、音声が認識されなかった旨の警告を示す画面（図示せず）を表示する（＃４０８）。

【０１００】

第１実施形態の認識辞書テーブルＴＢ３には、例えば「ウエスギイーメール」、「イーメールウエスギ」のように、２つの単語（「ウエスギ」、「イーメール」）の前後の配置が逆になった認識語句ＷＮが登録される。これにより、ユーザが「ウエスギ」および「イーメール」のどちらを先に発声しても、音声を認識することができる。また、２つの単語で検索対象を検索することができるので、単語を１つずつ発声する場合よりも早く検索対象を探し出すことができる。

〔第２実施形態〕

次に、画像形成装置１の第２実施形態について図１６～図２２を参照して説明する。なお、第１実施形態における画像形成装置１と同一の構成要素には、同一の符号を付してその詳細な説明は適宜省略する。以下同様である。

【０１０１】

図１６は登録情報テーブルＴＢ４の例を示す図、図１７は機能語句情報テーブルＴＢ５の例を示す図、図１８はボックス登録画面ＧＭ６の例を示す図、図１９は読み方登録画面ＧＭ７の例を示す図、図２０は認識辞書テーブルＴＢ６の例を示す図、図２１はボックス選択画面ＧＭ８の例を示す図、図２２はボックス選択画面ＧＭ９の例を示す図である。

【０１０２】

以下、ユーザが、「ボックス機能」を利用して、スキャナ部１０ｇで読み取った画像データを「上杉」という名前のボックスに保存する場合を例に説明する。また、ユーザは、保存先のボックスを検索する際に、検索対象のボックスのボックス名とボックスタイプとを連続して発声するものとする。

【０１０３】

図１６に、第２実施形態における登録情報テーブルＴＢ４を示す。登録情報テーブルＴ

Ｂ４には、画像形成装置１に登録されている各ユーザのボックスに関する情報であるボックス情報ＤＴ４が記憶され管理されている。ボックス情報ＤＴ４の「登録番号」フィールドには、他のボックス情報ＤＴ４と区別するための登録番号（識別番号）ＢＴが示される。この登録番号ＢＴは、各ボックスのボックス番号でもある。「ＢＯＸ名」フィールドには、ボックス名ＷＵ１１が示される。「読み（ＢＯＸ名）」フィールドには、「ＢＯＸ名」フィールドに示されるボックス名の読み方ＷＵ１２が示される。「検索文字」フィールドには、ボックスをボックス名で検索するための検索文字ＷＵ１３が示される。「ＢＯＸタイプ」フィールドには、ボックスの種類（グループ名）が示される。

【０１０４】

その他、図１６には示されていないが、登録情報テーブルＴＢ４には、各ボックスの登録場所または送信種別なども示されることがある。例えば、そのボックスが登録されているのが画像形成装置１ではなく通信回線を通じて画像形成装置１に接続されているサーバである場合には、そのボックスの登録場所としてサーバ名が表示される。また、データの送信方法など（例えば、ＦＴＰ、ＳＭＢ、ＷｅｂＤＡＶネットワークＢＯＸなど）も表示される。

【０１０５】

図１７に、第２実施形態における機能語句情報テーブルＴＢ５を示す。機能語句情報テーブルＴＢ５には、図１７に示すように、予め画像形成装置１に設定されているボックスの種類についての情報である機能語句情報ＤＴ５が記憶され管理されている。「ＢＯＸタイプ」フィールドには画像形成装置１に備えられているボックスの種類を示す語句が示され、「読み（ＢＯＸタイプ）」フィールドには「ＢＯＸタイプ」フィールドに示される語句の読み方ＷＳが示される。

【０１０６】

次に、第２実施形態において、ユーザがボックスを登録するときの手順について、図１８および図１９の各画面を参照しながら説明する。

【０１０７】

ユーザは操作パネル１０ｆのＢＯＸボタンＴＢを押下することによって、ボックス登録画面ＧＭ６（図１８参照）を表示させる。ここで、ユーザは、ボックス番号、ボックス名、パスワード、検索文字、およびボックスタイプなどを登録する。それぞれの項目は、第１実施形態と同一の方法で入力することができる。それぞれの項目を入力しＯＫボタンを押下すると、読み方登録画面ＧＭ７（図１９参照）が表示される。ここで、ユーザは、第１実施形態と同一の方法で読み方を入力する。

【０１０８】

読み方を入力して、ＯＫボタンを押下すると、ボックス登録画面ＧＭ６および読み方登録画面ＧＭ７での入力内容がボックス情報ＤＴ４として、総合情報記憶部３０に記憶されている登録情報テーブルＴＢ４に登録される。

【０１０９】

次に、第２実施形態における認識辞書テーブルＴＢ６について説明する。

【０１１０】

図２０に示すように、第２実施形態における認識辞書テーブルＴＢ６に登録される語句番号ＢＴが３００～１９９９の認識語句ＷＮつまり複合語句ＷＦは、登録情報テーブルＴＢ４の「読み（ＢＯＸ名）」フィールドに示される語句ＷＵ１２と、機能語句情報テーブルＴＢ５の「読み（ＢＯＸタイプ）」フィールドに示される語句ＷＳとの、２つの語句ＷＵ１２，ＷＳについて、それらの順序の異なる全ての組合せによって生成される。

【０１１１】

語句番号ＢＴが２０００～２０１０の認識語句ＷＮつまりコマンド語句ＷＣは、画面番号欄ＧＲＮに示される画面番号の画面に表示されるコマンドを示す、予め登録されている語句である。

【０１１２】

複合語句ＷＦの生成方法は、第１実施形態の場合と同一である。なお、第２実施形態に

おいても、複合語句ＷＦを生成するための語句の組合せは、ユーザが適宜変更することができる。例えば、ボックス登録画面ＧＭ６または読み方登録画面ＧＭ７などに、複合語句ＷＦを生成するために組合せる対象（語句）を選択するための選択ボタンを設けておき、ユーザが選択できるようにしておく。組合せとしては、ボックス名とボックスタイプとの組合せの他に、ボックス番号とボックス名との組合せまたは検索文字とボックスタイプとの組合せなどが挙げられる。

【０１１３】

次に、ユーザが音声によって画像データの保存先のボックスを検索する場合における画像形成装置１の音声認識処理について、図２１、図２２の各画面、および図１３のフローチャートを参照しながら参照する。

【０１１４】

ユーザが操作パネル１０ｆのＢＯＸボタンＴＢを押すと、図２１に示すようなボックス選択画面ＧＭ８が表示される。ボックス選択画面ＧＭ８には、登録情報テーブルＴＢ４に記憶されているそれぞれのボックス情報ＤＴ４に対応付けられているボックスアイコンＢＭ（ＢＭ１～６）が表示される。それぞれのボックスアイコンＢＭには、対応付けられているボックス情報ＤＴ４に示される登録番号およびＢＯＸ名などが示される。ここで、ユーザが音声認識ボタンＫＢを押すと（図１３の＃４０１でＹｅｓ）、操作パネル１０ｆのマイクＭＫがオンになり（＃４０２）、音声の入力が可能となる。ユーザは、ボックスタイプ（ここでは「共有」）および保存先のボックスのボックス名（ここでは「上杉」）を示す語句（「キョウユウ、ウエスギ」）を発声することによって、音声を入力する。

【０１１５】

画像形成装置１は、入力された音声について（＃４０３）、ボックス選択画面ＧＭ８に対応付けられた認識辞書テーブルＴＢ６を用いて音声認識処理を行う（＃４０４）。つまり、入力された音声を認識対象語句（「キョウユウウエスギ」）に変換し、認識辞書テーブルＴＢ６の中から「キョウユウウエスギ」と一致する認識語句ＷＮを検索する。認識対象語句と一致する認識語句ＷＮが検索されると、音声が認識される。第２実施形態では、語句番号３２５の認識語句情報ＤＴ６に示される認識語句ＷＮが「キョウユウウエスギ」と一致する。

【０１１６】

音声が認識されると（＃４０５でＹｅｓ）、認識対象語句と一致する認識語句ＷＮの語句番号ＢＴと同じ登録番号ＢＴのボックス情報ＤＴ４に対応付けられているボックスアイコンＢＭのみが表示されたボックス選択画面ＧＭ９（図２２参照）を表示する（＃４０６）。

【０１１７】

例えば、ボックスアイコンＢＭ５は、登録番号３２５のボックス情報ＤＴ４に対応付けられている。画像形成装置１は、ボックス選択画面ＧＭ９においてユーザが「オーケー」と発声しその音声が認識されると、登録情報テーブルＴＢ４に記憶されている登録番号３２５のボックスＤＴ４を参照し、ボックス名が「上杉」であるボックスの保存内容を示した確認画面（図示せず）を表示する（＃４０７）。ユーザは、ボックスの保存内容を確認した後、保存ボタンを押すなどして画像データをボックスに保存することができる。

【０１１８】

認識対象語句と一致する認識語句ＷＮが検索されなかった場合、つまり音声が認識されなかった場合は、音声が認識されなかった旨の警告を示す画面（図示せず）を表示する（＃４０８）。

〔第３実施形態〕

次に、画像形成装置１の第３実施形態について、図２３～図２８を参照して説明する。

【０１１９】

図２３は登録情報テーブルＴＢ７の例を示す図、図２４は文書保存画面ＧＭ１０の例を示す図、図２５は読み方入力画面ＧＭ１１の例を示す図、図２６は警告表示画面ＧＭ１２の例を示す図、図２７は認識辞書テーブルＴＢ８の例を示す図、図２８は文書利用画面Ｇ

Ｍ１３の例を示す図である。

【０１２０】

以下、ユーザが、画像形成装置１に登録されているボックスに保存されている文書データ（単に「文書」と記載することがある）を検索する場合を例に説明する。また、ユーザは、文書が保存されているボックスのボックス名と文書名とを連続して発声することによって、文書の検索を行うものとする。

【０１２１】

図２３に、第３実施形態における登録情報テーブルＴＢ７を示す。登録情報テーブルＴＢ７には、画像形成装置１に登録されている各ユーザのボックスに保存されている文書に関する情報である保存文書情報ＤＴ７が記憶され管理されている。

【０１２２】

保存文書情報ＤＴ７の「登録番号」フィールドには、他の保存文書情報ＤＴ７と区別するための登録番号（識別番号）ＢＴが示される。この登録番号ＢＴは、各文書の文書番号でもある。

【０１２３】

「文書名」フィールドには、文書名ＷＵ２１が示される。「読み（文書名）」フィールドには、「文書名」フィールドに示される文書名の読み方ＷＵ２２が示される。「検索文字」フィールドには、文書を文書名で検索するための検索文字ＷＵ２３が示される。「ＢＯＸ名」フィールドには、文書を保存しているボックスのボックス名ＷＵ１１が示される。「読み（ＢＯＸ名）」フィールドには、第２実施形態でユーザによって登録された登録情報テーブルＴＢ４の「読み（ＢＯＸ名）」フィールドに示されるボックス名と同じ読み方ＷＵ１２が示される。

【０１２４】

なお、ボックス名ＷＵ１１および読み方ＷＵ１２は、登録情報テーブルＴＢ７においては、システム語句ＷＳとして登録される。

【０１２５】

次に、第３実施形態において、ユーザが文書をボックスに保存するときの手順について、図２４～図２６の各画面を参照しながら説明する。

【０１２６】

ユーザは、第２実施形態の手順によってボックスを検索して、または所定の操作を行って、文書を保存させるためのボックスを表示させる。そして、保存ボタンなどを押下することによって、図２４に示すような文書保存画面ＧＭ１０を表示させる。ここで、ユーザは、登録番号、文書名、および文書名の読み方を登録する。それぞれの項目は、基本的に第１実施形態と同一の方法で入力することができる。例えば、文書名の読み方は、文書名（読み方）ボタンＢＴＮ７を押下することによって表示される読み方入力画面ＧＭ１１（図２５参照）から入力することができる。ここで、入力された文書名が既にそのボックスに保存されている文書の文書名の読み方として登録されている場合は、図２６に示すような再入力を促すための警告表示画面ＧＭ１２を表示するようにしてもよい。

【０１２７】

警告表示画面ＧＭ１２が表示された場合は、ユーザは、読み方入力画面ＧＭ１１に戻って、再度文書名の読み方を入力すればよい。なお、第１実施形態および第２実施形態においても、第３実施形態のように、ユーザが入力する読み方が既に登録されている場合に、警告表示画面ＧＭ１２のような警告を示す画面を表示するようにしてもよい。

【０１２８】

登録番号、文書名、および文書名の読み方を入力して、実行ボタンを押下すると、文書保存画面ＧＭ１０での入力内容がボックス情報ＤＴ４として、総合情報記憶部３０に記憶されている登録情報テーブルＴＢ７に登録される。

【０１２９】

次に、第３実施形態における認識辞書テーブルＴＢ８について説明する。

【０１３０】

図２７に示すように、第３実施形態における認識辞書テーブルＴＢ８に登録される語句番号ＢＴが５００～２９９９の認識語句情報ＤＴ８に示される認識語句ＷＮつまり複合語句ＷＦは、登録情報テーブルＴＢ７の「読み（文書名）」フィールドに示される語句ＷＵ２２と、「読み（ＢＯＸ名）」フィールドに示される語句ＷＳとの、２つの語句ＷＵ２２，ＷＳについて、それらの順序の異なる全ての組合せとして生成される。

【０１３１】

語句番号ＢＴが３０００～３０１０の認識語句情報ＤＴ８に示される認識語句ＷＮつまりコマンド語句ＷＣは、画面番号欄ＧＲＮに示される画面番号の画面に表示されるコマンドを示す、予め登録されている語句である。

【０１３２】

複合語句ＷＦの生成方法は、第１実施形態と同一である。なお、第３実施形態においても、複合語句ＷＦを生成するための語句の組合せは、ユーザが適宜変更することができる。例えば、ボックス名と文書名との組合せの他に、ボックスタイプと文書名との組合せなどが挙げられる。

【０１３３】

次に、ユーザが音声によってボックスに保存されている文書を検索する場合における画像形成装置１の音声認識処理について、図２１、図２８の各画面、および図１３のフローチャートを参照しながら参照する。

【０１３４】

文書の検索は、ボックス選択画面ＧＭ８から行うことができる。つまり、ボックス選択画面ＧＭ８には、ボックスを検索するための認識辞書テーブルＴＢ６および文書を検索するための認識辞書テーブルＴＢ８が対応付けられている。

【０１３５】

よって、ユーザは操作パネル１０ｆのＢＯＸボタンＴＢを押して、ボックス選択画面ＧＭ８（図２１参照）を表示させる。ユーザが音声認識ボタンＫＢを押すと（図１３の＃４０１でＹｅｓ）、操作パネル１０ｆのマイクＭＫがオンになり（＃４０２）、音声の入力が可能となる。ユーザは、ボックス名（ここでは「上杉」）および検索対象の文書名（ここでは「契約書１」）を示す語句（「ウエスギ、ケイヤクショイチ」）を発声することによって、音声を入力する。

【０１３６】

画像形成装置１では、入力された音声について（＃４０３）、ボックス選択画面ＧＭ８に対応付けられた認識辞書テーブルＴＢ６を用いて音声認識処理を行う（＃４０４）。つまり、入力された音声を認識対象語句（「ウエスギケイヤクショイチ」）に変換し、認識辞書テーブルＴＢ８の中から「ウエスギケイヤクショイチ」と一致する認識語句ＷＮを検索する。認識対象語句と一致する認識語句ＷＮが検索されると、音声が認識される。第３実施形態では、語句番号５４２の認識語句情報ＤＴ８に示される認識語句ＷＮが「ウエスギケイヤクショイチ」と一致する。

【０１３７】

音声が認識されると（＃４０５でＹｅｓ）、認識対象語句と一致する認識語句ＷＮの語句番号と同じ登録番号の保存文書情報ＤＴ７に対応付けられている文書アイコンＤＭのみが表示された文書利用画面ＧＭ１３（図２８参照）を表示する（＃４０６）。例えば、文書アイコンＤＭは、登録番号５４２の保存文書情報ＤＴ７に対応付けられている。よって、画像形成装置１は、文書利用画面ＧＭ１３において表示されるコマンド名をユーザが発声しその音声が認識されると、登録情報テーブルＴＢ７に記憶されている登録番号５４２の保存文書情報ＤＴ７を参照し、文書名が「契約書１」である文書の内容を示した確認画面（図示せず）を表示する（＃４０７）。ユーザは、文書の内容を確認した後、所定のボタンを押して文書の編集などを行うことができる。

【０１３８】

認識対象語句と一致する認識語句ＷＮが検索されなかった場合、つまり音声が認識されなかった場合は、音声が認識されなかった旨の警告を示す画面（図示せず）を表示する（

＃４０８）。
【０１３９】
第１～第３実施形態では、読み方をユーザが入力し登録するようにしたが、宛先の登録名、ＢＯＸ名、または文書名に英語名を入力した場合は、英語名が入力された時点で自動的にその読み方が入力されるようにしてもよい。登録名に数字だけが入力された場合も、自動的に読み方が入力されるようにしてもよい。
〔第４実施形態〕
次に、画像形成装置１の第４実施形態について、図２９および図３０を参照して説明する。
【０１４０】
図２９は複合語句生成テーブルＴＢ９ａ、９ｂの例を示す図、図３０は読み方登録画面ＧＭ１４の例を示す図である。
【０１４１】
第１～第３実施形態では、認識辞書テーブルに登録される認識語句ＷＮを、画像形成装置１に備えられている機能に関する語句であるシステム語句ＷＳとユーザが任意で登録した語句であるユーザ語句ＷＵとのみを組み合わせて生成するようにした。しかし、ユーザは、複数の単語を発声する際、それら複数の単語に助詞および動詞を付けて発声する場合がある。そこで、このような言葉のゆらぎを考慮して、システム語句ＷＳとユーザ語句ＷＵとの組合せに、助詞および動詞を付加した語句を認識語句ＷＮとするようにしてもよい。
【０１４２】
この場合は、複合語句ＷＦを生成するにあたって、予め、図２９に示すような複合語句生成テーブルＴＢ９ａ、９ｂをハードディスク１０ｄなどに格納しておく。日本語の複合語句ＷＦを生成する場合は、システム語句ＷＳおよびユーザ語句ＷＵをそれぞれ複合語句生成テーブルＴＢ９ａに当てはめて組合せを生成する。英語の複合語句ＷＦを生成する場合は、英語のシステム語句ＷＳおよびユーザ語句ＷＵをそれぞれ複合語句生成テーブルＴＢ９ｂに当てはめて組合せを生成する。
【０１４３】
例えば、第４実施形態では、ユーザが宛先を登録する場合において登録名の読み方を入力させる際に、図３０に示すような読み方登録画面ＧＭ１４を表示する。ここでユーザは、システム語句ＷＳおよびユーザ語句ＷＵのみからなる複合語句ＷＦを生成させる場合（連続単語モードの場合）は、連続単語ボタンＢＴＮ８を押下し、助詞および動詞を付加した複合語句ＷＦを生成させる場合（簡易文書モードの場合）は、簡易文書ボタンＢＴＮ９を押下する。
【０１４４】
簡易文書ボタンＢＴＮ８が押下された場合は、複合語句生成部３４ｂは助詞および動詞を付加した複合語句ＷＦを生成するための処理を行う。例えば、複合語句生成部３４ｂは、複合語句生成テーブルＴＢ９ａを用いて、「イーメールのウエスギさんを探す」、「イーメールのウエスギさんを検索する」、「ウエスギさんのイーメールを探す」、または「ウエスギさんのイーメールを検索する」などの複合語句ＷＦを生成する。複合語句生成テーブルＴＢ９ｂを用いた場合は、「It looks for UESUGI of Email.」、「It looks for UESUGI 's Email.」、「UESUGI of Fax is retrieved.」、「UESUGI's Email is retrieved.」などの英語の複合語句ＷＦが生成される。
【０１４５】
第４実施形態では、簡易文書ボタンＢＴＮ８が押下された場合に、助詞および動詞を含めた認識語句を生成するようにしたが、助詞および動詞を含めない認識語句（第１～第３実施形態の複合語句）と助詞および動詞を含めた複合語句ＷＦ（第４実施形態の複合語句）との２種類の複合語句ＷＦを生成し、複合辞書テーブルに登録するようにしてもよい。
【０１４６】
図３１は第１～第３実施形態における画像形成装置１全体の処理の流れの例を説明する

ためのフローチャート、図３２は第４実施形態における画像形成装置１全体の処理の流れの例を説明するためのフローチャートである。

【０１４７】

次に、第１～第３実施形態における画像形成装置１全体の処理について、図３１のフローチャートを参照して説明する。

【０１４８】

ユーザが操作パネル１０ｆを操作して、種々の登録画面（宛先登録画面ＧＭ１、ボックス登録画面ＧＭ６、または文書保存画面ＧＭ１０）を表示させると、画像形成装置１は、複合語句ＷＦを生成するための語句の組合せの設定を読み込む（＃５０１）。次に、ユーザが所定の読み方登録画面（読み方登録画面ＧＭ２、読み方登録画面ＧＭ７、または読み方登録画面ＧＭ１４）で、宛先名、ボックス名、または文書名などの読み方を入力した場合に、入力された読み方が既に画像形成装置１に登録されているものでないか否かを判別する。そして、入力された読み方が既に登録されているものではない場合は（＃５０３でＹｅｓ）、システム語句ＷＳおよびユーザ語句ＷＵの２つの語句を前後の配置を逆にして組み合わせることにより複合語句ＷＦを生成し、認識辞書テーブルに登録する（＃５０４）。入力された読み方が既に画像形成装置１に登録されているものである場合は（＃５０３でＮｏ）、警告を示す画面を表示する（＃５０５）。

【０１４９】

次に、第４実施形態における画像形成装置１全体の処理について、図３２のフローチャートを参照して説明する。

【０１５０】

＃６０１～＃６０３の処理は、図３１のフローチャートと同じである。画像形成装置１は、入力された読み方が既に登録されているものではない場合は（＃６０３でＹｅｓ）、連続単語モードまたは簡易文書モードのいずれが選択されたかを判別する。そして、連続単語モードが選択された場合は（＃６０４でＹｅｓ）、第１～第３実施形態と同様に、システム語句ＷＳおよびユーザ語句ＷＵの２つの語句を前後の配置を逆にして組み合わせることにより複合語句ＷＦを生成し、認識辞書テーブルに登録する（＃６０５）。簡易文書モードが選択された場合は（＃６０４でＮｏ）、認識語句生成テーブル９を用いて、システム語句ＷＳおよびユーザ語句の２つの語句の前後の配置を逆にした組み合わせに助詞および動詞を含めた複合語句ＷＦを生成し、認識辞書テーブルに登録する（＃６０６）。

【０１５１】

上述の実施形態において、スキャナ部１０ｇは本発明における「画像読取装置」に相当し、マイクＭＫは本発明における「音声入力手段」に相当し、認識語句情報ＤＴ３、ＤＴ６、ＤＴ８のそれぞれは本発明における「語句情報」に相当し、認識辞書テーブルＴＢ３、ＴＢ６、ＴＢ８のそれぞれは本発明における「辞書」に相当し、宛先情報ＤＴ１の「読み（登録名）」フィールドおよび機能語句情報ＤＴ２の「読み（機能名）」フィールドに示される読み方、ボックス情報ＤＴ４の「読み（ＢＯＸ名）」フィールドおよび機能語句情報ＤＴ５の「読み（ＢＯＸタイプ）」フィールドに示される読み方、および保存文書情報ＤＴ７の「読み（文書名）」フィールドおよび「読み（ＢＯＸタイプ）」フィールドに示される読み方、のそれぞれは、本発明における「複数の所定の語句」に相当し、複合語句生成部３４ｂは本発明における「複合語句生成手段」に相当し、複合語句書込み部３４ｄは本発明における「複合語句書込み手段」に相当し、認識対象語句は本発明における「発声語句」に相当し、音声認識処理部３５は本発明における「音声認識手段」に相当し、画像処理部３７は本発明における「画像処理手段」に相当する。

【０１５２】

また、宛先情報ＤＴ１、ボックス情報ＤＴ４、保存文書情報ＤＴ７のそれぞれは本発明における「関連情報」に相当し、総合情報記憶部３０は本発明における「関連情報記憶手段」に相当する。

【０１５３】

また、機能語句情報記憶部３１は本発明における「機能語句記憶手段」に相当し、情報

書込み部３３は本発明における「任意語句記憶手段」に相当し、宛先情報ＤＴ１の「読み（登録名）」フィールド、ボックス情報ＤＴ４の「読み（ＢＯＸ名）」フィールド、または保存文書情報ＤＴ７の「読み（文書名）」フィールドに示される語句のそれぞれは本発明における「ユーザ語句」に相当する。

【０１５４】

また、辞書管理部３４および音声認識処理部３５は、本発明における「音声認識処理装置」に相当する。

【０１５５】

上に述べた各実施形態では、２つの単語の組合せから複合語句ＷＦを生成するようにしたが、３つ以上の単語を組み合わせた複合語句ＷＦを生成するようにしてもよい。この場合も、それら３つ以上の単語を並べ替えて全ての組合せを生成することにより（つまり全ての順列を生成することにより）、複合語句ＷＦを生成すればよい。

【０１５６】

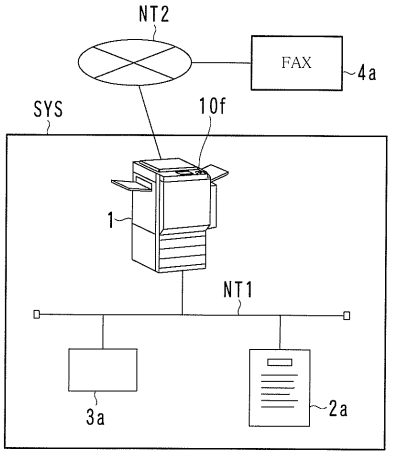
その他、ネットワークシステムＳＹＳおよび画像形成装置１の構成、機能、各データが示す内容、テーブルの内容、処理の内容または順序などは、本発明の趣旨に沿って適宜変更することができる。

【符号の説明】

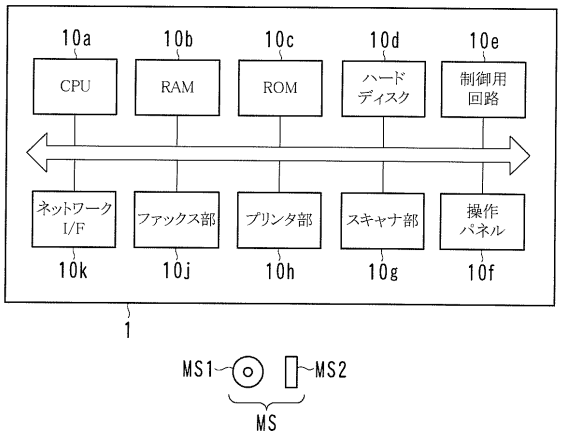
【０１５７】

１　画像形成装置（画像処理装置）
１０ｇ　スキャナ部（画像読取装置）
３０　総合情報記憶部（関連情報記憶手段）
３１　機能語句情報記憶部（機能語句記憶手段）
３２　表示制御部
３３　情報書込み部（任意語句記憶手段）
３４ａ　辞書データベース（辞書記憶手段）
３４ｂ　複合語句生成部（複合語句生成手段）
３４ｄ　複合語句書込み部（複合語句登録手段）
３５　音声認識処理部（音声認識手段）
３７　画像処理部（画像処理手段）
ＭＫ　マイク（音声入力手段）
ＴＰ　タッチパネル（表示装置）
ＧＭ１　宛先登録画面（操作画面）
ＧＭ２　読み方登録画面（操作画面）
ＧＭ３　検索文字入力画面（操作画面）
ＧＭ４　宛先選択画面（操作画面）
ＧＭ５　宛先選択画面（操作画面）
ＤＴ１、ＤＴ４、ＤＴ７　宛先情報、ボックス情報、保存文書情報（関連情報）
ＴＢ３、ＴＢ６、ＴＢ８　認識辞書テーブル（辞書）
ＷＵ　ユーザ語句
ＷＳ　システム語句（機能語句）
ＷＣ　コマンド語句

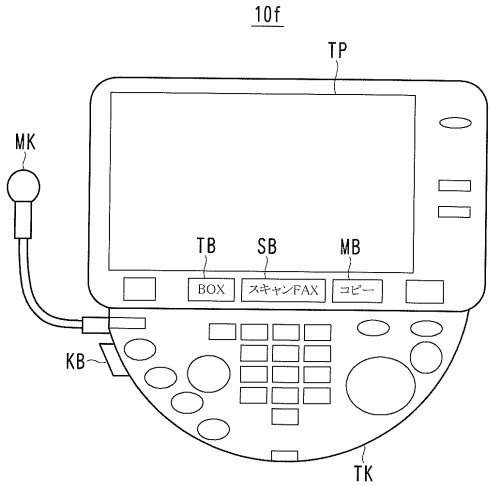
【図１】



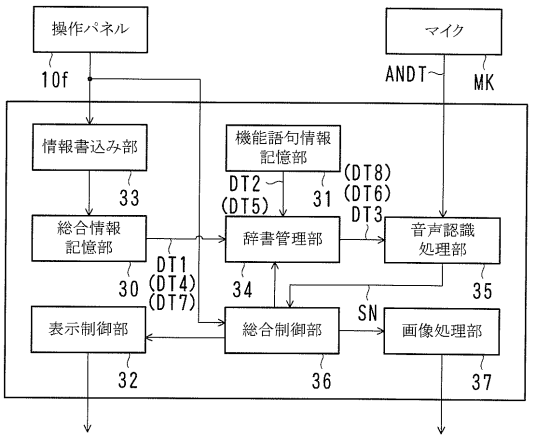
【図２】



【図３】



【図４】



【図５】

TB1(30)

BT	WU1	WU2	WU3	WU4		
登録番号	登録名	読み(登録名)	検索文字	宛先種別	宛先	
001	武田	タケダ	た	E-mail	takeda@xxx.xx.jp	DT1
002	鈴木	スズキ	さ	FAX	06-XXXX-XXXX	DT1
003	佐藤	サトウ	さ	SMB	server2	DT1
004	加藤	カトウ	か	FTP	server1	DT1
005	山本	ヤマモト	や	E-mail	yamamoto@xxx.xx.jp	DT1
006	上杉	ウエスギ	あ	E-mail	uesugi@xxx.xx.jp	DT1
⋮	⋮	⋮	⋮	⋮	⋮	

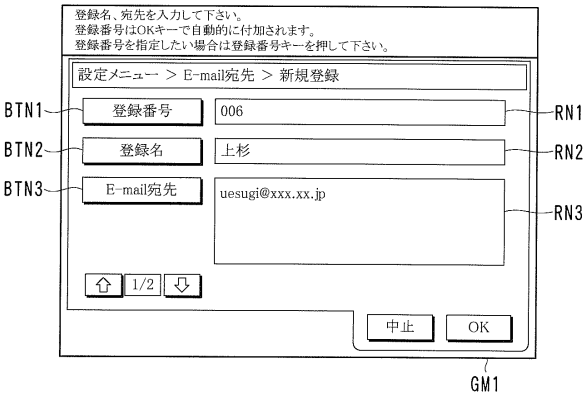
【図６】

TB2(31)

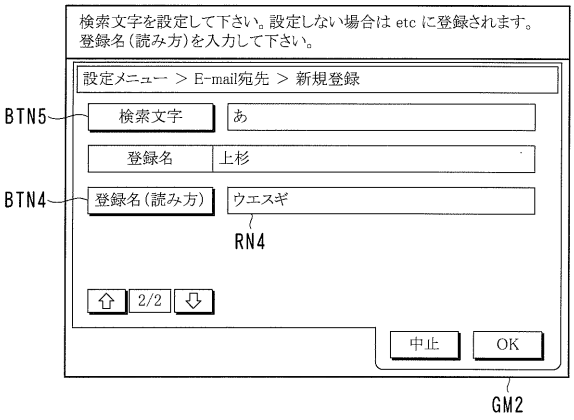
WS

機能名	読み(機能名)	
E-mail	イーメール	DT2
SMB	エスエムビー	DT2
FTP	エフティーピー	DT2
FAX	ファックス	DT2

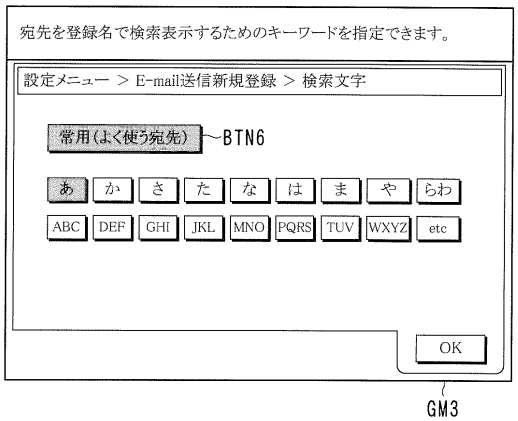
【図７】



【図８】



【図９】

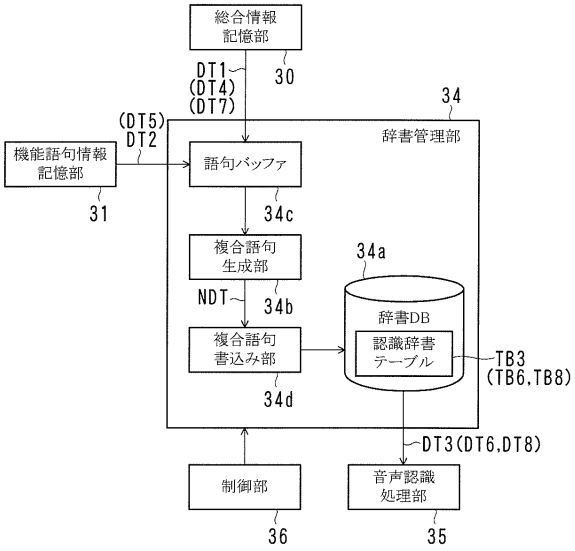


【図１０】

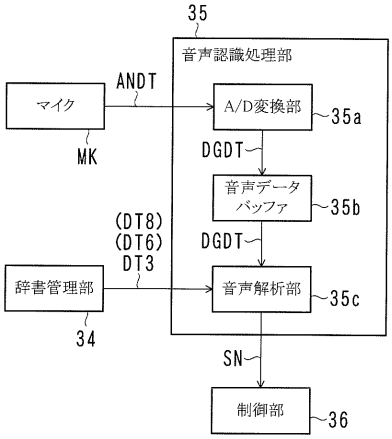
TB3(34a)

GRN	画面番号	999	
BT	語句番号	認識語句	WN
DT3	001	タケダイーメール	WF
		イーメールタケダ	WF
DT3	002	スズキファックス	WF
		ファックススズキ	WF
DT3	003	サトウエスエムビー	WF
		エスエムビーサトウ	WF
	⋮	⋮	
DT3	006	ウエスギイーメール	WF
		イーメールウエスギ	WF
	⋮	⋮	
DT3	1000	イーメール	WC(WS)
DT3	1001	エスエムビー	WC(WS)
DT3	1002	エフティーピー	WC(WS)
DT3	1003	ファックス	WC(WS)
DT3	1004	トップガメンエ	WC
DT3	1005	オーケー	WC
DT3	1006	ウエ	WC
DT3	1007	シタ	WC

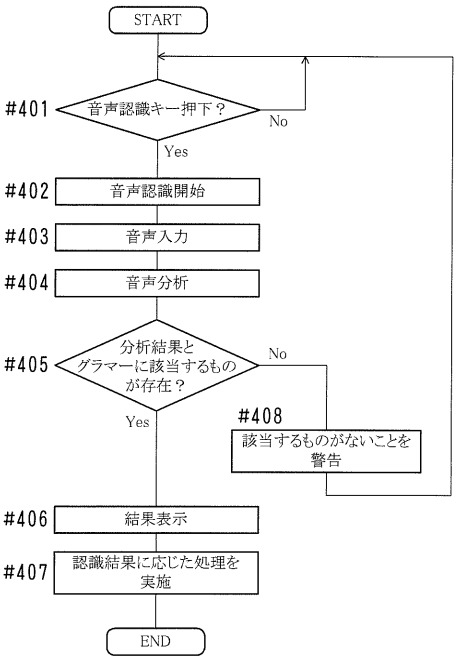
【図１１】



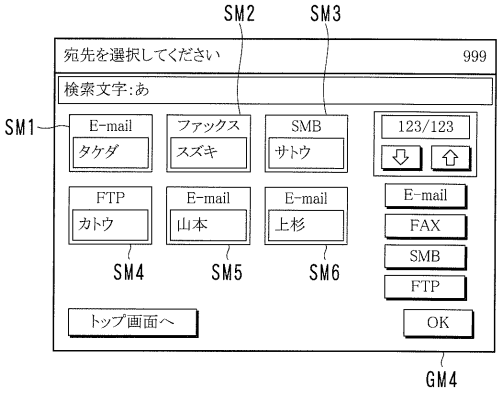
【図１２】



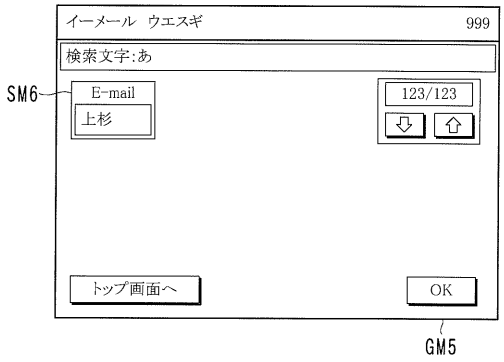
【図１３】



【図１４】



【図１５】



【図１６】

TB4(30)

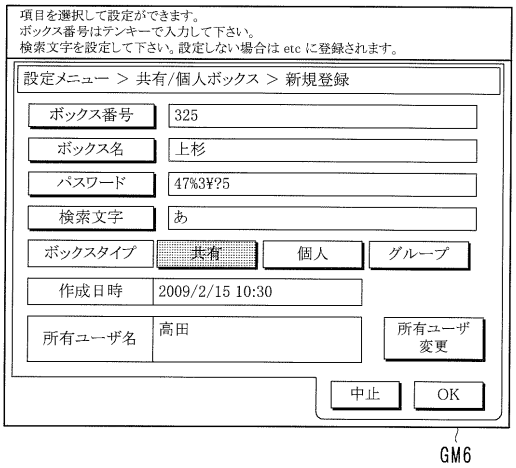
BT	WU11	WU12	WU13	
登録番号	BOX名	読み(BOX名)	検索文字	BOXタイプ
300	徳田	トクダ	た	共有 ~DT4
301	原田	ハラダ	は	個人 ~DT4
302	吉原	ヨシハラ	や	システム ~DT4
⋮	⋮	⋮	⋮	⋮
325	上杉	ウエスギ	あ	共有 ~DT4
⋮	⋮	⋮	⋮	⋮

【図１７】

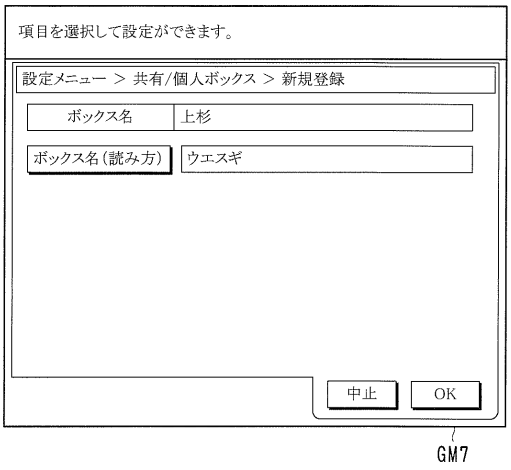
TB5(31)

	WS
BOXタイプ	読み(BOXタイプ)
共有	キョウユウ ~DT5
個人	コジン ~DT5
システム	システム ~DT5
グループ	グループ ~DT5

【図１８】



【図１９】

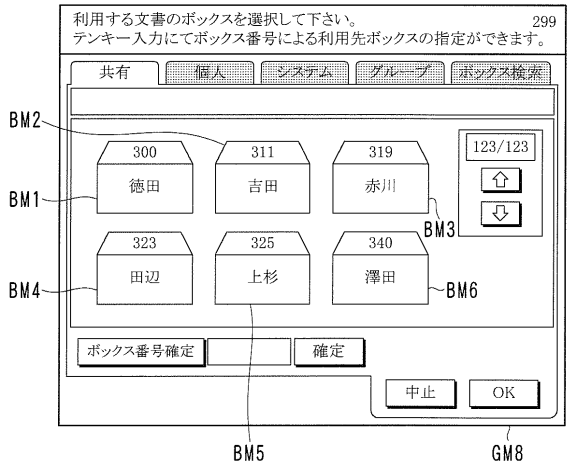


【図２０】

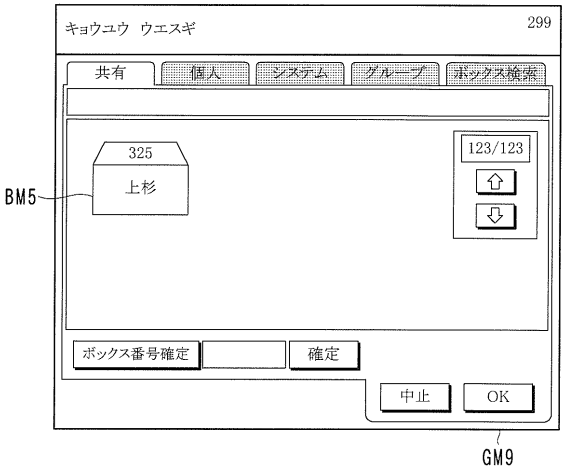
TB6(34a)

GRN	画面番号	299	
BT	語句番号	認識語句	WN
DT6	300	トクダキョウユウ	WF
		キョウユウトクダ	WF
DT6	301	ハラダコジン	WF
		コジンハラダ	WF
DT6	302	ヨシハラシステム	WF
		システムヨシハラ	WF
	⋮	⋮	WF
DT6	325	ウエスギキョウユウ	WF
		キョウユウウエスギ	WF
	⋮	⋮	
DT6	2000	キョウユウ	WC(WS)
DT6	2001	コジン	WC(WS)
DT6	2002	システム	WC(WS)
DT6	2003	グループ	WC(WS)
DT6	2004	ボックスケンサク	WC
DT6	2005	ボックスバンゴウシテイ	WC
DT6	2006	カクテイ	WC
DT6	2007	チェウシ	WC
DT6	2008	オーケー	WC
DT6	2009	ウエ	WC
DT6	2010	シタ	WC

【図２１】



【図２２】

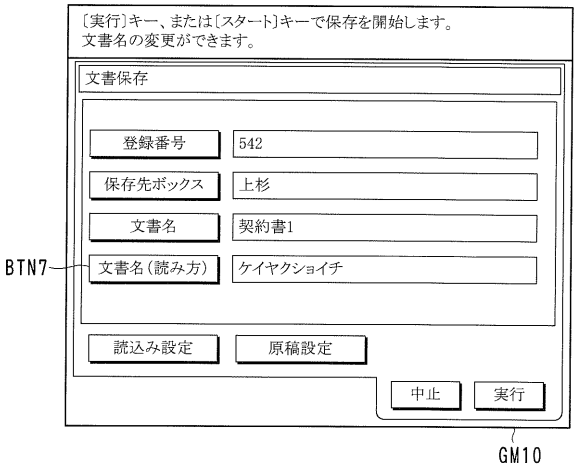


【図２３】

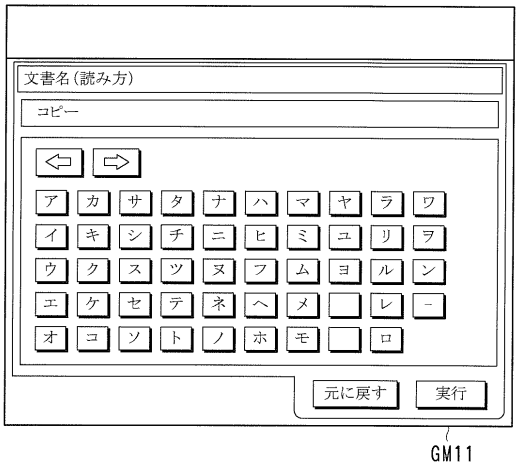
TB7(30)

BT	WU21	WU22	WU23	WS(WU11)	WS(WU12)	
登録番号	文書名	読み(文書名)	検索文字	BOX名	読み(BOX名)	
500	予定表	ヨテイヒョウ	や	徳田	トクダ	DT7
501	明細書	メイサイショ	ま	原田	ハラダ	DT7
502	企画書1	キカクショイチ	か	吉原	ヨシハラ	DT7
⋮	⋮	⋮	⋮	⋮	⋮	
542	契約書1	ケイヤクショイチ	か	上杉	ウエスギ	DT7
⋮	⋮	⋮	⋮	⋮	⋮	

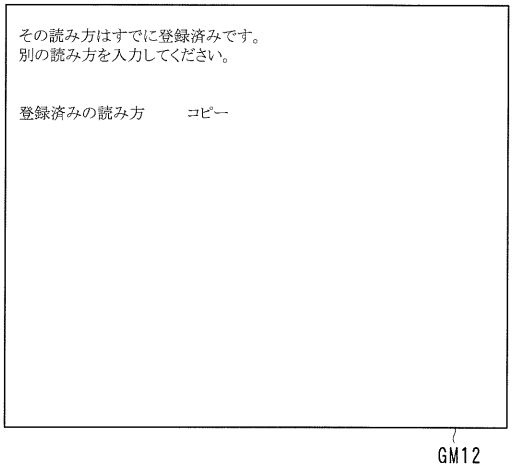
【図２４】



【図２５】



【図２６】

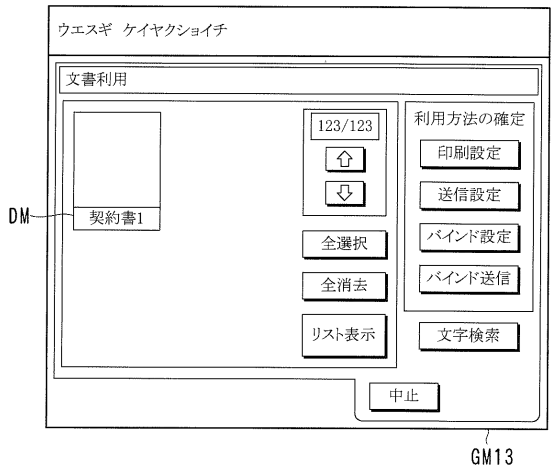


【図２７】

TB8(34a)

GRN	画面番号	299	
BT	語句番号	認識語句	WN
DT8	500	トクダヨテイヒョウ	WF
		ヨテイヒョウトクダ	WF
DT8	501	ハラダメイサイショ	WF
		メイサイショハラダ	WF
DT8	502	ヨシハラキカクショイチ	WF
		キカクショイチヨシハラ	WF
	⋮	⋮	
DT8	542	ウエスギケイヤクショイチ	WF
		ケイヤクショイチウエスギ	WF
	⋮	⋮	
DT8	3000	キョウユウ	WC(WS)
DT8	3001	コジン	WC(WS)
DT8	3002	システム	WC(WS)
DT8	3003	グループ	WC(WS)
DT8	3004	ボックスケンサク	WC
DT8	3005	ボックスバンゴウシテイ	WC
DT8	3006	カクテイ	WC
DT8	3007	チュウシ	WC
DT8	3008	オーケー	WC
DT8	3009	ウエ	WC
DT8	3010	シタ	WC

【図２８】



【図２９】

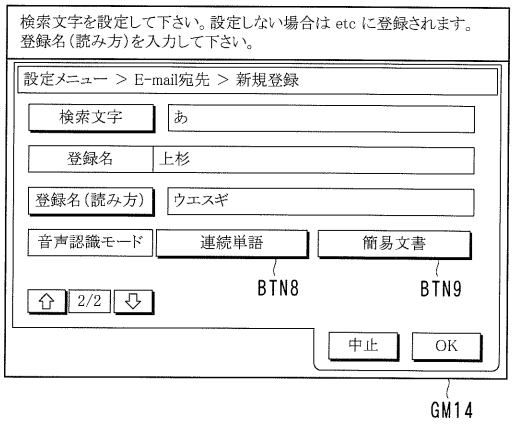
TB9a

WS		WU	
(システム語句)	の	(ユーザ語句)	さんを探す
(システム語句)	の	(ユーザ語句)	さんを検索する
(ユーザ語句)	さんの	(システム語句)	さんを探す
(ユーザ語句)	さんの	(システム語句)	さんを検索する

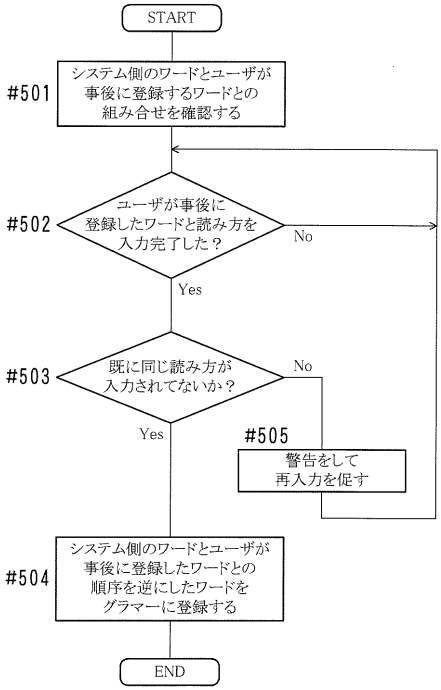
TB9b

WU			WS
It looks for	(ユーザ語句)		(システム語句)
It looks for	(ユーザ語句)	's	(システム語句)
(ユーザ語句)	of	(システム語句)	is retrieved.
(ユーザ語句)	's	(システム語句)	is retrieved.

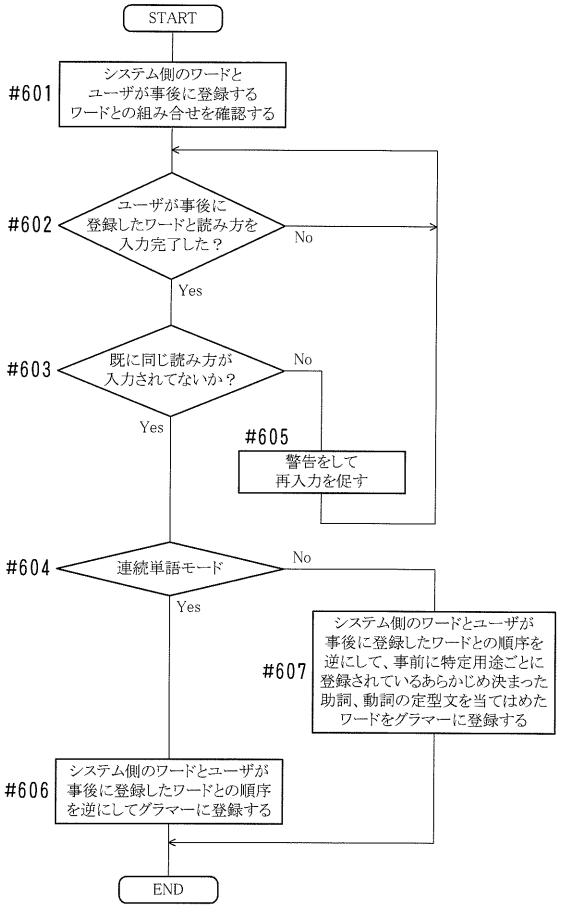
【図３０】



【図３１】



【図３２】



【図３３】

TB10

イーメール
ファックス
エスエムビー
エフティーピー
イーメールタケダ
ファックススズキ
エスエムビーサトウ
エフティーピーカトウ
イーメールヤマモト
イーメールウエスギ
トップガメンエ
オーケー
ウエ
シタ

フロントページの続き

(56)参考文献 特開２００４－２９４８７２（ＪＰ，Ａ）
特開２００８－２６８５７１（ＪＰ，Ａ）
特開２００５－３３８２７４（ＪＰ，Ａ）
特開２００１－０２７５４０（ＪＰ，Ａ）
特開２００４－３０９６５４（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１５／０６
Ｇ０６Ｆ　３／１６
Ｇ１０Ｌ　１５／００