

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6655331号
(P6655331)

(45) 発行日 令和2年2月26日(2020.2.26)

(24) 登録日 令和2年2月5日(2020.2.5)

(51) Int.Cl.

F 1

G 06 K 9/68 (2006.01)

G 06 K 9/68

B

G 06 K 9/03 (2006.01)

G 06 K 9/03

J

G 06 F 40/126 (2020.01)

G 06 F 17/22

6 1 7

請求項の数 6 (全 18 頁)

(21) 出願番号

特願2015-186801 (P2015-186801)

(22) 出願日

平成27年9月24日 (2015.9.24)

(65) 公開番号

特開2017-62584 (P2017-62584A)

(43) 公開日

平成29年3月30日 (2017.3.30)

審査請求日

平成30年7月31日 (2018.7.31)

(73) 特許権者 398058588

D y n a b o o k 株式会社

東京都江東区豊洲五丁目6番15号

(74) 代理人 110001737

特許業務法人スズエ国際特許事務所

(72) 発明者 筒井 秀樹

東京都港区芝浦一丁目1番1号 株式会社
東芝内

審査官 笠田 和宏

最終頁に続く

(54) 【発明の名称】電子機器及び方法

(57) 【特許請求の範囲】

【請求項 1】

第1言語の文字を認識するための辞書が記憶される第1辞書データと、

前記第1言語とは異なる第2言語の文字を認識するための辞書が記憶される第2辞書データと、

少なくとも前記第1言語の文字列および前記第2言語の文字列が描画される画像データを含む複数の画像データが格納されるデータベースと、

前記画像データに描画される前記複数の文字列を前記第1辞書データと前記第2辞書データを用いて文字認識すると共に、前記データベースの検索を実行するプロセッサと、を具備し、

前記プロセッサは、

前記データベースから前記画像データを読み取り、前記第1辞書データを用いて、前記画像データに含まれる複数の文字列を認識して、その認識した前記複数の文字列に対応する第1文字コード列を生成し、

前記第2辞書データを用いて、前記画像データに含まれる前記複数の文字列を認識して、その認識した前記複数の文字列に対応する第2文字コード列を生成し、

前記第1文字コード列と前記複数の文字列とを対応付けた第1インデックスデータと、前記第2文字コード列と前記複数の文字列とを対応付けた第2インデックスデータとを含む検索用のインデックスデータを生成して前記データベースに記憶し、

キーワードの入力に対し前記検索用のインデックスデータを検索して、前記キーワード

と前記第1文字コード列又は前記第2文字コード列が一致する場合は、前記第1文字コード列又は前記第2文字コード列に対応する画像データを前記データベースから読み出す電子機器。

【請求項2】

前記第1インデックスデータは、前記第1文字コード列と、前記第1文字コード列の位置に対応する前記画像データ内の第1領域の座標とを少なくとも含み、

前記第2インデックスデータは、前記第2文字コード列と、前記第2文字コード列の位置に対応する前記画像データ内の第2領域の座標とを少なくとも含む請求項1記載の電子機器。

【請求項3】

10

前記プロセッサは、

前記キーワードが前記第1文字コード列および第2文字コード列に含まれる場合、前記第1領域と前記第2領域が重複しているか否かを判定し、

重複している割合が閾値より大きいと判定した場合、前記第1領域と前記第2領域をマージすると共に、前記第1領域と前記第2領域をハイライトにした前記画像データを表示する請求項2記載の電子機器。

【請求項4】

前記プロセッサは、前記第1領域と前記第2領域とをマージするために、前記第1領域と前記第2領域の中間の領域を算出するか、又は前記第1領域と前記第2領域を包含する領域を算出する請求項3記載の電子機器。

20

【請求項5】

第1言語の文字列および前記第1言語とは異なる第2言語の文字列が少なくとも描画される画像データを含む複数の画像データが格納されるデータベースから前記画像データを読み取り、

前記第1言語の文字を認識するための辞書が記憶される第1辞書データを用いて、前記画像データに含まれる複数の文字列を認識して、その認識した前記複数の文字列に対応する第1文字コード列を生成し、

前記第2言語の文字を認識するための辞書が記憶される第2辞書データを用いて、前記画像データに含まれる前記複数の文字列を認識して、その認識した前記複数の文字列に対応する第2文字コード列を生成し、

30

前記第1文字コード列と前記複数の文字列とを対応付けた第1インデックスデータと、前記第2文字コード列と前記複数の文字列とを対応付けた第2インデックスデータとを含む検索用のインデックスデータを生成して前記データベースに記憶し、

キーワードの入力に対し前記検索用のインデックスデータを検索して、前記キーワードと前記第1文字コード列又は前記第2文字コード列が一致する場合は、前記第1文字コード列又は前記第2文字コード列に対応する画像データを前記データベースから読み出す方法。

【請求項6】

コンピュータにより実行されるプログラムであって、前記プログラムは、

第1言語の文字列および前記第1言語とは異なる第2言語の文字列が少なくとも描画される画像データを含む複数の画像データが格納されるデータベースから前記画像データを読み取る手順と、

前記第1言語の文字を認識するための辞書が記憶される第1辞書データを用いて、前記画像データに含まれる複数の文字列を認識する手順と、

その認識した前記複数の文字列に対応する第1文字コード列を生成する手順と、

前記第2言語の文字を認識するための辞書が記憶される第2辞書データを用いて、前記画像データに含まれる前記複数の文字列を認識する手順と、

その認識した前記複数の文字列に対応する第2文字コード列を生成する手順と、

前記第1文字コード列と前記複数の文字列とを対応付けた第1インデックスデータと、前記第2文字コード列と前記複数の文字列とを対応付けた第2インデックスデータとを含

40

50

む検索用のインデックスデータを生成して前記データベースに記憶する手順と、

キーワードの入力に対し前記検索用のインデックスデータを検索して、前記キーワードと前記第1文字コード列又は前記第2文字コード列が一致する場合は、前記第1文字コード列又は前記第2文字コード列に対応する画像データを前記データベースから読み出す手順と、

を前記コンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

ここに記載される実施形態は、文字を認識する技術に関する。

10

【背景技術】

【0002】

ノートや紙の資料等をスキャンすることによって画像データを生成し、この画像データを用いて文字を認識する光学文字認識（O C R）の技術が利用されている。この技術により、画像内の文字がテキスト（文字コード）に変換されるので、そのテキストに対する編集や検索を容易に行うことができる。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開平9-218923号公報

20

【発明の概要】

【発明が解決しようとする課題】

【0004】

ところで、画像に含まれる文字には、複数の言語の文字が含まれることがある。このような複数の言語の文字を含む画像に対して文字認識処理が施された場合、想定されていない言語の文字を正しく認識できないので、画像をキーワード（文字列）で検索することができない可能性がある。そのため、複数の言語の文字を含む画像をキーワードで検索できる新たな機能の実現が必要である。

【0005】

本発明は、複数の言語の文字を含む画像をキーワードで検索できる電子機器及び方法を提供することを目的とする。

30

【課題を解決するための手段】

【0006】

実施形態によれば、電子機器は、第1言語の文字を認識するための辞書が記憶される第1辞書データと、前記第1言語とは異なる第2言語の文字を認識するための辞書が記憶される第2辞書データと、少なくとも前記第1言語の文字列および前記第2言語の文字列が描画される画像データを含む複数の画像データが格納されるデータベースと、前記画像データに描画される前記複数の文字列を前記第1辞書データと前記第2辞書データを用いて文字認識すると共に、前記データベースの検索を実行するプロセッサと、を具備する。前記プロセッサは、前記データベースから前記画像データを読み取り、前記第1辞書データを用いて、前記画像データに含まれる複数の文字列を認識して、その認識した前記複数の文字列に対応する第1文字コード列を生成し、前記第2辞書データを用いて、前記画像データに含まれる前記複数の文字列を認識して、その認識した前記複数の文字列に対応する第2文字コード列を生成し、前記第1文字コード列と前記複数の文字列とを対応付けた第1インデックスデータと、前記第2文字コード列と前記複数の文字列とを対応付けた第2インデックスデータとを含む検索用のインデックスデータを生成して前記データベースに記憶し、キーワードの入力に対し前記検索用のインデックスデータを検索して、前記キーワードと前記第1文字コード列又は前記第2文字コード列が一致する場合は、前記第1文字コード列又は前記第2文字コード列に対応する画像データを前記データベースから読み出す。

40

50

【図面の簡単な説明】**【0007】**

【図1】実施形態に係る電子機器の外観を示す例示的な斜視図。

【図2】同実施形態の電子機器のシステム構成を示す例示的なブロック図。

【図3】同実施形態の電子機器によって実行されるデジタルノートブックアプリケーションプログラムの機能構成を示す例示的なブロック図。

【図4】同実施形態の電子機器によって、複数の言語環境のための複数の辞書データを用いて画像内の文字が認識される例を説明するための図。

【図5】同実施形態の電子機器によって生成されるインデックスデータの例を示す図。

【図6】同実施形態の電子機器によって、インデックスデータを用いてキーワード検索が行われる例を説明するための図。 10

【図7】同実施形態の電子機器によって表示される検索結果の例を示す図。

【図8】同実施形態の電子機器によって表示される検索結果の別の例を示す図。

【図9】同実施形態の電子機器によって認識された文字に対応する領域の例を示す図。

【図10】同実施形態の電子機器によって認識された文字に対応する領域の別の例を示す図。

【図11】同実施形態の電子機器によって実行される文字認識処理の手順の例を示すフローチャート。 20

【図12】同実施形態の電子機器によって実行される検索処理の手順の例を示すフローチャート。

【図13】同実施形態の電子機器によって実行されるマージ処理の手順の例を示すフローチャート。 20

【発明を実施するための形態】**【0008】**

以下、実施の形態について図面を参照して説明する。

【0009】

図1は、一実施形態に係る電子機器の外観を示す斜視図である。この電子機器は、例えば、ペン又は指によって入力可能な携帯型電子機器である。この電子機器は、タブレットコンピュータ、ノートブック型パーソナルコンピュータ、スマートフォン、PDA等として実現され得る。以下では、この電子機器がタブレットコンピュータ10として実現されている場合を想定する。タブレットコンピュータ10は、タブレット又はスレートコンピュータとも称される携帯型電子機器であり、図1に示すように、本体11とタッチスクリーンディスプレイ17とを備える。タッチスクリーンディスプレイ17は、本体11の上面に重ね合わせるように取り付けられている。 30

【0010】

本体11は、薄い箱形の筐体を有している。タッチスクリーンディスプレイ17には、フラットパネルディスプレイと、フラットパネルディスプレイの画面上のペン又は指の接触位置を検出するように構成されたセンサとが組み込まれている。フラットパネルディスプレイは、例えば、液晶表示装置(LCD)であってもよい。センサとしては、例えば、静電容量方式のタッチパネル、電磁誘導方式のデジタイザなどを使用することができる。以下では、デジタイザとタッチパネルである2種類のセンサの双方がタッチスクリーンディスプレイ17に組み込まれている場合を想定する。 40

【0011】

デジタイザ及びタッチパネルの各々は、フラットパネルディスプレイの画面を覆うように設けられる。このタッチスクリーンディスプレイ17は、指を使用した画面に対するタッチ操作のみならず、ペン100を使用した画面に対するタッチ操作も検出することができる。ペン100は例えば電磁誘導ペンである。

【0012】

ユーザは、外部オブジェクト(ペン100又は指)を使用してタッチスクリーンディスプレイ17上で、手書きにより複数のストロークを入力する手書き入力操作を行うことも 50

できる。手書き入力操作中においては、画面上の外部オブジェクトの動きの軌跡、つまり手書き入力操作によって手書きされるストロークの軌跡がリアルタイムに描画され、これによって各ストロークの軌跡が画面上に表示される。外部オブジェクトが画面に接触されている間の外部オブジェクトの動きの軌跡が1ストロークに相当する。多数のストロークの集合、つまり多数の軌跡の集合が、手書きの文字又は図形などを構成する。

【0013】

図2は、タブレットコンピュータ10のシステム構成を示す図である。

タブレットコンピュータ10は、図2に示されるように、CPU101、システムコントローラ102、主メモリ103、グラフィクスコントローラ104、BIOS-ROM105、不揮発性メモリ106、無線通信デバイス107、エンベデッドコントローラ(EC)108、カメラ109、等を備える。 10

【0014】

CPU101は、タブレットコンピュータ10内の各種コンポーネントの動作を制御するプロセッサである。CPU101は、ストレージデバイスである不揮発性メモリ106から主メモリ103にロードされる各種ソフトウェアを実行する。これらソフトウェアには、オペレーティングシステム(OS)201、及び各種アプリケーションプログラムが含まれている。アプリケーションプログラムには、デジタルノートブックアプリケーションプログラム202が含まれている。このデジタルノートブックアプリケーションプログラム202は文書作成機能を有する。作成される文書には、例えば、ソフトウェアキーボードを用いて入力された文字、タッチスクリーンディスプレイ17を用いて手書き入力された文字や図形、等が含まれる。また、デジタルノートブックアプリケーションプログラム202は、このような文書に、不揮発性メモリ106等に保存されている画像、ネットワークを介して受信した画像、タブレットコンピュータ10上で生成された画像(例えば、カメラ109を用いて撮影された画像、画面をキャプチャした画像、ペイントアプリケーションを用いて作成された画像、等)を貼り付けることもできるし、そのような画像を文書の一つとして扱うこともできる。 20

【0015】

さらに、デジタルノートブックアプリケーションプログラム202は、画像内の文字や手書きされた文字をテキスト(文字コード)に変換する文字認識機能と、文書をキーワード(文字列)で検索する検索機能とを有している。デジタルノートブックアプリケーションプログラム202は、この検索機能により、ソフトウェアキーボードを用いて入力された文字(文字コード)だけでなく、文字認識機能によってテキスト(文字コード)に変換された画像内の文字や手書きされた文字も検索することができる。 30

【0016】

また、CPU101は、BIOS-ROM105に格納された基本入出力システム(BIOS)も実行する。BIOSは、ハードウェア制御のためのプログラムである。

【0017】

システムコントローラ102は、CPU101のローカルバスと各種コンポーネントとの間を接続するデバイスである。システムコントローラ102には、主メモリ103をアクセス制御するメモリコントローラも内蔵されている。また、システムコントローラ102は、PCI EXPRESS規格のシリアルバスなどを介してグラフィクスコントローラ104との通信を実行する機能も有している。 40

【0018】

グラフィクスコントローラ104は、本タブレットコンピュータ10のディスプレイミニタとして使用されるLCD17Aを制御する表示コントローラである。このグラフィクスコントローラ104によって生成される表示信号はLCD17Aに送られる。LCD17Aは、表示信号に基づいて画面イメージを表示する。このLCD17A上にはタッチパネル17B及びデジタイザ17Cが配置されている。タッチパネル17Bは、LCD17Aの画面上で入力を行うための静電容量式のポインティングデバイスである。指が接触される画面上の接触位置及び接触位置の動き等はタッチパネル17Bによって検出される。 50

デジタイザ 17C は LCD 17A の画面上で入力を行うための電磁誘導式のポイントティングデバイスである。ペン 100 が接触される画面上の接触位置、接触位置の動き、接触圧力、等はデジタイザ 17C によって検出される。

【0019】

無線通信デバイス 107 は、無線 LAN 又は 3G / LTE 移動通信などの無線通信を実行するように構成されたデバイスである。EC108 は、電力管理のためのエンベデッドコントローラを含むワンチップマイクロコンピュータである。EC108 は、ユーザによるパワー ボタンの操作に応じて本タブレットコンピュータ 10 を電源オン又は電源オフする機能を有している。

【0020】

ところで、画像内の文字や手書きされた文字が認識される場合には、認識のための言語環境が予め決められていることが想定される。例えば、OS201 等が日本語で使用されることが指定されている場合、文字認識には日本語環境のための辞書データが用いられる。また、例えば、OS201 等が英語で使用されることが指定されている場合、文字認識には英語環境のための辞書データが用いられる。

【0021】

しかしながら、画像には、複数の言語（例えば、日本語と英語）の文字が含まれることがあり、また、手書きされる文字にも、複数の言語の手書き文字が含まれることがある。このような画像内の複数の言語の文字や複数の言語の手書き文字を、1つの言語環境のための辞書データを用いて認識した場合には、認識精度が低下するので、画像や手書き文書をキーワードで検索した場合の検索精度も低下する。

【0022】

なお、ユーザが文字認識のための言語環境を指定することも想定される。しかし、例えば、画像や手書き文書毎に、ユーザが文字認識のための言語環境を指定する操作を行うことは、ユーザにとって非常に煩雑である。

【0023】

そのため、本実施形態のデジタルノートブックアプリケーションプログラム 202 は、複数の言語環境のための複数の辞書データを用いて、画像内の文字や手書きされた文字を認識し、その認識結果を不揮発性メモリ 106 等に保存する。この認識結果を用いることにより、デジタルノートブックアプリケーションプログラム 202 は、画像内の文字や手書きされた文字がいずれの言語の文字であるかに関わらず、画像や手書き文書をキーワード（文字列）で精度良く検索することができる。

【0024】

図 3 は、本タブレットコンピュータ 10 によって実行されるデジタルノートブックアプリケーションプログラム 202 の機能構成の例を示す。デジタルノートブックアプリケーションプログラム 202 は、例えば、文字認識部 31、辞書選択部 32、格納処理部 33、検索部 34、表示制御部 35、マージ部 36、等を備える。なお、以下では、説明を分かりやすくするために、画像内の文字を認識及び検索の対象とした場合について説明する。

【0025】

画像は、例えば、データベース 42 に格納された画像データ 42A に基づく画像である。データベース 42 には、複数の言語の文字が描画された複数の画像が格納されている。この画像は、例えば、本やノートのページ、資料、貼り紙、ホワイトボード等を撮影又はスキャンして得られた画像、ウェブページやドキュメントが表示された画面のような、文字が表示された画面をキャプチャして得られた画像、等である。この画像に含まれる文字には、複数の言語の文字、例えば、日本語の文字である平仮名と英語の文字であるアルファベットなどが含まれ得る。

【0026】

辞書選択部 32 は、画像内の文字を認識するための複数の辞書データ 41 を選択する。辞書選択部 32 は、例えば、不揮発性メモリ 106 等に格納されている全ての辞書データ

10

20

30

40

50

41を選択してもよいし、予め決められたルールに基づいて複数の辞書データ41を選択してもよい。辞書選択部32は、例えば、OS201で指定されている言語環境のための辞書データ（例えば、日本語環境のための辞書データ）41と、他のよく使われている言語に対応する言語環境のための辞書データ（例えば、英語環境のための辞書データ）41とを選択する。

【0027】

辞書データ41は、文字認識のための文字認識辞書データを含む。この文字認識辞書データは、各々が文字とその文字の特徴量とを含む複数のエントリを含む。また、辞書データ41には、複数の文字のまとまりを認識するための単語知識のデータが含まれていてよい。

10

【0028】

なお、日本語の言語環境では、日本語の文字である平仮名、カタカナ、漢字、数字、及び記号に加えて、アルファベット等の他の言語の文字も用いられることがあるので、OS201やアプリケーションプログラム等はこれらすべての文字を扱うことができるよう構成されている。例えば、日本語の言語環境で取り扱われるファイル（文書ファイル、画像ファイル、ウェブページのファイル、等）には、日本語の文字である平仮名、カタカナ、漢字、数字、及び記号に加えて、アルファベット等の他の言語の文字も記述されることが多い。そのため、日本語環境のための辞書データ41には、平仮名、カタカナ、漢字、数字、及び記号に対応するエントリだけでなく、アルファベット等の他の言語の文字に対応するエントリも含まれている。

20

【0029】

また、英語の言語環境では、英語の文字であるアルファベット、数字、記号等が用いられるので、OS201やアプリケーションプログラム等はこれらの文字を扱うことができるよう構成されている。例えば、英語の言語環境で取り扱われるファイルには、アルファベット、数字、記号等が記述されることが多い。そのため、英語環境のための辞書データ41には、アルファベットや数字、記号に対応するエントリが含まれている。

【0030】

辞書選択部32は、選択された複数の辞書データ41の各々を文字認識部31に出力する。

【0031】

30

文字認識部31は、第1言語環境のための辞書データ41を用いて、複数の言語の文字を含む複数の画像の内の、第1画像に含まれる少なくとも1つの第1文字を少なくとも1つの第1文字コードに変換し、第2言語環境のための辞書データ41を用いて、少なくとも1つの第1文字を少なくとも1つの第2文字コードに変換する。格納処理部33は、少なくとも1つの第1文字コードと、当該第1文字コードに対応する第1画像内の第1領域との組をデータベース42に格納し、少なくとも1つの第2文字コードと、当該第2文字コードに対応する第1画像内の第2領域との組とをデータベース42に格納する。

【0032】

そして、検索部34は、ユーザによって少なくとも1つの第3文字コードが入力され、当該第3文字コードが第1文字コードに含まれる場合、又は第3文字コードが第2文字コードに含まれる場合、複数の画像から第1画像を抽出する。表示制御部35は、抽出された第1画像をLCD17Aの画面に表示する。

40

【0033】

表示制御部35は、第3文字コードが前記第1文字コードに含まれる場合、ハイライトされた第1領域を含む第1画像を表示し、第3文字コードが第2文字コードに含まれる場合、ハイライトされた第2領域を含む第1画像を表示する。また、表示制御部35は、第3文字コードが第1文字コードに含まれ、且つ第3文字コードが第2文字コードに含まれる場合、第1領域と第2領域の中間の領域、又は第1領域と第2領域とを包含する領域をハイライトして第1画像を表示する。

【0034】

50

より具体的には、文字認識部31は、辞書選択部32によって出力された言語環境毎の辞書データ41を用いて、画像内の文字を認識する。すなわち、文字認識部31は、1つの言語環境のための辞書データ41を用いて、画像に含まれる少なくとも1つの文字を少なくとも1つの文字コードに変換し、これを複数の辞書データ41のそれぞれについて繰り返し実行する。文字認識部31は、例えば、画像から文字と推定される領域を抽出し、その領域内の画像特徴量を算出する。文字認識部31は、辞書データ41を用いて、算出された画像特徴量との類似度が最も高い特徴量を有する文字（文字コード）を決定する。これにより、画像内の文字をテキスト（文字コード）に変換することができる。

【0035】

なお、文字認識部31は、文字認識辞書データを用いて、算出された特徴量との類似度が高い特徴量を有する複数の文字（文字コード）候補を決定してもよい。決定された文字候補と、画像上の近傍（上下左右、等）に描画された文字候補又は画像上の近傍で認識された文字とは、単語のような意味のある文字列を構成している可能性が高い。そのため、文字認識部31は、単語知識の辞書データ41を用いて、文字候補から、近傍の文字（又は文字候補）との組み合わせで単語が構成されるような文字候補を選択することにより、画像内の文字をテキスト（文字コード）に変換する。

【0036】

格納処理部33は、このような文字認識結果を用いてインデックスデータ42Bを生成し、データベース42に格納する。格納処理部33は、例えば、第1言語環境のための辞書データ41を用いて得られた文字認識結果を用いて、少なくとも1つの第1文字コードと、当該第1文字コードに対応する画像内の第1領域との組を不揮発性メモリ106等に格納し、第2言語環境のための辞書データ41を用いて得られた文字認識結果を用いて、少なくとも1つの第2文字コードと、当該第2文字コードに対応する画像内の第2領域との組を不揮発性メモリ106等に格納する。また、格納処理部33は、文字が認識された画像のデータをデータベース42に格納してもよい。データベース42は、不揮発性メモリ106や、ネットワークを介して接続されるストレージ等に格納されている。

【0037】

図4を参照して、複数の言語環境のための複数の辞書データ41を用いて、画像内の文字が認識される例を説明する。図4に示す画像51には、英語の文字（アルファベット）からなる文字列“Hello”51A及び“World”51Bと、日本語の文字（平仮名）からなる文字列“こんにちは”51Cとが描画されている。

【0038】

この画像51内の文字を日本語環境のための辞書データ41で認識した場合の認識結果52は、以下の通りである。画像51内の英語の文字列“Hello”51Aは、英語の文字コード列“Hello”52Aとして正しく認識される。画像51内の英語の文字列“World”51Bは、日本語の文字コード列“んみむ”52Bとして誤って認識される。そして、画像51内の日本語の文字列“こんにちは”51Cは、日本語の文字コード列“こんにちは”52Cとして正しく認識される。

【0039】

上述したように、日本語環境のための辞書データ41には、アルファベットに対応するエントリも含まれているので、日本語の文字列51Cが正しく認識されるだけでなく、一部の英語の文字列51Aも正しく認識されることがある。

【0040】

また、この画像51内の文字を英語環境のための辞書データ41で認識した場合の認識結果53は、以下の通りである。画像51内の英語の文字列“Hello”51Aは、英語の文字コード列“Hello”53Aとして正しく認識される。画像51内の英語の文字列“World”51Bは、英語の文字コード列“World”53Bとして正しく認識される。そして、画像51内の日本語の文字列“こんにちは”51Cは、英語の文字コード列“F h I E b l F”53Cとして誤って認識される。

【0041】

10

20

30

40

50

英語環境のための辞書データ41には、日本語の文字に対応するエントリは含まれていないので、英語の文字列51A, 51Bが正しく認識される一方、日本語の文字列51Cは英語の文字コード列53Cとして誤って認識される。このように、画像51内の文字の言語と、辞書データ41の言語とが一致していない場合、認識精度は低くなる。

【0042】

格納処理部33は、これら複数の言語環境のための複数の辞書データ41を用いた文字認識結果52, 53を、認識の正誤にかかわらず、インデックスデータ42Bとしてデータベース42に保存する。インデックスデータ42Bは、例えば、認識された少なくとも一つの文字と、その文字に対応する領域（例えば、文字を囲む矩形領域）を示す座標との組を含む。

10

【0043】

図5は、インデックスデータ42Bの一構成例を示す。インデックスデータ42Bは、画像内の少なくとも1つの文字が認識された少なくとも1つの文字（文字コード）に対応する少なくとも一つのエントリを含む。ある文字に対応するエントリは、例えば、ID、画像ID、辞書、文字、領域を含む。「ID」は、その文字に付与された識別情報を示す。「画像ID」は、その文字が認識された画像に付与された識別情報を示す。なお、「画像ID」の代わりに、画像のファイル名（ファイルパス）のような画像を特定可能な情報が用いられてもよい。「辞書」は、その文字の認識に用いられた辞書データを示す。「辞書」には、「日本語」、「英語」のような言語（言語環境）の名称が設定される。なお、「辞書」には、各辞書データに付与された識別情報（辞書ID）が設定されてもよい。「文字」は、その文字（文字コード）を示す。「領域」は、その文字が認識された画像上の領域を示し、例えば、その文字を包含する矩形領域の左上端の座標と右下端の座標とによって表される。

20

【0044】

図5に示すインデックスデータ42Bの例には、図4に示した認識結果52A, 52B, 52Cにそれぞれ対応するエントリ71, 72, 73と、認識結果53A, 53B, 53Cにそれぞれ対応するエントリ74, 75, 76とが含まれている。日本語環境のための辞書データ41を用いた認識処理と、英語環境のための辞書データ41を用いた認識処理が行われたことにより、少なくとも一方の認識処理によって、画像内の文字列51A, 51B, 51Cが正しく認識され、正しく認識された文字列52A, 52C, 53A, 53Bに基づくエントリ71, 73, 74, 75が生成されている。これにより、画像をキーワードで精度良く検索することができる。

30

【0045】

なお、このインデックスデータ42Bには、誤って認識された文字列52B, 53Cのエントリ72, 76が含まれている。しかし、そのような誤って認識された文字列52B, 53Cがキーワードとして入力される可能性は低いので、文字列52B, 53Cのエントリ72, 76がデータベース42に格納されていたとしても、検索精度に影響する可能性は低い。

【0046】

ユーザによってクエリ（以下、キーワードとも称する）が入力された場合、検索部34は、データベース42に格納されたインデックスデータ42Bを用いて、入力されたキーワードに対応する画像内の文字を検索する。このクエリは、例えば、ユーザがソフトウェアキーボードやキーボードを用いて入力した少なくとも1つの文字コード（第3文字コード）である。検索部34は、例えば、インデックスデータ42Bの「文字」の値を、入力されたキーワードで全文検索することにより、キーワードにマッチする文字を含む画像を決定することができる。

40

【0047】

図6は、ユーザが入力したキーワード（文字コード列）で、データベース42が検索される例を示す。以下では、インデックスデータ42Bに、図5に示した複数のエントリ71～76が含まれていることを想定する。

50

【0048】

まず、ユーザが文字列（文字コード列）“こんにちは”55を入力した場合について説明する。検索部34は、インデックスデータ42Bの複数のエントリ71～76から、この文字列55と一致する文字列52Cのエントリ73を抽出する。検索部34は、抽出されたエントリ73の「画像ID」に示される値に基づいて、画像データ42Aを読み出す。そして、表示制御部35は、図7に示すように、抽出されたエントリ73の「領域」に示される座標を用いて、文字列52Cに対応する矩形領域511にハイライト処理を施し、そのハイライトされた領域511を含む画像51を画面に表示する。

【0049】

次いで、ユーザが文字列“Hello”56を入力した場合について説明する。検索部34は、インデックスデータ42Bの複数のエントリ71～76から、この文字列56と一致する、文字列52Aのエントリ71と文字列53Aのエントリ74とを抽出する。検索部34は、抽出されたエントリ71, 74の「画像ID」に示される値に基づいて、画像データ42Aを読み出す。

10

【0050】

マージ部36は、抽出された2つのエントリ71, 74が同一の「画像ID」の値を含み、それら2つのエントリ71, 74によって示される「領域」が、同一の文字列を囲む領域であると推定される場合、2つのエントリ71, 74によって示される「領域」をマージする。文字列52Aを囲む領域と文字列53Aを囲む領域とは、大部分が重複し、同一の文字列を囲む領域であると推定されるので、マージ部36は、文字列52Aを囲む領域と文字列53Aを囲む領域とがマージされた領域512を算出する。そして、表示制御部35は、図8に示すように、マージされた領域512にハイライト処理を施し、そのハイライトされた領域512を含む画像51を画面に表示する。ハイライトは、例えば、領域511に対応する枠線の描画、領域511内の背景色の変更、等である。これにより、入力された文字列55に一致する部分をユーザに分かりやすく提示することができる。

20

【0051】

上述したように、日本語環境のための辞書データ41には、平仮名、カタカナ、漢字、数字、及び記号に対応するエントリだけでなく、アルファベット等の他の言語の文字に対応するエントリも含まれる。また、英語環境のための辞書データ41には、アルファベットや数字、記号に対応するエントリが含まれる。そのため、例えば、画像内のアルファベットの文字列が、日本語環境のための辞書データ41を用いた文字認識処理と、英語環境のための辞書データ41を用いた文字認識処理の両方で正しく認識されることがある。

30

【0052】

しかし、日本語環境のための辞書データ41と英語環境のための辞書データ41とでは、文字認識のための辞書データ、すなわち、辞書データで規定される文字の特徴量が異なる場合がある。そのため、認識された画像上の文字（文字列）を囲む領域は、2つの辞書データ41による認識結果で相違する可能性がある。

【0053】

図9及び図10は、2つの言語環境のための辞書データ41を用いた文字認識処理の結果の例を示す。

40

図9に示す例では、文字を装飾する要素61A, 61B, 61Cが文字の一部として認識されなかつたために、それら要素61A, 61B, 61Cが、認識された文字列610を囲む領域611内に含まれていない。一方、図10に示す例では、文字を装飾する要素61A, 61B, 61Cが文字の一部として認識されたので、それら要素61A, 61B, 61Cが、認識された文字列610を囲む領域612内に含まれている。これは、例えば、図9に示した例で用いられた辞書データ41が、文字を装飾する要素を含まないフォント（例えば、ゴシック体）を考慮した文字特徴量を規定していたのに対して、図10に示した例で用いられた辞書データ41が、文字を装飾する要素を含むフォント（例えば、Times New Roman）を考慮した文字特徴量を規定していたことによるものである。

50

【0054】

そして、これら2つの文字認識結果がインデックスデータ42Bとしてデータベース42に格納され、データベース42がキーワード（ここでは、“Japan”）で検索された場合、検索部34は、インデックスデータ42Bに含まれる複数のエントリから、「文字列」の値にそのキーワードが設定されている2つのエントリ（以下、第1エントリと第2エントリとも称する）を抽出する。これら2つのエントリはいずれも、画像61から認識された文字列“Japan”610に対応するエントリである。

【0055】

このような2つのエントリが抽出され、対応する画像61上の2つの領域611, 612の各々にハイライト処理が施された場合、1つの文字列610に対して、ハイライトされた2つの領域611, 612が表示される。そのため、画像61上に、例えば、1つの文字列610を囲むように、数ピクセルだけずれた2つの矩形領域611, 612の枠線が描画されるので、ユーザに見づらさを感じさせる可能性がある。

10

【0056】

そのため、マージ部36は、このような2つの領域611, 612をマージする。より具体的には、検索により抽出された2つのエントリが同一の「画像ID」の値を含むエントリである場合、マージ部36は、第1エントリ内の「領域」の値に基づく第1領域611と、第2エントリ内の「領域」の値に基づく第2領域612とが重複しているか否かを判定する。

【0057】

20

2つの領域611, 612が重複している場合、マージ部36は、第1領域611と第2領域612の平均面積に対する、重複した面積の割合を算出する。第1領域611と第2領域612とは、この算出された割合が大きいほど同一の文字の領域に対応している可能性が高いので、マージされることが望ましい。そのため、マージ部36は、算出された割合がしきい値よりも大きい場合、第1領域611と第2領域612とをマージする。このしきい値は、例えば、0.9に設定される。つまり、第1領域611と第2領域612とが重複した領域の面積が、第1領域611と第2領域612の平均面積の90パーセント以上を占める場合に、第1領域611と第2領域612とがマージされる。マージ部36は、2つの領域をマージするために、例えば、それら2つの領域611, 612の中間の領域を算出するか、又はそれら2つの領域611, 612を包含する領域を算出する。

30

【0058】

図9及び図10に示す例では、第2領域612が第1領域611を包含しているので、マージ部36は、マージ結果として、例えば、第2領域612を出力する。表示制御部35は、このマージ結果に基づいて、第2領域612にハイライト処理が施された画像61を画面に表示する。

【0059】

なお、第1領域611と第2領域612とが同一の領域（座標）である場合であっても、その同一の領域611, 612に二重にハイライト処理が施されると、例えば、濃いハイライトによって、表示される画像61が見づらくなる可能性がある。したがって、マージ部36は、領域611と領域612とが同一の領域である場合にも、領域611, 612をマージしてもよい。

40

【0060】

次いで、図11のフローチャートを参照して、本タブレットコンピュータ10によって実行される文字認識処理の手順の例を説明する。

【0061】

まず、文字認識部31は、文字を含む画像のデータを読み込む（ロックB11）。この画像内の文字には、複数の言語の文字、例えば、日本語の文字である平仮名と英語の文字であるアルファベットとが含まれ得る。

【0062】

次いで、辞書選択部32は、画像内の文字を認識するための辞書データ41を選択する

50

(ブロック B 1 2)。文字認識部 3 1 は、選択された辞書データ 4 1 を用いて、画像内の文字を認識する (ブロック B 1 3)。辞書データ 4 1 は、各々が文字とその文字の特徴量とを含む複数のエントリを含む。文字認識部 3 1 は、例えば、画像内の文字と推定される領域を抽出し、その領域内の特徴量を算出し、辞書データ 4 1 を用いて、算出された特徴量との類似度が最も高い特徴量を有する文字を決定することにより、画像内の文字を認識する。

【 0 0 6 3 】

格納処理部 3 3 は、この認識結果を用いてインデックスデータ 4 2 B を生成し、データベース 4 2 に格納する (ブロック B 1 4)。インデックスデータ 4 2 B の構成は、図 5 を参照して上述した通りである。

10

【 0 0 6 4 】

次いで、辞書選択部 3 2 は、画像内の文字を認識するための別の辞書データ 4 1 があるか否かを判定する (ブロック B 1 5)。別の辞書データ 4 1 がある場合 (ブロック B 1 5 の Y E S)、ブロック B 1 2 に戻り、その別の辞書データ 4 1 を用いた文字認識処理を続行する。また、別の辞書データ 4 1 がない場合 (ブロック B 1 5 の N O)、処理を終了する。

【 0 0 6 5 】

図 1 2 のフローチャートは、本タブレットコンピュータ 1 0 によって実行される検索処理の手順の例を示す。この検索処理は、文字認識処理によって構築されたデータベース 4 2 を、ユーザによって入力されたキーワード (文字列) で検索する処理である。

20

【 0 0 6 6 】

まず、検索部 3 4 は、キーワードが入力されたか否かを判定する (ブロック B 2 1)。ユーザは、タブレットコンピュータ 1 0 の画面上に表示されるソフトウェアキー ボードや、タブレットコンピュータ 1 0 に有線又は無線で接続されるキーボード (図示せず) 等を用いて、検索のためのキーワードとなる文字列を入力する。キーワードが入力されていない場合 (ブロック B 2 1 の N O)、ブロック B 2 1 に戻り、キーワードが入力されたか否かが再度判定される。

【 0 0 6 7 】

キーワードが入力されている場合 (ブロック B 2 1 の Y E S)、検索部 3 4 は、データベース 4 2 内のインデックスデータ 4 2 B から、キーワードに対応する文字 (文字列) を含むエントリを抽出する (ブロック B 2 2)。検索部 3 4 は、データベース 4 2 から、抽出されたエントリ内の「画像 I D 」に対応する画像データ 4 2 A を読み出す (ブロック B 2 3)。そして、検索部 3 4 は、抽出されたエントリの中に、同一の画像 I D を含む複数のエントリがあるか否かを判定する (ブロック B 2 4)。同一の画像 I D を含む複数のエントリがある場合 (ブロック B 2 4 の Y E S)、マージ部 3 6 は、検索結果を適切に表示するためのマージ処理を実行する (ブロック B 2 5)。マージ処理の手順の詳細については、図 1 3 を参照して後述する。そして、表示制御部 3 5 は、マージ部 3 6 によるマージ結果を用いて、キーワードに対応する画像上の領域をハイライトする (ブロック B 2 6) 。

30

【 0 0 6 8 】

同一の画像 I D を含む複数のエントリがない場合 (ブロック B 2 4 の N O)、表示制御部 3 5 は、抽出されたエントリに示される領域に基づいて、キーワードに対応する画像上の領域をハイライトする (ブロック B 2 7)。

40

【 0 0 6 9 】

そして、表示制御部 3 5 は、キーワードに対応する領域がハイライトされた画像 (画像のリスト) を L C D 1 7 A の画面に表示する (ブロック B 2 8)。

【 0 0 7 0 】

図 1 3 のフローチャートを参照して、本タブレットコンピュータ 1 0 によって実行されるマージ処理の手順の例を説明する。

【 0 0 7 1 】

50

まず、マージ部 3 6 は、同一の画像 ID を含む 2 つのエントリ（第 1 エントリ及び第 2 エントリ）によって示される、第 1 領域と第 2 領域とが重複しているか否かを判定する（ブロック B 3 1）。マージ部 3 6 は、第 1 エントリ内の「領域」の座標情報に基づく第 1 矩形領域と、第 2 エントリ内の「領域」の座標情報に基づく第 2 矩形領域とが重複しているか否かを判定する。2 つの領域が重複していない場合（ブロック B 3 1 の N O）、マージ処理を終了する。

【 0 0 7 2 】

一方、2 つの領域が重複している場合（ブロック B 3 1 の Y E S）、マージ部 3 6 は、第 1 領域と第 2 領域の平均面積に対する、重複した面積の割合を算出する（ブロック B 3 2）。第 1 領域と第 2 領域とは、この算出された割合が大きいほど同一の文字の領域に対応している可能性が高いので、マージされることが望ましい。そのため、マージ部 3 6 は、算出された割合がしきい値よりも大きいか否かを判定する（ブロック B 3 3）。そして、算出された割合がしきい値よりも大きい場合（ブロック B 3 3 の Y E S）、マージ部 3 6 は、第 1 領域と第 2 領域とをマージする（ブロック B 3 4）。

【 0 0 7 3 】

なお、上述した例では、画像内の文字を認識する場合について説明したが、タッチスクリーンディスプレイ 1 7 上で、手書きにより入力された少なくとも 1 つのストロークを含む手書き文書から、文字を認識することもできる。手書きにより入力されたストロークは、画像データではなく、各ストロークの軌跡の座標列とストローク間の順序関係とを示す時系列情報として記憶媒体に保存される。この時系列情報は、少なくとも 1 つのストロークにそれぞれ対応する少なくとも 1 つのストロークデータを含む。各ストロークデータは、ある 1 つのストロークに対応し、このストローク上の点それぞれに対応する座標データ系列（時系列座標）を含む。これらストロークデータの並びの順序は、ストロークそれが手書きされた順序つまり筆順に相当する。

【 0 0 7 4 】

辞書データ 4 1 には、このようなストロークデータ（手書き文書）から文字を認識するための文字認識辞書データが含まれていてもよい。この文字認識辞書データは、例えば、各々が文字とその文字を構成するストロークの特徴量とを含む複数のエントリを含む。

【 0 0 7 5 】

文字認識部 3 1 は、例えば、入力されたストロークのストロークデータを用いて特徴量を算出し、辞書データ（文字認識辞書データ）4 1 を用いて、算出された特徴量との類似度が最も高い特徴量を有する文字を決定する。これにより、ストロークデータをテキスト（文字コード）に変換することができる。

【 0 0 7 6 】

格納処理部 3 3 は、このような文字認識結果を用いてインデックスデータ 4 2 B を生成し、データベース 4 2 に格納する。インデックスデータ 4 2 B は、例えば、認識された少なくとも一つの文字と、その文字に対応する手書き文書内の領域（例えば、文字に対応するストロークを包含する矩形領域）を示す座標との組を含む。

【 0 0 7 7 】

検索部 3 4 は、上述した画像内の文字を検索する場合と同様に、データベース 4 2 に格納されたインデックスデータ 4 2 B を用いて、入力されたキーワードに対応する手書き文書内の文字を検索する。

【 0 0 7 8 】

さらに、検索のためのキーワードが手書きで入力されてもよい。その場合、文字認識部 3 1 が手書きで入力されたキーワード（ストローク）を文字コードに変換した後に、検索部 3 4 は、その文字コードを用いて、画像内の文字や手書き文書内の文字を検索する。

【 0 0 7 9 】

また、上述した例では、主に、日本語環境のための辞書データ 4 1 と英語環境のための辞書データ 4 1 とが用いられる場合について説明したが、他の言語環境のための辞書データ 4 1 でも同様に、画像内の文字を認識し、その認識結果を用いてキーワードによる検索

10

20

30

40

50

を行うことができる。また、同じ言語であっても、国や地域によって異なる言語環境の辞書データ41が用いられてもよい。

【0080】

以上説明したように、本実施形態によれば、複数の言語の文字を含む画像をキーワードで検索することができる。文字認識部31は、第1言語環境のための辞書データ41を用いて、複数の言語の文字を含む複数の画像の内の、第1画像に含まれる少なくとも1つの第1文字を少なくとも1つの第1文字コードに変換し、第2言語環境のための辞書データ41を用いて、少なくとも1つの第1文字を少なくとも1つの第2文字コードに変換する。格納処理部33は、少なくとも1つの第1文字コードと、当該第1文字コードに対応する第1画像内の第1領域との組をデータベース42に格納し、少なくとも1つの第2文字コードと、当該第2文字コードに対応する第1画像内の第2領域との組とをデータベース42に格納する。10

【0081】

これにより、画像内の文字がいずれの言語の文字であるかに関わらず、画像内の文字が認識され、その認識結果に基づく文字コードと領域の組であるインデックスデータ42Bが生成される。そして、このインデックスデータ42Bを用いることにより、複数の言語の文字を含む画像を、ユーザによって入力されたキーワードで検索することができる。

【0082】

なお、図11から図13のフローチャートで説明した本実施形態の処理手順は全てソフトウェアによって実行することができる。このため、この処理手順を実行するプログラムを格納したコンピュータ読み取り可能な記憶媒体を通じてこのプログラムを通常のコンピュータにインストールして実行するだけで、本実施形態と同様の効果を容易に実現することができる。20

【0083】

本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。30

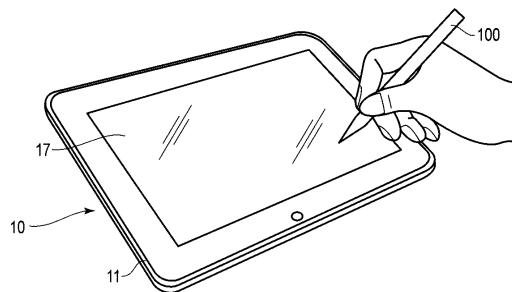
【符号の説明】

【0084】

17A...LCD、202...デジタルノートブックアプリケーションプログラム、31...文字認識部、32...辞書選択部、33...格納処理部、34...検索部、35...表示制御部、36...マージ部、41...辞書データ、42...データベース、42A...画像データ、42B...インデックスデータ。

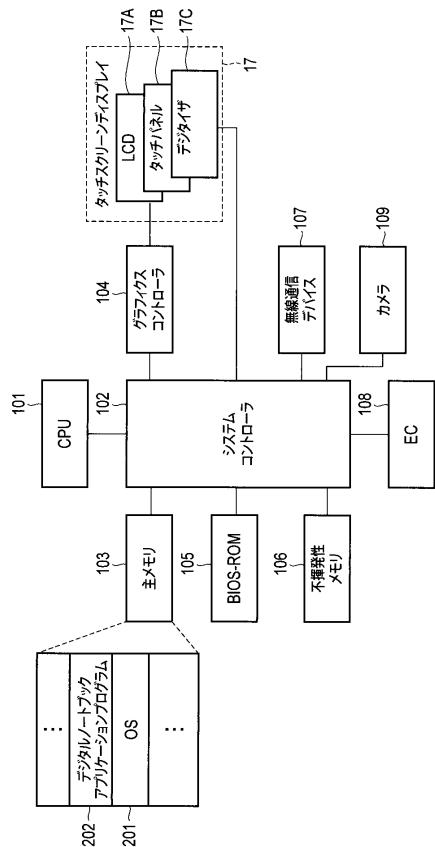
【図1】

図1



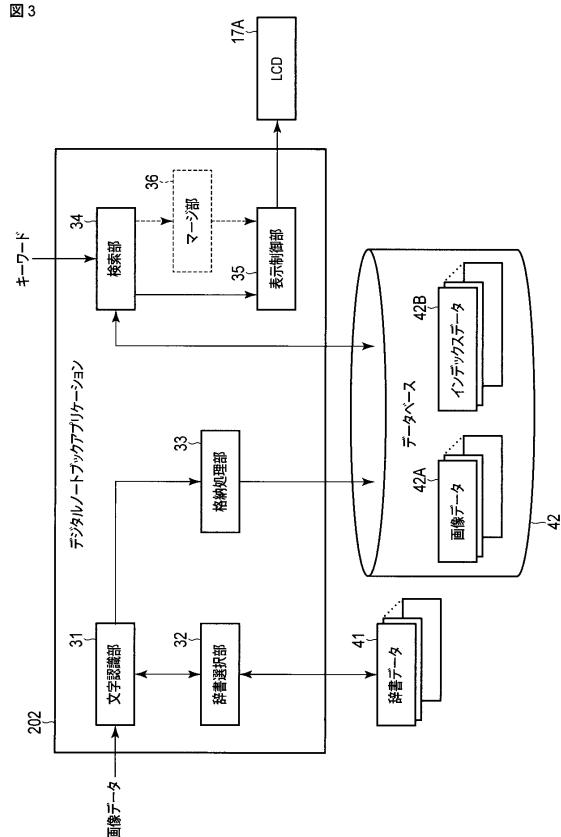
【図2】

図2



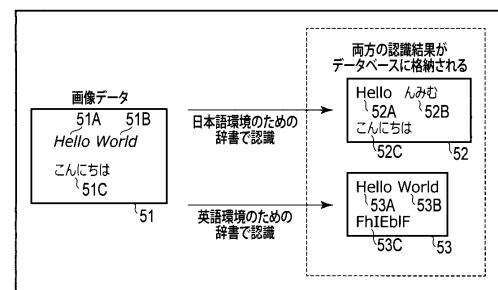
【図3】

図3



【図4】

図4



【図5】

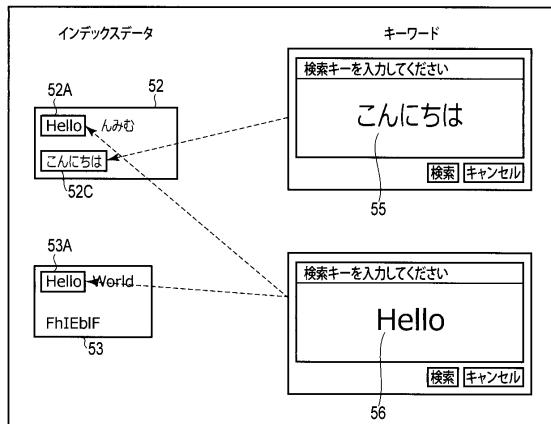
図5

ID	画像ID	辞書	文字	領域	...
0001	0001	日本語	Hello	(x1,y1)–(x2,y2)	...
0002	0001	日本語	んみむ	(x3,y3)–(x4,y4)	...
0003	0001	日本語	こんにちは	(x5,y5)–(x6,y6)	...
0004	0001	英語	Hello	(x7,y7)–(x8,y8)	...
0005	0001	英語	World	(x9,y9)–(x10,y10)	...
0006	0001	英語	FhIEbIF	(x11,y11)–(x12,y12)	...
...	:

42B

【図6】

図6



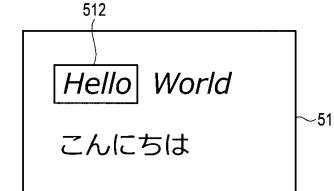
【図7】

図7



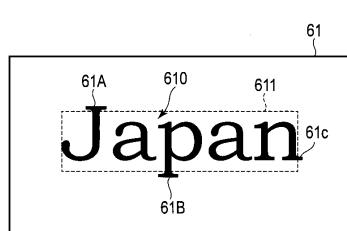
【図8】

図8



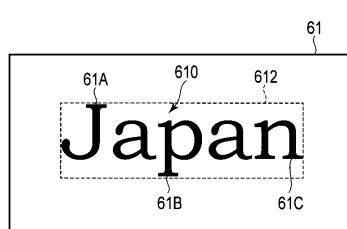
【図9】

図9



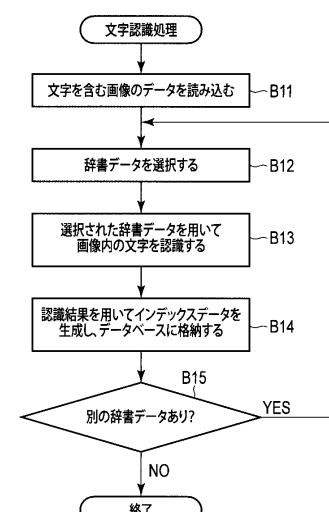
【図10】

図10

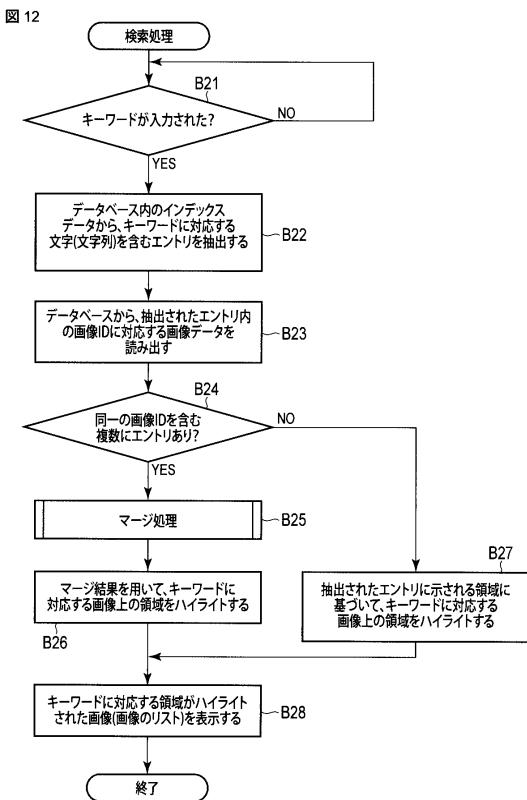


【図11】

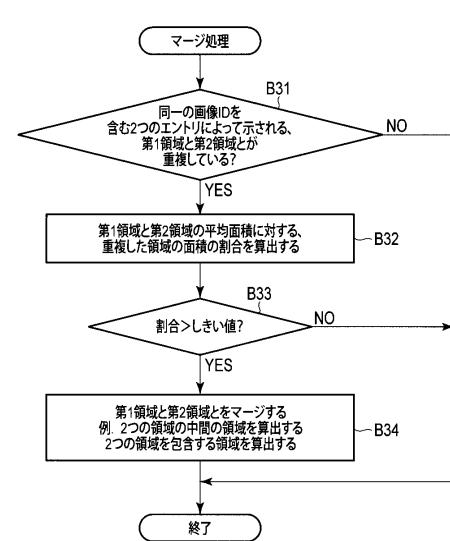
図11



【図12】



【図13】



フロントページの続き

(56)参考文献 特開2006-260115(JP,A)

特開2014-197341(JP,A)

特開2005-208687(JP,A)

特開2010-146509(JP,A)

(58)調査した分野(Int.Cl., DB名)

IPC	G 06 F	1 6 / 0 0	-	1 6 / 9 5 8
		1 7 / 2 0	-	1 7 / 2 8
	G 06 K	9 / 0 0	-	9 / 8 2