

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2010 年 9 月 23 日(23.09.2010)

PCT

(10) 国際公開番号
WO 2010/106642 A1

- (51) 国際特許分類:
G06F 17/30 (2006.01)
- (21) 国際出願番号: PCT/JP2009/055177
- (22) 国際出願日: 2009 年 3 月 17 日(17.03.2009)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人 (米国を除く全ての指定国について): 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 Kanagawa (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 岩倉 友哉 (IWAKURA, Tomoya) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内 Kanagawa (JP). 岡本 青史 (OKAMOTO, Seishi) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中 4 丁目 1 番 1 号 富士通株式会社内 Kanagawa (JP).
- (74) 代理人: 原田一男 (HARADA, Kazuo); 〒2200004 神奈川県横浜市西区北幸二丁目 1 番 6 号 鶴見ビル 2 階 Kanagawa (JP).

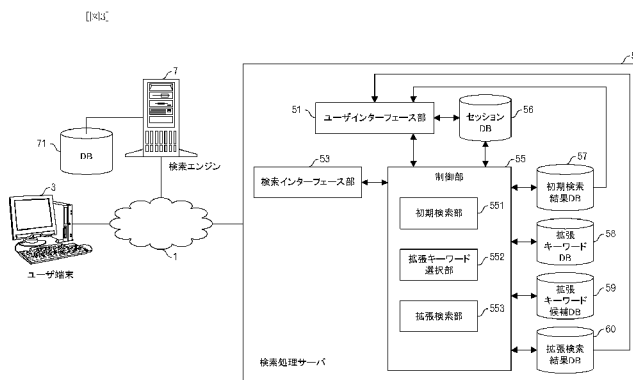
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))

(54) Title: SEARCH PROCESSING METHOD AND APPARATUS

(54) 発明の名称: 検索処理方法及び装置



3 USER TERMINAL
7 SEARCH ENGINE
5 SEARCH PROCESSING SERVER
51 USER INTERFACE UNIT
56 SESSION DB
53 SEARCH INTERFACE UNIT
55 CONTROL UNIT
551 INITIAL SEARCH SECTION
552 EXTENDED KEYWORD SELECTION SECTION
553 EXTENDED SEARCH SECTION
57 DB OF RESULT OF INITIAL SEARCH
58 DB OF EXTENDED KEYWORD
59 DB OF EXTENDED KEYWORD CANDIDATE
60 DB OF RESULT OF EXTENDED SEARCH

(57) Abstract: The result of initial search is acquired by search using a search keyword inputted by a searcher. On the other hand, an extended search keyword associated with the search keyword is extracted, and the number of corresponding documents or the frequency of appearance in the result of the initial search is counted. When the number of corresponding documents or the frequency of appearance is not larger than a predetermined value (including 0), the extended search keyword is adopted, and the result of the extended search is acquired by search using the search keyword and the extended search keyword adopted. Then, the result of the initial search and that of the extended search are presented to the searcher. Thereby, such a document as is buried in the result of the initial search can be presented to the searcher.

(57) 要約: 検索者により入力された検索キーワードで検索を行って初期検索結果を取得する。一方、検索キーワードに関連付けられた拡張検索キーワードを抽出して、初期検索結果における該当文書件数又は出現頻度を計数する。そして、当該該当文書件数又は出現頻度が所定値 (0 を含む) 以下である場合には当該拡張検索キーワードを採用し、検索キーワードと採用された拡張

検索キーワードで検索を行って、拡張検索結果を取得する。そして、初期検索結果と拡張検索結果を、検索者に提示する。これによって、初期検索結果に埋もれてしまうような文書を検索者に提示できるようになる。

WO 2010/106642 A1

明 細 書

検索処理方法及び装置

技術分野

[0001] 本技術は、キーワード検索技術に関する。

背景技術

[0002] 従来、例えば「リンゴ」というキーワードでデータベース検索を実施すると、図1に示すような検索結果が得られるとする。図1から分かるように、上位1000件の検索結果には、「リンゴジュース」と「リンゴジャム」についての文書ばかりが含まれる。通常検索エンジンは、検索結果として返す文書の件数を制限しており、検索者は例えば1001番目以降の文書、例えば「リンゴパイ」を含む文書にはアクセスできない。

[0003] また、予めキーワード毎に関連語を格納する関連辞書から、入力キーワードに関連する関連語を抽出し、抽出関連語をさらに用いて検索を行うような従来技術も存在する。抽出関連語は、予め定められた優先度などにに基づき抽出される。しかしながら、図1のような検索結果が得られるような状況において「ジュース」や「ジャム」といったような関連語を付加して別途検索を行っても、図2に示すように、ほとんど図1の検索結果と同様の結果が得られるので、検索者はあまり有効な付加情報が得られていない。

[0004] また、文書の検索結果としてヒット件数が多い場合に、検索語の近傍（文書内の近傍）から出現頻度順に追加検索語を抽出し、各単語の関連性を階層化して表示して、その中から選択された単語を検索条件に追加して絞込検索を行うといった技術も存在している。しかしながら、文書内の近傍の単語を追加検索語として抽出するので、入力キーワードだけでも得られる文書が抽出される可能性が高い。

特許文献1:特開2001-75981号公報

特許文献2:日本特許第3428554号公報

発明の開示

発明が解決しようとする課題

[0005] 従って、本技術の目的は、入力検索キーワードから得られる検索結果の内容とは異

なる内容の検索結果を自動的に提示できるようにするための新規な技術を提供することである。

課題を解決するための手段

[0006] 本検索処理方法は、検索キーワードを受信するステップと、受信した検索キーワードで、文書に関するデータを格納するデータベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む初期検索結果を取得し、初期検索結果格納部に格納するステップと、各キーワードに対応付けて拡張検索キーワードを格納する拡張検索キーワード格納部から、受信した検索キーワードに対応付けられている拡張検索キーワードを抽出するステップと、抽出された各拡張検索キーワードで初期検索結果格納部を検索して、抽出された各拡張検索キーワードについて該当文書件数又は出現頻度を計数し、各拡張検索キーワードに対応付けて該当文書件数又は出現頻度を集計結果格納部に格納するステップと、集計結果格納部に格納されている拡張検索キーワードのうち該当文書件数又は出現頻度が少ない順で上位所定数の拡張検索キーワードの各々又は該当文書件数又は出現頻度が所定値以下の拡張検索キーワードの各々と受信した検索キーワードとの組み合わせで、データベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む拡張検索結果を取得し、拡張検索結果格納部に格納するステップと、初期検索結果格納部に格納されている初期検索結果の少なくとも一部と、拡張検索結果格納部に格納されている拡張検索結果の少なくとも一部とを出力する出力ステップとを含む。

図面の簡単な説明

- [0007] [図1]図1は、従来技術を説明するための図である。
- [図2]図2は、従来技術を説明するための図である。
- [図3]図3は、実施の形態に係るシステム概要を示す図である。
- [図4]図4は、メインの処理フローを示す図である。
- [図5]図5は、初期検索処理の処理フローを示す図である。
- [図6]図6は、検索キーワード入力画面の一例を示す図である。
- [図7]図7は、セッションDBに格納されるデータの一例を示す図である。

[図8]図8は、初期検索結果DBに格納されるデータの一例を示す図である。

[図9]図9は、拡張キーワード選択処理の処理フローを示す図である。

[図10]図10は、拡張キーワードDBに格納されているデータの一例を示す図である。

[図11]図11は、拡張キーワード候補DBに格納されるデータの一例を示す図である。

[図12]図12は、ソート後の拡張キーワード候補リストの一例を示す図である。

[図13]図13は、拡張検索処理の処理フローを示す図である。

[図14]図14は、拡張検索結果DBに格納されるデータの一例を示す図である。

[図15]図15は、検索結果生成処理の処理フローを示す図である。

[図16]図16は、検索結果表示画面の一例を示す図である。

[図17]図17は、コンピュータの機能ブロック図である。

発明を実施するための最良の形態

[0008] 図3に本技術の実施の形態に係るシステムの概要図を示す。例えばインターネットやLAN(Local Area Network)などのコンピュータネットワーク1には、複数のユーザ端末3と、本実施の形態の主要な処理を実施する検索処理サーバ5と、周知のデータベース検索を実施する検索エンジン7とが接続されている。

[0009] 検索エンジン7は、データベース(DB)71に接続されており、このDB71には、文書そのものを蓄積している場合もあれば、例えばネットワーク1に接続されている多数のサーバが保持している多数の文書(Webページ・データなど)のインデックスデータ等を保持している。検索エンジン7及び検索エンジン7が管理しているDB71の構成は、本実施の形態の主要部ではなく周知なので、これ以上述べない。

[0010] また、検索処理サーバ5は、ユーザ端末3とのインターフェースとなるユーザインターフェース部51と、検索エンジン7とのインターフェースとなる検索インターフェース部53と、制御部55と、セッションDB56と、初期検索結果DB57と、拡張キーワードDB58と、拡張キーワード候補DB59と、拡張検索結果DB60とを有する。制御部55は、ユーザインターフェース部51と検索インターフェース部53と連携して動作する。

[0011] また、制御部55は、セッションDB56に格納されているデータを用いて処理を行い、初期検索結果DB57に初期検索結果を格納する初期検索部551と、初期検索結果DB57及び拡張キーワードDB58に格納されているデータを用いて処理を行い、

処理結果を拡張キーワード候補DB59に格納する拡張キーワード選択部552と、セッションDB56及び拡張キーワードDB58に格納されているデータを用いて処理を行い、処理結果を拡張検索結果DB60に格納する拡張検索部553とを有する。

[0012] ユーザインターフェース部51は、ユーザ端末3から受信したデータをセッションDB56に登録し、初期検索結果DB57及び拡張検索結果DB60に格納されているデータを用いて検索結果表示データを生成し、ユーザ端末3に送信する。

[0013] 次に、図4乃至図16を用いて図3に示したシステムの動作を説明する。まず、図4を用いて処理の概要を示す。最初に、初期検索処理を実施する(ステップS1)。初期検索処理については、後に詳細に述べるが、通常の実行処理と同様である。次に、拡張キーワード選択処理を実施する(ステップS3)。拡張キーワード選択処理については、後に詳細に述べるが、入力検索キーワードに加えて用いる拡張キーワード候補が選択される。但し、ステップS3はステップS1と並列に処理を実施してもよい。また、シリアルに実施する場合には、ステップS1より先にステップS3を実施しても良い。その後、拡張検索処理を実施する(ステップS5)。拡張検索処理については、後に詳細に述べるが、入力検索キーワードと拡張キーワード候補との組み合わせについて検索処理を実施するものである。最後に、検索結果生成処理を実施する(ステップS7)。検索結果生成処理については、後に詳細を述べるが、初期検索結果と拡張検索結果とを例えば1画面で上下又は左右に並べて表示するためのWebページ・データを生成してユーザ端末3に送信して表示する処理である。

[0014] 次に、図5を用いて初期検索処理を説明する。まず、検索処理サーバ5のユーザインターフェース部51は、ユーザ端末3から検索キーワードを受信する(ステップS11)。例えば、ユーザ端末3からのアクセスに応じてユーザインターフェース部51は、検索キーワード入力画面を表示するためのWebページ・データをユーザ端末3に送信し、ユーザ端末3は、当該Webページ・データを受信し、図6に示すような検索キーワード入力画面を表示する。図6の画面例は、入力欄601及び検索ボタン602を有している。ここでは、入力欄601に検索キーワード「リンゴ」を入力して、検索ボタン602をクリックしたものとする。ユーザ端末3は、検索キーワードを検索要求として検索処理サーバ5に送信する。

- [0015] 次に、ユーザインターフェース部51は、セッションIDを発行し、当該セッションIDと検索キーワードと端末情報(例えば端末IDやIPアドレスなど)とを、セッションDB56に登録する(ステップS13)。セッションDB56には、例えば図7に示すようなデータが登録される。図7の例では、セッションIDと、キーワードと、端末ID(IPアドレスなどの場合もある)とが登録されるようになっている。また、ユーザインターフェース部51は、制御部55に処理を指示する。
- [0016] 例えば新たなレコードがセッションDB56に登録されると、制御部55の初期検索部551は、セッションDB56から新たに登録された検索キーワードを読み出し、検索インターフェース部53に対して、当該検索キーワードによる検索を検索エンジン7に実施させるように要求する。検索インターフェース部53は、初期検索部551からの要求に応じて、検索キーワードを含む検索要求を検索エンジン7に送信する。検索エンジン7は、検索処理サーバ5から検索キーワードを含む検索要求を受信し、例えばDB71に対して検索処理を実施し、検索結果の上位M件分のデータを検索処理サーバ5に送信する。検索処理サーバ5の検索インターフェース部53は、検索結果の上位M件分のデータを受信し、制御部55の初期検索部551に出力する。初期検索部551は、検索インターフェース部53から検索結果の上位M件分のデータを取得する(ステップS15)。検索結果は、例えば該当文書のタイトルとURL(Uniform Resource Locator)などを含む。
- [0017] そして、初期検索部551は、タイトル及びURLを含む検索結果のデータを、現セッションIDに対応付けて、初期検索結果DB57に格納する(ステップS17)。初期検索結果DB57には、例えば図8のようなデータが登録される。図8の例では、検索結果IDと、タイトルと、URLとが登録されるようになっている。なお、本例では、上位100件分のデータが登録されたことになる。なお、初期検索結果DB57は、セッションID毎に図8のようなテーブルを格納している。
- [0018] このように通常どおり入力された検索キーワードに基づく検索結果を得ることができる。
- [0019] 次に、図9を用いて拡張キーワード選択処理について説明する。拡張キーワード選択部552は、拡張キーワードDB58から検索キーワードに対応する拡張キーワード候

補を取得する(ステップS51)。例えば図10に示すようなデータが拡張キーワードDB 58に格納されている。図10の例では、入力キーワードと、対応する拡張キーワード候補群とが登録されるようになっている。例えば、「リンゴ」という入力キーワードに対して「ジュース」「ジャム」「あめ」「パイ」といった拡張キーワード候補が登録されている。

[0020] 次に、拡張キーワード選択部552は、取得した拡張キーワード候補のうち、未処理の拡張キーワード候補を1つ特定する(ステップS53)。そして、特定された拡張キーワード候補で初期検索結果DB57を検索し、初期検索結果DB57に格納されている初期検索結果における該当文書件数を計数し、拡張キーワード候補DB59に格納する(ステップS55)。例えば、初期検索結果DB57中の文書のうち、タイトルに、特定された拡張キーワード候補を含んでいる文書を該当すると判断し、その件数を計数する。例えば、図11に示すようなデータが、拡張キーワード候補DB59に格納されている。図11の例では、拡張キーワード候補毎に、該当文書件数が登録されるようになっている。なお、該当文書件数ではなく、出現回数を計数して登録するようにしても良い。すなわち、例えばタイトル等が長くて1件の文書のタイトル等に同じ単語が2度以上出現する場合には、出現回数は多くなる。なお、ステップS55を1回実行すると、図11の1行が登録される。

[0021] そして、拡張キーワード選択部552は、未処理の拡張キーワード候補が存在しているか判断する(ステップS57)。未処理の拡張キーワード候補が存在している場合には、ステップS53に戻る。一方、全ての拡張キーワード候補を処理した場合には、計数結果(図11)に基づき拡張キーワード候補を昇順にソートし、ソート結果を拡張キーワード候補DB59に格納する(ステップS59)。図11の例では、「あめ」が最も計数値が少なく、「ジュース」が最も計数値が多い。従って、図12に示すような拡張キーワード候補のリストが得られ、拡張キーワード候補DB59に登録される。そして処理は元の処理に戻る。

[0022] なお、ステップS55の代わりに、計数値が「0」であるかを判断し、計数値が「0」であれば、拡張キーワード候補DB59に登録するようにしても良い。このようにすれば、全く初期検索結果には含まれていないような拡張キーワード候補を特定できるので、初期検索結果とは全く異なる内容の検索結果を以下の処理で取得できるようになる。こ

のような場合にも、計数値が「0」である拡張キーワード候補を、図12に示すようなリストとして保持する。

[0023] 次に、拡張検索処理について図13を用いて説明する。最初に、拡張検索部553は、カウンタ*i*を1に初期化すると共に、*N*に採用拡張キーワード数を設定する(ステップS61)。そして、*i*番目の拡張キーワード候補を拡張キーワード候補DB59から取得する(ステップS63)。そして、拡張検索部553は、セッションDB56から処理に係る検索キーワード(すなわち現セッションIDに対応付けられた検索キーワード)を読み出し、検索インターフェース部53に対して、「検索キーワード&拡張キーワード候補」による検索を検索エンジン7に実施させるように要求する。検索インターフェース部53は、拡張検索部553からの要求に応じて、「検索キーワード&拡張キーワード候補」を含む検索要求を検索エンジン7に送信する。検索エンジン7は、検索処理サーバ5から「検索キーワード&拡張キーワード候補」を含む検索要求を受信し、例えばDB71に対して検索処理を実施し、検索結果(すなわち拡張検索結果)の上位*M*件分のデータを検索処理サーバ5に送信する。検索処理サーバ5の検索インターフェース部53は、拡張検索結果の上位*M*件分(初期検索処理の際の*M*と同じでも異なっても良い。)のデータを受信し、制御部55の拡張検索部553に出力する。拡張検索部553は、検索インターフェース部53から拡張検索結果の上位*M*件分のデータを取得する(ステップS65)。拡張検索結果は、例えば該当文書のタイトルとURL(Uniform Resource Locator)などを含む。

[0024] そして、拡張検索部553は、取得した拡張検索結果を拡張検索結果DB60に格納する(ステップS67)。拡張検索結果DB60に格納されるデータの一例を図14に示す。図14の例では、ステップS65を実施する毎に発行される拡張検索結果ID毎に、拡張クエリの条件である「検索キーワード&拡張キーワード候補」と、タイトル及びURLとが登録されるようになっている。拡張検索結果ID毎に、*M*件分のデータが登録される。本実施の形態では、セッションID毎に図14に示したようなデータテーブルを、拡張検索結果DB60に格納する。

[0025] そして、拡張検索部553は、*i*が*N*より小さいか判断する(ステップS69)。*i*が*N*より小さい場合には、*i*を1インクリメントしてステップS63に戻る。一方、*i*が*N*以上であれば、

元の処理に戻る。

- [0026] このように、初期検索にはあまり含まれない又は全く含まれない内容を有し、検索者に対して提示すべき拡張検索結果を取得する。
- [0027] 次に、図15を用いて検索結果生成処理について説明する。制御部55は、拡張検索部553の処理が終了すると、ユーザインターフェース部51に対して処理が完了したセッションIDを出力して検索結果生成処理を実施させる。
- [0028] ユーザインターフェース部51は、保持している結果表示用フォームを読み出し(ステップS71)、処理完了に係る検索キーワードを例えばセッションDB56から読み出し、結果表示フォーム中の提示箇所に設定する(ステップS73)。また、初期検索結果DB57から、処理完了に係るセッションIDに対応する初期検索結果を読み出し、表示可能な件数分だけ、各該当文書のタイトルをハイパーリンク化して該当URLにアクセスできるようにして、結果表示用フォーム中の提示箇所に設定する(ステップS75)。また、カウンタ*i*を1に初期化すると共に、*N*を採用された拡張キーワード数に設定する(ステップS77)。
- [0029] そして、ユーザインターフェース部51は、拡張検索結果領域の*i*番目の枠を、結果表示用フォーム内に確保すると共に(ステップS79)、*i*番目の拡張クエリの条件(すなわち、「検索キーワード&拡張キーワード候補」とその拡張検索結果を、拡張検索結果DB60から読み出し、当該拡張クエリを結果表示用フォーム中のその提示枠に設定すると共に、表示可能な件数分だけ、各該当文書のタイトルをハイパーリンク化して該当URLにアクセスできるようにして、結果表示用フォーム中の提示箇所に設定する(ステップS81)。
- [0030] そして、ユーザインターフェース部51は、*i*が*N*より小さいか判断する(ステップS83)。*i*が*N*より小さい場合には、*i*を1インクリメントして(ステップS85)ステップS79に戻る。一方、*i*が*N*以上であれば、上で述べた処理で生成された結果表示用フォームによる検索結果表示用ページ・データを、現セッションIDに係るユーザ端末3に送信する(ステップS87)。
- [0031] これに対して、ユーザ端末3は、検索処理サーバ5から検索結果表示用ページ・データを受信し、表示装置に例えば図16に示すような表示を行う。図16の例では、検

索キーワードが、ページの冒頭部分1501に提示され、その下の左側に検索キーワードによる初期検索結果の提示枠1502が含まれ、この提示枠1502の右側に拡張検索結果の提示枠1510及び1520が設けられている。提示枠1504には、初期検索結果の各該当文書(Webページなど)が列挙されており、ハイパーリンクで各該当文書(例えばWebページ)の具体的内容を参照できるようになっている。

[0032] 同様に、提示枠1511には、拡張クエリの条件「リンゴ&あめ」が提示されており、具体的な拡張検索結果は提示枠1512に列挙されている。同様に、提示枠1521には、拡張クエリの条件「リンゴ&パイ」が提示されており、具体的な拡張検索結果は提示枠1522に列挙されている。いずれも「リンゴ」だけで検索した際には、提示されない又はされにくい検索結果を含んでいる。

[0033] 該当文書の提示件数や拡張検索結果の提示枠数については表示領域の広さなどに依存する部分もあるが、任意である。複数ページに分割して提示するようにしても良いが、図6のように比較対比できるような提示態様が好ましい。なお、1つのウィンドウで全ての情報を提示する場合もあれば、例えば初期検索結果を第1のウィンドウで、拡張検索結果を第2のウィンドウで分けて提示するようにしても良い。さらに、拡張キーワード毎にウィンドウを分けても良い。このようにすれば、それぞれについてより多くの検索結果を一度に表示させることができるようになる。

[0034] 以上述べたように、本実施の形態のような処理を実施すれば、検索者が入力した入力検索キーワードだけでは通常検索者に提示されないような内容を、拡張検索結果としてユーザに提示できるようになるので、検索者に新たな気付きや視点を与えることができるようになる。特に、入力検索キーワードだけでは検索結果中下位にランキングされてしまって埋もれてしまう文書を浮き上がらせることができるようになる。

[0035] 以上本技術の一実施の形態を説明したが、本技術はこれに限定されるものではない。例えば、図3の検索処理サーバ5の機能ブロック図は一例であって、必ずしも実際のプログラムモジュール構成と一致しない場合もある。さらに、図3で示したネットワーク接続関係についても一例であって、例えば検索エンジン7を、検索処理サーバ5のバックグラウンドで動作するサーバとして、ユーザ端末3から見て検索処理サーバ5の背後にネットワーク接続させても良い。さらに、サーバ・クライアント構成の例を示し

たが、ユーザ端末3と検索処理サーバ5とが一体となったスタンドアロン型構成であっても良い。さらに検索エンジン7も一体化される場合もある。

[0036] さらに、処理フローについても処理結果が変わらない限りにおいてステップの順番を入れ替えたり、並列実行するようにしても良い。

[0037] なお、上で述べた例では入力検索キーワードが1つの場合を説明したが、2つ以上であっても基本的には同じ処理で対処可能である。すなわち、複数の入力検索キーワードのそれぞれについて拡張キーワード候補を抽出して、上で述べた処理を実施すればよい。

[0038] さらに、画面構成例を図示したが、画面構成例も同様の情報を提示できれば、他の画面構成を採用することも可能である。

[0039] なお、上で述べた検索処理サーバ5や検索エンジン7、ユーザ端末3は、コンピュータ装置であって、図17に示すように、メモリ2501とCPU2503とハードディスク・ドライブ(HDD)2505と表示装置2509に接続される表示制御部2507とリムーバブル・ディスク2511用のドライブ装置2513と入力装置2515とネットワークに接続するための通信制御部2517とがバス2519で接続されている。オペレーティング・システム(OS: Operating System)及び本実施例における処理を実施するためのアプリケーション・プログラムは、HDD2505に格納されており、CPU2503により実行される際にはHDD2505からメモリ2501に読み出される。必要に応じてCPU2503は、表示制御部2507、通信制御部2517、ドライブ装置2513を制御して、必要な動作を行わせる。また、処理途中のデータについては、メモリ2501に格納され、必要があればHDD2505に格納される。本技術の実施例では、上で述べた処理を実施するためのアプリケーション・プログラムはコンピュータ読み取り可能なリムーバブル・ディスク2511に格納されて頒布され、ドライブ装置2513からHDD2505にインストールされる。インターネットなどのネットワーク及び通信制御部2517を経由して、HDD2505にインストールされる場合もある。このようなコンピュータ装置は、上で述べたCPU2503、メモリ2501などのハードウェアとOS及び必要なアプリケーション・プログラムとが有機的に協働することにより、上で述べたような各種機能を実現する。

[0040] 以上本実施の形態をまとめると以下のようなようになる。

[0041] 本検索処理方法は、検索キーワードを受信するステップと、受信した検索キーワードで、文書に関するデータを格納するデータベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む初期検索結果を取得し、初期検索結果格納部に格納するステップと、各キーワードに対応付けて拡張検索キーワードを格納する拡張検索キーワード格納部から、受信した検索キーワードに対応付けられている拡張検索キーワードを抽出するステップと、抽出された各拡張検索キーワードで初期検索結果格納部を検索して、抽出された各拡張検索キーワードについて該当文書件数又は出現頻度を計数し、各拡張検索キーワードに対応付けて該当文書件数又は出現頻度を集計結果格納部に格納するステップと、集計結果格納部に格納されている拡張検索キーワードのうち該当文書件数又は出現頻度が少ない順で上位所定数の拡張検索キーワードの各々又は該当文書件数又は出現頻度が所定値以下の拡張検索キーワードの各々と受信した検索キーワードとの組み合わせで、データベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む拡張検索結果を取得し、拡張検索結果格納部に格納するステップと、初期検索結果格納部に格納されている初期検索結果の少なくとも一部と、拡張検索結果格納部に格納されている拡張検索結果の少なくとも一部とを出力する出力ステップとを含む。

[0042] このように入力検索キーワードに関連する拡張検索キーワードのうち入力検索キーワードによる検索結果(すなわち初期検索結果)にあまり出てこない(又は全く出てこない)ような拡張検索キーワードを、拡張検索に用いることによって、入力検索キーワードのみによる検索結果とは異なる内容や観点の検索結果を検索者に提示できるようになる。

[0043] また、上で述べた出力ステップにおいて、初期検索結果の少なくとも一部と、拡張検索結果の少なくとも一部とを異なる画面領域で表示するためのデータを生成するようにしてもよい。上下又は左右に並べて見せることによって、検索者は異なる検索結果を対比して把握することができ、より目的に近い文書を見出す、又は検索の新たな方向を見出すことができるようになる。なお、同じウィンドウにおいて異なる領域を分けるようにしても良いが、例えば初期検索結果をメインウィンドウに表示させ、拡張検

索結果を1又は複数のサブウィンドウに表示させるようにしてもよい。

[0044] さらに、上で述べた所定値を「0」に設定するようにしても良い。入力検索キーワードによる初期検索結果では全く提示できないような異なる内容の検索結果を提示できるようになる。なお、初期検索結果に含まれる範囲は関連度が上位所定件数分のみであり、その範囲に拡張検索キーワードが含まれていないだけであって、上位所定件数より下位の文書においては拡張検索キーワードは含まれている。

[0045] なお、上で述べたような処理をハードウェアに実施させるためのプログラムを作成することができ、当該プログラムは、例えばフレキシブル・ディスク、CD-ROM、光磁気ディスク、半導体メモリ、ハードディスク等のコンピュータ読み取り可能な記憶媒体又は記憶装置に格納される。なお、処理途中のデータについては、コンピュータのメモリ等の記憶装置に一時保管される。

請求の範囲

- [1] 検索キーワードを受信するステップと、
受信した前記検索キーワードで、文書に関するデータを格納するデータベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む初期検索結果を取得し、初期検索結果格納部に格納するステップと、
、
各キーワードに対応付けて拡張検索キーワードを格納する拡張検索キーワード格納部から、受信した前記検索キーワードに対応付けられている拡張検索キーワードを抽出するステップと、
抽出された各前記拡張検索キーワードで前記初期検索結果格納部を検索して、抽出された各前記拡張検索キーワードについて該当文書件数又は出現頻度を計数し、各前記拡張検索キーワードに対応付けて前記該当文書件数又は出現頻度を集計結果格納部に格納するステップと、
前記集計結果格納部に格納されている前記拡張検索キーワードのうち前記該当文書件数又は出現頻度が少ない順で上位所定数の前記拡張検索キーワードの各々又は前記該当文書件数又は出現頻度が所定値以下の前記拡張検索キーワードの各々と受信した前記検索キーワードとの組み合わせで、前記データベースを前記検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む拡張検索結果を取得し、拡張検索結果格納部に格納するステップと、
前記初期検索結果格納部に格納されている前記初期検索結果の少なくとも一部と、前記拡張検索結果格納部に格納されている前記拡張検索結果の少なくとも一部とを出力する出力ステップと、
を、コンピュータに実行させるための検索処理プログラムを格納するコンピュータ読み取り可能な記録媒体。
- [2] 前記出力ステップにおいて、
前記初期検索結果の少なくとも一部と、前記拡張検索結果の少なくとも一部とを異なる画面領域で表示するためのデータを生成する
請求項1記載のコンピュータ読み取り可能な記録媒体。

- [3] 前記所定値が0である請求項1又は2記載のコンピュータ読み取り可能な記録媒体。
- [4] 検索キーワードを受信するステップと、
受信した前記検索キーワードで、文書に関するデータを格納するデータベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む初期検索結果を取得し、初期検索結果格納部に格納するステップと、
、
各キーワードに対応付けて拡張検索キーワードを格納する拡張検索キーワード格納部から、受信した前記検索キーワードに対応付けられている拡張検索キーワードを抽出するステップと、
抽出された各前記拡張検索キーワードで前記初期検索結果格納部を検索して、抽出された各前記拡張検索キーワードについて該当文書件数又は出現頻度を計数し、各前記拡張検索キーワードに対応付けて前記該当文書件数又は出現頻度を集計結果格納部に格納するステップと、
前記集計結果格納部に格納されている前記拡張検索キーワードのうち前記該当文書件数又は出現頻度が少ない順で上位所定数の前記拡張検索キーワードの各々又は前記該当文書件数又は出現頻度が所定値以下の前記拡張検索キーワードの各々と受信した前記検索キーワードとの組み合わせで、前記データベースを前記検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む拡張検索結果を取得し、拡張検索結果格納部に格納するステップと、
前記初期検索結果格納部に格納されている前記初期検索結果の少なくとも一部と、前記拡張検索結果格納部に格納されている前記拡張検索結果の少なくとも一部とを出力する出力ステップと、
を含み、コンピュータに実行される検索処理方法。
- [5] 前記出力ステップにおいて、
前記初期検索結果の少なくとも一部と、前記拡張検索結果の少なくとも一部とを異なる画面領域で表示するためのデータを生成する
請求項4記載の検索処理方法。

[6] 検索キーワードを受信し、受信した前記検索キーワードで、文書に関するデータを格納するデータベースを検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む初期検索結果を取得し、初期検索結果格納部に格納する手段と、

各キーワードに対応付けて拡張検索キーワードを格納する拡張検索キーワード格納部から、受信した前記検索キーワードに対応付けられている拡張検索キーワードを抽出する手段と、

抽出された各前記拡張検索キーワードで前記初期検索結果格納部を検索して、抽出された各前記拡張検索キーワードについて該当文書件数又は出現頻度を計数し、各前記拡張検索キーワードに対応付けて前記該当文書件数又は出現頻度を集計結果格納部に格納する手段と、

前記集計結果格納部に格納されている前記拡張検索キーワードのうち前記該当文書件数又は出現頻度が少ない順で上位所定数の前記拡張検索キーワードの各々又は前記該当文書件数又は出現頻度が所定値以下の前記拡張検索キーワードの各々と受信した前記検索キーワードとの組み合わせで、前記データベースを前記検索エンジンに検索させ、当該検索エンジンから該当文書の少なくとも一部のテキストデータを含む拡張検索結果を取得し、拡張検索結果格納部に格納する手段と、

前記初期検索結果格納部に格納されている前記初期検索結果の少なくとも一部と、前記拡張検索結果格納部に格納されている前記拡張検索結果の少なくとも一部とを出力する出力手段と、

を有する検索処理装置。

[7] 前記出力手段が、

前記初期検索結果の少なくとも一部と、前記拡張検索結果の少なくとも一部とを異なる画面領域で表示するためのデータを生成する

請求項6記載の検索処理装置。

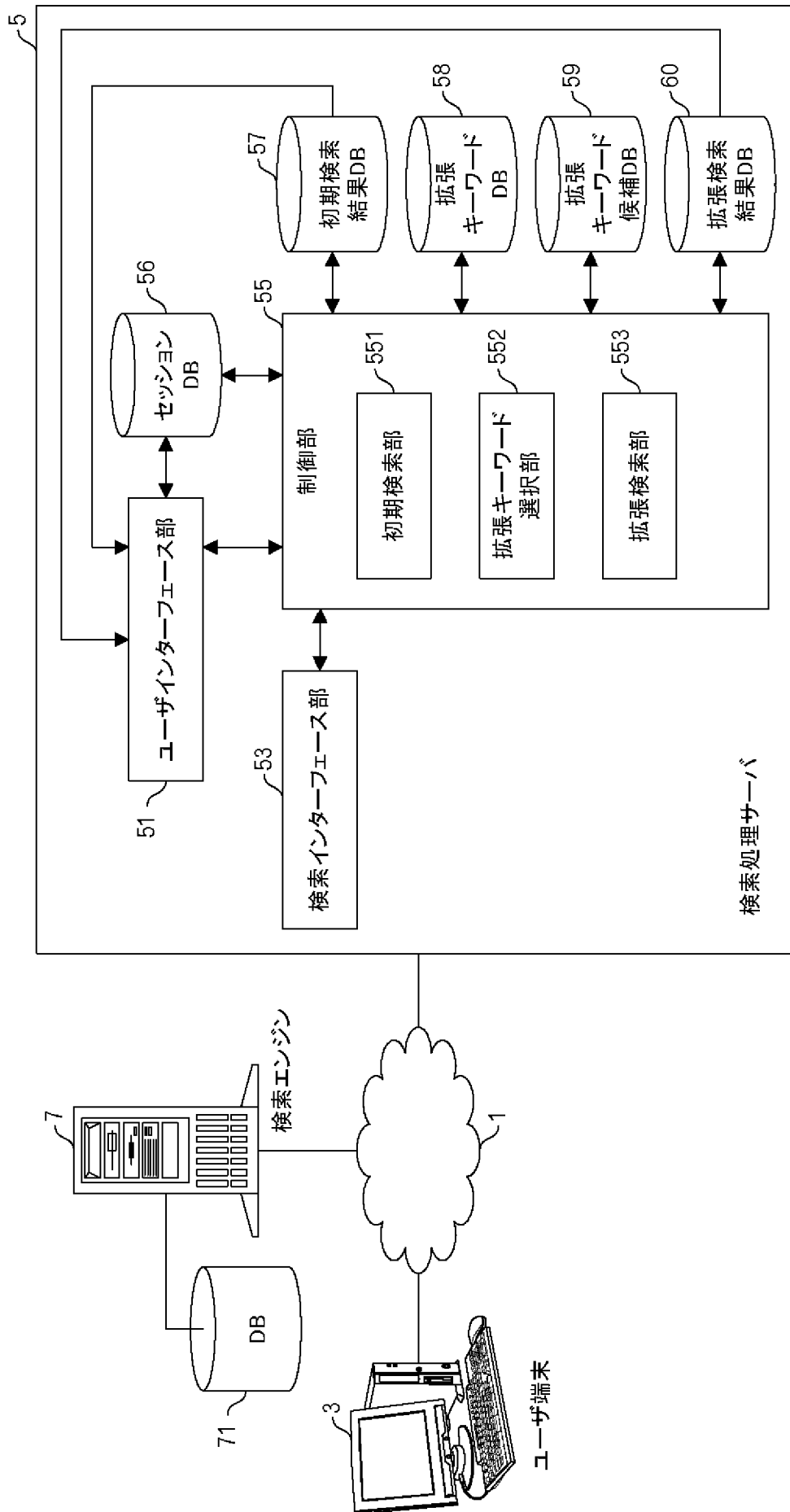
[図1]

検索結果ID	検索結果(リンゴ)
1	リンゴジュースの魅力
2	リンゴジュースとリンゴジャムの不思議
⋮	⋮
1000	リンゴジュースの歴史
1001	リンゴパイの歴史
⋮	⋮

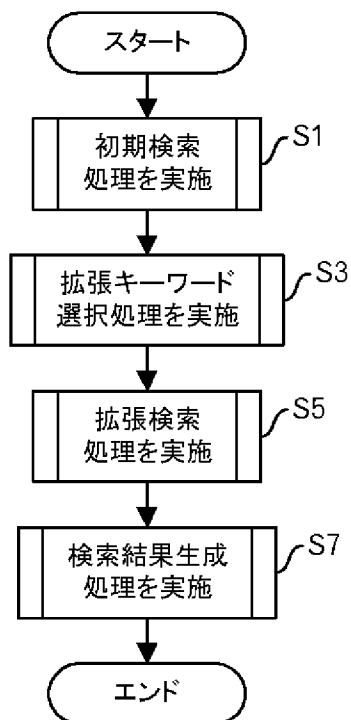
[図2]

検索結果ID	検索結果(リンゴ & ジュース)
1	リンゴジュースの魅力
2	リンゴジュースとリンゴジャムの不思議
⋮	⋮
1000	リンゴジュースの歴史
1001	リンゴジュースの歴史
⋮	⋮

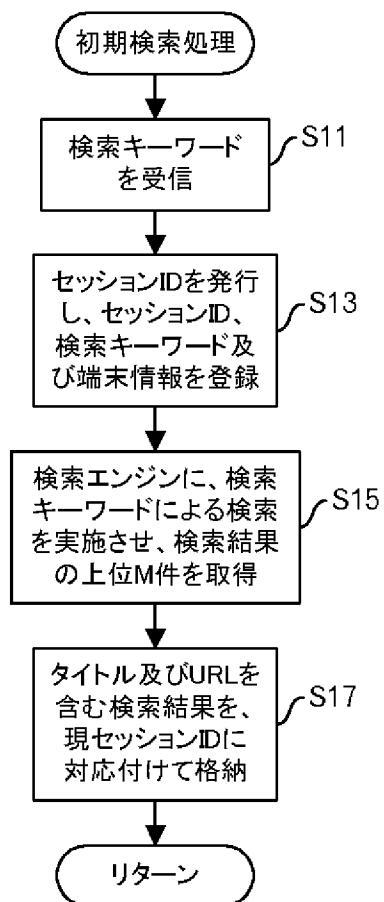
[図3]



[図4]



[図5]



[図6]

Figure 6 shows a search interface. It includes a label "検索キーワード:" (Search keyword:) and a text input field containing the character "リンゴ" (Apple). To the right of the input field is a button labeled "検索" (Search). A bracket labeled "601" points to the input field, and a bracket labeled "602" points to the search button.

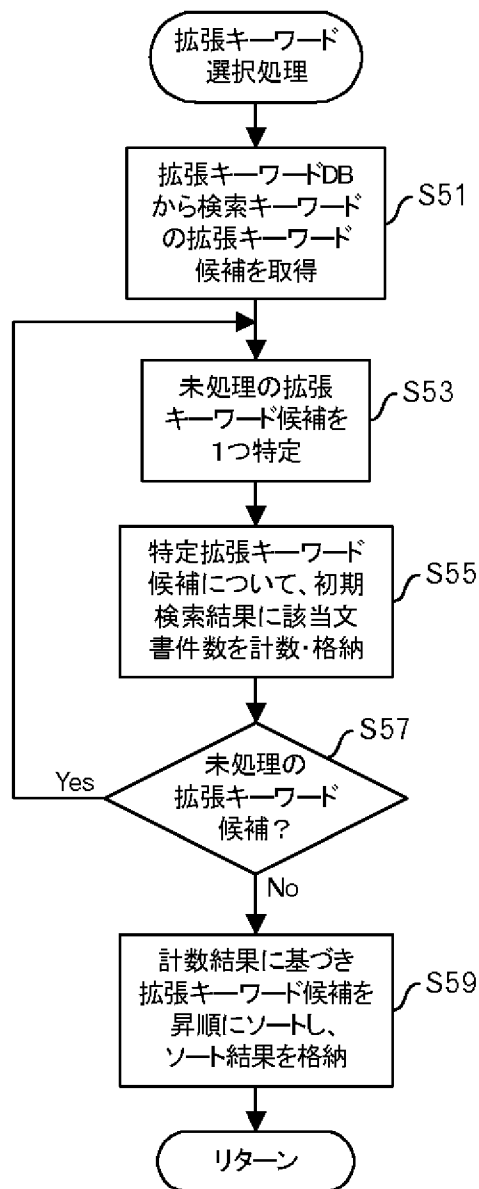
[図7]

セッションID	キーワード	端末ID
1	ミカン	端末1
2	リンゴ	端末2
⋮	⋮	⋮

[図8]

検索結果ID	タイトル	URL
1	リンゴジュースの魅力	apple-juice.jp
2	リンゴジュースとリンゴジャムの不思議	apple-juice-and-jelly.jp
3	リンゴジュースの歴史	apple-juice-history.jp
⋮	⋮	⋮
100	リンゴパイの歴史	apple-pie-history.jp

[図9]



[図10]

入力キーワード	拡張キーワード候補
リンゴ	ジュース, ジャム, あめ, パイ
パスタ	ペペロンチーノ, ミートソース
⋮	⋮

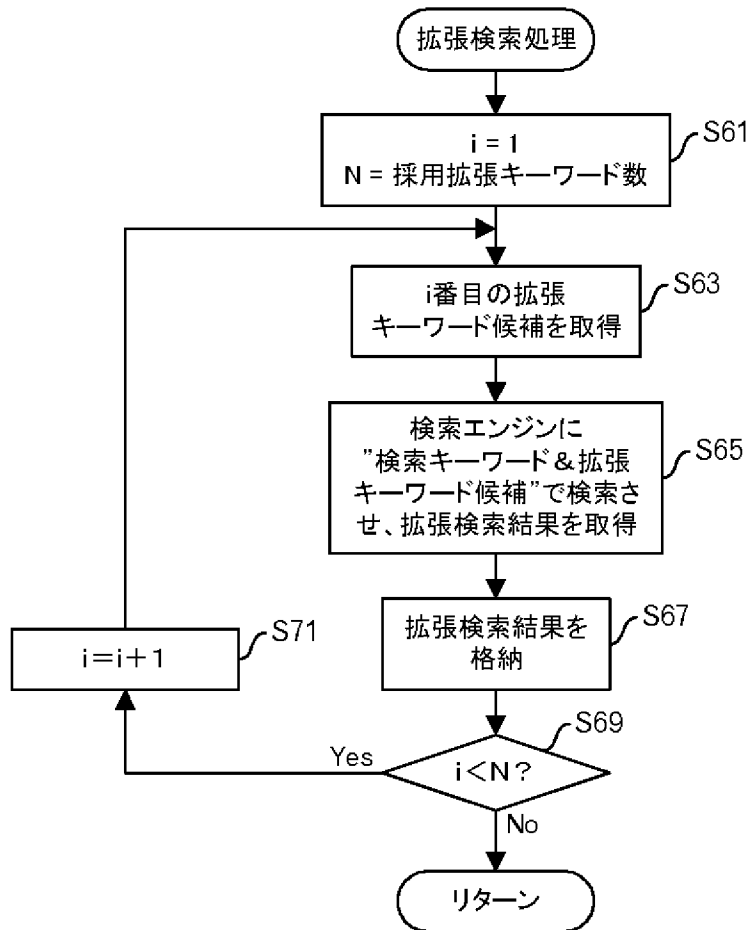
[図11]

拡張キーワード候補	件数
ジュース	55
ジャム	44
パイ	1
あめ	0

[図12]

ソート結果
あめ
パイ
ジャム
ジュース

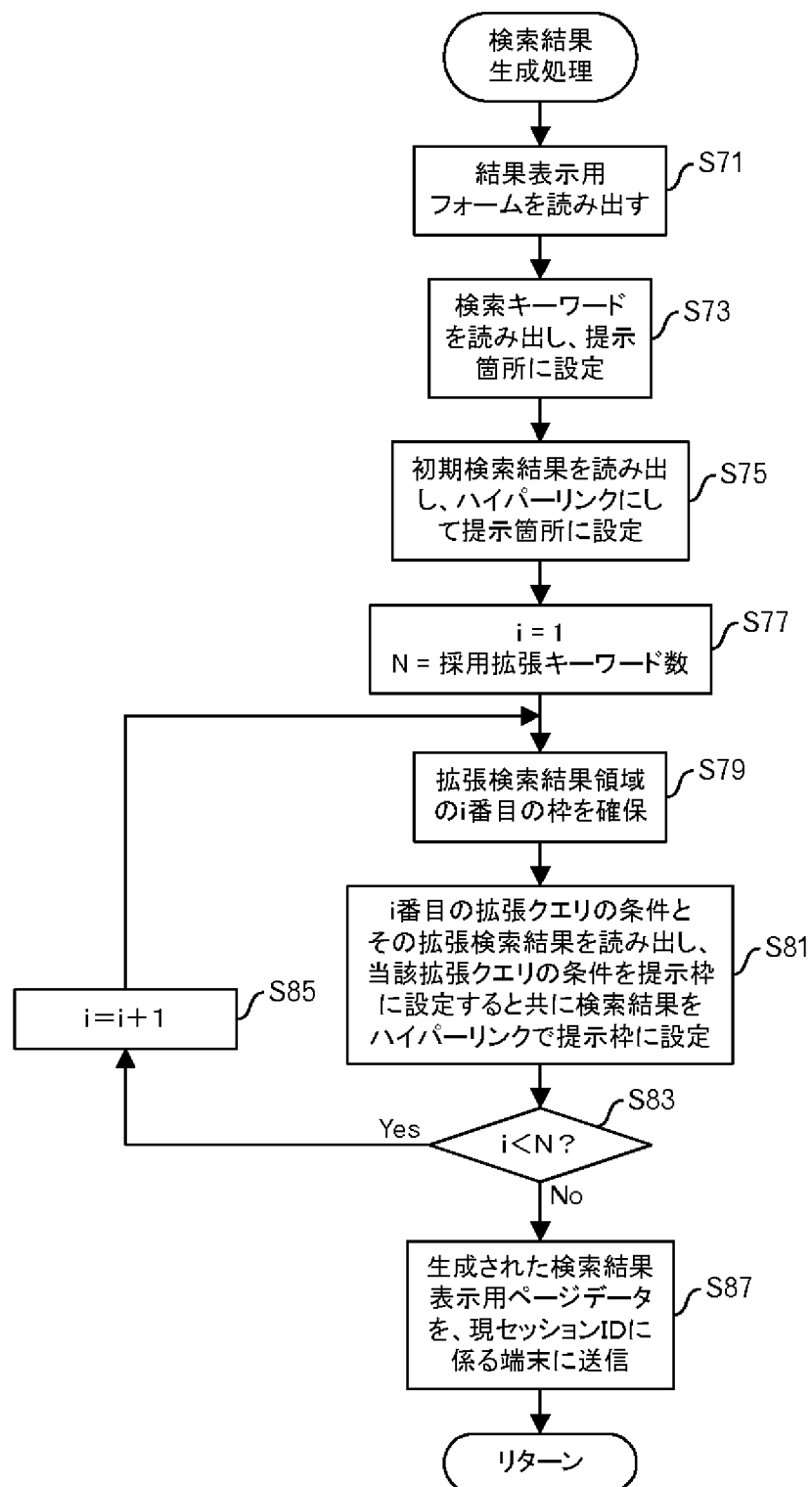
[図13]



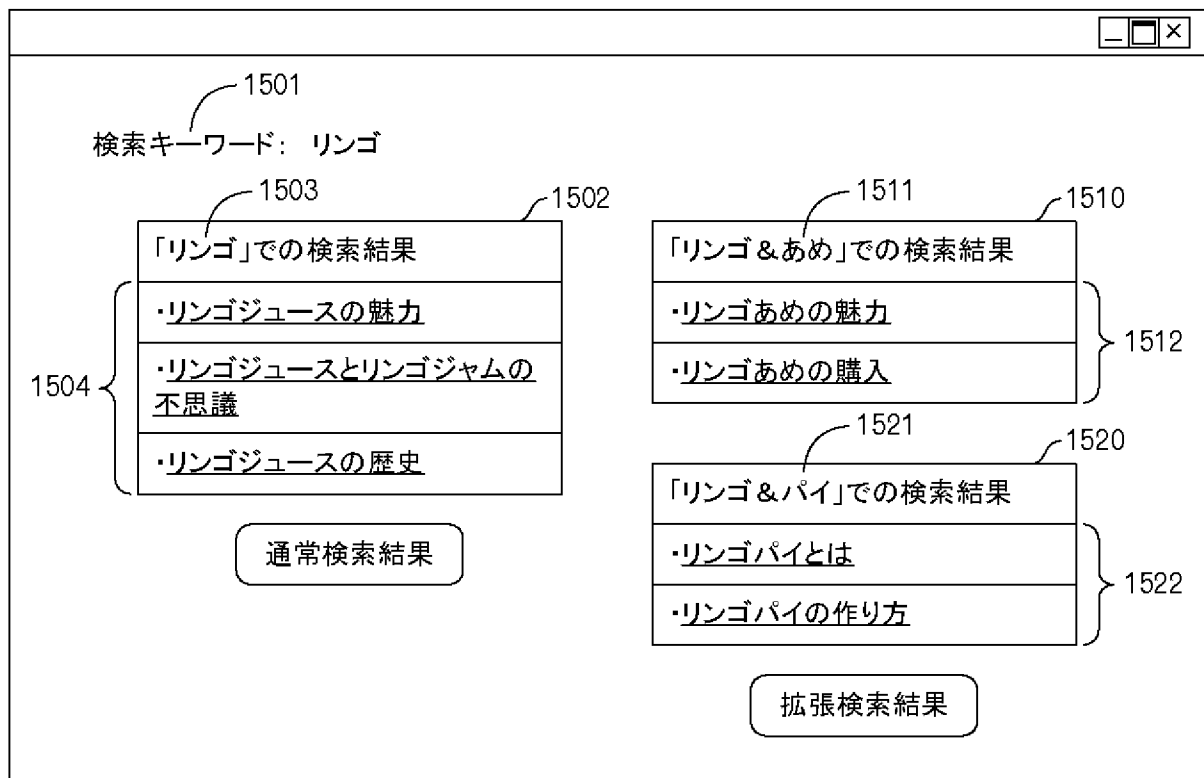
[図14]

拡張検索結果ID	拡張クエリ	タイトル	URL
1	リンゴ & あめ	リンゴあめの魅力	apple-candy.jp
		リンゴあめの購入	buy-apple-candy.jp
		⋮	⋮
2	リンゴ & パイ	リンゴパイとは	apple-pie.jp
		リンゴパイの作り方	cook-apple-pie.jp
		⋮	⋮

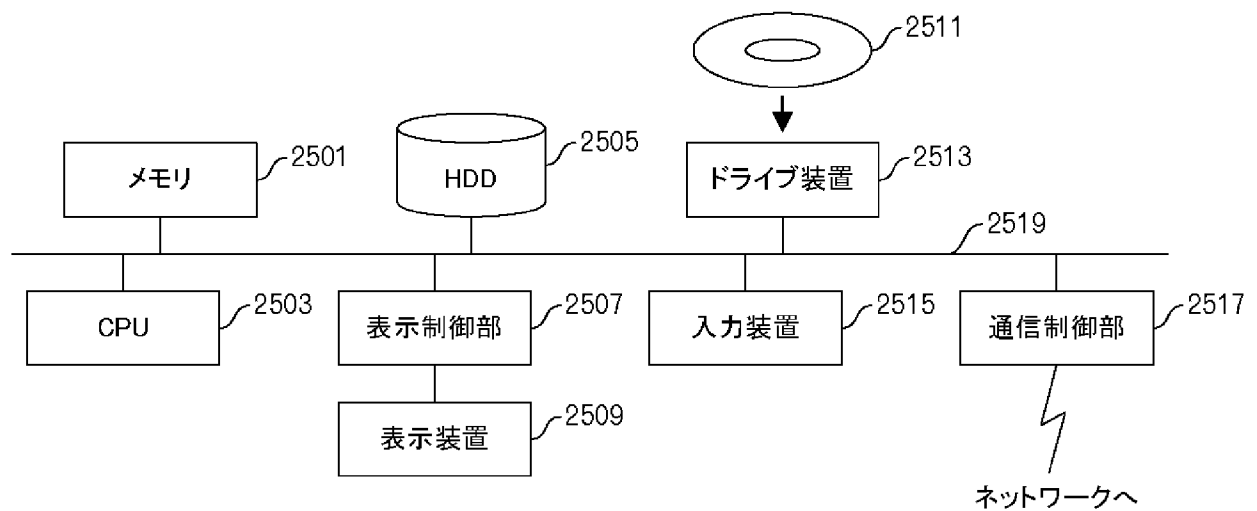
[図15]



[図16]



[図17]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/055177

A. CLASSIFICATION OF SUBJECT MATTER

G06F17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2009
Kokai Jitsuyo Shinan Koho	1971-2009	Toroku Jitsuyo Shinan Koho	1994-2009

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-164662 A (Justsystem Corp.), 10 June, 2004 (10.06.04), Par. Nos. [0064] to [0078]; Fig. 3 (Family: none)	1-7
A	JP 2001-75981 A (NTT Data Corp.), 23 March, 2001 (23.03.01), Full text; all drawings (Family: none)	1-7
A	JP 2007-133769 A (Ricoh Co., Ltd.), 31 May, 2007 (31.05.07), Par. Nos. [0023] to [0032]; Figs. 2 to 4 (Family: none)	1-7

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
09 April, 2009 (09.04.09)

Date of mailing of the international search report
21 April, 2009 (21.04.09)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/055177

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-110664 A (Ricoh Co., Ltd.), 08 April, 2004 (08.04.04), Par. Nos. [0004] to [0007]; Fig. 17 (Family: none)	1-7
A	JP 2004-118262 A (Toshiba Corp.), 15 April, 2004 (15.04.04), Full text; all drawings (Family: none)	1-7

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06F17/30(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06F17/30

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 0 9 年
日本国実用新案登録公報	1 9 9 6 - 2 0 0 9 年
日本国登録実用新案公報	1 9 9 4 - 2 0 0 9 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2004-164662 A（株式会社ジャストシステム）2004.06.10, 段落【0064】－【0078】，第3図（ファミリーなし）	1-7
A	JP 2001-75981 A（株式会社エヌ・ティ・ティ・データ）2001.03.23, 全文，全図（ファミリーなし）	1-7
A	JP 2007-133769 A（株式会社リコー）2007.05.31, 段落【0023】－【0032】，第2－4図（ファミリーなし）	1-7

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

0 9 . 0 4 . 2 0 0 9

国際調査報告の発送日

2 1 . 0 4 . 2 0 0 9

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

吉田 誠

5 M

3 6 5 9

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2004-110664 A (株式会社リコー) 2004. 04. 08, 段落【0004】 －【0007】, 第17図 (ファミリーなし)	1-7
A	JP 2004-118262 A (株式会社東芝) 2004. 04. 15, 全文, 全図 (ファミ リリーなし)	1-7