

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5060020号
(P5060020)

(45) 発行日 平成24年10月31日(2012.10.31)

(24) 登録日 平成24年8月10日(2012.8.10)

(51) Int.Cl.

F 1

G 0 6 F 17/30 (2006.01)

G 0 6 F 17/30 2 1 0 D
G 0 6 F 17/30 2 1 0 C
G 0 6 F 17/30 3 2 0 B
G 0 6 F 17/30 3 4 0 B

請求項の数 4 (全 14 頁)

| | | | |
|-----------|-------------------------------|-----------|---|
| (21) 出願番号 | 特願2004-317285 (P2004-317285) | (73) 特許権者 | 392026693 株式会社エヌ・ティ・ティ・ドコモ |
| (22) 出願日 | 平成16年10月29日(2004.10.29) | | 東京都千代田区永田町二丁目11番1号 |
| (65) 公開番号 | 特開2006-127325 (P2006-127325A) | (74) 代理人 | 100066980 弁理士 森 哲也 |
| (43) 公開日 | 平成18年5月18日(2006.5.18) | | |
| 審査請求日 | 平成19年10月3日(2007.10.3) | (74) 代理人 | 100109380 弁理士 小西 恵 |
| 審判番号 | 不服2010-18713 (P2010-18713/J1) | | |
| 審判請求日 | 平成22年8月19日(2010.8.19) | (74) 代理人 | 100103850 弁理士 田中 秀▲てつ▼ |
| | | (74) 代理人 | 100112863 弁理士 阪間 和之 |
| | | (72) 発明者 | 田中 希世子 東京都千代田区永田町二丁目11番1号 株式会社エヌ・ティ・ティ・ドコモ内 |

最終頁に続く

(54) 【発明の名称】 コンテンツ発見装置

(57) 【特許請求の範囲】

【請求項1】

検索キーワードによるコンテンツの検索を行う検索手段と、
前記検索手段による検索結果と、過去の検索結果とをグループ分け対象コンテンツとして選択するグループ分対象選択手段と、

前記グループ分対象選択手段により選択された検索結果を類似グループにグループ分けするグループ分手段と、

前記グループ分手段によるグループ分けによって生成された類似グループから当該類似グループを特徴づけるキーワードを抽出しこれを関連キーワードとして前記類似グループに対応付ける関連キーワード対応付手段と、

前記関連キーワード対応付手段により対応付けられた関連キーワードを提示する提示手段と

を備え、

前記検索手段は、前記提示手段により提示された関連キーワードが選択された場合に、該選択された関連キーワードを新たな前記検索キーワードとしてコンテンツの再検索を行うことを特徴とするコンテンツ発見装置。

【請求項2】

検索に用いられた検索キーワードの履歴を記憶するキーワード履歴記憶手段をさらに備え、

前記グループ分対象選択手段は、前記キーワード履歴記憶手段に記憶されている検索キ

ーワードによる検索結果をグループ分け対象コンテンツとして選択することを特徴とする請求項1に記載のコンテンツ発見装置。

【請求項3】

前記キーワード履歴記憶手段から抽出すべき検索キーワードの数を記憶する検索キーワード数記憶手段をさらに備え、

前記グループ分対象選択手段は、

前記検索キーワード数記憶手段に記憶されている数だけの検索キーワードを前記キーワード履歴記憶手段から抽出し、該抽出された検索キーワードによる検索結果をグループ分け対象コンテンツとして選択することを特徴とする

請求項2に記載のコンテンツ発見装置。

10

【請求項4】

検索キーワードによる検索結果のうち、グループ分け対象コンテンツとして選択すべき検索結果の数を記憶する検索結果数記憶手段をさらに備え、

前記グループ分対象選択手段は、

前記検索キーワードによる検索結果のうち、前記検索結果数記憶手段に記憶されている数だけの検索結果をグループ分け対象コンテンツとして選択することを特徴とする

請求項1から3の何れか1項に記載のコンテンツ発見装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、キーワードを用いてコンテンツを発見するコンテンツ発見装置に関する。

20

【背景技術】

【0002】

従来、インターネット上のコンテンツの中からユーザが所望のコンテンツを検索するために、種々の検索エンジンが提供されている。検索エンジンには、メタ検索エンジンやディレクトリ型検索エンジン等が存在する。

メタ検索エンジンとは、複数の検索エンジンに対して横断的に検索を行ったうえ、個々の検索エンジンが表示する検索結果をまとめて表示する検索エンジンであり、独自の検索用の情報データベースは保持しておらず、他の複数の検索エンジンのデータベースを利用して検索結果の表示を行う。

30

【0003】

統合型のメタ検索エンジンでは、各検索エンジンから取得した検索結果に対して、メタ検索エンジンが独自の手法によって解析及び情報の分類を行い、検索キーワードと個々の検索結果であるWebページとの関連性を考慮して整理された結果を表示する。解析の手法は検索エンジンごとに特徴があるが、その目的とすること、及び、最終的な表示内容は同一である。すなわち、キーワードによって検索結果を絞り込むことを目的としており、キーワードによって絞り込まれた検索結果が表示される。

【0004】

ディレクトリ型検索エンジンとは、検索用の情報データベースに情報を登録する際に、登録者が情報の内容を判断し、それに応じてあらかじめカテゴリー分けを行っているものである。大分類、中分類、小分類というように情報が階層化されているのが特徴である。ユーザはカテゴリーをたどって深い階層に進んでいくことで検索結果を絞り込むことができる。

40

【0005】

従来の検索エンジンを用いたコンテンツの検索手順について、図5を参照しながら具体的に説明する。インターネット上の所望のコンテンツを検索する場合、まず、ユーザはインターネットに接続された端末装置を操作して検索エンジンを起動し、ディスプレイにキーワード入力画面101を表示する(図5(a))。そして、ユーザが、キーワード入力画面101に検索キーワードとしてキーワード1を入力すると、検索エンジンは、キーワード1にマッチする(すなわち、コンテンツの中にキーワード1が含まれている)コンテ

50

ンツ全体 103 をコンテンツの内容やリンク数等をもとに検索エンジン独自の手法でグループ分けし、コンテンツのグループ 103 - 1, 103 - 2, 103 - 3 を生成する。検索エンジンは、生成されたグループ 103 - 1, 103 - 2, 103 - 3 毎に、グループを特徴付けるキーワード（以下「関連キーワード」と呼ぶ）として、キーワード 2, 2', 2'' をつける（図 5 (b)）。

【0006】

そして、端末装置のディスプレイには、図 5 (c) に示すコンテンツ表示画面 102 が表示される。コンテンツ表示画面 102 は、キーワード表示部 102 - 1 とコンテンツ表示部 102 - 2 とから構成され、キーワード表示部 102 - 1 には検索キーワードであるキーワード 1 と、グループ毎の関連キーワードであるキーワード 2, 2', 2'' とが表示される。また、コンテンツ表示部 102 - 2 には、検索キーワードであるキーワード 1 にマッチする全コンテンツが表示される。

10

【0007】

次に、ユーザが、キーワード表示部 102 - 1 に表示されている関連キーワードのうちキーワード 2 を選択すると、検索エンジンは、キーワード 2 で特徴付けられる全コンテンツ 103 - 1 を、キーワード 1 の検索結果に対して行ったのと同様に、コンテンツの内容やリンク数等をもとに検索エンジン独自の手法でグループ分けし、コンテンツのグループ 105 - 1, 105 - 2 を生成する。そして、検索エンジンは、グループ 105 - 1, 105 - 2 毎に各グループを特徴付けるキーワードとしてキーワード 3, 3' をつける（図 5 (d)）。

20

【0008】

以上の処理に基づいて、端末装置のディスプレイには、図 5 (e) に示すコンテンツ表示画面 102 が表示される。同図に示すように、コンテンツ表示画面 102 のキーワード表示部 102 - 1 には、図 5 (c) に示すキーワード表示部 102 - 1 で表示されていたキーワード群に加えて、キーワード 2 の配下に、キーワード 2 で特徴付けられるコンテンツのグループ分けによって得られたキーワード群（キーワード 3, 3'）が表示される。また、コンテンツ表示部 102 - 2 には、キーワード 1 にマッチしたコンテンツの中で、キーワード 2 の特徴をもつグループに含まれる全コンテンツ 15 - 1 の一覧が表示される。

このような処理を繰り返すことで、コンテンツ表示画面 102 中のキーワード表示部 102 - 1 に表示されるキーワード群の階層が深くなり、コンテンツ表示部 102 - 2 に表示されるコンテンツの数は少なくなり、検索結果の絞込みが行われる。

30

【0009】

図 6 は、上述した検索エンジンの機能を有する検索装置の構成の一例を示した図である。同図に示すように、検索装置は、ユーザからのキーワードの入力や検索結果の表示等の入出力を管理する入出力部、入力されたキーワードをもとにネットワークへアクセスしてコンテンツを検索する Web 検索部、取得したコンテンツの URL (Uniform Resource Locator)、コンテンツの実体、及び、概要等の、コンテンツに関する情報を管理するコンテンツ管理部、及び、コンテンツの内容などに応じて類似コンテンツのグループ分けを行うグループ管理部から構成される。

40

【0010】

図 7 は、図 6 に示す検索装置が行う検索処理手順の一例を示すフロー図である。同図を参照しながら、従来の検索装置が行う検索処理手順を説明する。まず、入出力部がユーザから検索キーワードを取得すると（ステップ S301）、Web 検索部にて、検索キーワードをもとにインターネット等のネットワークを介してコンテンツを取得する（ステップ S302）。コンテンツ管理部では、Web 検索部が取得したコンテンツの URL や、コンテンツの実体データや、コンテンツの内容から取得される特徴情報等のコンテンツ情報を格納する（ステップ S303）。グループ管理部では、コンテンツ管理部によって格納されたコンテンツ情報をもとに、コンテンツ群を複数のグループに分け、グループを特徴付けるキーワードを抽出する（ステップ S304）。入出力部では、ユーザに対してグル

50

ープ分けされた検索結果を表示する（ステップS305）。入出力部がユーザのグループ選択の操作を受け付けると（ステップS306；yes）、グループ管理部では選択済みグループに該当するコンテンツをグループ分け対象コンテンツとして設定し（ステップS307）、再度グループ分けを行い（ステップS304）、そのグループ分けされた結果を入出力部がユーザへ表示する（ステップS305）。また、入出力部がユーザのコンテンツ選択の操作を受け付けると（ステップS308；yes）、コンテンツ管理部は管理しているコンテンツ情報の中から選択済みコンテンツに関する情報を抽出し、入出力部が当該コンテンツに関する情報をユーザへ提示する（ステップS309）。

【0011】

このような、検索キーワードによる検索結果からキーワードを抽出したり、検索結果をグループ分けするといった技術内容は、例えば、特許文献1、2に開示されている。

【特許文献1】特開平09-259133号公報

【特許文献2】特開2000-76279号公報

【発明の開示】

【発明が解決しようとする課題】

【0012】

統合型のメタ検索エンジンやディレクトリ型検索エンジンでは、情報群に対してカテゴリー分けを繰り返し行っていくことで情報を絞りこみ、ユーザが検索の目的としている「欲しい情報」に近づくことを目的としている。しかしこのような絞りこみ型の検索エンジンでは、最初の検索キーワードによって取得可能な情報がある程度限定され、以降の検索処理は最初に限られた情報の中での探索となる。このため、ユーザの検索の目的がはっきりしていない場合や、目的や欲しい情報を象徴している検索キーワードの設定が適切でなかった場合には、所望の情報を取得することは難しくなる。また、ある程度一定の興味や目的に応じてユーザにより入力される検索キーワードでは、情報の検索範囲や方向性がある程度決まってしまうため、ユーザが取得可能な情報がマンネリ化する恐れがあり、検索を行いながら新しい興味の発見や気づきにつながるような情報を取得することは難しい。

本発明は、上記問題点に鑑みてなされたものであり、検索キーワードに限定されない多様なコンテンツを発見することが可能なコンテンツ発見装置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】

【0013】

上記課題を解決するために、請求項1に記載の発明は、検索キーワードによるコンテンツの検索を行う検索手段と、前記検索手段による検索結果と、過去の検索結果とをグループ分け対象コンテンツとして選択するグループ分対象選択手段と、前記グループ分対象選択手段により選択された検索結果を類似グループにグループ分けするグループ分手段と、前記グループ分手段によるグループ分けによって生成された類似グループから当該類似グループを特徴づけるキーワードを抽出しこれを関連キーワードとして前記類似グループに対応付ける関連キーワード対応付手段と、前記関連キーワード対応付手段により対応付けられた関連キーワードを提示する提示手段とを備え、前記検索手段は、前記提示手段により提示された関連キーワードが選択された場合に、該選択された関連キーワードを新たな前記検索キーワードとしてコンテンツの再検索を行うことを特徴とするコンテンツ発見装置を提供する。

【0014】

この構成によれば、コンテンツ発見装置は、検索が繰り返される毎に今までの検索結果をグループ分け対象コンテンツとして取り込んでグループ分けを行い、グループを特徴づける多様な関連キーワードをユーザに提示し、提示された関連キーワードが選択された場合に、選択された関連キーワードを検索キーワードとしてコンテンツの検索を繰り返すことができる。このため、ユーザは、検索キーワードに限定されない多様なコンテンツを発見することが可能となる。

【0015】

請求項 2 に記載の発明は、請求項 1 に記載のコンテンツ発見装置において、検索に用いられた検索キーワードの履歴を記憶するキーワード履歴記憶手段をさらに備え、前記グループ分対象選択手段は、前記キーワード履歴記憶手段に記憶されている検索キーワードによる検索結果をグループ分け対象コンテンツとして選択することを特徴とする。

この構成によれば、過去に検索に用いられた検索キーワードの履歴を記憶させておくことによって、過去の検索キーワードの履歴を正確に管理することができ、グループ分け対象コンテンツとして選択すべき検索結果を把握するのが容易になる。また、過去の検索キーワードの履歴をユーザに提示することができる。

【 0 0 1 6 】

請求項 3 に記載の発明は、請求項 2 に記載のコンテンツ発見装置において、前記キーワード履歴記憶手段から抽出すべき検索キーワードの数を記憶する検索キーワード数記憶手段をさらに備え、前記グループ分対象選択手段は、前記検索キーワード数記憶手段に記憶されている数だけの検索キーワードを前記キーワード履歴記憶手段から抽出し、該抽出された検索キーワードによる検索結果をグループ分け対象コンテンツとして選択することを特徴とする。

10

この構成によれば、検索キーワード数記憶手段を用いてキーワード履歴記憶手段から抽出される検索キーワードの数を制限することができ、グループ分け対象コンテンツの数が膨大となるのを防ぐことができる。また、適切な検索キーワードを優先的に抽出することができる。

【 0 0 1 7 】

20

請求項 4 に記載の発明は、請求項 1 から 3 の何れか 1 項に記載のコンテンツ発見装置において、検索キーワードによる検索結果のうち、グループ分け対象コンテンツとして選択すべき検索結果の数を記憶する検索結果数記憶手段をさらに備え、前記グループ分対象選択手段は、前記検索キーワードによる検索結果のうち、前記検索結果数記憶手段に記憶されている数だけの検索結果をグループ分け対象コンテンツとして選択することを特徴とする。

この構成によれば、グループ分け対象コンテンツの数を制限することができ、グループ分け対象コンテンツの数が膨大となり、処理負荷がかかるという不具合を解消することができる。

【 発明の効果 】

30

【 0 0 2 0 】

本発明によれば、コンテンツ発見装置は、検索が繰り返される毎に今までの検索結果をグループ分け対象コンテンツとして取り込んでグループ分けを行い、グループを特徴づける関連キーワードをユーザに提示し、提示された関連キーワードが選択された場合に、選択された関連キーワードを検索キーワードとしてコンテンツの検索を繰り返すことができる。このため、ユーザは、検索キーワードに限定されない多様なコンテンツを発見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 】

【 0 0 2 1 】

次に、図面を参照しながら、本発明を実施するための最良の形態について説明する。なお、以下の説明において参照する各図においては、他の図と同等部分に同一符号が付されている。

40

図 1 は、本発明の実施の形態に係るコンテンツ発見装置 10 の機能構成を例示する図である。同図に示すように、コンテンツ発見装置 10 は、ユーザからのキーワードの入力や検索結果の表示等の入出力を管理する入出力部 11 と、入力されたキーワードをもとにネットワークへアクセスしてコンテンツを検索する Web 検索部 12 と、取得したコンテンツの URL やコンテンツの実体、及び、コンテンツの内容に基づく特徴情報等のコンテンツに関する情報を管理するコンテンツ管理部 13 と、コンテンツの内容に応じて類似コンテンツのグループ分けを行い関連キーワードを抽出するグループ管理部 14 と、検索キーワードやその検索結果から抽出される関連キーワード、及び、ユーザが選択した関連キー

50

ワード等の、キーワードの履歴情報を管理するキーワード管理部 15 とを含んで構成される。ここで、コンテンツ発見装置 10 は、プログラムを実行し全体を制御する CPU (Central Processing Unit) と、プログラム及びデータを記憶する記憶装置と、入出力装置とを備えた一般的なコンピュータである。

【 0 0 2 2 】

次に、コンテンツ発見装置 10 の機能構成について詳細に説明する。Web 検索部 12 は、ネットワークを介して他の装置と通信を行うための通信インターフェース、CPU 及びプログラムで構成される。Web 検索部 12 は、ユーザが入力した検索キーワード、及び、ユーザが選択した関連キーワードについての情報を入出力部 11 から取得する。Web 検索部 12 は、入出力部 11 から取得した検索キーワードにマッチするコンテンツを、ネットワークを介して検索する。なお、Web 検索部 12 は、ユーザが選択した関連キーワードを入出力部 11 から取得した場合には、当該関連キーワードを検索キーワードとして扱い、検索を行う。

10

【 0 0 2 3 】

コンテンツ管理部 13 は、CPU、プログラム、及び、記憶装置を含んで構成される。コンテンツ管理部 13 は、Web 検索部 12 で得られたコンテンツの実体、インターネットにおけるコンテンツの所在を表す URL、及び、コンテンツの内容から抽出される特徴情報などを含むコンテンツ情報を、記憶装置に設けられたコンテンツ情報 DB (database ; データベース) 16 に保存する。

【 0 0 2 4 】

キーワード管理部 15 は、CPU、プログラム、及び、記憶装置で構成される。キーワード管理部 15 は、記憶装置に設けられたキーワード履歴記憶 DB 18 にキーワード履歴情報を保存する。当該キーワード履歴情報には、検索キーワードによる検索結果をグループ分けして生成されたグループ毎の関連キーワードと、検索キーワードの履歴 (ユーザが入力した検索キーワード、及び、ユーザがこれまでに選択した関連キーワード群) とが含まれる。

20

【 0 0 2 5 】

キーワード管理部 15 はグループ分対象選択機能 151 を備えている。グループ分対象選択機能 151 は、Web 検索部 12 による検索結果と、過去の検索結果とをグループ分け対象コンテンツとして選択する。具体的には、グループ分対象選択機能 151 は、まず、キーワード履歴記憶 DB 18 から検索に使用された検索キーワードの履歴を取得し、当該取得した検索キーワードによる検索結果をグループ分対象コンテンツとしてコンテンツ情報 DB 16 から抽出する。そして、グループ分対象選択機能 151 は、抽出した過去の検索キーワードによる検索結果と、入出力部 11 から今回入力された検索キーワードに基づいて Web 検索部 12 により検索された検索結果とをマージし、当該マージした検索結果を最終的なグループ分け対象コンテンツとしてグループ管理部 14 へ引き渡す。なお、入出力部 11 から今回入力された検索キーワードが既にキーワード履歴記憶 DB 18 に登録されており、当該今回入力された検索キーワードが過去の検索キーワードと共にキーワード履歴記憶 DB 18 から抽出される構成としている場合には、グループ分対象選択機能 151 は、キーワード履歴記憶 DB 18 から抽出した検索キーワードによる検索結果をそのまま最終的なグループ分け対象コンテンツとして選択し、グループ管理部 14 へ引き渡す。

30

40

【 0 0 2 6 】

ここで、グループ分け対象コンテンツの数を制御するためのデータベースとして、検索キーワード数記憶 DB 19 及び検索結果数記憶 DB 20 が記憶装置に設けられている。検索キーワード数記憶 DB 19 には、キーワード履歴記憶 DB 18 から抽出すべき検索キーワードの数が記憶されている。

グループ分対象選択機能 151 は、検索キーワード数記憶 DB 19 に記憶されている数だけの検索キーワードをキーワード履歴情報 DB 18 から抽出し、当該抽出された検索キーワードによる検索結果をグループ分け対象コンテンツとして選択する。

50

【 0 0 2 7 】

なお、キーワード履歴記憶DB18からグループ分け対象コンテンツを選択する場合のルールとしては、例えば、直近にキーワード履歴記憶DB18に登録された検索キーワードから優先的に抽出する。このようにすれば、相対的に古い検索キーワードを除外することができ、現在のユーザの興味に近いコンテンツを効率的に発見することができる。また、例えば、今回の検索キーワードに意味が近い過去の検索キーワードを優先的に抽出するようにしてもよい。

【 0 0 2 8 】

なお、検索キーワード数記憶DB19に記憶されている数だけの検索キーワードがキーワード履歴記憶DB18に記憶されていない場合には、グループ分対象選択機能151は

10

、キーワード履歴記憶DB18に記憶されている全ての検索キーワードを抽出する。
また、検索結果数記憶DB20には、検索キーワードによる検索結果のうち、グループ分け対象コンテンツとして選択すべき検索結果の数記憶されている。ここでの選択すべき検索結果の数は、各検索キーワード(過去の検索キーワード、今回の検索キーワード)による検索結果各々から選択すべき検索結果の数であってもよいし、検索結果がマージされた後の全体の検索結果から選択すべき検索結果の数であってもよい。グループ分対象選択機能151は、検索結果数記憶DB20を参照して、検索キーワードによる検索結果のうち、検索結果数記憶DB20に記憶されている数だけの検索結果をグループ分け対象コンテンツとして選択する。

【 0 0 2 9 】

20

グループ管理部14は、CPU、プログラム、及び、記憶装置で構成される。グループ管理部14は、グループ分機能141及び関連キーワード対応付機能142を備えている。グループ分機能141は、グループ分対象選択機能151によりグループ分け対象コンテンツとして選択された検索結果各々について、コンテンツ情報中の特徴情報を解析し、コンテンツの特徴を利用したグループ分けを行い、コンテンツのグループを生成する。

【 0 0 3 0 】

関連キーワード対応付機能142は、生成された各グループに対して、当該グループを特徴づける関連キーワードを対応付ける。具体的には、関連キーワード対応付機能142は、各グループに属するコンテンツの中から各グループを特徴付ける関連キーワードを抽出する。そして、関連キーワード対応付機能142は、各コンテンツがどのグループに別

30

けられたかをグループ、関連キーワード及びコンテンツとの対応情報として、記憶装置に設けられている、グループとコンテンツの対応情報データベース(以下「対応情報DB」という)17に保存する。

【 0 0 3 1 】

入出力部11は、例えば、キーボード、ディスプレイ、及び、入出力インターフェースで構成される。入出力部11は、入力機能111と提示機能112とを備えている。入力機能111は、ユーザが入力した検索キーワード及び画面上で選択された関連キーワードを受け付ける。提示機能112は、キーワード管理部15やグループ管理部14等からの指示により、キーワードやコンテンツの一覧をディスプレイに表示する。

【 0 0 3 2 】

40

図2には、提示機能112によりコンテンツ発見装置10のディスプレイに表示されるコンテンツ表示画面402の一例を示す。同図に示すように、コンテンツ表示画面402は、キーワード表示部402-1とコンテンツ表示部402-2とから構成される。キーワード表示部402-1には、検索キーワード(ここでは「スポーツ」と、グループ毎の関連キーワード(ここでは「野球」、「運動」、「秋」と)が表示される。また、コンテンツ表示部402-2には、グループ毎のコンテンツ一覧が表示される。

【 0 0 3 3 】

次に、図3に示す処理フローを参照しながら、コンテンツ発見装置10が行うコンテンツ発見の処理の流れを説明する。

まず、コンテンツ発見装置10の入出力部11がユーザから検索キーワードを取得する

50

と(ステップS601)、キーワード管理部15は検索キーワードをキーワード履歴記憶DB18に登録する(ステップS602)。Web検索部12は、検索キーワードをもとに、インターネット等のネットワークを介してコンテンツを取得する(ステップS603)。

【0034】

コンテンツ管理部13では、Web検索部12が取得したコンテンツを解析し、コンテンツのURL、コンテンツの実体データ、及び、コンテンツの内容から取得される特徴情報等を含むコンテンツ情報をコンテンツ情報DB16に格納する(ステップS604)。コンテンツの内容から特徴情報を抽出する方法としては、例えばHTML(Hyper Text Markup Language)のようなテキストデータからなるコンテンツに対して、形態素解析によ

10

ってコンテンツ中に出現する単語の頻度を数えることでコンテンツの特徴量として数値化するやり方がある。

【0035】

グループ管理部14では、キーワード管理部15に対してキーワード履歴記憶DB18に複数の検索キーワードの履歴があるかどうかを問い合わせる(ステップS605)。複数の検索キーワードの履歴がある場合には(ステップS605; yes)、キーワード管理部15のグループ分対象選択機能151は、キーワード履歴記憶DB18から過去の検索キーワードを抽出し、抽出した各検索キーワードに該当するコンテンツ情報をコンテンツ管理部13から取得する。グループ分対象選択機能151は、今回の検索キーワードと過去の検索キーワードとから取得された全コンテンツ情報をマージし、グループ分け対象

20

コンテンツとしてグループ管理部14へ引き渡す(ステップS606)。

【0036】

グループ管理部14のグループ分機能141は、グループ分け対象コンテンツのコンテンツ情報をもとにコンテンツ群を複数の類似グループに分け、グループとそこに属するコンテンツの対応情報を対応情報DB17に保存する。コンテンツ情報をもとに類似グループに分ける方法としては、数値化された特徴量を利用してクラスタリングによってコンテンツ間の類似度を求め、類似度の高い順にグループを作成するものがある。

【0037】

また、グループ管理部14の関連キーワード対応付機能142は、各グループからグループを特徴付けるキーワードである関連キーワードを抽出する(ステップS607)。関連キーワードを抽出する方法としては、グループ内の全コンテンツに共通して多く含まれる単語を関連キーワードとして抽出するやり方や、各グループ間での単語の出現頻度のばらつきを調べ、特定のグループにだけ高い頻度で出現する単語を関連キーワードとして抽出するやり方がある。また、コンテンツから抽出する以外に、グループ内の全コンテンツを特徴付ける概念を表すキーワードを新たに生成し、そのキーワードを関連キーワードとしてもよい。

30

キーワード管理部15は、グループ管理部14によって抽出された関連キーワードをキーワード履歴記憶DB18に保存する(ステップS608)。入出力部11の提示機能112は、検索キーワードおよび関連キーワードのリストと、グループ分けされた検索結果をユーザへ表示する(ステップS609)。

40

【0038】

入出力部11の入力機能111がユーザからグループ(又は、関連キーワード)選択の操作を受信すると(ステップS610; yes)、キーワード管理部15はユーザが選択したグループの関連キーワードを検索キーワードとして設定し(ステップS611)、当該検索キーワードをキーワード履歴記憶DB18へ保存する(ステップS602)。Web検索部12及びコンテンツ管理部13によって当該検索キーワードに該当するコンテンツの取得及びコンテンツ情報の格納が行われると(ステップS603、ステップS604)、キーワード管理部15は、キーワード履歴記憶DB18からこれまでにユーザが選択してきた検索キーワードを抽出し、各検索キーワードに対応するコンテンツ情報をコンテンツ管理部13から取得してマージし、より多くのコンテンツをグループ管理部14へ引

50

き渡す(ステップS606)。キーワード管理部15のグループ分対象選択機能151がキーワード履歴記憶DB18から抽出してマージする過去の検索キーワードの数は、検索キーワード数記憶DB19を用いて任意に設定することが可能であり、予めコンテンツ発見装置10側で抽出すべき検索キーワードの数を固定的に設定する場合や、ユーザからの操作によって随時抽出すべき検索キーワードの数を変更する場合もある。また、マージした結果のコンテンツ数が膨大となる場合などに、検索結果数記憶DB20を用いて、各キーワード毎に検索結果として抽出するコンテンツの数を任意に設定する場合もある。

【0039】

グループ管理部14は、キーワード管理部15から取得したコンテンツ情報に対してグループ分けを行い、各グループに対応する関連キーワードを抽出し(ステップS607)、入出力部11の提示機能112が検索キーワードおよび関連キーワードのリストとグループ分けされたコンテンツ群をユーザへ表示する(ステップS609)。

10

入出力部11がユーザから個別コンテンツ選択の操作を検知すると(ステップS612; yes)、提示機能112は、選択されたコンテンツに関するコンテンツ情報をコンテンツ管理部13から取得し、ユーザに当該コンテンツ情報を表示する(ステップS613)。

【0040】

次に、図4を参照しながら、本実施形態に係る動作例について説明する。まず、ユーザはコンテンツ発見装置10のキーボードを操作することにより、図4(a)に示す検索キーワード入力画面を表示する。ユーザが検索キーワード入力画面401に検索キーワードとしてキーワード1を入力すると、コンテンツ発見装置10の入出力部11は、キーワード1を検索キーワードとして受け付ける。Web検索部12は、キーワード1にマッチするコンテンツを検索する。

20

【0041】

次に、図4(b)に示すように、グループ管理部14のグループ分機能141は、キーワード1にマッチするコンテンツ全体403をコンテンツの内容でグループ別けし、グループ403-1, 403-2, 403-3を生成する。関連キーワード対応付機能142は、グループ403-1, 403-2, 403-3毎に、各グループを特徴付ける関連キーワード(ここでは、キーワード2、キーワード2'、キーワード2")を対応付ける。

【0042】

30

入出力部11の提示機能112は、ディスプレイに図4(c)に示すコンテンツ表示画面402を表示する。コンテンツ表示画面402のキーワード表示部402-1には、検索キーワード(キーワード1)とグループ毎の関連キーワード(キーワード2, 2', 2")とが表示される。またコンテンツ表示部402-2には、キーワード2で特徴付けられるコンテンツ一覧と、キーワード2'で特徴付けられるコンテンツ一覧と、キーワード2"で特徴付けられるコンテンツ一覧とがグループ分けされて表示される。

【0043】

次に、ユーザが、図4(c)に示すキーワード表示部402-1に表示されているキーワード2を選択する操作を行ったものとする。キーワード管理部15のグループ分対象選択機能151は、選ばれたキーワード2にマッチするコンテンツと、過去の検索キーワードであるキーワード1にマッチするコンテンツとをグループ分け対象コンテンツとしてマージする。グループ管理部14のグループ分機能141は、キーワード1にマッチするコンテンツとキーワード2にマッチするコンテンツとの全体405をコンテンツの内容でグループ別けし、3つのグループ405-1, 405-2, 405-3を生成する。関連キーワード対応付機能142は、グループ毎に各グループを特徴付けるキーワード3, 3', 3"を対応付ける(図4(d))。

40

【0044】

提示機能112は、ディスプレイに図4(e)に示すコンテンツ表示画面402を表示する。コンテンツ表示画面402のキーワード表示部402-1には検索キーワード(キーワード1, 2)と関連キーワード(キーワード3, 3', 3")とが表示される。また

50

、コンテンツ表示部 402 - 2 には、キーワード 3 で特徴づけられるコンテンツ一覧と、キーワード 2 ' で特徴付けられるコンテンツ一覧と、キーワード 2 " で特徴付けられるコンテンツ一覧とがグループ分けされて表示される。

【0045】

同様の操作を繰り返すと、コンテンツ表示画面 402 のキーワード表示部 402 - 1 には、ユーザが入力又は選択した n 個のキーワードによって取得されるコンテンツ全てを対象にグループ分けを行った場合の関連キーワード（キーワード (n + 1)、キーワード (n + 1) '、・・・）と、入力又は選択済みの n 個の全キーワード（キーワード 1、キーワード 2、・・・、キーワード n）のリストが表示される。また、コンテンツ表示部 402 - 2 には、関連キーワード（キーワード (n + 1)、キーワード (n + 1) '、・・・）各々に特徴付けられるコンテンツの一覧が表示される。

10

【0046】

以上説明したように、コンテンツ発見装置 10 は、ユーザから受信した検索キーワードをもとに検索を行い、取得したコンテンツのグループ分けを行ってグループを特徴付けるキーワードを関連キーワードとして抽出し、検索キーワード及び関連キーワードのリストと、グループ分けされたコンテンツ一覧とを表示する。そして、ユーザが選択した関連キーワードによって再検索を行い、過去の検索結果と共にマージをし、グループ化を行い、関連キーワードを抽出する、という処理を繰り返していくことで、最初に入力された検索キーワードによって限定された検索結果を絞り込んでいくのではなく、最初に入力された検索キーワードに関連した情報をどんどん検索結果として取り込み、検索結果を関連性のあるコンテンツ群に拡張していくことで、マンネリ化しない新たなコンテンツの発見を実現することができる。

20

【0047】

このように、検索を繰り返す毎に、より多くのコンテンツを対象にグループ分けを行い、関連キーワードの抽出を行うことができるため、多様な関連キーワードに関連する多様なコンテンツを取得することが可能となり、キーワード検索でありながら、検索キーワードに限定されない多様な検索結果を得ることができる。

従来の絞りこみ型の検索エンジンでは、最初の検索キーワードによって取得可能な情報がある程度限定され、以降の検索処理は最初に限られた情報の中での探索となるため、ユーザの検索の目的がはっきりしていない場合や、目的や欲しい情報を象徴している検索キーワードの設定が適切でない場合に、ユーザは所望の情報を取得することが難しかったが、本実施形態に係るコンテンツ発見装置 10 によればユーザは所望の情報を取得しやすくなる。また、ユーザが取得可能な情報がマンネリ化することがなく、検索を行いながら新しい興味の発見や気づきにつながるような情報を取得することが可能となる。

30

【0048】

[変形例]

以上、本発明の実施形態について説明したが、本発明は係る実施形態に限定されるものではなく、その技術思想の範囲内で様々な変形が可能である。変形例としては、例えば、以下のようなものが考えられる。

(1) 上述した実施形態においては、ユーザはコンテンツ発見装置 10 のキーボード及びディスプレイを用いて検索キーワードの入力や検索結果の確認を行うとして説明したが、ユーザは自己が所持する携帯電話機やパーソナルコンピュータからコンテンツ発見装置 10 に無線又は有線で接続して検索キーワードの送信や検索結果の受信を行うようにしてもよい。

40

【0049】

(2) 上述した実施形態においては、インターネット上のコンテンツを検索する実施形態について説明したが、これに限定されることはなく、例えば、コンテンツ発見装置 10 のデータベースに格納されている文書ファイル等のコンテンツを検索してもよい。

(3) 上述した実施形態においては、グループ分け対象となる過去の検索キーワードによる検索結果をコンテンツ情報 DB 16 から取得するとして説明したが、処理負荷がかから

50

ない場合には、過去の検索キーワードで再度インターネット検索を行うことにより、グループ分け対象となるコンテンツを取得してもよい。

(4) キーワード履歴記憶DB18に格納されている検索キーワードの履歴及び関連キーワードの履歴と、対応情報DB17に蓄積されているグループ(関連キーワード)とコンテンツの対応情報とを用いて、過去に表示したコンテンツ表示画面402画面を再度表示可能な仕組みを構築してもよい。

【0050】

(5) 上述した実施形態においては、関連キーワード(キーワードn)が選択された場合に、コンテンツ表示部402-2には、新たに抽出された関連キーワード(キーワード(n+1)、キーワード(n+1)'、...)に特徴付けられたコンテンツの一覧が表示されることとして説明したが、表示内容はこれに限定されることはない。例えば、コンテンツ表示部402-2に、今までの検索キーワード(キーワード1、キーワード2、...、キーワード(n-1))の少なくとも1つにマッチするコンテンツの中で、今回選択された関連キーワード(キーワードn)の特徴をもつコンテンツを表示してもよい。また、画面分割により複合的な表示を行うようにしてもよい。

10

【図面の簡単な説明】

【0051】

【図1】本発明の実施の形態に係るコンテンツ発見装置の機能構成を示すブロック図である。

【図2】同実施の形態に係るコンテンツ表示画面の一例を示す図である。

20

【図3】同実施の形態に係るコンテンツ発見装置が行う処理手順の一例を示すフロー図である。

【図4】同実施の形態に係るコンテンツ発見装置が行うコンテンツ発見の手順の一例を説明するための図である。

【図5】従来の検索エンジンを用いた検索手順の一例を説明するための図である。

【図6】従来の検索装置の機能構成の一例を示す図である。

【図7】従来の検索装置が行う検索処理手順の一例を示すフロー図である。

【符号の説明】

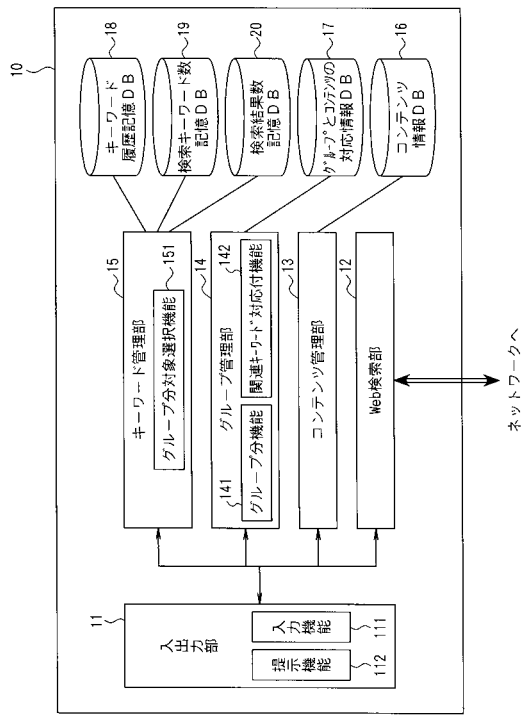
【0052】

- 10 コンテンツ発見装置
- 11 入出力部
- 111 入力機能
- 112 提示機能
- 12 Web検索部
- 13 コンテンツ管理部
- 14 グループ管理部
- 141 グループ分機能
- 142 関連キーワード対応付機能
- 15 キーワード管理部
- 151 グループ分対象選択機能
- 16 コンテンツ情報DB
- 17 対応情報DB
- 18 キーワード履歴記憶DB
- 19 検索キーワード数記憶DB
- 20 検索結果数記憶DB
- 401 検索キーワード入力画面
- 402 コンテンツ表示画面
- 402-1 キーワード表示部
- 402-2 コンテンツ表示部

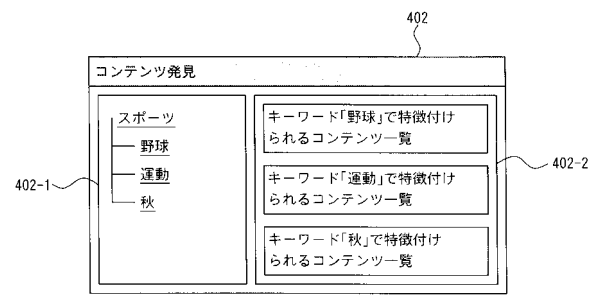
30

40

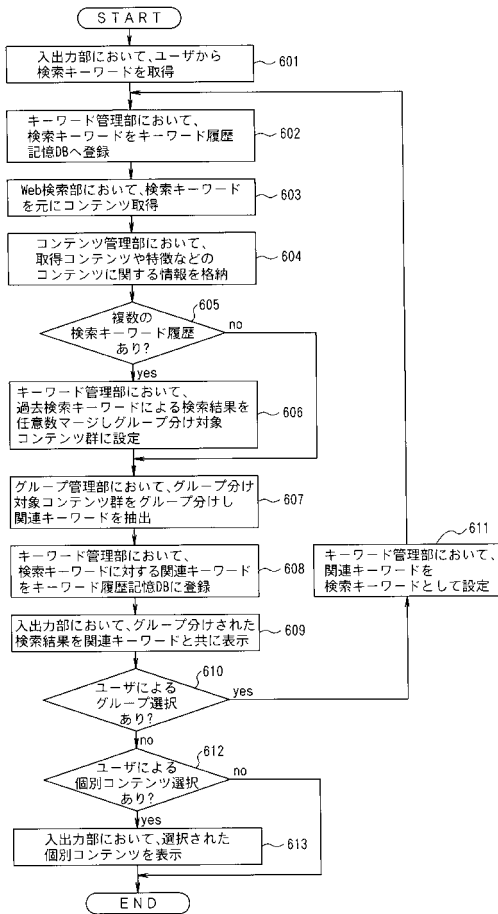
【図1】



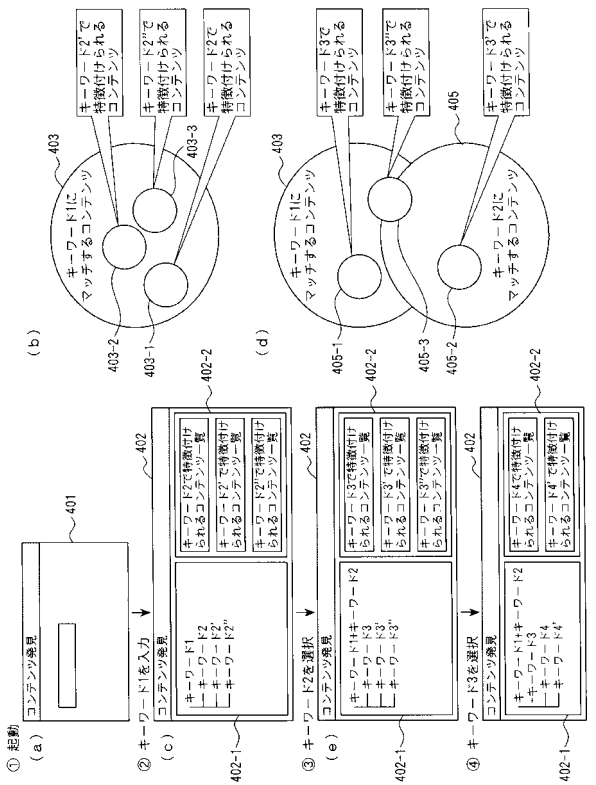
【図2】



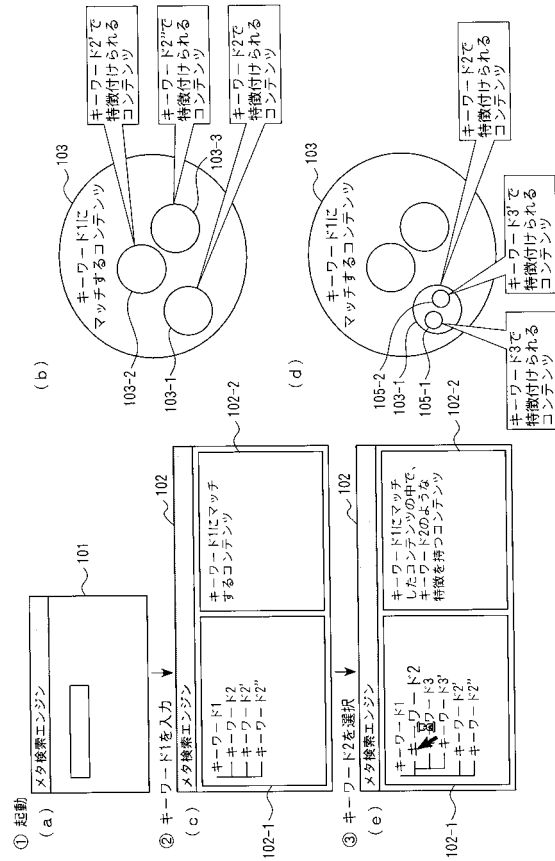
【図3】



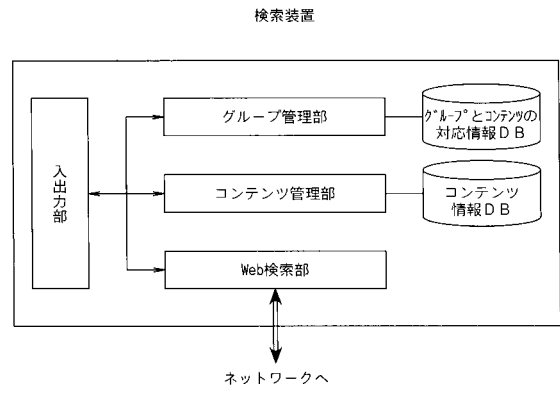
【図4】



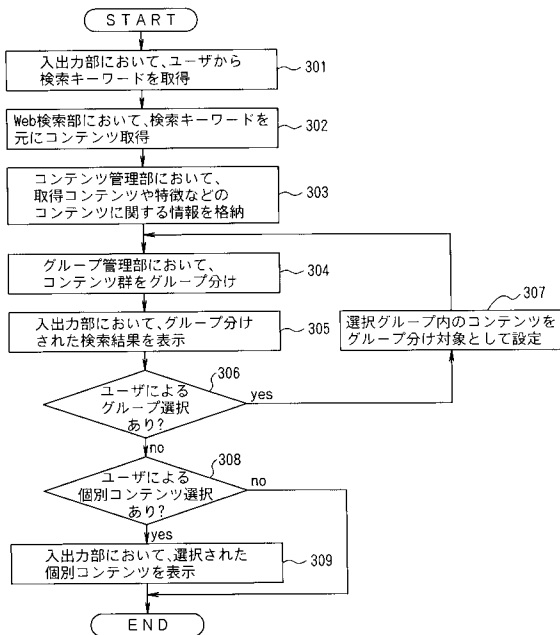
【図5】



【図6】



【図7】



フロントページの続き

(72)発明者 鈴木 偉元

東京都千代田区永田町二丁目11番1号 株式会社エヌ・ティ・ティ・ドコモ内

合議体

審判長 山崎 達也

審判官 原 秀人

審判官 石井 茂和

(56)参考文献 特開2000-76279(JP,A)

特開2004-287670(JP,A)

特開2004-192156(JP,A)

特開2002-351896(JP,A)

特開平11-338873(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F17/30