

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4623985号
(P4623985)

(45) 発行日 平成23年2月2日(2011.2.2)

(24) 登録日 平成22年11月12日(2010.11.12)

(51) Int.Cl. F I
HO4N 5/445 (2011.01) HO4N 5/445 Z

請求項の数 17 (全 32 頁)

(21) 出願番号	特願2004-87788 (P2004-87788)	(73) 特許権者	500046438
(22) 出願日	平成16年3月24日 (2004. 3. 24)		マイクロソフト コーポレーション
(65) 公開番号	特開2004-289848 (P2004-289848A)		アメリカ合衆国 ワシントン州 9805
(43) 公開日	平成16年10月14日 (2004.10.14)		2-6399 レッドモンド ワン マイ
審査請求日	平成19年3月26日 (2007. 3. 26)		クロソフト ウェイ
(31) 優先権主張番号	10/395, 679	(74) 代理人	100140109
(32) 優先日	平成15年3月24日 (2003. 3. 24)		弁理士 小野 新次郎
(33) 優先権主張国	米国 (US)	(74) 代理人	100089705
			弁理士 社本 一夫
		(74) 代理人	100075270
			弁理士 小林 泰
		(74) 代理人	100080137
			弁理士 千葉 昭男
		(74) 代理人	100096013
			弁理士 富田 博行

最終頁に続く

(54) 【発明の名称】 電子番組ガイド (EPG) データのフリーテキスト検索および属性検索

(57) 【特許請求の範囲】

【請求項 1】

電子番組ガイド (EPG) データを検索する方法であって、

(a) テキスト検索文字列を受け取るステップ、

(b) 前記テキスト検索文字列を、前記 EPG データのフォーマットに一致するフォーマットを有する検索可能な語に変換することによって、前記テキスト検索文字列を正規化するステップ、

(c) 前記検索可能な語を属性リスト中の属性と比較して、前記検索可能な語を テキスト語と属性語 とに分離するステップ、

(d) 前記テキスト語および前記属性語について前記 EPG データの検索を行うステップ、

(e) 前記テキスト検索文字列に合致する、前記 EPG データに含まれる番組のリストを作成するために、前記検索の結果の組み合わせとランク付けとフィルタリングとのうちの少なくとも 1 つを行うステップ、および

前記リストを表示するステップ

を含むことを特徴とする方法。

【請求項 2】

前記正規化するステップ (b) は、前記テキスト検索文字列中の記号を除去または変換することを含むことを特徴とする請求項 1 に記載の方法。

【請求項 3】

10

20

前記分離するステップ(c)はさらに、前記検索可能な語を各種の組合せでグループ化して、検索可能なフレーズのセットを実現することを含むことを特徴とする請求項1に記載の方法。

【請求項4】

前記分離するステップ(c)はさらに、前記グループ化の前に、前記検索可能な語の少なくとも2つの順序を入れ替えることを含むことを特徴とする請求項3に記載の方法。

【請求項5】

前記属性リストは、インデックス付けされた属性を、インデックス付けされた各属性の対象リソースおよび検索値に関係付けるインデックスを含むことを特徴とする請求項1に記載の方法。

10

【請求項6】

前記属性リストは、前記属性リストの属性に関連付けられた分岐ロジックを含むことを特徴とする請求項1に記載の方法。

【請求項7】

前記属性リストは、前記属性リストの属性についてのメタデータを含むことを特徴とする請求項1に記載の方法。

【請求項8】

前記検索の結果の組合せとランク付けとフィルタリングとのうちの1つを行うステップ(e)は、検索結果の各々に関連度のランク付けを割り当てることを備え、

前記作成された番組のリストは、前記割り当てられた関連度のランク付けに応じて番組をリストしている

20

ことを特徴とする請求項1に記載の方法。

【請求項9】

前記テキスト検索文字列を受け取るステップ(a)は、一度に一文字ずつテキスト検索文字列を受け取ることを含み、

新しい文字を受け取るたびに、前記新しい文字を前記テキスト検索文字列のすでに受け取っている文字に連結して新しいテキスト検索文字列を作成し、前記ステップ(b)~(f)を前記新しいテキスト検索文字列について実行する

ことを特徴とする請求項1乃至8のいずれかに記載の方法。

【請求項10】

30

前記正規化するステップ(b)は、スペルチェック、オートコレクト、語の分離、大文字/小文字の正規化、句読点の除去、記号の正規化、冠詞の除去、および前記テキスト検索文字列の複数形化のうちの1つを含むことを特徴とする請求項1に記載の方法。

【請求項11】

前記検索を行うステップ(d)は、単語をステミングしたバージョンの検索語を使用することを含むことを特徴とする請求項1に記載の方法。

【請求項12】

前記検索を行うステップ(d)は、単語を派生したバージョンの検索語を使用することを特徴とする請求項1に記載の方法。

【請求項13】

40

コンピュータに請求項1乃至12のうちのいずれか一つに記載された方法を実行させるための命令を記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

【請求項14】

電子番組ガイド(EPG)データベースを検索するための検索エンジンであって、テキスト検索文字列を受け付ける検索文字列レシーバと、

前記検索文字列レシーバに結合され、前記テキスト検索文字列を、前記EPGデータベース内のEPGデータのフォーマットに一致するフォーマットを有する検索可能な語に変換することによって、前記テキスト検索文字列を正規化するノーマライザと、

前記ノーマライザに結合された文字列インタープリタであって、前記検索可能な語を前記文字列インタープリタに通信的に結合された属性リスト内の属性と比較して、前記検索

50

可能な語をテキスト語と属性語とに分離する文字列インタープリタと、

前記文字列インタープリタに結合され、それぞれの前記テキスト語および前記属性語について前記EPGデータベースに照会して検索語を生成するクエリジェネレータと、

前記クエリジェネレータに結合され、前記検索語を使用して前記EPGデータの検索を行うデータマネージャと、

前記データマネージャに結合され、前記テキスト検索文字列に合致する、前記EPGデータに含まれる番組のリストを作成するために、前記検索の結果の組合せとランク付けとフィルタリングとのうちの1つを行う結果コンパイラと、

前記結果コンパイラに結合され、前記リストを表示する表示部とを備えることを特徴とする検索エンジン。

10

【請求項15】

前記結果コンパイラは、前記検索結果に、関連度のランク付けを割り当てる結果ランカーを備えることを特徴とする請求項14に記載の検索エンジン。

【請求項16】

前記文字列インタープリタは、検索制約条件テーブルへのアクセス権を有し、

前記クエリジェネレータは、前記EPGデータベースおよびウェブベースのEPGリソースに結合され、

さらに、

前記文字列インタープリタに含まれ、前記テキスト語と属性語とを各種の組合せでグループ化し、前記検索可能な語の少なくとも2つの順序を入れ替える文字列トランスフォーマと、

20

前記文字列インタープリタに含まれ且つ前記属性リストおよび前記検索制約条件テーブルに結合され、前記テキスト語と前記属性語とを分離する属性アイデンティファイアと、

前記結果コンパイラに含まれ且つ前記文字列インタープリタに結合され、前記文字列インタープリタでの前記検索結果がヒットであるか否かを判定する検索フィルタと、

前記結果コンパイラに含まれ、前記フィルタリングされた検索結果に関連度を割り当てる結果ランカーと、

前記結果コンパイラに含まれ、それらの割り当てられた関連度に従って前記フィルタリングされた検索結果をフィルタリングする包含フィルタと、

を備えることを特徴とする請求項14に記載の検索エンジン。

30

【請求項17】

前記文字列インタープリタが、前記属性リスト中の情報および前記検索制約条件テーブル中の情報を用いて、前記テキスト語と属性語とを各種の組合せでグループ化し、前記検索可能な語を前記テキスト語と前記属性語とに分離するように動作することを特徴とする請求項16に記載の検索エンジン。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般にはマルチメディアデータの通信に関し、詳細には、電子番組ガイド(EPG)データのフリーテキストおよび属性の検索に関する。

40

【背景技術】

【0002】

通常、デジタルテレビジョンシステムでは、ユーザは電子番組ガイド(EPG)の情報を見て番組を探す。EPGは、例えば1週間などのある期間中にユーザの地域のヘッドエンド(番組提供者のチャンネルラインアップ)から視聴できる番組に関連する番組とスケジュールのメタデータからなるデータセットである。ユーザは、キーワードとなる言葉を指定して番組名のテキスト検索を行うことにより番組を検索する。検索基準に一致する検索結果、通常は番組名であるが、これを「ヒット」と呼ぶ。

【0003】

テレビ番組は、番組名に含まれるキーワードだけでなく、番組の属性を検索することに

50

よっても見つけることができる。番組名は番組の属性の1つと考えることができるが、マルチメディアの分野では、用語「属性」は番組の名前以外の多くの特性を表す語としてより専門的な意味で使用される。そのため、属性には、番組とその放映についての無限の多数の特性が含まれる。番組とその属性は通常、E G P情報の巨大な階層の中に適合する。「カテゴリ」など一部の属性は、求める番組が構成要素である広範囲の番組の集合を記述することができ、すなわち、「カテゴリ」属性は、求める番組を包含する階層レベルを記述し、ユーザが「カテゴリ」属性の検索を行うとそのカテゴリのすべての番組がヒットする。例えば「西部劇」の検索では、西部劇を扱った映画およびテレビのジャンルに属する多くのヒットが得られる可能性がある。

【0004】

属性は番組名に固有の属性も記述することができ、すなわち、属性は、同じ番組の2つのバージョンを互いと異ならせるオプションを記述することができる。このより狭いタイプの属性は、特定の番組より「下」のE G P階層の要素、すなわち1つの番組に異なる特性(flavor)を与える要素を記述する。例えば、スペンサー・トレーシー主演の「The Mountain」の1バージョンは「白黒」および「モノラル」の属性を有し、一方、別のバージョンは「カラー」、「ドルビー(登録商標)デジタル」、「クローズキャプション」、および「格付けPG」の属性を有する。番組名のキーワード要素に限定されたテキスト検索では両方のバージョンがヒットとして得られ、一方、「ドルビー」の属性検索と組み合わせた番組名のキーワード要素のテキスト検索では、ドルビー(登録商標)デジタルとドルビー(登録商標)ステレオサウンドシステムのバージョンだけが得られる。

【0005】

従来の番組名のキーワード要素のテキスト検索は、ユーザインタフェース(UI)の検索欄にキーワードテキストを入力することによって行い、従来の属性検索は、UIの別の部分、または全く別個のUIにテキスト検索のキーワードとは別に属性の基準を(例えばチェックボックスに)入力することによって行う。別個のテキスト検索と属性検索とのヒットを両方の検索の最後にのみ論理的に組み合わせてヒットリストを生成する。

【0006】

図1に、従来のテキスト検索100と従来の属性検索102とを示し、ここでは、テキストキーワードと属性検索の条件は、1つまたは複数のUIにより表示される属性のチェックボックスとは別個の検索欄などに、別々に入力される。

【0007】

ブロック104でテキストキーワードを入力する。ブロック106で、キーワードの入力とは別に属性の条件を入力する。入力されたテキストキーワードを用いて通常は番組名に対してテキスト検索を行い(ブロック108)、入力された属性検索の条件を用いて属性検索を行う(ブロック110)。それぞれの検索により別個の結果が得られる(ブロック112、114)。フィルタリングロジック116により結果112、114を比較し、通常はテキスト検索100と属性検索102との両方を満たすヒットリスト118を生成する。

【発明の開示】

【発明が解決しようとする課題】

【0008】

通常、ユーザはUIの別の部分で属性を検索することを求められるので、多くのユーザは検索の入力を番組名の単語からなるキーワード式のクエリだけに限定し、また一方で、他のユーザは、デジタルテレビジョンシステムではそのように入力された属性を通常の番組名の検索語として解釈することを理解せずに、誤って番組名のテキスト検索のキーワードとして属性を入力してしまう。

【課題を解決するための手段】

【0009】

本主旨は、電子番組ガイド(EPG)データ用の検索エンジンおよびそれに関連する方

10

20

30

40

50

法を含む。例示的方法では、テキスト検索文字列 (text search string) を正規化して検索可能な語とし、それらの語をテキスト検索語または属性検索語として解釈する。検索文字列の語の解釈に従って、様々な複雑度の検索ロジックを使用するクエリを作成する。各テキスト語および各属性語についてEPGデータベースおよび/またはEPGのウェブリソースで1回または複数回の検索を行い、検索文字列の解釈に従って結果に関連度のランク付けを与える。組み合わせた検索結果は、ユーザに表示するためにランク付け、フィルタリング、およびグループ化することができる。結果は、ユーザが検索文字列の各文字を入力すると同時に漸進的に表示することもできる。

【0010】

一実施形態では、属性のリストを使用してテキスト語と属性語を分離することができる。例示的な属性リストはさらに、検索文字列中の各属性語について行われる検索の進行を指示する各属性のインデックスまたはメタデータも含むことができる。

【発明を実施するための最良の形態】

【0011】

概要

電子番組ガイド (EPG) データの迅速で、直感的、かつ正確な検索は、テレビジョンセットに付加する「セットトップボックス」システムと、テレビ放送の受信および録画機能を持つパーソナルコンピュータ (PC) システムとの両方を含むデジタルテレビジョンシステムのエンドユーザにとって重要である。例示的な主旨は、デジタルテレビジョンシステムでユーザのテキストと属性とのクエリを入力し、解釈する検索エンジンとそれに関連する方法について記載する。この迅速で正確な検索機構は、例えば、ユーザが入力した単一のテキスト列を解釈し、可能性の高いユーザの意図を推測し、ユーザが検索する内容の解釈に一致することを試みるヒット結果のセットを表示する。このような高度な検索機能は、PCのデジタルビデオレコーダ (DVR) およびセットトップボックスの市場が急速に拡大しているのに従い重要となっている。

【0012】

図2に、EPGデータベースまたは他のリソースにおける例示的検索200の概要を示す。例示的検索200は、ユーザが検索文字列202を入力する、あるいはユーザから文字列202を受け取るか、またはストレージから検索文字列202を取り出すことから開始する。例示的な検索エンジン204は、検索文字列202を受け取り、その文字列の語にいくつかの変形を行って、複雑度が様々に異なる、考えられるクエリ命令を生成する。

【0013】

検索文字列202の複雑度と不確実性とに応じて、単一の検索文字列202から1つまたは複数の独立した検索が発生する可能性があり、したがって、一実施形態では、文字列インタープリタ214が特定の検索文字列202に基づいて可能な限り少なく (可能であれば1つ) 可能な限り限定的なクエリを作成することを試みる。例えば、一実施形態では、文字列インタープリタ214は、検索文字列202中のどの語が属性検索語として検索すべき知られている属性であるかを判断することにより、この最小限の数の簡潔かつ/または限定的なクエリを実現する。検索文字列202中の残りの語は、可能な属性語または可能なテキスト語である可能性があるが、(例えば知られる属性のデータベースやリストを使用して) 知られる属性を判定すると、多くの場合クエリの範囲がかなり絞られる。このように、文字列インタープリタ214は、下記で図3との関連でより詳しく述べるように各種の方法とリソースを使用して、一連のテキスト検索206、属性検索208、条件付きテキスト検索210、および条件付き属性検索212を「検討」する。

【0014】

一実施形態では、検索文字列202の可能な意味の強制的な判定は、EPGデータベースおよび/またはオンライン辞書とランキング結果とで、検索文字列202中の正規化した各語に対する予備的な検索を行うことによって実現される。例えば、「Ryan」が検索語である場合は、その解釈がそのデータベースまたは辞書でより一般的であれば、「Saving Private Ryan」が「Ryan O'Neal」よりも優位にそ

10

20

30

40

50

の検索の可能性の高い解釈として選択される。別の実施形態では、文字列インタープリタ 214 は、組合せの 1 つが検索対象の EPG リソースに見つかるか、論理的に好ましいものとなるまで、検索語の様々な組合せを試みる。したがって、検索文字列 202 中の多数の属性語の中で不規則にはあるが唯一のテキスト語として出現する検索語「moon」および「paper」の場合は、恐らくは解釈「Paper Moon」が得られ、この解釈は、クエリを介して実際にヒットとして見つかった場合には高いランク付けになると考えられる。

【0015】

このように、文字列インタープリタ 214 は 1 つまたは複数の解釈に到達し、その解釈がクエリ 216 となる。クエリ 216 を使用して 1 回または複数回の検索を行い、その結果、組み合わせ、フィルタリングし、および / またはランク付けしたヒットリスト 222 が得られ、このリスト 222 がユーザに表示される。例示的な検索 200 は、1 つのみの検索文字列 202 から、テキストと属性とを区別する 1 回または複数回の検索が発生する仕組みを示す一例に過ぎない。検索文字列 202 を例えば文字列インタープリタ 214 によって構文解析、変形、および分析して、検索文字列 202 中の語のテキストと属性とを認識した解釈を達成する。

10

【0016】

一実施形態では、組み合わせ、フィルタリングし、および / またはランク付けしたヒットリスト 222 は、各種の検索パスから結果が返されると更新される。あるいは、検索文字列 202 を入力する、すなわち検索文字列 202 に各文字を追加して新規の検索および / または修正した検索を開始すると同時に、組み合わせ、フィルタリングし、および / またはランク付けしたヒットリスト 222 に漸進的に結果を表示していくことも可能である。

20

【0017】

例示的な検索エンジン

図 3 に、図 2 の例示的な検索エンジン 204 をより詳細に示す。例示的な検索エンジン 204 は、1 つまたは複数の EPG データベース 300 に通信的に結合されている。例示的な検索エンジン 204 が行う機能およびプロセスについて以下で説明するが、言うまでもなく、例示的な検索エンジン 204 は、ここで説明する機能およびプロセスに追加的な機能を有し、プロセスを行うことも可能であり、それには例えばデータベース内容のインデックス付け (indexing) が挙げられる。言うまでもなく、例示的な検索エンジン 204 の各種の実装は各種の人間言語を用いて動作する。したがって、例示的な検索エンジン 204 の一部の機能は、一部の言語には当てはまらない、あるいは必要でない場合がある。

30

【0018】

一般に、例示的な検索エンジン 204 は、まずアプリケーションプログラムインタフェース (API) などの検索文字列レシーバ 301 を公開することにより EPG データベースの検索を行うが、レシーバ 301 は検索式を受け取り、プログラムの生成された、あるいはユーザによって生成された制約条件のセットにアクセスする。部分的にその制約条件に基づいてクエリを作成すると、例示的な検索エンジン 204 は結果のセットを返す。

【0019】

検索文字列レシーバ 301 は、1 つまたは複数の UI を介して検索文字列 202 を受け付けることができる。EPG データ検索のための UI は、「two-foot」システムまたは「ten-foot」システムに合わせて最適化することができる。two-foot システムは、ユーザが PC システムに典型的な情報入力機能を有するデジタルテレビジョンシステムである。例えば、two-foot システムでは、フルキーボード、マウス、およびモニタを使用する。一方、ten-foot システムは、two-foot システムに比べてユーザの入力機能が限られたデジタルテレビジョンシステムである。ten-foot システムは、セットトップボックスによって生成された UI であり、検索文字列 202 は、作動させるキーが比較的少ないリモートコントロールを使用して入力される可能性がある。

40

50

【 0 0 2 0 】

一般的には ten - foot システムでは UI を中心としたデータ入力とナビゲーションがより難しいので、 ten - foot システムに検索を入力する際のユーザの意図を把握するのに必要なデータ量を最小限にし、一方で、検索の精度と性能とを最大にすることは有用である。

【 0 0 2 1 】

検索文字列レシーバ 3 0 1 は、言うまでもなく、任意数の入力欄、チェックボックス、スイッチ、および/またはデバイスからユーザ入力を受け付けることができる。ただし、入力が単に単一の検索文字列 2 0 2 である場合、例示的検索エンジン 2 0 4 は、検索文字列を「デコード」してテキスト検索の情報と属性検索との情報を分離し、適切なクエリを作成することができる。

10

【 0 0 2 2 】

検索文字列レシーバ 3 0 1 に通信的に結合されたノーマライザ 3 0 2 は、検索文字列 2 0 2 中の不規則性を変換して(すなわち「標準化されていない」語のフォーマットを変換して)、各種クエリで整合的に適用できる語にする。非標準的な入力とは、検索する E P G データベース 3 0 0 の形式 (f o r m a l i t y) に一致しないテキスト、文字、言語および/または句読法の形式を有する検索文字列 2 0 2 の語または一部である。非標準的な入力は、文字列インタープリタ 2 1 4 にとって一部または全体が解釈不能となるリスクを伴い、その形式が検索対象の E P G データベース 3 0 0 に合わない時にはヒットの判定の誤りが生じる可能性がある。

20

【 0 0 2 3 】

検索要求の一部として、検索文字列レシーバ 3 0 1 には、しばしば、例えば「The West Wing」などのユーザが入力した文字列が渡される。ノーマライザ 3 0 2 または例示的検索エンジン 2 0 4 の別のモジュールは、語の分離を用いて、その句を、句を構成する 1 つ 1 つの単語に分割することを試み、各単語について個別にテキスト検索を行う。この個々の単語に分離することには、ユーザがよく知られるタイトル中の単語を省略した時、あるいは並べ替えた時に一致を見つけられる利点があるが、ユーザが意図するよりも多くのヒットを生成する可能性もある。個々の単語検索語は、しばしばスペースとダッシュとによる境界で区切られている。ノーマライザ 3 0 2 は、引用符で囲まれた句の単語間にある余分な空間を取り除くこともできる。

30

【 0 0 2 4 】

ユーザは、大文字/小文字 (l e t t e r c a s i n g) を常に正しく使用して検索文字列 2 0 2 を入力するとは限らない。例えば、「The West Wing」を「the west wing」とタイプする可能性がある。また、E P G データベース情報は、「SUNDAY TICKET」のような句で「Sunday Ticket」や「sunday ticket」の代わりに非慣用的な大文字/小文字の使用を含むことがある。ノーマライザ 3 0 2 は、各検索語と検索対象の E P G データベースのテキストを、例えば現在のロケールの規則に従って一般的な大文字/小文字の使用(大文字または小文字)に変換することにより、大文字/小文字の正規化を行うことができる。

【 0 0 2 5 】

ノーマライザ 3 0 2 は、特定の記号およびマークの正規化も行うことができる。発音区別符号は、多くの国際的な言語のアルファベットに見られるアクセントまたはウムラウトなどの結合文字であり、通常は母音を伴う。本主題の一実施形態では、検索語と検索対象のデータベースのテキストとは、発音区別符号がないものとして比較する。正確な検索を促進するために、ノーマライザ 3 0 2 により記号および特殊なアルファベット文字を書き直す、かつ/または検索語に追加することもできる。例えば、ドイツ語のエスツェット「ß」などの文字は「ss」と書き直して、ドイツ語のエスツェットを含む検索が、検索対象のテキスト中にエスツェットの形または「ss」の形がある場合にヒットを生成するようにする。

40

【 0 0 2 6 】

50

記号の正規化には、より正確な一致を得るために検索語および検索対象のテキスト両方の句読記号などの記号を除去または変換することが伴う。例えば、「F . B . I」などの略語中のピリオドや「M * A * S * H」のアスタリスクを取り除く。「&」や「@」などの記号は、各自の単語の形態「and」、「at」に変換することができる。例えば「V e g a \$」の記号「\$」など、文字の代用とする記号を変換することができる。この記号の正規化により、追加的な検索が行われる可能性がある。例えば、例示的な検索エンジン 204 は、慣例的に 1 つだけの形を使用し、その慣用的な 1 つの形に従って検索語と検索対象のテキストとの両方を正規化するのではなく、ノーマライザ 302 が検索文字列 202 を処理した後に「V e g a s」と「V e g a \$」の両方を検索することができる。

【0027】

10

冠詞と、「a」、「the」、「is」、「are」、「el」、「la」、「los」、「les」などの頻出するがあまり記述的ではない単語は、ノーマライザ 302 により検索語から取り除いて処理の速度を上げることができる。一部の実施形態では、冠詞と頻出する単語は、その語が検索にある唯一の単語である場合、あるいは検索の最初のパスで結果が得られなかった場合には、除去されないか、または条件をつけて検索される。

【0028】

例示的な検索エンジン 204 は、検索の意味を何らかの形で変える標識 (token) があるかどうか検索文字列を事前にスキャンすることができる。例えば、ある単語の前にマイナス記号「-」が付いている場合は、その単語が一致結果に出現してはならないことを意味する。複数の語を引用符で囲まれた語句としてまとめることは、そのグループの語が与えられた順序で共に出現する時だけにそのグループを考慮すべきことを意味する。単語「AND」、「OR」および「NOT」は、特殊な形で使用して検索演算子に影響を与えることができる。これにより、検索に対するより多くのコントロールがユーザに与えられる。

20

【0029】

一実施形態では、ノーマライザ 302 または検索エンジン 204 の他のコンポーネントは、入力された検索文字列 202 にスペルチェックを適用することができる。いくつかのストラテジを用いて、ユーザが検索文字列 202 を入力する際に間違いを防止するのを助ける。ユーザには、スペリングのダイアログボックスあるいは「波線」による INTELLISENSE (登録商標) ハイライト機能などの別のテキストによる表示機構を介して、検索を実行する前に、検索テキスト中にスペルが間違った単語があることを知らせることができる。あるいは、検索の実行の最後までスペルチェックを遅らせて、正確な結果が見つからない場合にのみ行うこともできる。そしてスペルチェックは透過に行うことができる。ノーマライザ 302 によってスペルが誤っていると解釈された単語の修正後の形態は、自動的に検索文字列 202 に付加され、条件付きの検索として照会することができる。そして検索でスペルが誤ったその検索語を実際に含むヒットが得られた場合は、関連度のランキング (下記で説明する) を使用してその条件付き検索語のランクを下げることもできる。

30

【0030】

ノーマライザ 302 は、スペルチェックを簡略化した形態であるオートコレクトも使用することができる。各検索語を辞書と比較するのではなく、よくスペルを間違える単語の単純な検索テーブルを使用して、タイプされるのと同時に単語を修正する。オートコレクトが間違った判断を行う (ユーザが意図的に「スペルを間違えた」単語を修正する) 稀な事例では、ユーザは、1 回バックスペースキーを押してその自動的な修正を無効にすることができる。ユーザがタイプするのと同時に、ノーマライザ 302 により、既存のフルテキストインデックスに基づいて考えられる単語の完成形のポップアップリストを提示することもできる。これにより、ユーザが検索基準をより迅速に入力することが可能になる。

40

【0031】

文字列インタープリタ 214 は、ノーマライザ 302 と通信的に結合されている。一実施形態では、文字列インタープリタ 214 は、属性リスト 310 に通信的に結合された文

50

字列トランスフォーマ306および属性アイデンティファイア308を含む。文字列インタープリタ214は、検索制約条件のテーブル312とも通信的に結合することができる。

【0032】

文字列トランスフォーマ306は、検索文字列202中の検索語を単独で、または組み合わせでテストして、その検索文字列202の解釈に到達し、その文字列を再配置すべきかどうかを判定することができる。文字列トランスフォーマ306は、再配置および/またはスクランプリングの方式を使用して、検索文字列202中の各種語の組合せを試すことができ、様々な順序においてユーザが検索しようとしている可能な属性を識別することができる。したがって、「dolby the closed mile green captioning」からは、「dolby」、「the Green Mile」、および「closed captioning」を含む候補検索語を得ることができる。語の2つ以上を共にまとめる方がよいと思われる場合は、文字列トランスフォーマ306は、語を置き換えるか、かつ/または仮定したグループを検索文字列202の検索語に透過に追加することができる。これを行うために、文字列トランスフォーマ306は、属性リスト310および/または検索制約条件のテーブル312に照らして検索語をテストすることができる。

10

【0033】

文字列トランスフォーマ306は、検索語をステミングして検索範囲を広げることができる。「ステミング」とは、検索語から接頭辞および/または接尾辞を取り除いて、検索語の語根を元の検索語の代わりとすることを可能にするプロセスである。これにより、元の検索語のより多くの変異形を捉えるヒットが増す。例えば、「divine」という単語は「divin」から派生することができるので、「divinity」、「divination」、「diviner」、「divinity」等と一致する。文字列トランスフォーマ306は、「Porter」アルゴリズムなど知られるステミングアルゴリズムを使用することができる。ステミングには、検索語をその通りの形で照会した場合には捉えられなかったヒットを生成する利点があるが、しばしば関連性のないヒットを生成するという欠点がある。この作用を、文字列トランスフォーマ306中、あるいは結果コンパイラ(下記)など例示的検索エンジン204の別の部分にある関連度ランカー(ranker)を使用することにより軽減することができる。一部の実施形態では、ステミングを行うには、検索速度を低下させる可能性のあるデータの前処理ステップが必要となる。したがって、一部の実施形態では、ステミングは、例えば出現頻度のインデックスから決定される稀にしか出現しない単語などの特別な単語に限って行う。

20

30

【0034】

文字列トランスフォーマ306は、追加的な語と文字列の変形とを行うことができる。単語の派生は、単一の語から検索語のセットを生成するものである。関連する単語のリストを使用することができる。例えば、検索語「NFL」の場合は「フットボール」が検索文字列202に付加され、したがって「NFL」から追加的な「フットボール」の検索が生じることになる。語「30」が検索文字列202中に出現する場合には、語「thirty」が検索文字列202に追加される。これらの関連語は、検索するデータに関連する例外リスト、および/またはユーザによって定義された検索制約条件テーブル312、あるいはプログラムの決定された検索制約条件テーブル312に格納された例外リストから得ることができる。単語の派生も検索の速度を低下させる可能性があり、また関連性のない結果を生成する可能性があるが、ヒットが見つからなかったであろう事例でもヒットを生成する。関連度ランカーを使用して、関連度の低い結果をヒットリスト222の下部に配置することができる。

40

【0035】

複数形化も使用することができる。複数形化とは、元の単語の単数形または複数形に基づいて追加的な検索語を生成する単純な形態の単語ステミングである。例えば、検索語が「baby」である場合は、追加的な検索語「babies」を含めることができる。

50

【 0 0 3 6 】

オプションで、文字列インタープリタ 2 1 4 または検索エンジン 2 0 4 の別のコンポーネントが、他の文字列の解釈機能およびプロセスに自然言語の構文解析を追加することができる。このタイプの構文解析には、「When is Star Wars shown in Dolby-Digital? (「スターウォーズ」がドルビーデジタルで放映されるのはいつ?)」などの検索フレーズを、ユーザの検索要求に合った実際的な検索語に変換することが含まれる。自然言語の構文解析は非常に強力であるが、一般に、文法はユーザによる名詞/動詞の意味の選択に限定される。また、(ブール型の語検索に比べて)自然言語の検索文字列 2 0 2 では単語数が多くなることを考えると、自然言語検索は、リモートコントロールのキーパッドを使用して検索語を入力しなければならない t e n - f o o t システムには最適な選択肢でない可能性がある。最後に、自然言語の構文解析には、実装するのにより多くの処理能力を必要とする可能性がある。

10

【 0 0 3 7 】

文字列トランスフォーマ 3 0 6 および例示的検索エンジン 2 0 4 の他のモジュールによって行われるこれらの機能およびプロセスは一定の順序で説明したが、これらの機能およびプロセスを提示した順序は、処理のパイプラインにおける提案される使用順序を意味するものではない。多くの場合、これらの各機能およびプロセスの順序および/または実行は、生成される結果にとって重要である。

【 0 0 3 8 】

属性アイデンティファイア 3 0 8 は、検索語の候補あるいは可能性のある検索語を属性リスト 3 1 0 の属性と比較するなどの各種の方法により属性検索語とテキスト検索語とを分離する。属性アイデンティファイア 3 0 8 はまた、テキスト検索の結果からフィードバックを受け取り、検索文字列 2 0 2 の語に除外のプロセスを適用して属性を見つける。この場合、属性アイデンティファイア 3 0 8 は、検索語間に各種のブール型関係を仮定することができる。例えば、検索語「グリーン マイル ドルビー」は、「'グリーン' OR 'マイル' AND ドルビー」または「'グリーンマイル' AND ドルビー」との解釈が可能であるが、2 番目の仮定されたブール型関係のセットのために、ドルビー(登録商標)デジタルまたはドルビー(登録商標)ステレオ音声の番組「ザ グリーンマイル」のインスタンスだけを返す検索を生む可能性がある。

20

【 0 0 3 9 】

言うまでもなく、属性のリスト 3 1 0 は、文字列の解釈を高速に行うためにコンピュータメモリに格納することができる。属性リスト 3 1 0 でおよそ 2 0 0 個のテレビ番組属性のリストを使用する場合は、恐らくは、ユーザが入力するテレビ番組に関連する実質的にすべての属性を属性アイデンティファイア 3 0 8 で識別し、テキスト検索の語と分離することができる。

30

【 0 0 4 0 】

コンピュータプログラムリスティング

本主旨により使用する X M L テキストファイルリスティングの 1 セットを添付の表として提供し、参照により本明細書に組み込む。この X M L テキストファイルのリスティングは、例示的な属性リスト(図 3 のブロック 3 1 0 など)である。

40

【 0 0 4 1 】

例示的な X M L (e x t e n s i b l e m a r k u p l a n g u a g e) のデータ構造として表した例示的な属性リスト 3 1 0 の表現を表 1 ~ 3 の「例示的属性リスト」に示す。この表現は、テレビ番組属性、要素、および値の選択の例だけを含んでいる。したがって、典型的な属性リスト 3 1 0 は、これらの表よりも多くのエントリ、または少ないエントリ、および/または異なる要素、属性、および/または値を含むことができる。

【 0 0 4 2 】

表の例示的な X M L バージョンの属性リスト 3 1 0 によって示すように、属性リスト 3 1 0 は、属性検索語の候補を与えられた時に真/偽または「存在する/存在しない」による一致だけを提供する連続的な属性要素の配列に限定する必要はない。属性リスト 3 1 0

50

は、少なくとも一部の属性についてのメタデータ311を含むことができ、ある属性のメタデータは、その属性を、検索で使用する検索リソースと値とに関係付ける。メタデータ311は、属性を、検索制約条件のテーブル312の論理検索条件のセットにも関係付けることができる。また、属性リスト310は、属性リスト310の属性を、検索文字列202に合致するヒットを生成するために検索するのが妥当な他の関連する属性および値に関係付けるインデックス313を含むことができる。

【0043】

一部の実施形態では、インデックス313および/またはメタデータ311は、検索文字列202を作成した際のユーザの意図に合致するヒットを生成できる比較的複雑な分岐ロジックおよび/または論理検索条件のツリー構造を含むことができる。例えば、属性アイデンティファイア308が検索語「ステレオ」について属性リスト310を調べる場合は、この属性の名前そのものを指定するか、かつ/または番組のステレオ属性を含む可能性が高いEPG情報中の検索フィールドを指定することにより一致を見つけることができ、例えばデータベースレコードの「ScheduleEntryAttribute」フィールドを見ることにより、「ステレオ」の番組を見つけることができる。さらに、候補の検索語が属性の基準に合致することが判明し、EPGデータの検索フィールドを指定すると、属性リスト310は、ステレオの番組を見つける検索語、キー、またはカテゴリを表すことができる例えば「MSEPGA__STEREO」などの1つまたは複数の正確な値も生じることができる。

【0044】

属性リスト310のインデックス機能は複数分岐することができる。例えば、属性アイデンティファイア308が検索語の候補「スポーツ」について属性リスト310を調べる場合、属性リスト310は、「スポーツ」を属性として識別することができ、さらにその語が「カテゴリ」タイプの属性であると識別し、属性アイデンティファイア308に、関連する番組を求めて3つの他のカテゴリ、「ニュース」、「特別番組」、および「MSEPGC__SPORTS」をクロスチェックするように指示することができる。

【0045】

文字列トランスフォーマ306について上述した機能およびプロセスを展開し、属性リスト310を介して属性と関連するメタデータを識別することにより、文字列インタープリタ214は、検索文字列202を作成した際のユーザの意図に近似しようとする検索語のセットに到達する。すなわち、文字列インタープリタ214は、大半のユーザがその検索文字列202を作成した場合に求められる内容に一致するヒットを返す確率が高いテキスト検索語と属性検索語とのセットを実現する。例えば、ユーザが検索文字列「SFドルビーレターボックス」を作成した場合、例示的な検索エンジン204は、レターボックス形式によりドルビー（登録商標）ステレオ音声で放映されるSF番組のヒットリスト222を返す。ユーザが検索文字列「ゴールデンタイム ネットワーク コメディ」を作成した場合、例示的な検索エンジン204は、ネットワーク局のゴールデンタイムのコメディ番組のヒットリスト222を返す。ユーザが検索文字列「格付けPGのドラマ」を作成した場合、例示的な検索エンジン204は、格付けがPGのドラマ番組のヒットリスト222を返す。ユーザが検索文字列「Discovery」を作成した場合、例示的な検索エンジン204は、番組名に「discovery」を含む番組、および例えば「DISCOVERY CHANNEL（登録商標）」などチャンネル名に「discovery」を含むチャンネル両方のヒットリスト222を返す。

【0046】

クエリジェネレータ314は、文字列インタープリタ214に通信的に結合されている。クエリジェネレータ314は、テキストクエリフォーマミュレータ316と属性クエリフォーマミュレータ318とを含むことができる。一実施形態では、テキストクエリフォーマミュレータ316および属性クエリフォーマミュレータ318は、EPGデータベース300およびウェブベースのEPGリソース334で検索する、可能な限り少なく（多くの場合1つ）限定的なクエリを作成する。複数のクエリを作成する場合は、結果を組み合わせる

10

20

30

40

50

ことができる。

【0047】

文字列インタープリタ214から受け取る各検索語について、クエリジェネレータ314は、EPGデータベース300またはウェブベースのEPGリソース334を照会する際に一致を見つける方式に影響する演算子を適用することができる。一実施形態では、各演算子は、検索対象のテキスト中で検索語を見つけなければならない（または見つけてはならない）場所または相対的位置を指定する。いくつかの例示的な演算子を以下で説明する。「完全一致」演算子は、その検索語が検索対象のテキストの所与の部分に正確に一致することを指定する。すなわち、クエリジェネレータ314が特定タイプのフィールドを検索している場合、検索語はそのフィールドに正確に一致しなければならない。「前方一致（フレーズ）」演算子は、検索対象のテキストフィールドの冒頭で検索語を見つけることを指定する。「前方一致（単語）」演算子は、検索語が、検索対象のテキストフィールド中のある単語の冒頭であることを指定する。「contains」演算子は、検索語を検索対象のテキストフィールド中のどこで見つけてもよいことを指定する。

10

【0048】

上述の各演算子について、「逆演算子」（NOTの形）を使用して、個々の条件が真である場合には検索エンジン204がヒットを生成しないことを指定することができる。検索演算子「AND」および「OR」も使用して、検索対象のテキストフィールドにすべての検索語を見つけなければならないか、または検索語の一部のみが一致すればよいことを指定することができる。

20

【0049】

データマネージャ319は、クエリジェネレータ314が生成する各種の検索語を使用してEPGデータベース300で検索を行う。本主旨の一部の実施形態では、データマネージャ319は、SQL（構造化クエリ言語）データベースエンジンまたはJETデータベースエンジンであっても、それらを用いてもよく、その場合は、結果コンパイラ320の機能の一部を兼ね備えることができる。したがって、データマネージャ319は、その内部で複数の検索パスを実行し、それらのパスを組み合わせたことができる。そのため、EPGデータベース300は、いくつかの異なるフォーマット（例えばXML、バイナリ、JET、SQLなど）をとることができるが、それらのフォーマットには記憶に適したものと、ネットワークを通じた伝送に適したものとがある。伝送媒体の制約に応じて、伝送に最適なフォーマットは、記憶および/または検索に最適なフォーマットと異なる場合がある。上記のフォーマットの選択とそれぞれの相対的な長所と短所との一部について以下で述べる。

30

【0050】

一部のガイドサービスでは、定義されたスキーマに一致するようにフォーマットされたXMLを使用して複数日分のガイドをクライアントに送信している。XMLは、理解が容易で、人間が読むことができ、既存のツールを多数所有するので扱いが利便なフォーマットである。しかしXMLはあまりコンパクトなフォーマットではない。その理由から、eHomeガイドは送信する前に圧縮してCABファイルにし、受信完了後に解凍する。

【0051】

大きなXMLファイルを容易に検索することはできない。例えば、10日分のEPGデータは、6.6メガバイトのファイルを満たす可能性がある。このサイズのXMLファイルをドキュメントオブジェクトモデル（DOM）にロードするには、およそ90メガバイトのランダムアクセスメモリが必要となる。検索のたびにそのような大きなファイルをスキャンするのは実際的ではない。これらの理由から、XMLフォーマットは主に伝送と永続的な記憶とに使用することができる。多数のクライアントがEPGデータファイルをロードする時、検索のために、非テキストデータはメモリに格納され、すべてのテキストベースのデータはメモリマップファイルに書き込まれる。一部のEPGデータAPIは、データおよび/または検索パターンの変更に対応するためにはカスタムデータベースを実装しなければならない。

40

50

【0052】

E P Gデータベース300の別の選択肢はバイナリフォーマットである。バイナリフォーマットはしばしば非常にコンパクトであるが、一般にはその生成とブラウズを行うにはカスタムツールに依存する。XMLファイルと同様にバイナリファイルは必ずしも検索には最適でなく、検索のために代替の中間記憶フォーマットを必要とする。

【0053】

検索エンジン204は、E P Gデータベース300を照会するのに、データベースツールとともに広くサポートされているJ E Tデータベースエンジンを使用することもできる。M D A Cと称する再配布可能なA P Iは、S Q L、A D O、またはO L E D Bを使用してJ e tデータベースを作成し、照会するためのリッチなA P Iを提供する。E P Gデータベース300にJ e tデータベースを使用する利点の1つは、業界標準のS Q Lコマンドを使用した検索の容易さと、共通ソフトウェアを使用した閲覧の容易さである。J e tを使用した検索のパフォーマンスは一般に、インデックスを適切に生成し、記憶すれば、多くの他の方式よりも高速である(1つの例外はS Q Lサーバ)。また、J e tデータベースを単一のファイルに格納して、E P Gデータベース300全体の操作を容易にする。

10

【0054】

J e tデータベースエンジンと同様に、S Q Lは、広くサポートされるE P Gデータベース300向けのデータベースプラットフォームであり、ツールとA P Iとの大規模なライブラリを備える。現在のよく知られるバージョンのS Q Lサーバはサーバマシンでしか作動しないが、M S D Eと称するバージョンのS Q Lサーバ(あるいはS Q L D e s k t o p E d i t i o n)は、現在の多くのオペレーティングシステムのソフトウェアバージョンにインストールすることができる。適切にインデックス付けされたS Q Lデータベースの検索は、恐らくは他の知られるデータ検索技術よりも高速で、柔軟性がある。ただし、S Q Lデータベースは、ネットワークを通じた伝送は容易ではない。M S D Eを使用するには、E P Gデータを別のフォーマットからS Q Lデータベースにロードした後処理ステップが必要となる。また、サーババージョンとは異なり、M S D Eは、自然言語検索あるいはフルテキストのインデックス付けをサポートするとは限らない(他の方法を用いればどちらもエミュレートすることは可能)。

20

【0055】

多くの場合、E P Gデータベース300のネイティブフォーマットは、例示的検索エンジン204が必要とする最終的な形態ではないことに留意されたい。例えば、データインデックスを生成する必要があるか、あるいは検索のために最適化されているがネットワークを介して送信するのは難しい代替のデータベースにデータを読み込む必要がある場合がある。一実施形態では、例示的検索エンジン204は、検索をより高速にするために不必要なデータを破棄し、ユーザ固有のデータを併合することさえある。

30

【0056】

データインデックスを使用する場合、検索エンジン204は、フルテキストのインデックス付けを実装することができる。これには、逆引きの単語(reverse word)インデックスをテキストベースのE P Gデータの各要素の中に生成することが必要となる。フルテキストのインデックス付けにより、E P Gデータベース300がより大きくなり、インデックスの生成にかかる時間が増すが、検索のパフォーマンスが大幅に高まる。

40

【0057】

検索エンジン204は、署名のインデックス付けも使用することができる。フルテキストのインデックス付けと同様に、署名のインデックス付けは、可能な検索の一致を迅速に識別する一方式である。E P Gデータベース300では、各データ行(番組、チャンネルなど)は、各単語中に出現する文字を示す、ハッシュに似たバイナリの署名を有する。検索語と検索対象のテキストの単語とにこのハッシュに似た計算を行い、2進の算術演算を使用して結果を照合すると、その単語が検索語に一致するかどうかを迅速に概算することができる。次いで、一致した単語を含む行がヒットを生成すべきかどうかを判定するため

50

に、よりプロセッサ集約的な第2の検索を必要とする場合がある。

【0058】

上記で触れたように、本主旨は、単一の検索文字列202からテキスト検索語と属性検索語とを導き出す検索エンジン204の能力によって制限されない。例示的な検索エンジン204は、単一の検索文字列202からテキスト検索語と属性検索語とを導き出すのに加えて、UI中の各種の入力フィールドなど複数のソースから検索語を導き出すことができる。例えば、一部の製品では、「ステレオ」、「クローズドキャプション」、「開始時刻」、「公開年」などの属性をしばしば検索語として利用することができるが、これらは一部のUIではチェックボックスやラジオボタンを介して入力される。これらの属性は、例示的な検索エンジン204により検索語として検索APIに渡すことによって対処することもできる。この対処は、単一の検索文字列202から他の属性検索語とテキスト検索語とを導き出すことから検索エンジン204を制限しない。

10

【0059】

例示的な検索エンジン204は、ユーザが検索を行いたいフィールドを選択するようにユーザを促すことができる。そのため、例示的な検索エンジン204は、すべてのドメイン(番組名、説明、チャンネル、出演者など)を自動的に検索し、各自の予測される関連度によってヒットをランク付けするのに加えて、ユーザがフィールドおよび/または検索ドメインを手動で絞りこむことも可能にする。

【0060】

例示的な検索エンジン204は、データマネージャ319および文字列インタープリタ214に通信的に結合された結果コンパイラ320も含む。結果コンパイラ320中のヒットレシーバ322は、各種の検索モダリティ、または複数のデータベースを検索する単一の検索モダリティからヒットを収集する。複数の検索は同時に行うことができるので、ヒットレシーバ322は、さらにコンパイルを進めるまで一時的にヒットを格納することができる受け取りマニホールドまたはバッファを備えることができる。

20

【0061】

例示的な検索エンジン204では、検索コンパレータ/フィルタ324が、ヒットレシーバ322および文字列インタープリタ214と通信的に結合されている。検索コンパレータ/フィルタ324は、文字列インタープリタ214が確立した文字列の解釈に照らしてヒットをテストする。上述のように、文字列インタープリタ214は、文字列トランスフォーマ306および属性アイデンティファイア308を介して、検索文字列202を作成した際のユーザの意図の実際的な仮説または解釈に到達する。結果コンパイラ320では、ヒットが解釈に合致しない場合は、そのヒットをフィルタリングすることができる。すなわち、表示されるヒットリスト222から排除するか、表示されるヒットリスト222中でランク付けする目的で格下げすることができる。一実施形態では、検索コンパレータ/フィルタ324は、検索のセットを1つずつスキャンして、各結果が本当にヒットであるかどうかを判定する。これは、すべてのヒットを含めるのと同時に、すべてのミスを排除するのに十分な精度までクエリを分解(factor)できなかった場合に行われる可能性がより高い。解釈の一例として、元の検索文字列202が「ダック ドルビー」であり、文字列インタープリタ214が、「ドルビーデジタル方式のドナルドダック」の解釈よりも「トーマスドルビーによる『ハワード ザ ダック』」の解釈を支持する場合には、ドナルドダックのヒットは、ヒットリスト222から完全に排除されるか、または少なくともより低いランクを与えられる。

30

40

【0062】

結果ランカー326は、すべての検索が終了した時か、あるいは一部の実施形態ではヒットレシーバ322が結果をばらばらに受け取り、その結果が検索コンパイラ/フィルタ324を通じて渡された時に、関連度またはその他の基準によりヒットを表示のために編成することができる。検索エンジン204が生成したヒットはしばしばランク付けして、最も可能性が高い結果、すなわち検索文字列202を作成した際のユーザの意図解釈に最も近いヒット(文字列インタープリタ214によって仮定される)を最初に提示する方式

50

で、ヒット結果を表示のためにソートすることができる。例えば、各検索語の正確な一致には100%の関連度ランキングを与え、最も高い優先度を与えることができる。ステミングやスペル補正などの何らかの変形の後を行った時にしか検索対象のテキストと一致しないヒットには、通常はより低い関連度を与える。

【0063】

包含フィルタ328は、結果ランカー326と通信的に結合することができる。一部の變形形態では、包含フィルタ328は一定の方式でフィルタリングするようにプログラマ的に設定することができ、一方、他の變形形態ではユーザによって調整することができる。例えば、包含フィルタ328は、関連度が100%のヒットが少なくとも1つ見つかった場合には関連度が100%未満のヒット結果を完全に排除するように設定または調整することができる。他の場合、包含フィルタ328は、長い結果リストを切り捨て、その結果の一部が100%の関連度を有しても一定数の上位の結果だけを表示するように設定または調整することができる。さらに別の場合、包含フィルタ328は、一定のパーセントより上に割り当てられたランクを有するヒット結果の表示だけを許可することができる。したがって、検索が「Where the Red Fern Grows」に対するものであり、唯一のヒットが関連度ランクがわずかに15%の「The Red Skelton Hour」である場合、包含フィルタ328は、「一致する結果は見つからない」と表示する。

10

【0064】

包含フィルタ328は、結果の数を定義された値に制限するか、かつ/またはペアレンタルコントロールおよび/または検索を行うユーザに応じて結果を制限することも可能である。録画のコンテキストでは、包含フィルタ328は、今後放映がない結果を除くことができる。

20

【0065】

表示アップデータ330は結果コンパイラ320に含まれ、ヒットリスト222をユーザに提示する表示332を制御する。検索結果のヒットリスト222を提示するには多くの可能な方式がある。ヒットは、アルファベット順または日付(時間)順に表示することができる。結果は、関連度によってグループ化し、ランク付けすることもできる。一部の實施形態では、表示アップデータ330は、結果の情報センター(clearinghouse)であり、すなわち表示内容の記録を保持することができる。ヒット結果をばらばらに受け取る場合は、表示アップデータ330は、古い結果に新しい結果を組み合わせ、どの結果を表示するかを決定する統合機能を行うことができる。個別に行われる検索の際には、表示アップデータ330は、表示を区分して検索結果のグループを受け取る順序を表示するか、かつ/または表示を区分してどのヒットがどの検索語に関連するかを示すことができる。漸進的検索の際(下記でより詳しく説明する)には、表示アップデータ330は、ユーザが新しい数字を検索文字列202に追加する(例えばタイプおよび/または入力する)たびに、表示された結果のすべてまたは大半を破棄することができる。したがって、漸進的検索では、表示アップデータ330は、入力されている検索文字列202の意味が展開するのに従って検索結果を始めから再構成する。

30

【0066】

無論、一部の實施形態では、結果コンパイラ320の各種のモジュールはともに組み合わせても、異なる順序としてもよい。變形形態では、モジュールの一部を排除しても、他の従来のデータベースモジュールを追加してもよい。

40

【0067】

例示的検索エンジン204の變形形態を、実装で使用する利用可能なハードウェアおよび/またはソフトウェア環境の速度と処理能力に合うように適合させることができる。検索の実行には、ユーザが許容できる以上の時間がかかる可能性がある。検索エンジン204は、検索を実行しながら、ヒットが見つかりと非同期にヒットを返すことができる。あるいは、例示的検索エンジン204は、インテリジェントに検索を部分に分割し、検索が完了するまで結果のセットを周期的に更新することができる。例えば、ステミングと正規

50

化の機能を適用する前に正確な単語の一致のテキスト検索を行い（そして結果を返し）、後のパスで返すことができる。検索は時間を区分して行うこともでき、予定される放送時間に従って分割したグループごとに番組を検索する。

【 0 0 6 8 】

別のオプションでは、例示的検索エンジン 2 0 4 が追加的なデータストアを利用できる場合は、検索ドメインを拡大することができる。検索エンジン 2 0 4 は、各データストアにわたって検索を実行し、結果セットを最終的に確定する前に結果を併合することができる。これを使用して、例えば、ローカルの E P G データベース 3 0 0 の検索にウェブベースの E P G リソース 3 3 4 の検索を組み合わせることができる。また、ローカルの媒体に事前に記録された番組を検索ドメインに追加することもできる。そのような検索は、複数の A P I を通じて実行し、表示する前に併合することができる。

10

【 0 0 6 9 】

例示的検索エンジン 2 0 4 は、後の実行および/またはバックグラウンドでの実行のために検索基準を持続することもできる。ユーザは、検索文字列 2 0 2 を再入力せずに呼び戻すことが可能な検索を作成することに興味を持つ可能性がある。ユーザがすでに閲覧したヒットを追跡し、一部の実施形態では、容易に目を通すことができるように結果の一部を隠す、分離する、あるいは「グレイアウト（選択不可能）」することも有用である。E P G データベース 3 0 0 の情報が更新されるたびに永続的な検索を実行し、ユーザが興味を持つ予定されるイベントをユーザに知らせることができる。

【 0 0 7 0 】

検索基準は例えばマクロ中にシリアルライズし、自動的な記録、すなわち検索の履歴リストを定義するためのシリアルライズ機構として機能することができる。検索語の各シリアルライゼーションにトークンを割り当て、ユーザがトークンを選択することによりシリアルライゼーションを選択できるようにすることもできる。シリアルライズした検索は、電子メール、ウェブリンクを介して他のユーザと交換するか、かつ/または「クリックして記録する」リンクのために T V 放送中に埋め込むことができる。例示的検索エンジン 2 0 4 は、1 つまたは複数のシリアルライゼーションフォーマットを定義することもできる。検索および/またはシリアルライゼーションは、コンパクトに格納するためにバイナリ形態で格納するか、電子メールで容易に送信できるように A S C I I 形態で表すことができる。シリアルライゼーションは、単独で、または自動的な記録と連動して使用することができ、ここで、ユーザは検索を定義し、その検索基準に一致する今後のすべての番組（重複するインスタンスを除く）を記録するように指示することができる。

20

30

【 0 0 7 1 】

例示的な検索

図 4 に、本主旨による例示的な検索 4 0 0 を示し、これは、例示的検索エンジン 2 0 4 などの検索エンジンによって行うことができる。「d o L B y W E S T - r a t e d ' g '」などの比較的不明瞭な検索文字列 2 0 2 をユーザから、または例示的検索エンジン 2 0 4 を介して永続的なストレージから受け取る。この検索文字列 2 0 2 は、単一の簡潔なクエリに限定できない可能性がある。しかし、例示的検索エンジン 2 0 4 は、複数の検索を実行して可能性のあるヒットを見つける高い多機能性を持つ。例示的ノーマライザ 3 0 2 は、この文字列を、標準化した例示的な形態、すなわち正規化した検索文字列 4 0 2 「D O L B Y W E S T R A T E D G」に変形する。この場合、ノーマライザ 3 0 2 は、「d o L B y」中の一部の文字の大文字/小文字を変えて、例えば E P G データベース 3 0 0 で検索するテキストの大文字/小文字の使用に一致するテキスト検索語を実現している。ノーマライザ 3 0 2 は、単語間のスペースと「W E S T」と「r a t e d」の間のダッシュを境界として使用して検索語を分離する。ノーマライザ 3 0 2 は、「g」から引用符を除去し、検索語の文字列全体を 1 種類の文字、すなわち大文字にしている。

40

【 0 0 7 2 】

正規化した検索文字列 4 0 2 を例示的な文字列インタープリタ 2 1 4 に渡し、インタープリタ 2 1 4 は、検索文字列 2 0 2 を作成した際のユーザの意図のかなり確定的な解釈に

50

達することができる。かなり確定的な解釈は、ユーザが入力した属性検索語が、例えば属性リスト310を利用してテキスト検索語から明確に区別できる場合に到達することができる。文字列インタプリタ214は、テキストの語と属性の語との確定的な分離に達することができる場合があるが、その場合、文字列インタプリタ214は、代替の検索および/または条件付きの検索を行うために検索語をクエリジェネレータ314に渡すことができる。すなわち、ある語についてあいまい性がある場合、その語をテキスト検索に使用し、また属性検索にも使用して、例えばそれぞれのタイプの検索でいくつのヒットが得られるかを調べることができる。したがって、検索語「DOLBY」は、「The Dolby Digital Experience」などの番組名中の単語、「Thomas Dolby」など番組制作者の属性、あるいは「DOLBY（登録商標）デジタル」など番組の性質的な属性である可能性がある。例示的なクエリジェネレータ314は、これらの可能性をそれぞれ、2つの属性検索406（DOLBY（登録商標）デジタルとDOLBY（登録商標）ステレオサウンド）、代替のテキスト検索410（番組名中の語としての「dolby」）、および代替の属性検索412（「Dolby」という名前の出演者および関係者）としてテストすることができる。属性としての「DOLBY」の関連度ランキングが非常に高いので、クエリジェネレータ314は、基本的なテキスト検索406として「DOLBY」を検索しないことを決定することができる。ただし、「DOLBY」がテキスト検索で見つかる場合は、代替のテキスト検索410で見つかる。したがって、代替の検索または条件付き検索は、より高い優先度を持つヒットが検索によって返されない場合の予備として使用するか、かつ/または、代替の検索410、412は、検索文字列202についてより多くの情報を得る、すなわち検索文字列202の解釈を強化するため、あるいは新しい検索手段を求めるための探知または修正機構として使用することができる。

【0073】

正規化された検索文字列402の検索語「WEST」は、番組名中の単語である可能性があり、または一部の方式では番組に関連付けられた人の属性である可能性があるため、代替の属性検索412でも使用することができる。ただし「WEST」が文字列インタプリタ214による属性リスト310のスキャンで見つからない場合、文字列インタプリタ214は、「WEST」は番組名検索のための主要なテキスト語であろうとの強い仮定を形成することができる。

【0074】

文字列インタプリタ214は、検索語「RATED」が検索語「G」より前にあり、そのため2つの語は「RATED G」と解釈される可能性が高く、それに従って検索されることを検出する。単一の語「RATED」は代替のテキスト検索410で照会することができるが（すなわち番組名中に「rated」という単語を含む番組）、そのような代替の検索は、「RATED」と「G」がよく知られる属性の適切な順序で共に検索文字列202に見つかる場合には、低い関連度ランクを与えられる可能性が高い。

【0075】

クエリジェネレータ314および/または検索エンジン204は、各種のテキスト検索406、属性検索408、代替のテキスト検索410、および代替の属性検索412を逐次行う。同時に行える検索、あるいは検索文字列202の長さや意味によって決まる論理的順序で行うことが可能な検索の数には制限がない。場合によっては、例示的検索エンジン204は、すべての検索を行い、すべての検索を通る「第1のパス」の後に関連度に従って結果をリストする。他の場合、例示的検索エンジン204は、文字列インタプリタ214またはクエリジェネレータ314によって設定された論理条件のツリーに従う。この論理条件は、例えば「テキスト検索406でAが得られた場合は、検索語Xを使用して属性検索408および代替のテキスト検索410に進むが、テキスト検索406でBが得られた場合は、3回の代替の属性検索412に進む」等である。このロジックはさらに、Aがヒットの量であり、Bが特定のテキストの一致を含む特定のヒットであることを規定することができる。

10

20

30

40

50

【0076】

各タイプの検索は、その特定の検索の結果を含む中間のヒットリストを返すことができる。検索語「WEST」に対するテキスト検索のヒットリスト414は、「How the West was Won」、「West Side Story」、「The Wild Wild West」、「The West Wing」などの番組名を含む可能性がある。「WEST」のような比較的曖昧な単一の検索語の場合、クエリジェネレータ314は、恐らくは汎用的な「in-string（文字列中）」演算子を使用して、単語「WEST」を含む番組名を見つける。

【0077】

属性検索のヒットリスト415は、「DOLBY（登録商標）ステレオサウンド」、「DOLBY（登録商標）デジタル」、「rated G」の属性を有する多数の番組名を含むか、あるいは返すことを試みる。クエリジェネレータ314は、これらの属性がかなり一般的であることを知っているため、一般的な属性に対する属性検索408を割愛し、代わりに、テキスト検索406の条件および/またはフィルタとしてそれらの属性を含める。すなわち、検索語「WEST」を含むテキスト検索（406）のヒットは、そのヒットが少なくとも1つの「Dolby」の属性および「G」の格付けにも関連付けられていない場合には許可されない。

【0078】

「WEST」のテキスト検索406では多数の種々のヒットが得られるので、検索エンジン204は、「DOLBY」、「RATED」、および「G」について代替のテキスト検索410を行うものと考えられ、恐らくは各検索のヒットの計数を行ってさらなる情報を収集する。恐らく、「Dolby West」という題の番組があった場合、検索エンジン204は、その正確な番組名の一致を鑑みて最初の解釈を変更する。すなわち、検索エンジン204は、当初は隣接する検索語をテキスト検索語と属性検索語との混合と推定した。

【0079】

代替の属性検索412は、例えば出演者と関係者との名前を属性として使用する方式で追加的な結果を返すことができる。したがって、代替属性検索の語「DOLBY」および「WEST」は、Thomas Dolbyの映画、Steve Dolbyの映画、Dominic Westの映画、およびShane Westの映画、の4カテゴリを返す。代替の属性検索語「G」も、格付け「G」に関連付けられた多数の番組名を返すが、それらはそのサイズのため、あるいはその検索がすでに属性検索408に包含されているためにフィルタリングして排除される。

【0080】

最後に、各種検索406、408、410、412によるすべてのヒットリスト414、416、418、420を結果コンパイラ320で処理する。検索コンパレータ/フィルタ324は、受け取ったヒットと、文字列インタープリタ214によって確立された検索文字列202の解釈とを比較することができる。結果ランカー326は、各ヒットに関連度のランク付けを与えることができる。包含フィルタ328は、累積したヒット、ランク付けが低いヒット、および/またはオーバーフロー（多すぎる）ヒットと、検索文字列202の解釈に合わないヒットとを取り除くことができる。図のフィルタリング後のヒットリスト422で、文字列インタープリタ214がすでに検索語間のブール型関係を判定している場合、すなわち、番組名は語「WEST」を含み、かつその番組はDOLBY（登録商標）デジタルであり、かつ格付けが「G」であるという検索文字列202の解釈に達している場合は、「How the West was Won」および比較的少ない他の番組名だけが表示されることになる。無論、例示的検索エンジン204が到達した解釈がユーザの意図に近くない（ユーザが有用な結果がないと判定することで明らかになる）場合、特に多くの実装ではプロセス全体がほぼ瞬時に行われるのでユーザは常に検索を修正することができる。

【0081】

10

20

30

40

50

例示的な漸進的検索

図5に、本主旨による例示的な漸進的検索を示す。例示的検索エンジン204などの検索エンジンのモジュールを構成するのに利用できるハードウェアおよび/またはソフトウェアは強力で高い検索速度を提供できるので、例示的検索エンジン204の一実施形態では漸進的検索を行うことが可能である。

【0082】

漸進的検索では、ユーザがUIに文字をタイプあるいは入力すると、結果を収集すると同じ速さで検索結果を動的に更新する。各文字が入力される（または削除される）と、新たな一連の検索が開始される。非漸進的検索では、ユーザは、「Enter」キーを押すか、かつ/または何らかの他の方式で検索文字列202が完成したことを通知することによって手動で検索を開始することが必要である。

10

【0083】

図5に、数週間分のEPG情報と、デジタルビデオ録画システムのローカルのハードドライブにローカルに録画した番組の集合を有するEPGデータベース300で検索をするために、検索文字列202「The West Wing Dolby Election」を入力する際のスナップショットを5つ示している。ブロック502では単語「The」が入力され、例示的検索エンジン204がゼロ個の結果504を掲示する。この理由としては、ノーマライザ302が「The」を冠詞と認識し、冠詞のみに基づく検索を抑制したことが考えられる。しかし、「The」の「T」が入力された時に、非常に高速のシステムであれば、例えば「Mr. T」など番組名中に「T」を単独で含む番組を返す。

20

【0084】

ブロック506では単語「The West」が入力されており、例示的検索エンジン204は、「The West Wing」、「How the West was Won」などの番組名を含む50個の結果508を返す。

【0085】

ブロック510では単語「The West Wing」が入力されており、ヒット数が50から10個の結果512に絞られる。恐らくはこの時点で番組名が確定し、10個の結果は恐らくは、それぞれが各自の番組ID番号を有するテレビと映画との「The West Wing」のエピソードである。

【0086】

ブロック514では単語「The West Wing Dolby」が入力されており、これを満たすヒットはさらに5個の結果516に絞られる。先の10個のエピソードのうち、DOLBY（登録商標）デジタルでない、または何らかの他のDolby属性を持つものが除外されている。

30

【0087】

ブロック518では、単語「The West Wing Dolby Election」が入力されている。これで結果520は2つだけになる。この2つのヒットは、恐らくはそのエピソードの説明フィールド中の単語として「Dolby」属性と、「Election」属性を有する「The West Wing」のエピソードからなる。

【0088】

検索文字列202が入力されるのと同時に各種のヒット結果が表示されるので、例えば、ブロック506で単語「The West」だけが入力された時など、プロセスの初期に希望の「The West Wing」のエピソードが表示332に表示され、ユーザは、直ちにその希望のエピソードをハイライト表示し、選択することができるはずである。

40

【0089】

例示的方法

図6に、本主旨によるフリーテキストおよび属性の検索を行う例示的方法600を示す。この方法600は、図2、3および5に示す例示的検索エンジン204などのモジュールによって行うことができる。流れ図で、動作は各ブロックに要約して記載している。こ

50

の動作は、プロセッサによる実行が可能なハードウェア命令および/またはマシン可読命令(ソフトウェアまたはファームウェア)で実行することができる。

【0090】

ブロック602で、検索文字列を受け取る。この文字列は、1つまたは複数のUIを通じてユーザから受け取るか、またはストレージから格納されたマクロあるいはシリアライズされた検索として受け取ることができる。検索文字列は、一度に一文字ずつ受け取ることもでき、新しい文字が受け取られ、長くなっていく検索文字列に追加されるたびに新しい検索のセットを漸進的に行うことができる。受け取る検索文字列は、属性がテキスト形態で書かれたテキスト文字列であると予測される。ただし、受け取った検索文字列は正規化して、非テキストの要素を標準化されたテキストまたは標準化された検索記号に変換するので、受け取る検索文字列は、ASCIIセットに割り当てられたアルファベット文字、記号、トークンおよび/または国際的なアルファベット文字の任意の組合せであってよい。一部の變形形態では、検索文字列は、チェックボックスへの入力、ソフトキー入力、クリックされたアイコンなども含むことができる。

10

【0091】

ブロック604で、検索文字列を正規化する。正規化は、受け取った文字列を特定のEPG情報ソースを検索するために理解できる検索可能な語、すなわち検索するEPG情報のデータフォーマットに一致するフォーマットを有する語に変換するものである。この目的は、テキスト、大文字/小文字、および文字の形式の単なる違いをなくして、その形式の違いがなければヒットを生成するはずの検索語とEPG情報との一致を妨げないようにすることである。したがって、正規化は、検索語および恐らくは対象とするデータベース情報両方のスペルチェック、オートコレクト、語の分離、大文字/小文字の標準化、記号の正規化、冠詞の除去、句読点の除去などからなるツールキット全体を指す。例えば、「mash」と「M*A*S*H」とは、検索語、検索対象の情報、またはその両方でどちらの形態が見つかっても常に一致する。

20

【0092】

ブロック606で、検索文字列をテキスト語と属性語とに分離する。この分離は、検索文字列の解釈に影響し、検索文字列の解釈と同時に行われる。テキスト語と属性語とへの分離の結果、正規化によって得られる検索可能な語の何らかの解釈と、何らかの變形と置換とが行われる。平均的な検索文字列は通常は、わずか5~6個あるいはそれ以下の検索語なので、様々な大きさのグループ化とすべての可能な順序でその検索語のすべての可能な組合せを試みる方式に従うことができる。したがって、検索文字列をスクランブルして、検索語の様々なグループ化と組合せを試すことができる。

30

【0093】

各種の検索語と検索フレーズ(グループ化)との候補を生成すると、それらを既存の属性リストに照らしてテストまたは解釈して、どの検索語およびフレーズが知られている属性であるかを判定することができる。検索文字列に行われる各種の検索は、同時に行うのではなく、上述の検索文字列の解釈の際に検索の制約条件として設定される条件付きの検索ロジックに従って逐次あるいは反復的に行われることがあるので、検索語の候補は、すでにその検索文字列に行われている可能性のある検索結果があれば、それに照らしてテストすることもできる。

40

【0094】

検索フレーズ中の単語を迅速にスキャンすることにより、ユーザが意図した他の詳細事項も明らかにすることができ、これは解釈の別の要素を構成する。「in」、「and」、「is」、「the」など短い冠詞や接続語が多数ある場合、検索文字列は自然言語入力として解釈される。ただし、数個の簡潔な語だけを含む短い検索文字列は、プール型の検索として解釈される可能性が高い。

【0095】

ブロック608で、各テキスト語および各属性語について1回または複数回の検索を行う。存在するテキスト語と属性語、およびそれらの相互の関係に応じて、比較的複雑な検

50

索クエリを確立することができる。ありそうもないが、説明のための例として、検索文字列「西部劇 1980以後 SF俳優(western after 1980 sci-factors)」は、「1980年より後に製作され、SF映画にも主演している俳優のみが出演する西部劇の検索」と解釈される。このような検索には、「西部劇」のカテゴリを検索し、減法のロジックを使用して1980年以前に製作された西部劇と、SF映画にも主演した俳優が出演していない西部劇とを除外することが伴う。多くのテキスト語および属性語の検索は、受け取る結果に応じて、同時に行うことも、逐次または繰り返し行うこともできる。したがって、例示的検索エンジン204など例示的方法600を行う検索エンジンは、検索のマニホールドとして機能することができ、「同時にあらゆる方向に」行われる検索を開始する。

10

【0096】

ブロック610で、表示のために検索結果を組み合わせる、ランク付けする、および/またはフィルタリングする。組合せは通常、属性語とテキスト語とを分離する際に仮定されるプール型の関係に従って行われる。すなわち、検索結果を組み合わせる方式は、部分的には、テキスト語と属性語とを分離する際、およびクエリを形成する際に行われる検索文字列の解釈によって事前に決定される。結果に関連度のランク付けを与える方式も、プロセスの早い段階で行われる解釈によって決まる。フィルタリングは、事前に設定されたユーザの表示の嗜好によって決まるか、または関連度のランク付けが閾値以下の結果を表示から除外する方式によって決まることことができる。

20

【0097】

ブロック612で検索結果をユーザに表示する。結果は、すべての検索が終わってから表示することができるが、受け取り、コンパイルするのと同時に表示してもよい。漸進的検索では、結果は、検索文字列が入力され、入力される各文字により新しい検索セット全体が開始されるのと同じ速度で表示される。

【0098】

上述の例示的方法600は、本主旨の一例に過ぎない。この方法は限定的なものではなく、ユーザが利用可能な多くの変形形態を備えるプロセスの一例を説明するものである。

【0099】

例示的コンピューティングデバイス

図7に、本主旨の態様を実施するための環境として適した例示的コンピュータ700を示す。例示的コンピュータ700の構成要素には、これらに限定しないが、処理装置720、システムメモリ730、およびシステムメモリ730を含む各種のシステム構成要素を処理装置720に結合するシステムバス721が含まれる。システムバス721は、各種のバスアーキテクチャの任意のものを使用したメモリバスまたはメモリコントローラ、ペリフェラルバス、およびローカルバスを含む数タイプのバス構造のいずれでもよい。これらに限定しないが、例として、このようなアーキテクチャには、ISA(Industry Standard Architecture)バス、MCA(Micro Channel Architecture)バス、EISA(Enhanced ISA)バス、VESA(Video Electronics Standards Association)バス、およびメザンバスとも称されるPCI(Peripheral Component Interconnect)バスが含まれる。

30

40

【0100】

例示的コンピュータ700は通常各種のコンピュータ可読媒体を含む。コンピュータ可読媒体は、例示的コンピュータ700がアクセスすることができる任意の利用可能媒体でよく、揮発性および不揮発性の媒体、取り外し可能および取り外し不能媒体が含まれる。これに限定しないが、例としてコンピュータ可読媒体は、コンピュータ記憶媒体と通信媒体を含むことができる。コンピュータ記憶媒体には、コンピュータ可読命令、データ構造、プログラムモジュール、またはその他のデータなどの情報を記憶するための任意の方法または技術に実装された揮発性および不揮発性の媒体、取り外し可能および取り外し不能の媒体が含まれる。コンピュータ記憶媒体には、これらに限定しないが、RAM、ROM

50

、EEPROM、フラッシュメモリまたは他のメモリ技術、CD-ROM、デジタル多用途ディスク(DVD)または他の光ディスクストレージ、磁気カセット、磁気テープ、磁気ディスクストレージまたは他の磁気記憶装置、あるいは、所望の情報の記憶に用いることができ、例示的コンピュータ700によるアクセスが可能な任意の他の媒体が含まれる。通信媒体は、通常、搬送波または他の搬送機構などの変調データ信号にコンピュータ可読命令、データ構造、プログラムモジュール、または他のデータを実施し、任意の情報伝達媒体を含む。用語「変調データ信号」とは、信号中に情報を符号化する方式でその特性の1つまたは複数を設定または変化させた信号を意味する。例として、通信媒体には、配線ネットワークまたは直接配線接続などの配線媒体と、音響、RF、赤外線、および他の無線媒体などの無線媒体が含まれるが、これらに限定しない。上記の媒体のいずれの組合

10

【0101】

システムメモリ730は、読み取り専用メモリ(ROM)731およびランダムアクセスメモリ(RAM)732など、揮発性および/または不揮発性メモリの形態のコンピュータ記憶媒体を含む。起動時などに例示的コンピュータ700内の要素間の情報転送を助ける基本ルーチンを含むBIOS(basic input/output system)733は、通常ROM731に記憶される。RAM732は通常、処理装置720から即座にアクセス可能な、かつ/または現在処理装置720によって操作中のデータおよび/またはプログラムモジュールを含む。これに限定しないが、例として、図7にはオペレーティングシステム734、アプリケーションプログラム735、例示的検索エンジン204、他のプログラムモジュール736、およびプログラムデータ737を示している。例示的検索エンジン204は、ソフトウェアとしてランダムアクセスメモリ732内に図示しているが、例示的検索エンジン204の他の実装は、ハードウェアであっても、ソフトウェアとハードウェアとの組合せであってもよい。

20

【0102】

例示的コンピュータ700は、この他の取り外し可能/取り外し不能、揮発性/不揮発性のコンピュータ記憶媒体も含むことができる。単なる例として、図7には、取り外し不能、不揮発性の磁気媒体の読み取りまたは書き込みを行うハードディスクドライブ741、取り外し可能、不揮発性の磁気ディスク752の読み取りまたは書き込みを行う磁気ディスクドライブ751、およびCD-ROMや他の光学媒体などの取り外し可能、不揮発性の光ディスク756の読み取りまたは書き込みを行う光ディスクドライブ755を示す。例示的動作環境で使用できるこの他の取り外し可能/取り外し不能、揮発性/不揮発性のコンピュータ記憶媒体には、これらに限定しないが、磁気テープカセット、フラッシュメモリカード、デジタル多用途ディスク、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどが含まれる。ハードディスクドライブ741は通常、インタフェース740などの取り外し不能のメモリインタフェースを通じてシステムバス721に接続され、磁気ディスクドライブ751および光ディスクドライブ755は通常、インタフェース750など取り外し可能なメモリインタフェースによってシステムバス721に接続される。

30

【0103】

上記で説明し、図7に示すドライブとそれに関連付けられたコンピュータ記憶媒体は、例示的コンピュータ700のコンピュータ可読命令、データ構造、プログラムモジュール、およびその他のデータの記憶を提供する。例えば図7では、ハードディスクドライブ741にオペレーティングシステム744、アプリケーションプログラム745、他のプログラムモジュール746、およびプログラムデータ747を記憶している。これらのコンポーネントは、オペレーティングシステム734、アプリケーションプログラム735、他のプログラムモジュール736、およびプログラムデータ737と同じものでも、異なるものでもよいことに留意されたい。ここではオペレーティングシステム744、アプリケーションプログラム745、他のプログラムモジュール746、およびプログラムデータ747には、それらが少なくとも異なるコピーであることを表すために異なる参照符号

40

50

をつけている。ユーザは、キーボード762、および一般にはマウス、トラックボール、タッチパッドと称するポインティングデバイス761などの入力装置を通じて例示的コンピュータ700にコマンドと情報を入力することができる。他の入力装置(図示せず)には、マイクロフォン、ジョイスティック、ゲームパッド、衛星放送受信アンテナ、スキャナなどがある。これらおよび他の入力装置は、システムバスに結合されたユーザ入力インタフェース760を通じて処理装置720に接続することが多いが、パラレルポート、ゲームポート、あるいはユニバーサルシリアルバス(USB)など他のインタフェースおよびバス構造によって接続することも可能である。モニタ791または他タイプの表示装置も、ビデオインタフェース790などのインタフェースを介してシステムバス721に接続される。コンピュータは、モニタに加えて、スピーカ797やプリンタ796など他の周辺出力装置も含むことができ、それらは出力周辺インタフェース795を通じて接続することができる。

10

【0104】

コンピュータ700は、リモートコンピュータ780など1つまたは複数のリモートコンピュータとの論理接続を使用するネットワーク環境で動作することができる。リモートコンピュータ780はパーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピアデバイス、あるいはその他の一般的なネットワークノードでよく、図7にはメモリ記憶装置781しか示していないが、通常は例示的コンピュータ700との関連で上記で挙げた要素の多くまたはすべてを含む。図7に示す論理接続には、ローカルエリアネットワーク(LAN)771とワイドエリアネットワーク(WAN)773が含まれるが、他のネットワークを含むことも可能である。このようなネットワーキング環境は、オフィス、企業内のコンピュータネットワーク、イントラネット、およびインターネットに一般的に見られる。

20

【0105】

LANネットワーキング環境で使用する場合、例示的コンピュータ700はネットワークインタフェースまたはアダプタ784を通じてLAN771に接続される。WANネットワーキング環境で使用する場合、コンピュータ700は通常、インターネットなどのWAN773を通じて通信を確立するためのモデム772またはその他の手段を含む。モデム772は内蔵型でも外付け型でもよく、ユーザ入力インタフェース760または他の適切な機構を介してシステムバス721に接続することができる。ネットワーク環境では、例示的コンピュータ700との関連で図示したプログラムモジュール、またはその一部は遠隔のメモリ記憶装置に格納することができる。これに限定しないが、例として図7ではリモートアプリケーションプログラム785がメモリ装置781に常駐している。図のネットワーク接続は例示的なものであり、コンピュータ間に通信リンクを確立する他の手段を使用できることは理解されよう。

30

【0106】

言うまでもなく、本趣旨は、適切なハードウェアを有するセットトップボックス、ゲーム機、およびその他多くの種類の処置装置など多くの他の環境でも使用することができる。それらの環境はそれぞれ、例示的コンピュータ700について上記で述べた構成要素のすべてではないが一部を含むことができる。

40

【0107】

結論

上述の本趣旨は、ハードウェア、ソフトウェア、またはハードウェアおよびソフトウェア両方に実施できることに留意されたい。一定の実施形態では、例示的システムとそれに関連する方法は、コンピュータによって実行されるプログラムモジュールなどのコンピュータ実行可能命令に概ね即して説明することができる。一般に、プログラムモジュールには、特定のタスクを行うか、または特定の抽象データ型を実装するルーチン、プログラム、オブジェクト、コンポーネント、データ構造などが含まれる。本趣旨は、通信ネットワークを通じてリンクされた遠隔の処理装置によって無線通信を通じてタスクを行う分散通信環境でも実施することができる。無線ネットワークでは、メモリ記憶装置を含むローカ

50

ルおよびリモート両方の通信デバイスの記憶媒体にプログラムモジュールを置くことができる。

【0108】

上述の説明では、EPGデータのフリーテキストおよび属性の検索のための例示的システムおよび方法について述べた。本発明について、構造的特徴および/または方法論的動作に固有の術語で記載したが、添付の特許請求の範囲に規定する趣旨は、必ずしもここに記載する特定の機能または動作に限定されないことを理解されたい。それら特定の機能および動作は、特許権を請求する本発明を実施する例示的形態として開示する。

【0109】

【表 1】

```

<?xml version="1.0" encoding="utf-8" standalone="yes"?>
<searchInfo>
  <categoryAndAttributeMappings separatorChar="|">
    <mapping keywords="closed captioned|closed captioning|closed
captions" field="ScheduleEntryAttribute" value="MSEPGA_CC" />
    <mapping keywords="stereo" field="ScheduleEntryAttribute"
value="MSEPGA_STEREO" />
    <mapping keywords="repeat|rerun|re run"
field="ScheduleEntryAttribute" value="MSEPGA_REPEAT" />
    <mapping keywords="live" field="ScheduleEntryAttribute"
value="MSEPGA_LIVE" />
    <mapping keywords="subtitled|subtitles"
field="ScheduleEntryAttribute" value="MSEPGA_SUBTITLED" />
    <mapping keywords="season premiere|season premier"
field="ScheduleEntryAttribute" value="MSEPGA_SEASONPREMIERE" />
    <mapping keywords="season finale" field="ScheduleEntryAttribute"
value="MSEPGA_SEASONFINALE" />
    <mapping keywords="series premiere|season premier"
field="ScheduleEntryAttribute" value="MSEPGA_SERIESPREMIERE" />
    <mapping keywords="series finale" field="ScheduleEntryAttribute"
value="MSEPGA_SERIESFINALE" />
    <mapping keywords="premiere" field="ScheduleEntryAttribute"
value="MSEPGA_PREMIERE" />
    <mapping keywords="finale" field="ScheduleEntryAttribute"
value="MSEPGA_FINAL" />
    <mapping keywords="letterbox" field="ScheduleEntryAttribute"
value="MSEPGA_LETTERBOX" />
    <mapping keywords="hdtv" field="ScheduleEntryAttribute"
value="MSEPGA_HDTV" />
    <mapping keywords="sap" field="ScheduleEntryAttribute"
value="MSEPGA_SAP" />
    <mapping keywords="dolby" field="ScheduleEntryAttribute"
value="MSEPGA_DOLBY" />
    <mapping keywords="dvs" field="ScheduleEntryAttribute"
value="MSEPGA_DVS" />
    <mapping keywords="rated pg 13|rated pg13"
field="ProgramRatingAttribute" value="MSEPGR_PG13" />
    <mapping keywords="rated pg" field="ProgramRatingAttribute"
value="MSEPGR_PG" />
    <mapping keywords="rated nc 17|rated nc17"
field="ProgramRatingAttribute" value="MSEPGR_NC17" />
    <mapping keywords="rated g" field="ProgramRatingAttribute"
value="MSEPGR_G" />
    <mapping keywords="rated r" field="ProgramRatingAttribute"
value="MSEPGR_R" />
    <mapping keywords="rated x" field="ProgramRatingAttribute"
value="MSEPGR_X" />
    <mapping keywords="rated ao" field="ProgramRatingAttribute"
value="MSEPGR_AO" />
    <mapping keywords="paid programming" field="ProgramAttribute"
value="MSEPGA_PAIDPROGRAMMING" />
    <mapping keywords="miniseries" field="ProgramAttribute"
value="MSEPGA_MINISERIES" />

```

【 0 1 1 0 】

【表 2】

```

    <mapping keywords="action adventure series" field="Category"
value="Series-701" />
    <mapping keywords="action adventure movie|action adventure movies"
field="Category" value="CMOVIE_ACTIONADV" />
    <mapping keywords="action adventure" field="Category"
value="CMOVIE_ACTIONADV|Series-701" />
    <mapping keywords="comedy|comedies" field="Category" value="Movies-
504|Series-703|Kids-303" />
    <mapping keywords="comedy series" field="Category" value="Series-
703" />
    <mapping keywords="comedy movie" field="Category" value="Movies-504"
/>
    <mapping keywords="documentary|documentaries" field="Category"
value="Educational-103" />
    <mapping keywords="drama" field="Category"
value="CMOVIE_DRAMA|Series-705" />
    <mapping keywords="educational series" field="Category"
value="Series-706" />
    <mapping keywords="educational" field="Category" value="Series-
706|Kids-304|Educational-1" />
    <mapping keywords="movies" field="Category" value="MSEPGC_MOVIES" />
    <mapping keywords="películas" field="Category" value="MSEPGC_MOVIES"
/>
    <mapping keywords="music" field="Category" value="Movies-508|Series-
709|Special-803" />
    <mapping keywords="news" field="Category" value="MSEPGC_NEWS" />
    <mapping keywords="science fiction movie|science fiction movies"
field="Category" value="Movies-511" />
    <mapping keywords="science fiction" field="Category" value="Movies-
511" />
    <mapping keywords="fantasy movie|fantasy movies" field="Category"
value="Movies-506" />
    <mapping keywords="fantasy" field="Category" value="Movies-506" />
    <mapping keywords="soap operas|soaps" field="Category"
value="Series-711" />
    <mapping keywords="sports" field="Category" value="News-605|Special-
805|MSEPGC_SPORTS" />
    <mapping keywords="adults only|xxx" field="Category" value="Movies-
502|Lifestyle-401" />
    <mapping keywords="kids movies" field="Category" value="Movies-503"
/>
    <mapping keywords="kids series" field="Category" value="Series-702"
/>
    <mapping keywords="kids" field="Category" value="Series-702|Movies-
503|Kids-3" />
    <mapping keywords="musicals" field="Category" value="Movies-508" />
    <mapping keywords="mystery movies|mysteries|mystery"
field="Category" value="Movies-509" />
    <mapping keywords="love stories|romance" field="Category"
value="Movies-510" />
    <mapping keywords="suspense" field="Category" value="Movies-512" />
    <mapping keywords="westerns|western" field="Category" value="Movies-
513" />
    <mapping keywords="animated|cartoon" field="Category" value="Kids-
302" />
    <mapping keywords="baseball|mlb" field="Category" value="CBASEBALL"
/>

```

【 0 1 1 1 】

【表 3】

```

    <mapping keywords="basketball|nba" field="Category"
value="CBASKETBALL" />
    <mapping keywords="boxing" field="Category" value="CBOXING" />
    <mapping keywords="football|nfl" field="Category" value="CFOOTBALL"
/>
    <mapping keywords="golf|pga|lpga|usga" field="Category"
value="CGOLF" />
    <mapping keywords="hockey|nhl" field="Category" value="CHOCKEY" />
    <mapping keywords="outdoor sports" field="Category" value="COUTDOOR"
/>
    <mapping keywords="racing" field="Category" value="CRACING" />
    <mapping keywords="soccer" field="Category" value="CSOCCER" />
    <mapping keywords="tennis" field="Category" value="CTENNIS" />
    <mapping keywords="sports news" field="Category" value="News-605" />
    <mapping keywords="sports" field="Category"
value="MSEPGC_SPORTS|News-605" />
    <mapping keywords="business news|financial news" field="Category"
value="News-601" />
    <mapping keywords="interviews|interview" field="Category"
value="News-602" />
    <mapping keywords="public affairs" field="Category" value="News-604"
/>
    <mapping keywords="weather" field="Category" value="News-606" />
    <mapping keywords="cooking" field="Category" value="Series-704" />
    <mapping keywords="game shows|game show" field="Category"
value="Series-707" />
    <mapping keywords="how to|howto" field="Category" value="Series-708"
/>
    <mapping keywords="talk shows|talk show" field="Category"
value="Series-712" />
    <mapping keywords="travel" field="Category" value="Series-713" />
    <mapping keywords="religious" field="Category" value="Special-
804|Lifestyle-409" />
    <mapping keywords="holidays|holiday" field="Category"
value="Special-802" />
    <mapping keywords="biography" field="Category" value="Educational-
102" />
    <mapping keywords="science" field="Category" value="Educational-105"
/>
    <mapping keywords="collectibles" field="Category" value="Lifestyle-
402" />
    <mapping keywords="exercise" field="Category" value="Lifestyle-404"
/>
    <mapping keywords="health" field="Category" value="Lifestyle-405" />
    <mapping keywords="home garden" field="Category" value="Lifestyle-
406" />
    <mapping keywords="international" field="Category" value="Lifestyle-
407" />
    <mapping keywords="outdoors" field="Category" value="Lifestyle-408" />
    <mapping keywords="primetime|prime time" field="TimeOfDay"
value="1140|1380" />
    <mapping keywords="morning" field="TimeOfDay" value="300|720" />
    <mapping keywords="afternoon" field="TimeOfDay" value="720|1080" />
    <mapping keywords="evening" field="TimeOfDay" value="1020|1200" />
    <mapping keywords="late night" field="TimeOfDay" value="1380|300" />
</categoryAndAttributeMappings>
</searchInfo>

```

【図面の簡単な説明】

【0112】

【図1】従来技術によるテキスト検索および属性検索の図である。

【図2】本主旨による例示的検索の図である。

【図3】例示的検索エンジンのブロック図である。

【図4】図3の例示的検索エンジンが行う例示的検索の図である。

【図5】図3の例示的検索エンジンが行う例示的な漸進的検索の図である。

【図6】本主旨による例示的方法の流れ図である。

【図7】本主旨を実施する例示的なコンピューティングデバイス環境のブロック図である。

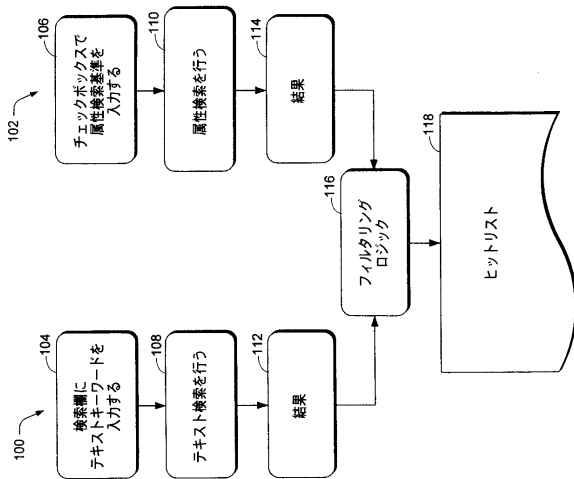
【符号の説明】

【0113】

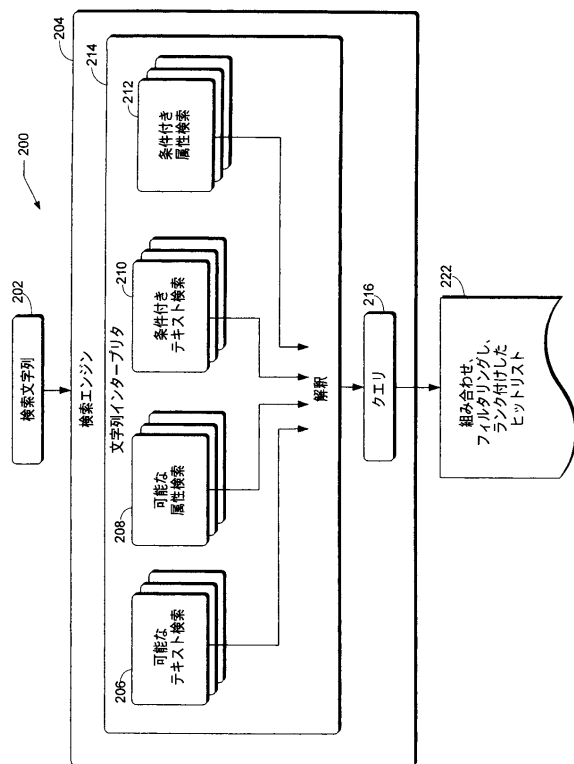
100	テキスト検索	
102	属性検索	
200	検索	10
202、402	検索文字列	
204	検索エンジン	
206	テキスト検索	
208	属性検索	
210	条件付きテキスト検索	
212	条件付き属性検索	
214	文字列インタープリタ	
216	クエリ	
222	ヒットリスト	
300	データベース	20
301	検索文字列レシーバ	
302	ノーマライザ	
306	文字列トランスフォーマ	
308	属性アイデンティファイア	
310	属性リスト	
311	メタデータ	
312	検索制約条件テーブル	
313	インデックス	
314	クエリジェネレータ	
316	テキストクエリフォーミュレータ	30
318	属性クエリフォーミュレータ	
319	データマネージャ	
320	結果コンパイラ	
322	ヒットレシーバ	
324	検索コンパレータ/フィルタ	
326	結果ランカー	
328	包含フィルタ	
330	表示アップデート	
334	EPGリソース	
410、412	代替の検索	40
700、780	コンピュータ	
720	処理装置	
721	システムバス	
730	システムメモリ	
731	ROM	
732	RAM	
733	BIOS	
734、744	オペレーティングシステム	
735、745、785	アプリケーションプログラム	
736、746	プログラムモジュール	50

- 737、747 プログラムデータ
- 741 ハードディスクドライブ
- 750 インタフェース
- 751 磁気ディスクドライブ
- 752 磁気ディスク
- 755 光ディスクドライブ
- 756 光ディスク
- 760 ユーザ入力インタフェース
- 761 ポインティングデバイス
- 762 キーボード
- 770 アダプタ
- 771 LAN
- 772 モデム
- 773 WAN
- 790 ビデオインタフェース
- 791 モニタ
- 796 プリンタ
- 797 スピーカ

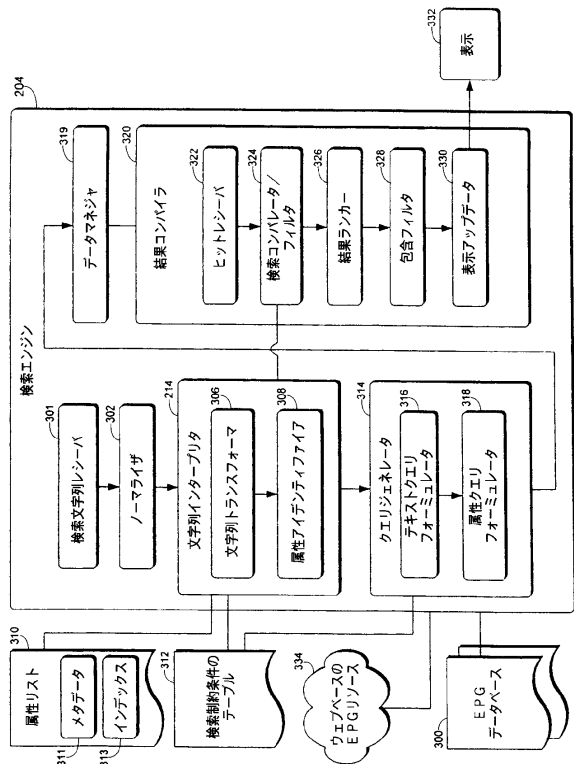
【図1】



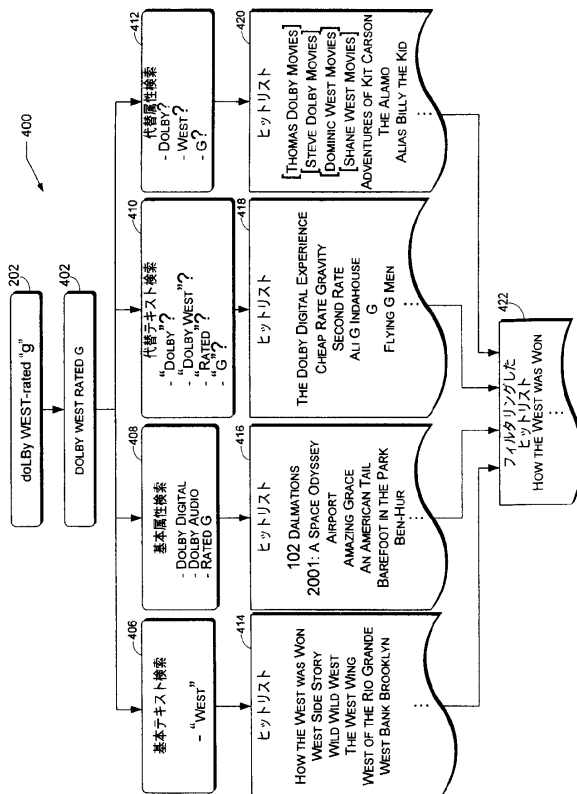
【図2】



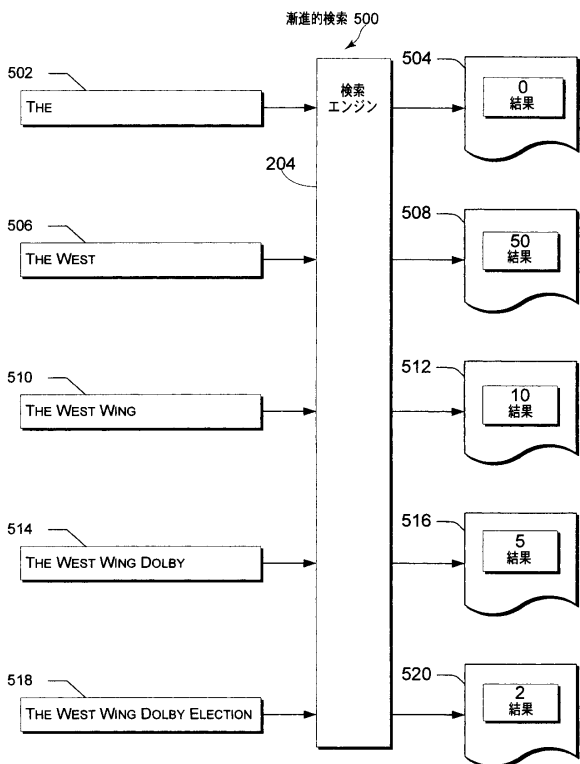
【図3】



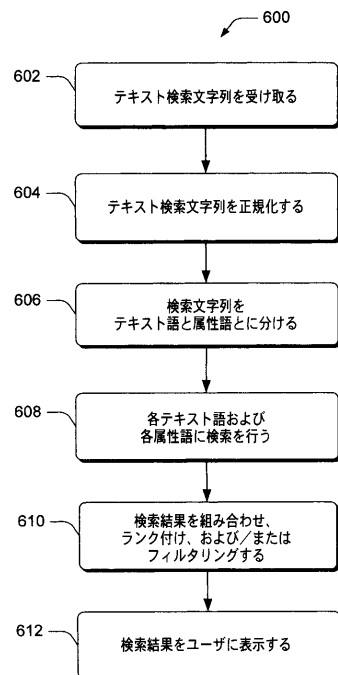
【図4】



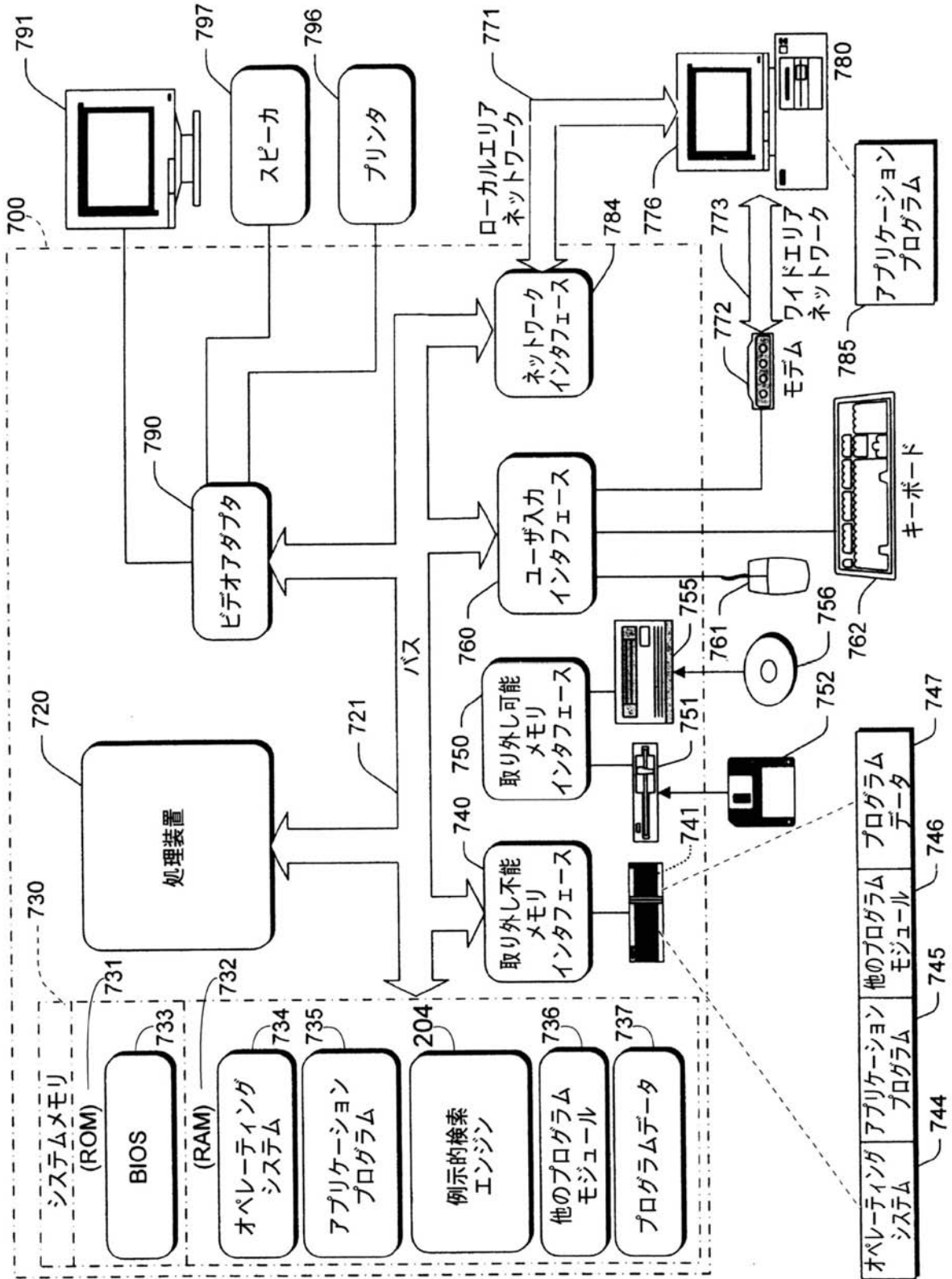
【図5】



【図6】



【図7】



フロントページの続き

(74)代理人 100091063

弁理士 田中 英夫

(74)代理人 100153028

弁理士 上田 忠

(74)代理人 100120112

弁理士 中西 基晴

(74)代理人 100113974

弁理士 田中 拓人

(72)発明者 スコット ディー・サンダース

アメリカ合衆国 9 4 5 8 7 カリフォルニア州 ユニオン シティ ジャスミン コート 3 4
4 9 5

審査官 仲間 晃

(56)参考文献 特開2002-112186(JP,A)

特開2000-200282(JP,A)

特開2003-150624(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04N 5/38 - 5/46

7/10

7/14 - 7/173

7/20 - 7/22