

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4730869号
(P4730869)

(45) 発行日 平成23年7月20日(2011.7.20)

(24) 登録日 平成23年4月28日(2011.4.28)

(51) Int. Cl.	F I
HO4N 5/445 (2011.01)	HO4N 5/445 Z
HO4N 5/44 (2011.01)	HO4N 5/44 Z
HO4N 7/173 (2011.01)	HO4N 7/173 630

請求項の数 9 (全 23 頁)

(21) 出願番号	特願2001-204032 (P2001-204032)	(73) 特許権者	000005049
(22) 出願日	平成13年7月4日(2001.7.4)		シャープ株式会社
(65) 公開番号	特開2003-18492 (P2003-18492A)		大阪府大阪市阿倍野区長池町2番2号
(43) 公開日	平成15年1月17日(2003.1.17)	(74) 代理人	100078282
審査請求日	平成20年7月3日(2008.7.3)		弁理士 山本 秀策
		(72) 発明者	乾 隆夫
			大阪府大阪市阿倍野区長池町2番2号
			シャープ株式会社内
		審査官	矢野 光治

最終頁に続く

(54) 【発明の名称】 テレビジョン番組情報表示装置

(57) 【特許請求の範囲】

【請求項1】

テレビジョン番組の放送波と共に配信されるか、またはネットワークを介して配信される電子番組表データを取得する番組表取得手段と、

取得した電子番組表データをデータベースとして格納する記憶手段と、

該データベースに含まれる番組情報を表示する番組情報表示手段と、

該番組情報表示手段に表示された番組情報から、利用者によって任意の番組が指定される番組指定手段と、

該データベースに含まれる番組情報から番組の特徴を抽出する特徴抽出手段と、

複数の番組の特徴を比較して番組間の類似度を求める類似度算出手段と、

該特徴抽出手段が求めた複数の番組の特徴に基づいて、該類似度算出手段を用いて番組間の類似度を比較し、前記番組指定手段にて指定された番組と関連がある類似番組を検索する番組検索手段と、

検索結果を表示する検索結果表示手段と、

該番組情報表示手段、該番組指定手段、該番組検索手段および該検索結果表示手段を制御して、該番組情報表示手段により番組情報を表示させ、該番組指定手段にて利用者が指定した番組と類似度の高い番組を、該番組検索手段により検索させて、検索結果を該検索結果表示手段に表示させる関連番組表示制御手段とを有し、

該特徴抽出手段が抽出する番組の特徴は、該番組情報中に含まれる番組内容の説明文章に含まれる単語にそれぞれ対応する意味ベクトル値の総和を求めることにより得られた番

組内容の説明文章の意味ベクトル値を正規化して得られる意味ベクトル値であり、
該意味ベクトル値は、各概念に対して1つの次元をそれぞれ割り当てた多次元ベクトル
値であり、

該記憶手段は、

番組情報中に含まれる番組内容を説明する番組内容の説明文章と、対応する番組の番組
IDとを組にして格納する説明文格納手段と、

個々の該番組内容の説明文章を構成する単語と、その単語を含む番組内容の説明文章に
対応する番組の番組IDのリストとを組にして格納する単語格納手段と、

単語の意味を表す意味ベクトル値をその単語と共に格納する単語意味格納手段とを有す
ることを特徴とする番組情報表示装置。

10

【請求項2】

前記記憶手段は、

番組内容の説明文章を単語に分解する単語抽出手段と、

前記説明文格納手段に格納されている全ての番組内容の説明文章における単語の分布を
元に、個々の単語の意味を前記意味ベクトル値により表すことにより単語の意味の数値化
を行う意味数値化手段と

を有することを特徴とする請求項1に記載の番組情報表示装置。

【請求項3】

前記特徴抽出手段は、

番組情報中に含まれる番組内容の説明文章を単語に分解する単語抽出手段と、

分解した個々の単語の意味の数値化として、前記記憶手段の単語意味格納手段で単語の
意味を数値化した意味ベクトル値を前記データベースから読み取る処理を行う単語意味取
得手段と、

該記憶手段の意味数値化手段での数値化により得られる、個々の単語の意味に対応する
意味ベクトル値に対する演算により、個々の単語から構成される番組内容の説明文章の意
味を表す意味ベクトル値を求める文章意味取得手段と、

番組内容の説明文章の意味を数値化したものを文章の特徴として抽出する文章特徴抽出
 手段とを有することを特徴とする請求項1または請求項2に記載の番組情報表示装置。

20

【請求項4】

前記番組検索手段は、

複数の番組の特徴を元に番組間の類似度を求める類似度取得手段と、

類似度が大きい順に複数の番組を並べる類似度順整列手段とを有することを特徴とする
 請求項1乃至請求項3のいずれかに記載の番組情報表示装置。

30

【請求項5】

前記記憶手段は、番組のジャンル情報を格納するジャンル情報格納手段を有し、

前記特徴抽出手段は、番組のジャンル情報に前記意味ベクトルを割り当てることにより
該番組のジャンル情報を数値化して番組の特徴に加えるジャンル情報加味手段を有するこ
とを特徴とする請求項1乃至請求項4のいずれかに記載の番組情報表示装置。

【請求項6】

前記ジャンル情報加味手段は、

ジャンル別に固定的な重みの値を格納したジャンル別固定重み格納手段と、固定的な重
 みの値を用いて、番組毎にジャンル情報を元にした重みを求めるジャンル別重み加算手段
 とを有することを特徴とする請求項5に記載の番組情報表示装置。

40

【請求項7】

前記ジャンル情報加味手段は、

前記データベースに含まれる全番組の数に対するジャンルに属する番組の数の割合を利
用してジャンル別に動的な重みの値を算出するジャンル別重み算出手段と、

算出した重みの値を用いて、番組毎にジャンル情報を元にした重みを求めるジャンル別
 重み加算手段とを有することを特徴とする請求項5に記載の番組情報表示装置。

【請求項8】

50

前記ジャンル情報加味手段は、ジャンル別の固定的な重みの値とデータベースを用いて算出した動的な重みの値とを合成する重み合成手段を有し、

合成した重みの値を用いて重み付けを行ったジャンル情報を番組の特徴に加味することを特徴とする請求項 6 または請求項 7 に記載の番組情報表示装置。

【請求項 9】

前記番組指定手段は、

前記番組情報表示手段に表示された番組情報から、利用者が複数の任意の番組を指定する複数番組指定手段と、

指定された複数の番組の番組情報に重みを付ける特徴合成手段と、

利用者が複数の任意の番組の重み付けを指定する番組重み指定手段とを有し、

前記特徴合成手段は、指定された複数の番組の番組情報に指定された重みを付ける請求項 1 乃至請求項 8 のいずれかに記載の番組情報表示装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、電子番組表 (Electrocal Program Guide: 以下、EPG データと称する) データを用いてテレビジョン番組の番組情報を表示する装置であって、番組を検索して提示する機能を有する番組情報表示装置に関する。

【0002】

【従来の技術】

テレビジョン放送等において、各番組の放送スケジュールは予め決められている。近年では、その放送スケジュールが EPG データという形で、各種電子媒体等によって配信されている。EPG データは、例えばテレビジョン番組の放送波と共に、その垂直帰線期間を利用して配信されており、従って、配信される EPG データは、各番組を受信する放送受信手段にて取得することができる。また、EPG データは、サービス事業者から、インターネットなどのネットワークを介して定期的または不定期的に配信されており、電話回線などの通信手段によっても取得することができる。取得された EPG データは、パーソナルコンピュータの表示装置に表示される。

【0003】

EPG データに基づくテレビジョン番組の番組情報は、テレビジョン装置などの番組情報表示装置によって表示される。番組情報表示装置は、EPG データに基づいて、テレビジョン番組のタイトル名、番組内容の説明文章、放送日時および放送局名を表示するが、所望の番組を検索するために、各種の番組検索手段が提案されており、例えば、番組のジャンル種別、出演者名等の付加的な情報に基づいて、特定の番組を検索できるようになっている。このような番組検索手段は、利用者による利便性を向上させるために、改良されている。

【0004】

例えば、特開 2000 - 67074 号公報には、利用者が番組を予約すると、EPG データに含まれるその番組の番組内容説明文章が表示されて、番組内容説明文章に含まれる単語をキーワードとして指定することにより、指定された番組と関連する番組を検索する構成が開示されている。このように、番組を予約する際に、予約された番組の番組説明文章内の単語を指定することにより、予約した番組に関連した番組を検索することができるために、利用者は、関連する番組を特定するためのキーワードを設定する必要がなく、操作性が向上する。

【0005】

さらに、特開平 11 - 8808 号公報には、番組を特徴付ける属性情報を定めて、その属性情報に基づいて番組を予め分類し、属性情報に関する条件を指定することにより、特定の番組を検索する構成が開示されている。このような構成によって、特定の番組を高速で検索することができる。

【0006】

10

20

30

40

50

【発明が解決しようとする課題】

しかしながら、これらの番組検索手段では、さらに的確に特定の番組を検索するためには、条件の指定がさらに複雑化するおそれがある。

【0007】

例えば、特開2000-67074号公報に開示された構成では、EPGデータに含まれる番組内容説明文章から、利用者が検索に適したキーワードを選定する必要があるが、キーワードは、利用者が適切に選定するとは限らず、従って、利用者が選定した単語が必ずしもキーワードとして適切にならず、目的とする番組が検索されないおそれがある。

【0008】

また、特開平11-8808号公報に開示された構成では、利用者は、予め定められた各種の属性情報と番組との関係、すなわち、各種の属性情報に対してどのような条件を指定すれば、所望する番組を検索することができるかを理解する必要がある。利用者が、このような理解が十分でない場合には、検索すべき番組に対して、属性情報に関する条件をどのように指定すれば良いか分からないおそれがある。さらに、利用者が検索する番組として適切と考える番組の属性情報が、設定されていない場合もある。例えば、利用者が「明るい雰囲気」の番組を検索したいと考えても、「明るい雰囲気」が属性情報として設定されていない場合があり、この場合には、利用者が検索の条件を指定することは容易でない。

【0009】

テレビジョン番組の番組情報を表示する装置の利用者は、専門的な知識を有する特別な人々ではなく、一般家庭においてテレビジョン番組の視聴を行う一般的な人々である。このため、ここで例に挙げたように、番組を検索する際の条件の指定方法が複雑化することにより、結果的に利用者にとって容易に利用することができなくなるおそれがある。

【0010】

本発明は、このような従来技術の課題を解決するものであり、特別な知識を持たない一般の人々でも、EPGデータから特定の番組を容易に検索することができる番組情報表示装置を提供することを目的とする。

【0011】

【課題を解決するための手段】

本発明の番組情報表示装置は、テレビジョン番組の放送波と共に配信されるか、またはネットワークを介して配信される電子番組表データを取得する番組表取得手段と、取得した電子番組表データをデータベースとして格納する記憶手段と、該データベースに含まれる番組情報を表示する番組情報表示手段と、該番組情報表示手段に表示された番組情報から、利用者によって任意の番組が指定される番組指定手段と、該データベースに含まれる番組情報から番組の特徴を抽出する特徴抽出手段と、複数の番組の特徴を比較して番組間の類似度を求める類似度算出手段と、該特徴抽出手段が求めた複数の番組の特徴に基づいて、該類似度算出手段を用いて番組間の類似度を比較し、前記番組指定手段にて指定された番組と関連がある類似番組を検索する番組検索手段と、検索結果を表示する検索結果表示手段と、該番組情報表示手段、該番組指定手段、該番組検索手段および該検索結果表示手段を制御して、該番組情報表示手段により番組情報を表示させ、該番組指定手段にて利用者が指定した番組と類似度の高い番組を、該番組検索手段により検索させて、検索結果を該検索結果表示手段に表示させる関連番組表示制御手段とを有し、該特徴抽出手段が抽出する番組の特徴は、該番組情報中に含まれる番組内容の説明文章に含まれる単語にそれぞれ対応する意味ベクトル値の総和を求めることにより得られた番組内容の説明文章の意味ベクトル値を正規化して得られる意味ベクトル値であり、該意味ベクトル値は、各概念に対して1つの次元をそれぞれ割り当てた多次元ベクトル値であり、該記憶手段は、番組情報中に含まれる番組内容を説明する番組内容の説明文章と、対応する番組の番組IDとを組にして格納する説明文格納手段と、個々の該番組内容の説明文章を構成する単語と、その単語を含む番組内容の説明文章に対応する番組の番組IDのリストとを組にして格納する単語格納手段と、単語の意味を表す意味ベクトル値をその単語と共に格納する単語意

10

20

30

40

50

味格納手段とを有し、そのことにより上記目的が達成される。

【 0 0 1 2 】

前記記憶手段は、番組内容の説明文章を単語に分解する単語抽出手段と、前記説明文章格納手段に格納されている全ての番組内容の説明文章における単語の分布を元に、個々の単語の意味を前記意味ベクトル値により表すことにより単語の意味の数值化を行う意味数值化手段とを有する構成とすることができる。

【 0 0 1 3 】

前記特徴抽出手段は、番組情報中に含まれる番組内容の説明文章を単語に分解する単語抽出手段と、分解した個々の単語の意味の数值化として、前記記憶手段の単語意味格納手段で単語の意味を数值化した意味ベクトル値を前記データベースから読み取る処理を行う単語意味取得手段と、該記憶手段の意味数值化手段での数值化により得られる、個々の単語の意味に対応する意味ベクトル値に対する演算により、個々の単語から構成される番組内容の説明文章の意味を表す意味ベクトル値を求める文章意味取得手段と、番組内容の説明文章の意味を数值化したものを文章の特徴として抽出する文章特徴抽出手段とを有する構成とすることができる。

【 0 0 1 4 】

前記番組検索手段は、複数の番組の特徴を元に番組間の類似度を求める類似度取得手段と、類似度が大きい順に複数の番組を並べる類似度順整列手段とを有する構成とすることができる。

【 0 0 1 5 】

前記記憶手段は、番組のジャンル情報を格納するジャンル情報格納手段を有し、前記特徴抽出手段は、番組のジャンル情報に前記意味ベクトルを割り当てることにより該番組のジャンル情報を数值化して番組の特徴に加えるジャンル情報加味手段を有する構成とすることができる。

【 0 0 1 6 】

前記ジャンル情報加味手段は、ジャンル別に固定的な重みの値を格納したジャンル別固定重み格納手段と、固定的な重みの値を用いて、番組毎にジャンル情報を元にした重みを求めるジャンル別重み加算手段とを有する構成とすることができる。

【 0 0 1 7 】

前記ジャンル情報加味手段は、前記データベースに含まれる全番組の数に対するジャンルに属する番組の数の割合を利用してジャンル別に動的な重みの値を算出するジャンル別重み算出手段と、算出した重みの値を用いて、番組毎にジャンル情報を元にした重みを求めるジャンル別重み加算手段とを有する構成とすることができる。

【 0 0 1 8 】

前記ジャンル情報加味手段は、ジャンル別の固定的な重みの値とデータベースを用いて算出した動的な重みの値とを合成する重み合成手段を有し、合成した重みの値を用いて重み付けを行ったジャンル情報を番組の特徴に加味する構成とすることができる。

【 0 0 1 9 】

前記番組指定手段は、前記番組情報表示手段に表示された番組情報から、利用者が複数の任意の番組を指定する複数番組指定手段と、指定された複数の番組の番組情報に重みを付ける特徴合成手段と、利用者が複数の任意の番組の重み付けを指定する番組重み指定手段とを有し、前記特徴合成手段は、指定された複数の番組の番組情報に指定された重みを付ける構成とすることができる。

【 0 0 2 1 】

以下に、本発明の作用について説明する。

【 0 0 2 2 】

本発明にあっては、取得したEPGデータに含まれる番組情報を番組情報表示手段に表示して、表示された中から利用者が任意の番組を番組指定手段により指定するだけで、指定された番組の特徴を検索条件として、指定された番組と関連性の高い類似番組を検索することが可能である。さらに、取得しているEPGデータに含まれる番組内容の説明文章お

10

20

30

40

50

よびジャンル情報等を利用して番組の特徴を抽出しているため、検索条件として適切な番組の特徴を求めることが可能である。さらに、ジャンル情報が番組に与える影響の大きさを重みとして、番組の特徴を修正し、検索条件をより適切なものに修正することが可能である。

【0023】

本発明にあっては、利用者が複数の任意の番組を指定することにより、複数の指定番組に共通した特徴を有する、関連性の高い番組を検索することが可能である。さらに、利用者が複数の番組に重みの値を指定して、指定された重みの値により番組の特徴を修正し、検索条件をより適切なものに修正することが可能である。

【0024】

【発明の実施の形態】

以下に、本発明の実施の形態について、図面に基づいて説明する。

【0025】

図1は、本発明の一実施形態である番組情報表示装置の構成を示すブロック図である。なお、図1の矢印は、本実施形態の番組情報表示装置を構成する各手段間のデータの流れを表したものである。また、図1の2本の破線矢印aおよびbは、図7および図8を用いて後述する本実施形態における処理の流れと、本実施形態の番組情報表示装置を構成する各手段との関係を、それぞれ表している。

【0026】

この番組情報表示装置は、配信される電子番組表（EPGデータ）データを取得する手段である番組表取得手段1と、番組表取得手段1等を制御する関連番組表示制御手段9とを有している。

【0027】

EPGデータは、例えば番組の放送開始日時、放送終了日時、放送局名、番組タイトル名、番組内容説明文章等の番組情報を含み、さらに、ジャンル情報等の属性情報も含んでいる。EPGデータは、放送局、テレビジョン番組情報を有するサービス事業者等から配信される。EPGデータの配信方法としては、例えばテレビジョン番組の放送波と共に垂直帰線期間を利用して配信する方法、ネットワークを介して定期的にまたは不定期に配信する方法等がある。番組表取得手段1は、例えば、放送波を受信するチューナ、ネットワークに接続されたCPU等によって構成される。

【0028】

関連番組表示制御手段9は、関連番組を検索して表示させる処理を行う際に、番組表取得手段等の各手段を制御し、番組取得手段1が取得したEPGデータは、関連番組表示制御手段9によって、解析されて、記憶手段（データベース）2に出力される。

【0029】

記憶手段2には、番組表取得手段1によって取得されたEPGデータがデータベース化されて記憶される。記憶手段2としては、例えば半導体メモリ、磁気ディスク等が用いられる。記憶手段2の詳しい構成については、図2を用いて後述する。記憶手段2がデータベースとして保持する番組情報は、関連番組表示制御手段9からの指示によって、番組情報表示手段3に出力される。

【0030】

関連番組表示制御手段9は、記憶手段2にデータベース化されて記憶されている複数の番組の番組情報を読み出して、番組情報表示手段3に表示する。番組情報表示手段3は、例えばCRT、液晶表示装置等によって構成されている。番組情報表示手段3によって、複数の番組に関する番組情報が表示されると、番組指定手段4によって、特定の番組の番組情報が指定される。

【0031】

この番組指定手段4としては、例えば、ボタン、スイッチ等の機械的な操作部によって、あるいは、番組情報表示手段3に番組情報と共に表示されるGUI（Graphycal User Interface）によって構成され、利用者が任意の番組を選択する際

10

20

30

40

50

に使用される。番組指定手段4の詳しい構成については、図6を用いて後述する。番組指定手段4によって特定の番組の番組情報が指定されると、関連番組表示制御手段9によって、指定された番組情報は、特徴抽出手段5に出力される。

【0032】

特徴抽出手段5は、番組指定手段4にて指定された番組の番組情報に基づいて、番組の特徴を求める。特徴抽出手段5のより詳しい構成については、図3を用いて後述する。特徴抽出手段5が抽出した番組の特徴は、番組検索手段7に出力される。

【0033】

番組検索手段7には、類似番組の検索を行う類似度算出手段6が設けられており、類似度算出手段6は、特徴抽出手段5によって求められた複数の番組の特徴を比較し、それらの番組の間にどの程度の類似度があるかを求める。類似度算出手段6が算出した番組間の類似度は、番組検索手段7に出力される。

10

【0034】

番組検索手段7は、特徴抽出手段5が求めた複数の番組の特徴に基づいて、類似度算出手段6を用いて番組間の類似度を比較し、特徴の似通った番組を検索する。番組検索手段6のより詳しい構成については、図5を用いて後述する。番組検索手段7が求めた検索結果は、検索結果表示手段8に出力される。

【0035】

検索結果表示手段8には、番組検索手段7によって得られた検索結果として、類似する番組の番組情報が、関連番組表示制御手段9によって表示される。この検索結果表示手段8は、例えばCRT、液晶表示装置等から構成される。

20

【0036】

図2は、本実施形態における記憶手段2に記憶されたデータベースをより詳しく説明するためのブロック図である。

【0037】

記憶手段2は、番組表取得手段1が取得した番組情報に含まれる全ての番組内容説明文章を格納する説明文章格納手段10と、番組情報に含まれる全ての番組のジャンル情報を格納するジャンル情報格納手段15とを有している。番組説明文章は、各番組の番組内容をそれぞれ説明した文章である。ジャンル情報は、各番組が属するジャンルが各番組毎にそれぞれ設定されており、ジャンル情報格納手段15は、各番組のジャンル情報を、各番組毎に格納する。

30

【0038】

説明文章格納手段10は、例えば半導体メモリ、磁気ディスク等によって構成されている。説明文章格納手段10に格納される番組内容説明文章は、番組情報表示手段3に出力されるとともに、説明文章格納手段10が格納している番組内容説明文章は、番組の特徴を抽出する際に特徴抽出手段5に出力される。

【0039】

図12は、説明文章格納手段10により番組内容説明文章を格納する方法の一例を示している。説明文章格納手段10に番組内容説明文章を格納する場合には、番組表取得手段1が取得した番組情報に含まれる全ての番組に対して、固有のID番号(0000001、0000002、...)をそれぞれ付与し、ID番号と、ID番号が付与された番組に対する番組内容説明文章とを1組としてそれぞれ格納している。番組内容説明文章は、各番組毎に付与されたID番号により管理される。

40

【0040】

なお、この例では、1つの番組における最小限の情報として、1つの番組について、番組のID番号と番組内容説明文章とを1組として格納しているが、利用者の利便性を考慮して、1つの番組について、番組の放送局名、放映日時等を、番組IDおよび番組内容説明文章と共に1組として格納するようにしてもよい。

【0041】

ジャンル情報格納手段15としては、例えば半導体メモリ、磁気ディスク等が用いられる

50

。ジャンル情報格納手段 15 が格納している番組が属するジャンルの種類は、番組の特徴を抽出する際に特徴抽出手段 5 によって利用される。

【 0 0 4 2 】

図 16 は、ジャンル情報格納手段 15 に番組毎のジャンル種類を格納する方法の一例を示す説明図である。ジャンル情報格納手段 15 に番組毎のジャンル種類を格納する場合にも、番組表取得手段 1 が取得した番組情報に含まれる全ての番組に対してそれぞれ付与された固有の ID 番号 (0 0 0 0 0 0 1、0 0 0 0 0 0 2、・・・) と、ID 番号が付与された番組に対するジャンル情報とを 1 組としてそれぞれ格納している。ジャンル情報は、各番組毎に付与された ID 番号により管理される。

【 0 0 4 3 】

なお、この例では、配信される EPG データにおいて、1 つの番組は、1 つのジャンル種類に属しているものとして、1 つの番組に対して 1 種類のジャンル種類のみが与えられている場合について説明したが、例えば 1 つの番組に 2 つ以上のジャンル情報が与えられている場合でも、全く同様な方法によって、ジャンル情報を扱うことができる。

【 0 0 4 4 】

記憶手段 2 には、単語抽出手段 11 が設けられている。単語抽出手段 11 は、説明文章格納手段 10 にて格納されている番組内容説明文章を、単語毎に切り分けて、単語格納手段 12 に格納するようになっている。

【 0 0 4 5 】

単語抽出手段 11 にて抽出された単語は、単語格納手段 12 に、その単語が含まれていた番組内容説明文章と対応付けて格納される。単語格納手段 12 としては、例えば半導体メモリ、磁気ディスク等によって構成される。

【 0 0 4 6 】

番組内容説明文章を単語毎に切り分ける方法としては、例えば形態素解析技術を用いることができる。形態素解析技術は、かな漢字変換処理等の日本語処理を行う際に広く利用されている技術であり、日本語の文章を解析して品詞毎に切り分けるようになっている。本実施形態では、文章の意味内容を反映する品詞として、名詞、動詞および形容詞が設定されている。なお、抽出される品詞の種類は、必ずしもこれらに限定されず、例えば形容動詞等の他の品詞を抽出するようにしてもよい。

【 0 0 4 7 】

図 13 は、単語格納手段 12 により単語を格納する方法の一例の説明図である。ここでは、全ての番組内容説明文章に含まれる 1 つの単語について、その単語を含む全ての番組の ID 番号を、番組 ID リストとしてまとめて格納している。

【 0 0 4 8 】

記憶手段 2 には、さらに、意味数値化手段 13 が設けられている。意味数値化手段 13 は、説明文章格納手段 10 に格納されている番組内容説明文章と、単語格納手段 12 に格納された単語とに基づいて、番組内容説明文章内に含まれる各単語毎に、各単語の意味内容に適した数値を割り当てる。

【 0 0 4 9 】

単語意味格納手段 14 は、意味数値化手段 13 により各単語毎に割り当てられた数値を、その単語と共に格納する。単語意味格納手段 14 としては、例えば半導体メモリ、磁気ディスク等を用いることができる。

【 0 0 5 0 】

本実施形態では、単語の意味内容を表す数値として、意味ベクトル値を使用しており、意味数値化手段 13 は、番組毎に、対応する番組内容説明文章から取り出された各単語の意味ベクトル値をそれぞれ求めて、各単語毎に意味ベクトル値が割り当てられる。

【 0 0 5 1 】

本実施形態においては、意味ベクトル値は、例えば 266 個の概念に分類しており、各概念に対して 1 つの次元をそれぞれ割り当てている。図 14 は、本実施形態における分類された 1 つの概念に含まれる種類の一例を、一部のみを抜粋して示している。

10

20

30

40

50

【 0 0 5 2 】

ある単語の意味を、意味ベクトル値を用いて表す場合には、その単語の意味が予め用意された各概念のそれぞれと関連があるか否かを判断し、関連がある場合には、その概念に対して値 1 を割り当てる。関連が無い場合には、その概念に対して値 0 を割り当てる。

【 0 0 5 3 】

例えば、概念として{人間、悲しい、芸術、科学、興奮、政治}の6つに分類した場合、「パイロット」という単語が、分類された各概念毎に、関連性についてそれぞれ判断されて「0」または「1」の値が設定される。単語「パイロット」は、概念「人間」とは関連性があるものとして「1」が設定され、概念「悲しい」とは関連性がないものとして「0」が設定される。以下、同様に、概念「芸術」とは関連がなく「0」が設定され、概念「科学」とは関連性があるものとして「1」が設定され、概念「興奮」とは関連性があるものとして「1」が設定され、概念「政治」とは関連性がないものとして「0」が設定される。その結果、単語「パイロット」の意味ベクトル値は、(1 , 0 , 0 , 1 , 1 , 0) となる。

10

【 0 0 5 4 】

なお、意味ベクトル値を構成する概念分類の数および種類は、本実施形態と同じである必要はなく、必要に応じて増減することができる。本実施形態では、基本的な単語については、意味ベクトル値が予め固定的に割り当てられており、割り当てられた意味ベクトル値は、変更されないようになっている。これに対して、意味ベクトル値が割り当てられた基本的な単語以外の単語が、番組内容説明文章から取り出されると、その単語の意味ベクトル値は、以下のようにして、動的に設定される。

20

【 0 0 5 5 】

この場合には、まず、番組に対する番組内容説明文章から抽出された各単語から、予め固定的に意味ベクトル値が割り当てられた基本的な単語をさらに抽出して、抽出された基本的な単語の意味ベクトル値だけを用いて、その番組の意味ベクトル値を求める。この場合、基本的な単語それぞれについて、その番組内容説明文章に含まれる出現回数(変数 $\times 3$)、および、全番組を、その基本的な単語が出現する番組の個数(変数 $\times 1$)で割った値(変数 $\times 2$)を求める。そして、求められた変数 $\times 3$ および $\times 2$ を、基本的な単語の意味ベクトル値に乘じることにより、その単語の意味ベクトル値 V_1 を求める。このような処理が、番組内容説明文書に含まれる全ての基本的な単語について実施され、各基本的な単語毎に意味ベクトル値が得られると、得られた意味ベクトル値が、すでに演算された意味ベクトル値に順次加算される。

30

【 0 0 5 6 】

このようにして、番組内容説明文章に含まれる全ての基本的な単語の意味ベクトル値が加算されることにより、全ての基本的な単語が含まれた番組内容説明文章における基本的な単語の意味ベクトル値の総和が得られる。この総和が、番組に対する仮の意味ベクトル値となる。

【 0 0 5 7 】

次に、このようにして基本的な単語の意味ベクトル値の総和が求められた番組内容説明文章に含まれている基本的な単語以外の全ての単語について、それぞれの単語の意味ベクトル値を求める。この場合には、全ての基本的な単語以外の単語について基本的な単語以外の単語を含む個々の番組について、その番組の仮の意味ベクトル値に、その番組が含む基本的な単語以外の単語の個数を掛けた値を、加算していく。

40

【 0 0 5 8 】

そうして得られた意味ベクトル値の総和を、その基本的な単語以外の単語の意味ベクトル値とし、単語意味格納手段 1 4 に格納する。

【 0 0 5 9 】

単語意味格納手段 1 4 は、意味数値化手段 1 3 が割り当てた単語毎の意味内容を表す数値をその単語と共に格納する。

【 0 0 6 0 】

50

図15は、単語意味格納手段14によって単語の持つ意味を格納する方法の例を説明するための図である。ここでは、1つの単語について、その単語の表記と共に、上述のようにして求めた単語の意味ベクトル値を格納している。

【0061】

図3は、本実施形態における特徴抽出手段3をより詳しく説明するためのブロック図である。特徴抽出手段3は、記憶手段2に設けられた単語意味格納手段14に番組IDとともに格納された各単語毎に割り当てられている意味内容を表す数値を読み取る単語意味取得手段16と、単語意味取得手段16にて読み取られた各単語毎に割り当てられた意味内容を表す数値に基づいて、文章全体の意味内容を表す数値を求める文章意味取得手段17と、説明文章格納手段10が格納している番組内容説明文章および番組指定手段4が選択した番組情報中に含まれる番組内容の説明文章を、単語毎に切り分けて取り出す単語抽出手段19とを有している。この単語抽出手段19は、上述した単語抽出手段11と同様の手段である。

10

【0062】

文章意味取得手段17では、単語の意味内容を表す数値として、上述した意味ベクトル値を使用する。さらに、文章を構成している各単語について、単語の意味を表す意味ベクトル値を全て加算した結果を、文章全体の意味内容を表す意味ベクトル値とする。なお、ベクトル値を単に加算しただけではベクトルの大きさが変わってしまい、ベクトル値を比較するのが困難となるため、本実施形態では、ベクトル値の加算を行った後、正規化を行ってベクトルの大きさを一定値(例えば10)に揃えるようになっている。この一定値は、必要に応じて適切な値に設定される。

20

【0063】

文章意味取得手段17にて求められた文章全体の意味を表す数値は、文章特徴抽出手段18に与えられる。文章特徴抽出手段18は、文章意味取得手段17が求めた文章全体の意味内容を表す数値と、後述するジャンル情報加味手段20が求めたジャンル情報に基づく意味を表す数値とを合成して、文章全体の特徴を求める。文章特徴抽出手段18が求めた文章全体の特徴は、番組検索手段7によって利用される。

【0064】

文章特徴抽出手段18では、意味内容を表す数値として上述した意味ベクトル値を使用する。文章全体の意味内容を表す数値も、ジャンル情報に基づく意味を表す数値も、いずれも意味ベクトル値で表されているので、両者のベクトル値を加算して正規化した結果を文章全体の特徴とする。従って、文章全体の特徴も、意味内容を表す数値と同様に、意味ベクトル値を使用して表される。

30

【0065】

ジャンル情報加味手段20は、ジャンル情報格納手段15が保持している番組毎のジャンル情報および番組指定手段4が選択した番組のジャンル情報を元に、ジャンル情報に基づく意味を表す数値を求めて番組の特徴に加える。

【0066】

図4は、本実施形態におけるジャンル情報加味手段20をより詳しく説明するためのブロック図である。

40

【0067】

ジャンル別重み算出手段21は、ジャンル情報格納手段15が保持している番組毎のジャンル情報を用いて、ジャンル別に重みの値を動的に求める。

【0068】

ジャンル別重み算出手段21では、あるジャンルに属する番組が、データベースに含まれる全番組に対してどれくらいの割合を占めているかによって、ジャンル別の重みの値を動的に求めるようになっている。例えば、あるジャンルに属する番組が、データベースに含まれる全番組に対して占める割合が大きい場合には、そのジャンルの重みの値を小さくする。これに対して、あるジャンルに属する番組が、全番組に対して占める割合が小さい場合には、そのジャンルの重みの値を大きくする。

50

【 0 0 6 9 】

例えば、ジャンルが「スポーツ」に属する番組は、ジャンルが「野球」に属する番組よりも、全番組に対する割合が一般的に大きい。このため、ジャンルが「スポーツ」に属する番組と関連する番組を求める際に、同じ「スポーツ」に属する番組の重みを大きくして検索時に優先しても、件数が増加するばかりであり、利用者が求める番組を検索するのが困難になる。これに対して、ジャンルが「野球」に属する番組と関連する番組を求める際には、同じ「野球」に属する番組の重みを大きくして検索時に優先することにより、徒に件数を増加させることなく、利用者が求める番組を求めることができる。

【 0 0 7 0 】

本実施形態では、ジャンル別の重みの値を 0 から 1 までの実数値で表す。そして、あるジャンルに属する番組の重みの値として、そのジャンルが全番組に占める割合をそのまま用いる。但し、ジャンル別の重みを求める方法として、本実施形態と同じ方法を用いる必要は無い。例えば、各ジャンルに属する番組が全番組に占める割合と重みの値とを対応させた換算表を用意し、その換算表を用いてジャンル別の重みを求める等の方法を用いることもできる。

10

【 0 0 7 1 】

ジャンル別重み算出手段 2 1 にて求められたジャンル別の動的な重みの値は、ジャンル別重み加算手段 2 2 に与えられている。ジャンル別重み加算手段 2 2 は、ジャンル別重み算出手段 2 1 が求めたジャンル別の動的な重みの値を元に、ジャンル情報格納手段 1 5 が保持している全番組および番組指定手段 4 が選択した番組に対して、ジャンル情報を元にした重みの値を求める。得られた値は、重み合成手段 2 3 に与えられる。

20

【 0 0 7 2 】

重み合成手段 2 3 は、ジャンル別重み加算手段 2 2 が求めた番組毎のジャンル情報を元にした重みの値と、ジャンル別固定重み格納手段 2 4 が保持しているジャンル別の固定的な重みの値とを合成する。重み合成手段 2 3 により合成された重みの値は、文章特徴抽出手段 1 8 によって利用される。

【 0 0 7 3 】

本実施形態では、ジャンル別の固定的な重みを表す値として、上述した意味ベクトル値を使用する。この意味ベクトル値は、各々のジャンルの意味を表す値を予め設定しておく。さらに、この意味ベクトル値は大きさが正規化されておらず、ベクトルの値の大きさを、番組の特徴を示すジャンルの重みとして用いる。

30

【 0 0 7 4 】

重み合成手段 2 3 は、ジャンル別の固定的な重みを表す意味ベクトル値に、番組毎のジャンル情報に基づいて求めたジャンル別の重みを表す実数値を掛けて、その結果の意味ベクトル値を合成した重みの値として文章特徴抽出手段 1 8 に出力する。

【 0 0 7 5 】

ジャンル別固定重み格納手段 2 4 は、ジャンル別に固定した重みの値を格納する。格納手段としては、例えば半導体メモリ、磁気ディスク等を用いることができる。ジャンル別固定重み格納手段 2 4 が格納するジャンル別に固定した重みの値は、重み合成手段 2 3 によって利用される。

40

【 0 0 7 6 】

ジャンル別に固定した重みの値を設定することにより、例えば「再放送」または「その他」等、番組内容と関連性が低いジャンルの種類があった場合に、それらのジャンルに属している番組が、番組内容とは無関係に、指定された番組と関連する類似番組として検索されてしまうのを防ぐことができる。

【 0 0 7 7 】

図 1 7 は、ジャンル別固定重み格納手段 2 4 によってジャンル別に固定した重みの値を格納する方法の例を説明するための図である。1つのジャンルの種類について、1つの意味ベクトル値を対応付けて格納している。

【 0 0 7 8 】

50

図5は、本実施形態における番組検索手段7をより詳しく説明するためのブロック図である。番組検索手段7は、類似度取得手段25を有している。類似度取得手段25は、特徴抽出手段5が求めた複数の番組の特徴を元に、類似度算出手段6を用いて、番組間の類似度を求める。

【0079】

本実施形態では、意味内容を表す数値として上述した意味ベクトル値を使用し、複数の番組の特徴についても、意味内容を表す数値と同様に、意味ベクトル値を使用する。類似度算出手段6は、2つの番組の特徴間の類似度を、意味ベクトル値の距離として求める。ベクトル間の距離は、一般的に、ベクトルの内積値として表すことができる。本実施形態でも、2つの番組の特徴間の類似度を、意味ベクトル値の内積値として求める。類似度取得手段25は、2つの番組間の特徴間の類似度を求める際に、予め両者の一ベクトル値の大きさを正規化して、例えば10に揃えておく。そして、類似度取得手段25は、類似度算出手段6を用いて意味ベクトル値の内積値を求める。例えば、類似度が最大の場合には内積値は10となり、類似度が最小の場合には内積値は0となる。

【0080】

類似度取得手段25によって取得された類似度は、類似度順整列手段26に与えられている。

【0081】

類似度順整列手段26は、類似度取得手段25が求めた複数の番組の類似度を元に、指定された番組と類似度が大きい順番に、番組を整列させる。類似度整列手段26によって指定番組との類似度が大きい順番に整列された番組の番組情報は、検索結果表示手段8に提供される。

【0082】

図6は、本実施形態における番組指定手段4をより詳しく説明するためのブロック図である。番組指定手段4は、複数番組指定手段27を有している。

【0083】

複数番組指定手段27は、関連番組表示制御手段9の指示によって、番組情報表示手段3が表示している番組情報に対して、利用者が複数の番組を選択する。

この複数番組指定手段27には、例えばボタンおよびスイッチ等の機械的な操作部が設けられる。複数番組指定手段27としては、番組情報表示手段3に番組情報と共に表示されるGUI等も用いることができる。

【0084】

番組重み指定手段28は、関連番組表示制御手段9の指示によって、複数番組指定手段27が選択した複数の番組に対して、利用者が重みの順位を付ける。この番組重み指定手段28には、例えばボタンおよびスイッチ等の機械的な操作部が設けられる。番組重み指定手段28としては、番組情報表示手段3に番組情報と共に表示されるGUI等も用いることができる。

【0085】

本実施形態では、利用者が番組を指定した順に、番組の重みの順位を付ける方法を採用する。この場合、最初に利用者に指定された番組が最も大きな重みを有し、以下、番組が指定される度に重みが小さくなっていく。これにより、利用者に複雑な操作を強いることなく、しかも、直感的に複数の番組に関連性が高い番組を所望の順番で検索することができる。但し、複数の番組の重みを指定する方法としては、本実施形態と同じ方法を用いる必要はない。例えば、番組を指定する度に、何段階かの重みを利用者がその都度設定する方法等も用いることができる。

【0086】

番組重み指定手段28の出力は、特徴合成手段29に与えられている。特徴合成手段29は、複数番組指定手段27が選択した複数の番組の番組情報に対して、番組重み指定手段28で求めた重みを付ける。特徴合成手段29により重みを付けられた複数の番組の番組情報は、特徴抽出手段5に出力される。

10

20

30

40

50

【 0 0 8 7 】

本実施形態では、番組の重みを 0 から 1 までの実数値で表し、選択された番組の番組情報と共に、特徴抽出手段 5 に出力する。特徴抽出手段 5 では、選択された各番組の番組情報を用いて各番組の特徴を意味ベクトル値として求めた後、その特徴を表す意味ベクトル値に対応する番組の重みを掛けることにより、選択した複数の番組全体を表す特徴を求める。

【 0 0 8 8 】

以下に、このような構成の本実施形態の番組情報表示装置における全体の処理流れを順に説明する。

【 0 0 8 9 】

本実施形態では、処理の流れは、大きく 2 つに分けることができる。一つは、E P G データを新たに取得した際に実行する番組表更新処理である。この処理は、図 1 に破線 a で示すように、番組表取得手段 1、記憶手段 2 および番組情報表示手段 3 を用いて実行される処理である。この番組表更新処理の詳細については、図 7 の流れ図を用いて後述する。

【 0 0 9 0 】

他の一つは、利用者が関連番組の検索を指示した際に実行する関連番組表示制御処理である。この処理は、図 1 に破線 b で示すように、関連番組表示手段 9 によって、番組指定手段 4、特徴抽出手段 5、番組検索手段 7 および検索結果表示手段 8 を制御して実行される処理である。この関連番組表示制御処理の詳細については、図 8 の流れ図を用いて後述する。

【 0 0 9 1 】

図 7 は、番組表更新処理の流れを説明するための図である。

【 0 0 9 2 】

まず、ステップ S 7 0 1 では、番組表取得手段 1 によって新たな E P G データを取得する。取得した E P G データの内容は、記憶手段 2 に出力される。次に、ステップ S 7 0 2 では、取得された E P G データを用いて記憶手段 2 に記憶されたデータベースを更新する。更新された番組内容の説明文章は、番組情報表示手段 3 に出力される。ステップ S 7 0 2 の処理の詳細な説明については、図 9 の流れ図を用いて後述する。次に、ステップ S 7 0 3 では、更新後のデータベースの内容に基づいて、番組情報表示手段 3 によって番組内容の説明文章を表示する。

【 0 0 9 3 】

図 9 は、ステップ S 7 0 2 におけるデータベースの更新処理の流れを説明するための流れ図である。

【 0 0 9 4 】

まず、ステップ S 9 0 1 では、番組表取得手段 1 から得た番組情報のうち、番組内容の説明文章を、説明文章格納手段 1 0 に格納する。次に、ステップ S 9 0 2 では、番組表取得手段 1 から得た番組情報のうち、番組のジャンルを表すジャンル情報をジャンル情報格納手段 1 5 に格納する。次に、ステップ S 9 0 3 では、単語抽出手段 1 1 によって、説明文章格納手段 1 0 から得た番組内容の説明文章を、その文章を構成する単語に分解する。そして、分解した単語を単語格納手段 1 2 に格納する。次に、ステップ S 9 0 4 では、意味数値化手段 1 3 によって、説明文章格納手段 1 0 から得た番組内容の説明文章と、単語格納手段 1 2 から得た各単語とを用いて、単語の意味を数値化した意味ベクトル値を求める。数値化した単語の意味は、単語意味格納手段 1 4 に格納する。

【 0 0 9 5 】

図 1 8 は、ステップ S 9 0 4 における単語の意味ベクトル値を求める処理の流れを説明するための流れ図である。この処理は、データベースに含まれる全番組の内容説明文章を構成する、全ての単語の意味ベクトル値を求めるためのものである。

【 0 0 9 6 】

まず、ステップ S 1 8 0 1 では、ベクトル V 1 の要素を全て 0 にして初期化する。ベクトル V 1 は、基本的な単語の意味ベクトル値から現在着目している番組の意味ベクトル値を

10

20

30

40

50

求めるために、一時的に数値を保存するために用いられる。

このステップS 1801からステップS 1808までは、番組単位での繰り返し処理であり、データベース2に含まれている全ての番組について順に、着目している番組を進めていく。次に、ステップS 1802では、現在着目している単語（基本的な単語）を含む番組の個数を変数X 1に記録する。

【0097】

このステップS 1802からステップS 1806までは、単語単位での繰り返し処理であり、全ての単語について順に、着目している単語を進めていく。次に、ステップS 1803では、データベース2に含まれている全番組数を変数X 1で割った値を、変数X 2に記録する。次に、ステップS 1804では、現在着目している番組において、現在着目している単語を含む個数を、変数X 3に記録する。

10

【0098】

次に、ステップS 1805では、現在着目している単語の意味ベクトル値に、変数X 2と変数X 3とを掛けて、その結果をベクトルV 1に加算する。次に、ステップS 1806では、ステップS 1802からステップS 1806までの処理を全ての単語に対して行ったか否かを判断する。そして、まだ処理が行われずに残っている単語があれば、ステップS 1802に戻り、全ての単語について処理が済んでいれば、ステップS 1807に進む。

【0099】

次に、ステップS 1807では、ベクトルV 1の値を、現在着目している番組のベクトル値とする。次に、ステップS 1808では、ステップS 1801からステップS 1808までの処理を全ての番組に対して行ったか否かを判断する。そして、まだ処理が行われずに残っている番組があれば、ステップS 1801に戻り、全ての番組について処理が済んでいれば、ステップS 1809に進む。

20

【0100】

次に、ステップS 1809では、ベクトルV 2の要素を全て0にして初期化する。ベクトルV 2は、上述のようにして求めた番組の意味ベクトル値から現在着目している単語（基本的な単語以外の単語）の意味ベクトル値を求めるために、一時的に数値を保存するために用いられる。このステップS 1809からステップS 1814までは、単語単位での繰り返し処理であり、全ての単語について順番に処理を進める。

【0101】

次に、ステップS 1810では、現在着目している番組において、現在着目している単語を含む個数を、変数X 3に記録する。このステップS 1810からステップS 1812までは、番組単位での繰り返し処理であり、データベース2に含まれている全ての番組について順番に処理を進める。次に、ステップS 1811では、現在着目している番組の意味ベクトル値に、変数X 3を掛けて、その結果をベクトルV 2に加算する。

30

【0102】

次に、ステップS 1812では、ステップS 1810からステップS 1812までの処理を全ての番組に対して行ったか否かを判断する。そして、まだ処理が行われずに残っている番組があれば、ステップS 1810に戻り、全ての番組について処理が済んでいれば、ステップS 1813に進む。次に、ステップS 1813では、ベクトルV 2の値を、現在着目している単語のベクトル値とする。次に、ステップS 1814では、ステップS 1809からステップS 1814までの処理を全ての単語に対して行ったか否かを判断する。そして、まだ処理が行われずに残っている単語があれば、ステップS 1809に戻り、全ての単語について処理が済んでいれば、処理を終了する。

40

【0103】

以上の処理によって、ある単語の意味ベクトル値を、その単語が全番組に対してどのように分布しているのか（偏っているのか）に対応させた値として求めることができる。例えば、多数の番組の番組内容説明文章に含まれる単語については、その意味ベクトル値は番組の意味内容から大きな影響を受けない。これに対して、特定の番組に集中して含まれる単語については、その意味ベクトル値は番組の意味内容から強い影響を受ける。よって、

50

番組内容を説明する際に、各々の単語がどのような意味で使用されているのかということ
を、意味ベクトルの値として反映させることができる。

【0104】

図8は、関連番組表示制御処理の流れを説明するための流れ図である。

【0105】

まず、ステップS801では、番組表示手段3から番組指定手段4に、利用者が指定した
番組の番組情報が出力される。番組指定手段4にて指定された番組の番組情報は、特徴抽
出手段5に出力される。このとき、複数番組指定手段27によって複数の番組が指定され
、番組重み指定手段28によって重みが指定された場合には、特徴合成手段29によつて
重み付けした複数の番組情報が特徴抽出手段5に出力される。次に、ステップS802で
は、特徴抽出手段5が、利用者が指定した番組の番組情報を用いて、その番組の特徴を求
める。求められた利用者が指定した番組の特徴は、番組検索手段7に出力される。このス
テップS802の処理の詳細な説明については、図10の流れ図を用いて後述する。

10

【0106】

次に、ステップS803では、特徴抽出手段5が、記憶手段2がデータベースとして記憶
している全ての番組に対して、その番組情報を用いて特徴を求める。データベースに含ま
れる全ての番組の特徴は、番組検索手段7に出力される。このステップS803の処理の
詳細な説明については、図10の流れ図を用いて後述する。

【0107】

次に、ステップS804では、番組検索手段7が、ステップS802で求められた利用者
が指定した番組の特徴を元に、ステップS803で求められた全ての番組の特徴を用いて
、データベースに含まれる全ての番組を、指定された番組と類似度が大きい順に整列さ
せる。番組を整列した結果は、検索結果表示手段8に出力される。このステップS804の
処理の詳細な説明については、図11の流れ図を用いて後述する。

20

【0108】

次に、ステップS805では、検索結果表示手段8が、指定番組との類似度が大きい順に
整列した番組を上位から順に関連番組として表示する。関連番組表示制御手段9は、上
の処理を制御する。

【0109】

図10は、ステップS802およびステップS803における番組の特徴を求める処理の
流れを説明するための流れ図である。

30

【0110】

まず、ステップS1001では、単語抽出手段19によって、利用者が指定した番組の番
組内容説明文章を、その文章を構成する単語に分解する。次に、ステップS1002では
、単語意味取得手段16によって、分解された単語からその意味を意味ベクトル値として
求め、さらに文章意味取得手段17によって、個々の単語の意味ベクトル値から説明文章
全体の意味を意味ベクトル値として求める。次に、ステップS1003では、ジャンル別
重み算出手段21によって、ジャンル情報格納手段15から得た番組毎のジャンル情報に
基づいて、ジャンル別の重みの値を動的に求める。次に、ステップS1004では、ジャ
ンル別重み加算手段22によって、その番組に対するジャンル別の重みを求める。そして
、重み合成手段23によって、求められたジャンル別の重みの値に、ジャンル別固定重み
格納手段24に保持されている固定した重みの値を合成した重みを求める。次に、ステ
ップS1005では、文章特徴抽出手段18によって、番組内容説明文章の意味と、合成さ
れたジャンル別の重みとを用いて、その番組の特徴を意味ベクトル値として求める。

40

【0111】

図11は、ステップS804における番組検索処理の流れを説明するための流れ図である
。

【0112】

まず、ステップ1101では、特徴抽出手段5が類似度算出手段6を用いて、利用者に指
定された番組の特徴とデータベースに含まれる番組との特徴とを意味ベクトル値に基づい

50

て比較し、番組間の類似度を意味ベクトル値の内積値として求める。次に、ステップS 1 1 0 2では、上記ステップS 1 1 0 1の処理を、データベースに含まれる全ての番組について繰り返す。次に、ステップS 1 1 0 3では、類似度順整列手段26によって、データベースに含まれる全ての番組について、利用者が指定した番組との類似度が大きい順番に整列する。

【0113】

【発明の効果】

以上詳述したように、本発明によれば、利用者が番組表示手段に表示された番組情報の中から単に任意の番組を指定するだけで、指定された番組と関連性の高い類似番組を、EPGデータから検索して表示することができる。この操作は、利用者が番組検索についての予備知識を必要とせず、簡単かつ直感的で理解し易い。従って、本発明の主要な利用者が、特別な知識を持たない一般家庭の人々であることを考慮すると、本発明は非常に利用し易いものとなっている。

10

【0114】

さらに、本発明によれば、利用者が番組表示手段に表示された番組情報の中から単に複数の任意の番組を指定するだけで、指定された複数の番組に共通した特徴を有する関連性の高い類似番組を、EPGデータから検索することができる。この操作は、複数の検索条件を指定する従来の検索方法と比べて、利用者が番組検索についての予備知識を必要とせず、簡単かつ直感的で理解し易い。従って、本発明の主要な利用者が、特別な知識を持たない一般家庭の人々であることを考慮すると、本発明は非常に利用し易いものとなっている。

20

【0115】

さらに、本発明によれば、利用者が複数の任意の番組を指定する際に、それらの番組の間に順番を指定することができる。これにより、その順番を各番組の特徴に対する重みの差として反映した形で、指定された複数の番組に共通した特徴を有する関連性の高い類似番組を、EPGデータから検索することができる。この操作は、複数の検索条件を指定する従来の検索方法と比べて、利用者が番組検索についての予備知識を必要とせず、簡単かつ直感的で理解し易い。従って、本発明の主要な利用者が、特別な知識を持たない一般家庭の人々であることを考慮すると、本発明は非常に利用し易いものとなっている。

30

【0116】

さらに、本発明によれば、利用者が指定した番組と関連性の高い類似番組を検索するために、取得しているEPGデータの内容に基づいて番組の特徴を求めることができる。これにより、年数が経過したり、季節が経過することによって、番組の編成内容に変化があっても、その変化に動的に対応して、常に適切な番組の特徴を求めることができる。

【0117】

さらに、本発明によれば、番組の特徴を求めるために、番組内容の説明文章と共にその番組の属するジャンル情報も利用することができる。従来は、このような異なる番組属性情報については、個別に検索条件を指定する必要があったが、本発明では、これらを数値化して統一した特徴として扱うことができる。これにより、複数の番組間での類似度を比較する単一の基準を持つことができ、関連性の高い番組を類似度の大きい順に整列して利用者に表示することができる。

40

【0118】

さらに、本発明によれば、番組の属するジャンル毎に、そのジャンル情報が番組に与える影響の強さを、固定の重みとして保持しており、この固定の重みを利用することができる。これにより、ジャンルの分類方法自体が必ずしも適切ではない場合があっても、関連性の高い類似番組を検索する際に、悪影響が生じることを防ぐことができる。

【0119】

さらに、本発明によれば、番組の属するジャンル毎に、そのジャンル情報が番組に与える影響の強さを、取得しているEPGデータの内容に基づいて動的に修正し、重みとして求めることができる。これにより、年数が経過したり、季節が経過することによって、番組

50

の編成内容に変化があっても、その変化に動的に対応して、常に適切な番組の特徴を求めることができる。

【図面の簡単な説明】

【図 1】本発明の一実施形態である番組情報表示装置の構成を説明するためのブロック図である。

【図 2】本発明の一実施形態である番組情報表示装置における、データベースのより詳しい構成を説明するためのブロック図である。

【図 3】本発明の一実施形態である番組情報表示装置における、特徴抽出手段のより詳しい構成を説明するためのブロック図である。

【図 4】本発明の一実施形態である番組情報表示装置における、ジャンル情報加味手段のより詳しい構成を説明するためのブロック図である。

10

【図 5】本発明の一実施形態である番組情報表示装置における、番組検索手段のより詳しい構成を説明するためのブロック図である。

【図 6】本発明の一実施形態である番組情報表示装置における、番組指定手段のより詳しい構成を説明するためのブロック図である。

【図 7】本発明の一実施形態である番組情報表示装置における、番組表更新処理の流れを説明するための流れ図である。

【図 8】本発明の一実施形態である番組情報表示装置における、関連番組表示制御処理の流れを説明するための流れ図である。

【図 9】本発明の一実施形態である番組情報表示装置における、データベース更新処理の流れを説明するための流れ図である。

20

【図 10】本発明の一実施形態である番組情報表示装置における、番組の特徴を求める処理の流れを説明するための流れ図である。

【図 11】本発明の一実施形態である番組情報表示装置における、番組検索処理の流れを説明するための流れ図である。

【図 12】本発明の一実施形態である番組情報表示装置における、説明文章格納手段による説明文章の格納方法の例を示す図である。

【図 13】本発明の一実施形態である番組情報表示装置における、単語格納手段による単語の格納方法の例を示す図である。

【図 14】本発明の一実施形態である番組情報表示装置における、概念分類の種類（一部）の例を示す図である。

30

【図 15】本発明の一実施形態である番組情報表示装置における、単語格納手段による単語が持つ意味の格納方法の例を示す図である。

【図 16】本発明の一実施形態である番組情報表示装置における、ジャンル情報格納手段による番組毎のジャンル種類の格納方法の例を示す図である。

【図 17】本発明の一実施形態である番組情報表示装置における、ジャンル別固定重み格納手段によるジャンル別に固定した重みの値の格納方法の例を示す図である。

【図 18】本発明の一実施形態である番組情報表示装置における、単語の意味ベクトル値を求める処理の流れを説明するための流れ図である。

【符号の説明】

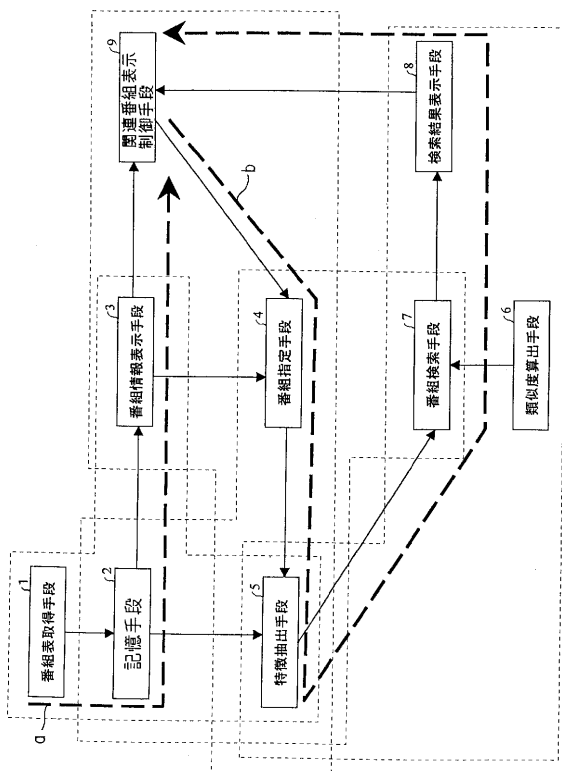
40

- 1 番組表取得手段
- 2 記憶手段
- 3 番組情報表示手段
- 4 番組指定手段
- 5 特徴抽出手段
- 6 類似度算出手段
- 7 番組検索手段
- 8 検索結果表示手段
- 9 関連番組表示制御手段
- 10 説明文章格納手段

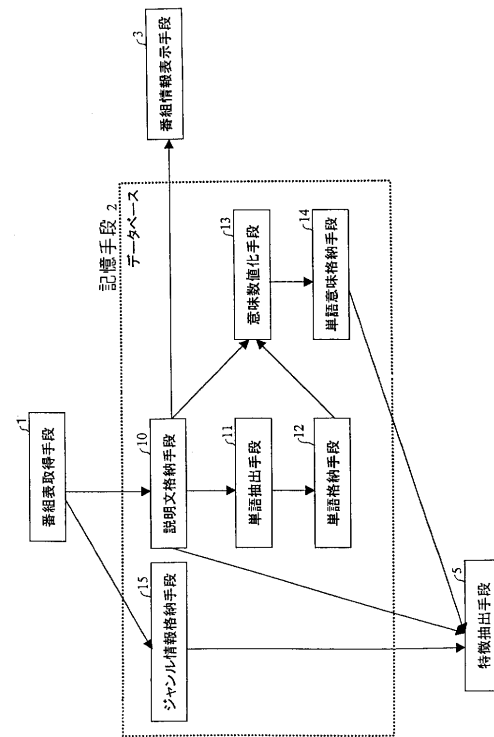
50

- 1 1 単語抽出手段
- 1 2 単語格納手段
- 1 3 意味数値化手段
- 1 4 単語意味格納手段
- 1 5 ジャンル情報格納手段
- 1 6 単語意味取得手段
- 1 7 文章意味取得手段
- 1 8 文章特徴抽出手段
- 1 9 単語抽出手段
- 2 0 ジャンル情報加味手段
- 2 1 ジャンル別重み算出手段
- 2 2 ジャンル別重み加算手段
- 2 3 重み合成手段
- 2 4 ジャンル別固定重み格納手段
- 2 5 類似度取得手段
- 2 6 類似度順整列手段
- 2 7 複数番組指定手段
- 2 8 番組重み指定手段
- 2 9 特徴行政手段

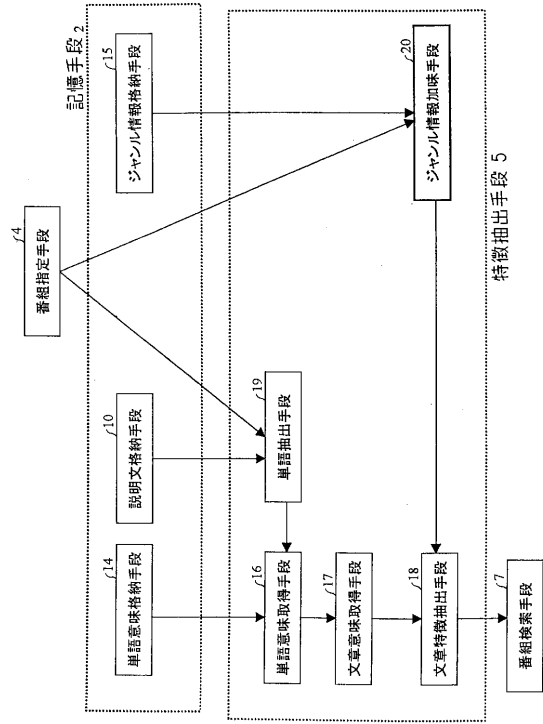
【図 1】



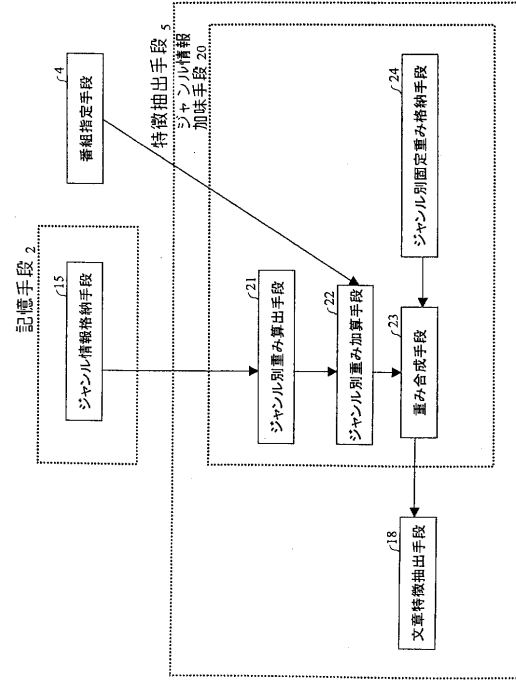
【図 2】



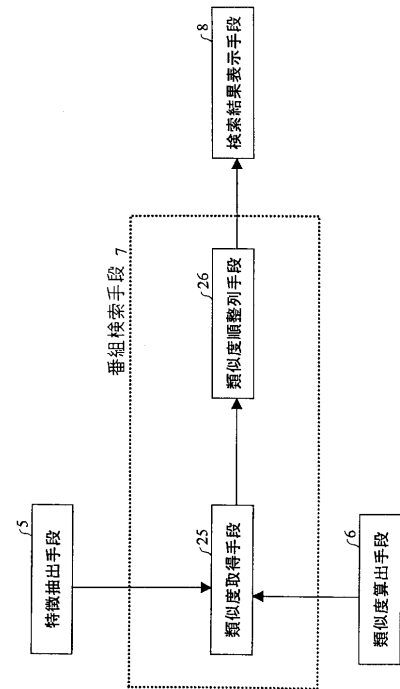
【図3】



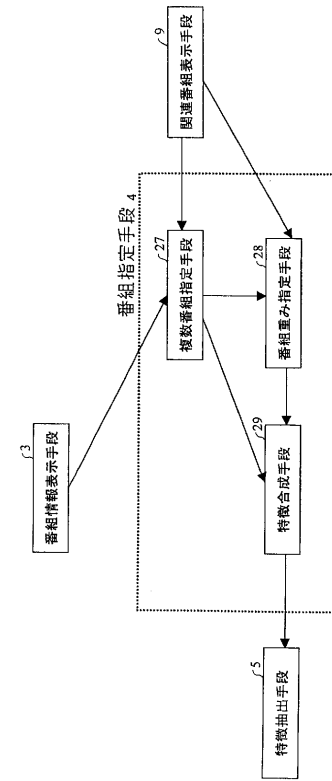
【図4】



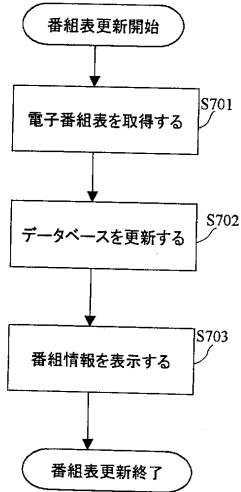
【図5】



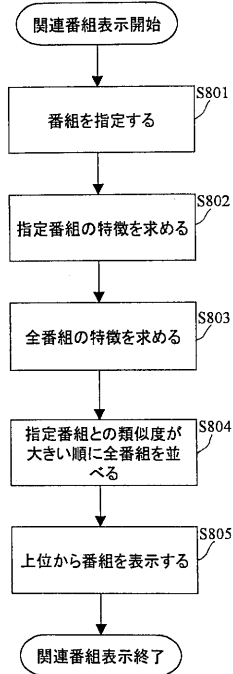
【図6】



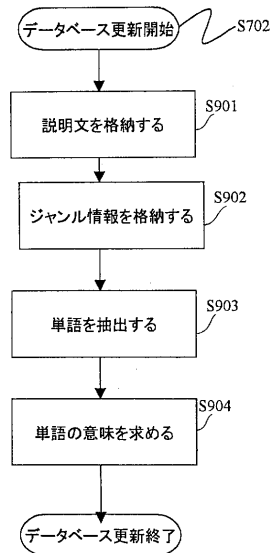
【図7】



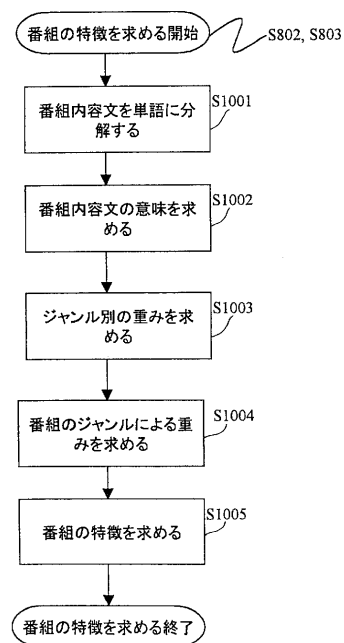
【図8】



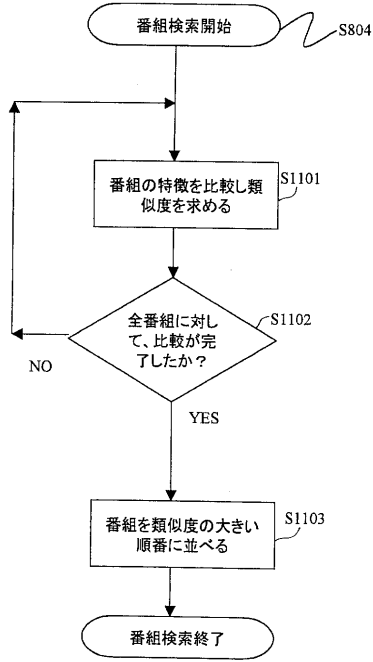
【図9】



【図10】



【図11】



【図12】

番組ID	番組内容説明文
0000001	大相撲中継 第4日目 今日の取組み...
0000002	映画「世界の不思議」(1990年アメリカ)...
0000003	国際ニュース 今週の話...
∴	∴

【図13】

単語表記	番組IDリスト
特集	0000003、0000005、0000006
あかるい	0000005
世代	0000004、0000006
∴	∴
∴	∴

【図14】

人間、人名、男性、女性、子供、大人、老人、家族、家庭、動物、水棲生物、鳥類、虫、微生物、植物、生命、生死、誕生、病氣、老い、殺生、...

【図15】

単語表記	意味ベクトル値
特集	(1,0,0,1,1,0,0,0,1,・・・)
あかるい	(0,0,0,0,1,1,1,0,1,・・・)
世代	(1,1,0,0,1,0,0,0,0,・・・)
⋮	⋮
⋮	⋮

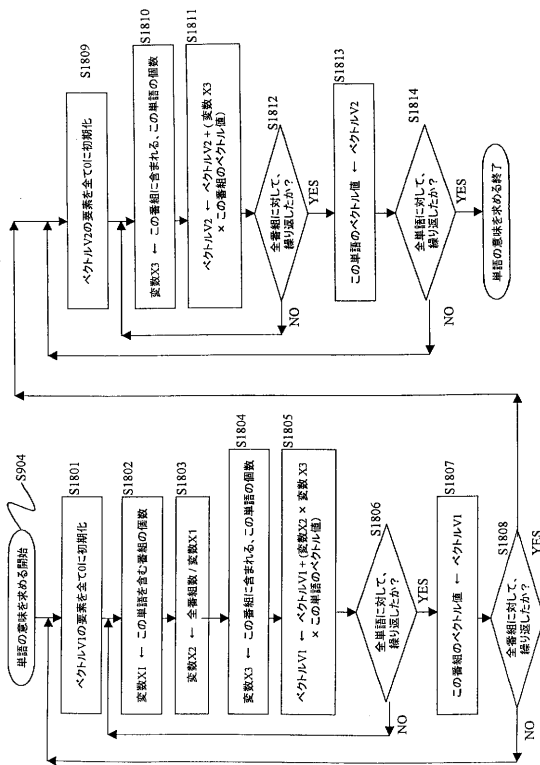
【図17】

ジャンル種類	意味ベクトル値
スポーツ	(1,0,0,1,1,0,0,0,1,・・・)
映画	(0,0,0,0,1,1,1,0,1,・・・)
ニュース	(1,1,0,0,1,0,0,0,0,・・・)
⋮	⋮
⋮	⋮

【図16】

番組ID	ジャンル種類
0000001	スポーツ
0000002	映画
0000003	ニュース
⋮	⋮

【図18】



フロントページの続き

- (56)参考文献 特開平 1 1 - 3 0 8 5 8 1 (J P , A)
特開 2 0 0 0 - 0 3 2 4 0 5 (J P , A)
特開 2 0 0 1 - 1 0 1 1 9 5 (J P , A)
特表平 0 8 - 5 0 6 9 4 2 (J P , A)
特開 2 0 0 0 - 2 7 8 6 2 3 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04N 5/44-5/445

H04N 7/16-7/173