

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-96560
(P2007-96560A)

(43) 公開日 平成19年4月12日(2007.4.12)

(51) Int. Cl.	F I	テーマコード (参考)
HO4N 17/00 (2006.01)	HO4N 17/00 Z	5C025
HO4N 5/44 (2006.01)	HO4N 5/44 Z	5C061

審査請求 未請求 請求項の数 6 O L (全 12 頁)

(21) 出願番号	特願2005-281085 (P2005-281085)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成17年9月28日 (2005.9.28)	(74) 代理人	100100310 弁理士 井上 学
		(72) 発明者	内部 こなぎ 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内
		(72) 発明者	隈井 裕之 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内
		(72) 発明者	森本 康嗣 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内
		Fターム(参考)	5C025 AA30 BA27 DA01 DA05 5C061 BB03 CC09

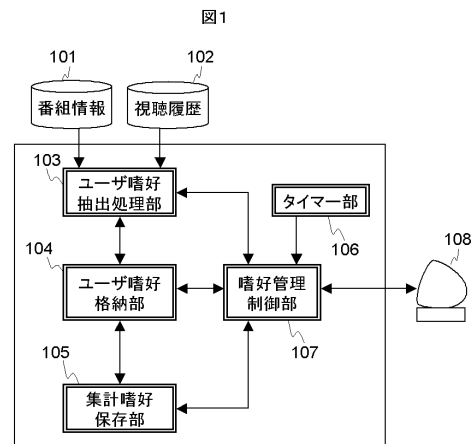
(54) 【発明の名称】 ユーザ嗜好抽出装置

(57) 【要約】

【課題】番組推薦などに使用するユーザ嗜好抽出装置において、特に季節などで異なるユーザ嗜好を適切に抽出、保存、再利用する。

【解決手段】番組情報101とユーザの視聴履歴102に基づいてユーザ嗜好情報の抽出を行う処理部103と、ユーザの嗜好情報に含めるタイマー部106から取得した時刻情報に基づいて嗜好情報を時系列的に逐次追記しながら格納するユーザ嗜好格納部104とを設ける。ユーザ嗜好格納部104は格納する嗜好情報を、予め定められた期間ごとに集計した集計嗜好情報を作成して格納する。更に、一定期間を経過したとき、当該期間の集計嗜好情報を格納部から移動し保存する集計嗜好保存部105を設け、集計嗜好保存部105に保存した集計嗜好情報は、ある指定時期にユーザ嗜好格納部に再格納する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

時刻情報を取得するタイマー部と、
番組情報と番組の視聴履歴情報からユーザの嗜好情報を抽出するユーザ嗜好抽出処理部と、
上記嗜好情報を格納するユーザ嗜好格納部とを有し、
上記格納部は上記時刻情報を用いて、上記嗜好情報を時系列的に格納し、予め定められた期間ごとに上記嗜好情報を集計することにより上記期間ごとの集計嗜好情報を作成することを特徴とするユーザ嗜好抽出装置。

【請求項 2】

上記集計嗜好情報を保存する集計嗜好保存部を更に有し、
上記ユーザ嗜好格納部は上記期間の開始または終了時刻と現在時刻との差があらかじめ定められた時間を超える場合に上記集計嗜好情報を上記集計嗜好保存部に移すことを特徴とする請求項 1 に記載のユーザ嗜好抽出装置。

【請求項 3】

上記ユーザ嗜好格納部は、上記現在時刻の情報に基づいて、予め定められた特定期間の上記集計嗜好情報を上記集計嗜好情報保存部から読み出し再格納することを特徴とする請求項 2 に記載のユーザ嗜好抽出装置。

【請求項 4】

上記ユーザ嗜好格納部は、上記集計嗜好情報に含まれる上記嗜好度がある一定期間変わらない単語を削除もしくは他の記憶領域に移動することを特徴とする請求項 2 又は 3 の何れかに記載のユーザ嗜好抽出装置。

【請求項 5】

上記ユーザ嗜好格納部は、該ユーザ嗜好格納部に格納されている集計嗜好情報と上記再格納された集計嗜好情報とから統合嗜好情報を作成し、該統合嗜好情報と上記番組情報とを用いて、番組推薦を行うことを特徴とする請求項 3 又は 4 の何れかに記載のユーザ嗜好抽出装置。

【請求項 6】

番組情報とユーザの視聴履歴情報に基づいてユーザプロファイリングを行うユーザ嗜好抽出装置であって、
上記番組情報および上記視聴履歴情報からユーザの嗜好情報を抽出するユーザ嗜好抽出処理部と、
該ユーザの嗜好情報を格納するユーザ嗜好格納部とを有し、
上記ユーザ嗜好抽出処理部は、番組ごとに断続的に視聴した各視聴時間の該番組の放映時間に対する比率を用いて嗜好度を算出し、
上記ユーザ嗜好格納部は、上記嗜好度を上記嗜好情報として格納することを特徴とするユーザ嗜好抽出装置。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、ユーザ嗜好に合致したテレビ番組をユーザに推薦するために、番組情報とユーザの視聴履歴からユーザの嗜好を抽出する装置に関するものである。

【背景技術】**【0002】**

テレビの多チャンネル化に伴い、新聞等のテレビ番組表や EPG（電子番組表）に掲載される大量の番組情報からユーザが好みの番組を選択することは困難となってきた。

【0003】

そのため、EPG（電子番組表）に含まれる番組情報とユーザの視聴履歴を用いてユーザ嗜好をより正確に抽出し、適切な番組をユーザに提示するために、ユーザ嗜好を構成する嗜好要素（タイトル、番組説明文に表れるキーワード、ジャンル、放送時間帯、チャン

10

20

30

40

50

ネル等)とその嗜好度(対象要素に対する嗜好の強さ)を精度良く求める技術が公開されている。主な技術として、ユーザの嗜好変化に対応する技術と、視聴の仕方(視聴中の操作)を利用する技術がある。

ユーザの嗜好変化に対応する技術としては、嗜好要素のグループ(ジャンル、キーワード等)や個々の嗜好に対して、長期と短期のそれぞれの期間ごとに、嗜好度を求め、両者から嗜好変化を求める技術がある(例えば、特許文献1参照)。

一方、ユーザの視聴の仕方を考慮した技術としては、番組の放送時間に対する視聴時間の割合に加え、視聴中のユーザの操作種類(例えば、チャンネルの切り替え)とその回数を用いて嗜好度を求める技術がある(例えば、特許文献2参照)。

【0004】

10

【特許文献1】特開2005-85366号公報

【0005】

【特許文献2】特開2003-23589号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

ユーザの嗜好には1年のうちのある期間や季節などに依存している、時期依存性の高い嗜好要素が多く存在する。例えば、ある特定の時期にしか行われぬスポーツなどが該当する。

【0007】

20

特許文献1では、期間ごとに嗜好要素に対する嗜好度を求め、嗜好情報を生成している。長期の嗜好情報については、嗜好度の更新速度が緩やかであり、長期の嗜好が反映されるよう考慮はしている。しかし、一年中コンスタントに放映され、継続的に視聴される番組に関わる嗜好要素の嗜好度に比べると、ある一定時期にしか放映されない番組の嗜好要素の嗜好度は、視聴頻度の減少に伴い減少してしまう。短期においてはその傾向は更なる強くなる。そのため、時期依存性の高い嗜好要素の嗜好度を捕らえきれない。

【0008】

また、ユーザの嗜好継続性は番組視聴中のユーザ操作情報に反映される。例えば、チャンネル切替回数や各チャンネルでの視聴時間などが該当する。特許文献2の装置では、番組視聴中のチャンネル切替回数が多い場合、視聴番組に対する嗜好は弱いとするなど、操作の嗜好への影響を考慮した嗜好度の計算を行っているが、視聴継続時間については考慮されていない。そのため、同じ放送時間長の番組を同じ時間視聴し、かつ視聴中のチャンネル切替回数が同数であった場合でも嗜好度は同じになり、継続的に見た時間の長さによる嗜好度合の差を算出できない。

30

【課題を解決するための手段】

【0009】

本発明において開示される発明のうち代表的なものの概要を説明すれば、下記のとおりである。本発明のユーザ嗜好抽出装置は、番組情報とユーザの視聴履歴情報に基づいてユーザ嗜好情報の抽出を行う。その嗜好情報に含まれる時刻情報に基づいて、嗜好情報を時系列的に逐次追記しながら格納部に格納する。格納した嗜好情報は、あらかじめ定められた期間ごとに集計して集計嗜好情報を作成して格納する。また、番組情報とユーザの視聴履歴情報に基づくユーザプロファイリングでは、ユーザが番組ごとに断続的に視聴した各視聴時間の放映時間に対する比率を用いて嗜好度を算出し、番組情報から得られる単語に対応付けて嗜好情報を抽出する。

40

【発明の効果】

【0010】

嗜好情報を予め定められた期間ごとに格納し利用することにより、時期依存性の高い嗜好要素について、最新のユーザ嗜好とともに嗜好抽出を行う時期に合致したユーザの嗜好を抽出できる。

【発明を実施するための最良の形態】

50

【0011】

本発明を実施するための一形態であるユーザ嗜好抽出装置を図1に示す。本実施例の構成は、ユーザの嗜好情報を抽出する際、時間要素と視聴継続時間を考慮することにより、時期依存性の高いユーザの嗜好情報を正確に反映することができる点を特徴としている。

【0012】

本装置は以下の構成を有する。ユーザ嗜好抽出処理部103、ユーザ嗜好格納部104、集計嗜好保存部105、タイマー部106、嗜好管理制御部107、ユーザインターフェース部108からなり、これらは、いわゆるパーソナルコンピュータなどの計算機に具備されている、メモリ、HDDなどの外部記憶、リアルタイムクロック、マイクロプロセッサ、コンソール、キーボードにより構成できる。また本発明の各部の動作は、上述したパーソナルコンピュータなどの計算機上にインストールされ動作するプログラムモジュールなどのソフトウェアによって実現できる。また、パーソナルコンピュータに限らず、同様の構成を内蔵するテレビ、ハードディスクレコーダ、携帯情報端末などのコンピュータ組込み機器などでも実施可能なことはいうまでもない。

10

【0013】

以下、本実施例を構成する各部について説明する。ユーザ嗜好抽出処理部103は、番組情報101の電子データと、ユーザの視聴番組とその視聴時間や録画した番組などの視聴履歴102の電子データからユーザ嗜好を抽出する。番組情報101は、テレビ受像機が受信するEPGやネットワークを介して得られるインターネット用EPGに含まれる、番組タイトル、ジャンル、出演者名、番組概要記述などのデータである。また、視聴履歴102はテレビ受像機や録画機器に対する操作履歴を記録保存したデータである。操作履歴は、操作の際に発生する信号値等を観察して逐次的に記録する。ユーザ嗜好抽出処理部103は、番組情報101を構成する番組タイトル、ジャンル、出演者名や、番組情報101に含まれる番組概要記述などのテキスト情報に対して、既知の手段である形態素解析などの言語処理を用いて単語を抽出する。抽出した単語に対して視聴履歴102から得られる視聴頻度や視聴時間を基にユーザの嗜好の強さ(嗜好度)を求める。

20

【0014】

なお、ユーザ嗜好情報には、ユーザによって入力されたユーザが好む番組タイトル、ジャンル、出演者名、キーワード等の情報を含めてもよい。

【0015】

ユーザ嗜好格納部104は、ユーザ嗜好抽出処理部103で作成された嗜好情報をメモリなど一次記憶装置上に時系列的に逐次格納する。また、ユーザの嗜好情報に含まれる時刻の情報に基づいて逐次追記して格納したユーザの嗜好情報を、予め定められた期間ごとに集計し、集計嗜好情報を作成する。更に、集計嗜好情報に期間の情報を含めて格納する。また、期間情報が現在時刻から一定期間を過ぎた集計嗜好情報を集計嗜好保存部105に移す。

30

【0016】

集計嗜好保存部105は、ユーザ嗜好格納部104が格納する集計嗜好情報で、その期間情報が一定期間を過ぎたものを保存する。

【0017】

タイマー部106は、現在時刻を取得し、ユーザ嗜好格納部104や集計嗜好保存部105に対し、所定期間の経過を通知する。計算機に内蔵されたリアルタイムクロック機能やアラーム機能により実現できる。

40

【0018】

嗜好管理制御部107は、ユーザ嗜好抽出処理部103、ユーザ嗜好格納部104、集計嗜好保存部105、タイマー部106の各部間のデータフロー等の制御を行う。

【0019】

まず、嗜好の強さ(嗜好度)を求める方法について説明する。嗜好度は、番組の視聴合計時間の長さにより評価する方法や視聴合計時間の番組の放映時間に対する比率を用いて計算する方法、視聴中のユーザの操作種類(例えば、チャンネルの切り替え)とその回数

50

を利用して求める方法など一般的な方法で求めてもよいが、本実施例では、視聴履歴と番組情報から得られる断続的に視聴した各視聴時間の、番組の放映時間に対する比率を用いて計算する方法について説明する。すなわち、ユーザの嗜好度を次式で求める。

【0020】

【数1】

$$S = \sum_{i=1}^m (T_i / T_p)^n \quad (\text{ただし } n > 1) \quad (\text{数1})$$

10

【0021】

ここで、 T_p は番組Pの放映時間、 T_i は断続的にm回に分割して視聴した際のi回目の視聴時間である。実際の計算では、視聴と録画の区別、録画の種類（ユーザ自身の操作による録画、キーワード入力による自動録画等）、再生有無や仕方等に依存した重み付けを行う場合に補正係数を乗じる。具体的には、ユーザ自身が意図して録画した番組の方が自動録画された番組より嗜好が強い、録画して再生された番組が再生されない番組より嗜好が強い、通常で再生された番組が倍速再生された番組より嗜好が強い、など操作の内容に応じて嗜好度計算時に重み付けを行う。この重み付けを行うことにより、ユーザの操作種類による嗜好への影響を考慮でき、より正確な嗜好度を計算することができる。

以下、図4を用いて具体的に説明する。図4はユーザの視聴履歴102の例を表す図である。ユーザは時刻 T_0 に放送開始した番組Aを時刻 T_1 から T_{a1} 分間視聴し、番組Bを T_{b1} 分間視聴し、番組Aを T_{a2} 分間視聴し、番組Bを T_{b2} 分間視聴し、番組Aを T_{a3} 分間視聴し、番組Bを T_{b3} 分間視聴した視聴履歴を示している。番組Aの放送時間、即ち放送開始時刻から放送終了時刻までの時間は T_{pa} 、番組Bの放送時間は T_{pb} であるとす。このとき、番組Aの番組情報から抽出される単語の嗜好度 S_a を次式で計算する。

20

【0022】

【数2】

$$S_a = \alpha((T_{a1}/T_{pa})^n + (T_{a2}/T_{pa})^n + (T_{a3}/T_{pa})^n) \quad (\text{数2})$$

30

【0023】

ただし、 α は補正係数、 n は1より大きい数とする。同様に番組Bから抽出される単語の嗜好度 S_b を次式で計算する。

【0024】

【数3】

$$S_b = \alpha((T_{b1}/T_{pb})^n + (T_{b2}/T_{pb})^n + (T_{b3}/T_{pb})^n) \quad (\text{数3})$$

40

【0025】

と n の値は数2と同じ値とする。ここで、数2及び数3は、嗜好度 S_a 及び S_b の計算に数1を適用した場合の計算式である。このとき、図4の視聴履歴において、番組Aと番組Bの放送時間が等しく、即ち $T_{pa} = T_{pb}$ であり、番組Aの視聴時間と番組Bの視聴時間が等しい、即ち $T_{a1} + T_{a2} + T_{a3} = T_{b1} + T_{b2} + T_{b3}$ であるとする。このとき番組Aに対する部分的な視聴時間 T_{a1} 、 T_{a2} 、 T_{a3} の最大値である T_{a1} は番組Bの部分的な視聴時間 T_{b1} 、 T_{b2} 、 T_{b3} の最大値 T_{b1} よりも十分大きな値であり、数2で計算される嗜好度 S_a は数3で計算される嗜好度 S_b よりも大き

50

い。

【0026】

例えば、番組A、番組Bともに30分番組であり、それぞれTa1が10分、Ta2が4分、Ta3が1分の合計15分、Tb1、Tb2、Tb3がそれぞれ5分の合計15分を視聴したとする。このとき補正係数 = 1、n = 2であったとすると、Sa = 0.128、Sb = 0.083（小数点以下3桁有効）となり番組Aの嗜好度が番組Bよりも高いと判定できる。

【0027】

断続的に視聴した各視聴時間の和を視聴時間とし、断続的な視聴時間を考慮しないで嗜好度を計算した場合、即ち次式で計算した場合には、各視聴時間が異なっても総合した視聴時間が同じであれば、嗜好度は同じになる。

【0028】

【数4】

$$S = \left(\left(\sum_{i=1}^m T_i \right) / T_p \right)^n \quad (\text{ただし } n > 1)$$

(数4)

10

【0029】

本発明の嗜好度の計算方式では、同じ放送時間の番組をトータルで同じ時間視聴したとしても、部分的な視聴時間が長い番組の方が嗜好度は高くなり、ユーザの嗜好をより正確に表すことができる。特にザッピング（チャンネルをよく変える）を繰り返すユーザの嗜好度を正確に表すことができる。

20

【0030】

なお、図4では番組Aと番組Bが異なる開始時刻、同じ番組の長さである例を示したが、上述した嗜好度の計算式は、開始時刻や番組の長さ、放送が重なっている番組の数に関係なく適用可能である。

【0031】

次に、本発明の嗜好情報の管理方法について説明する。ユーザ嗜好抽出処理部103は、各番組に対して上記数1を適用し、番組情報に含まれる単語に対する嗜好度として嗜好情報を抽出する。ユーザ嗜好格納部104は、ユーザ嗜好抽出処理部103が抽出した嗜好情報を時系列的に逐次追加しながら格納する。

30

【0032】

図2はユーザ嗜好格納部104がユーザ嗜好抽出処理部103で生成される嗜好情報を管理するテーブルの例を表す図である。嗜好情報201は時刻情報202とプロファイル203で構成され、プロファイルは番組情報101から抽出される単語204と単語の嗜好度205の一つ以上の組から構成される。時刻情報202は嗜好情報を同テーブルに格納開始した時刻に関する情報である。つまり、番組情報101から得られる番組の開始時刻または終了時刻である。また、視聴もしくは録画履歴から得られる番組の視聴開始時刻または視聴終了時刻としてもよい。この場合は番組編成者の指定した番組の放映の時期、時刻よりも、ユーザが実際に視聴した時期、時刻を優先することになる。例えば、春に放送された夏向きの内容の番組をユーザが録画し実際に夏に視聴した場合など、録画済み番組の視聴が大部分を占める用途では、ユーザの嗜好を反映させる点で視聴開始または視聴終了時刻を用いる方が適しているといえる。

40

【0033】

プロファイル203の単語204は、格納開始から現在時刻までに放映された番組の情報から抽出される単語であり、嗜好度205は、各番組に対して例えば数1により求められる単語に対する嗜好度の和や積を求め、正規化するなどして逐次的に求めた値である。視聴履歴が存在しない場合には、番組情報から単語を抽出し、その単語の嗜好度を0としてプロファイルに登録することもできる。ここで、嗜好情報201がある時刻T1での状

50

態としたとき、嗜好情報206はT1より後の時刻T2における状態であるとする。時刻T1のプロファイルに既に登録されている単語については、嗜好度の値の更新を行う。図2のプロファイルの嗜好度208は、時刻T1から時刻T2までの抽出されたユーザ嗜好が反映された値となっている。例えば、「温泉」に対する嗜好度は時刻T1においては0.7であったが、時刻T2においては、0.8となっており、「グルメ」の嗜好度は時刻T1では0.6であったのが、時刻T2においては、0.55となっている。「サッカー」の嗜好度は0のままである。また、時刻T1までの嗜好情報201のプロファイルの単語204に存在せず、時刻T1からT2までの間に新たに抽出された単語がある場合には、嗜好情報の単語情報206に新たな単語を追加し、単語情報207とする。例えば図2においては「園芸」が追加されている。単語を一意的に保持することで、使用メモリを効率的に利用できる。 10

【0034】

ユーザ嗜好格納部104はユーザ嗜好抽出処理部103で作成された、図2に示した構成の嗜好情報をメモリなど一次記憶装置上に格納する。さらに、ユーザ嗜好格納部104は、嗜好情報に含まれる時刻情報に従って、予め定めた期間ごとに集計した集計嗜好情報を作成し、集計嗜好情報を期間情報を含めてメモリなど一次記憶装置上に格納する。

【0035】

図3は集計嗜好情報の例を表す図である。集計嗜好情報301は、期間情報302と集計プロファイル303で構成される。期間情報302は予め定められた期間で集計された集計嗜好情報301の集計開始時刻や集計終了時刻を登録することができ、プロファイルと対応づけられている。このように予め定められた期間でプロファイルを集計管理することにより、ある期間におけるユーザの嗜好を正確に算出することができる。つまり、時期依存性の高い嗜好要素を精度よく捕らえることができる。また、ある程度まとまった単位でプロファイルを管理することにより、古いプロファイルの消去や別媒体間の移動などが簡便に行えるため、メモリを有効に使用したプロファイル管理が可能となる。 20

【0036】

集計嗜好情報を構成する集計プロファイル303は、時刻情報202から一定期間を経過した状態のときに、ユーザ嗜好格納部104で格納した嗜好情報201のプロファイル203を集計して作成される。この場合も、単語に対する嗜好度を逐次的に求めたように、各期間情報に該当する時間情報に対応するプロファイルから求められる。つまり、集計した個々の嗜好情報の和や積を正規化した値などとして求めた値を登録する。図3の集計嗜好情報の例においては、予め定めた期間を1ヶ月とし、期間情報には2005年5月1日～2005年5月31日（以下2005年5月と表記）、2005年6月1日～2005年6月30日（以下2005年6月と表記）などの集計開始及び終了日が期間情報として登録されている。期間情報2005年5月に対応する集計プロファイルは2005年5月1日00:00から2005年5月31日24:00まで逐次格納された状態の嗜好情報201である。もちろん、予め定める区分分割する期間は任意に設定可能である。 30

【0037】

本発明のユーザ嗜好抽出装置では、タイマー部106から一定期間の経過通知があった場合、通知時の集計嗜好情報の集計プロファイルの中から嗜好度が集計開始時の初期値と変わらない単語を抽出し、当該集計プロファイルの中の当該単語とその嗜好度を集計嗜好情報から削除したり、他の記憶領域に移動したりできる。直近のユーザ嗜好情報において負例として利用していた嗜好度が初期値のままである情報も、一定の期間が経過すると負例として用いる価値が低下する。このように嗜好度が初期値のままである情報を削除することで、嗜好情報の一定精度を保ちつつメモリを有効利用することができる。嗜好度が初期値のままである場合のみではなく、その変化量がある閾値を超えない場合を対象としてもよい。これにより、ユーザの嗜好度があまり変化していない情報も削除もしくは移動の対象となり、メモリを有効に活用することができる。もちろん、一定期間の始点を集計開始時ではなく、集計開始時からのある時点すなわちプロファイリング中のある時点とすることも可能であり、任意の時点を設定することができる。 40

【0038】

例えば、図3の集計嗜好情報において、嗜好度の初期値が0であった場合、2005年5月を期間情報とする集計嗜好情報の集計プロファイルの中で、単語として登録されているサッカーの嗜好度は0であるため、この単語の削除を行う。削除した単語のプロファイルが使用していたメモリ領域は新たな単語のプロファイルのために開放する。または、削除する代わりに他の安価な記憶領域、例えばディスクなどの2次記憶装置への移動を行ってもよい。嗜好度が0である単語のプロファイルを他の記憶領域に移動した場合には、集計嗜好情報を用いた番組推薦を行う場合などに、負例として再利用することができる。

【0039】

このように、一定期間経過後の単語の嗜好度の変化に着目して集計プロファイルを管理することが可能である。また、以下のように予め定めた期間ごとに定期的に更新することもできる。

【0040】

集計嗜好保存部105は、集計嗜好情報に含まれる期間情報に登録された集計開始時刻または集計終了時刻とタイマー部106から得られる現在時刻との差が、あらかじめ定められた時間を越えた場合に、対象となる集計嗜好情報をディスクなどの2次記憶装置に保存し、ユーザ嗜好格納部104から集計嗜好情報を削除する。現在時刻に近い情報をメモリ上に置くことにより、ユーザの最近の嗜好情報を容易に参照できる。後述するように本実施例では集計嗜好情報を用いて番組推薦を行うことを想定しているが、現在時刻に近い嗜好情報を用いることにより、最近のユーザの嗜好に合致した番組を推薦することが可能となる。また、メモリを有効に利用することもできる。

【0041】

図6は集計嗜好情報の保存処理の流れを説明するPAD図である。以下図6を用いながら、本発明のユーザ嗜好抽出装置の特徴の一つである、集計嗜好情報の集計嗜好保存部105への保存と、ユーザ嗜好格納部への再格納の処理について説明する。

【0042】

以下では、図3に示した集計嗜好情報において、指定期間を3ヶ月と定めた場合について説明する。尚、3ヶ月という指定は、1ヶ月を単位とする本願実施例においてユーザ嗜好格納部に最大3区分の集計嗜好情報を格納することを意味する。

【0043】

まず、現在時刻C0を取得する(S601)。今、取得した現在時刻が2005年8月1日00:00であったとする。次にユーザ嗜好格納部104に区分されて格納された各集計嗜好情報について、以下の処理を行う。集計嗜好情報の期間情報C1を取得する(S603)。次に、C0とC1が指定期間を超えるかどうかを判定する(S604)。判定の結果指定期間を超えていた場合には、当該期間情報C1を持つ集計嗜好情報を集計嗜好保存部105に保存し(S605)、ユーザ嗜好格納部104から削除する(S606)。たとえば、図3の例では、期間情報2005年5月の集計嗜好情報が現在時刻2005年8月1日00:00に対して指定期間3ヶ月を超えると判断され、集計嗜好保存部105に保存され、ユーザ嗜好格納部104からは削除される。

【0044】

次に集計嗜好保存部105に保存された集計嗜好情報の再格納処理について説明する。集計嗜好保存部105に保存された集計嗜好情報について、以下の処理を行う。集計嗜好保存部に保存されている集計嗜好情報の期間情報C2を取得する(S608)。タイマー部106から得られた現在時刻C0と、C2の期間情報とが予め定められた再格納の条件を満たすか判定し(S609)、再格納の条件を満たす場合には、対象となる集計嗜好情報を再びユーザ嗜好格納部に格納する(S612)。例えば、再格納の条件を「現在時刻と集計嗜好情報の期間情報に登録された時刻との差が1年の倍数」と定めると、集計嗜好保存部105に保存されている期間情報が2004年8月、2003年8月などの集計嗜好情報は、現在時刻が2005年8月1日00:00であるため条件に一致し、ユーザ嗜好格納部104に再格納される。つまり、ある特定の時期に関して、現在だけでなく過去

10

20

30

40

50

の嗜好情報も参照することができる。これにより、スポーツのようにある特定の時期にしか放映されず時期依存性の高い嗜好要素に対して、ユーザの嗜好を正確に抽出することができる。

【0045】

このとき、集計嗜好保存部105に、2004年8月を期間情報とする集計嗜好情報と、2003年8月を期間情報とする集計嗜好情報が保存されている場合には、これらの集計嗜好情報を更に集計したのちに、ユーザ嗜好格納部104に格納してもよい。集計はユーザ嗜好格納部104における嗜好情報の集計方法で行えばよい。あるいは最近の情報、上記の例では2004年8月の集計嗜好情報のみ格納してもよい。嗜好変化の速度に応じて、古い情報も用いるか、新しい情報のみを用いるかを選択することで、より正確な嗜好情報を得ることができる。

10

【0046】

さらに、上述のように現在時刻との関係で再格納データを指定するのではなく、例えば2004年7月、9月などの特定の区間を月単位などでユーザが直接指定した集計嗜好情報を再格納の対象にしてもよい。これにより、現在の時刻によらずユーザは例えば季節に応じた過去履歴を参照することができる。

【0047】

また、集計嗜好保存部105からユーザ嗜好格納部104への再格納に際しては、ユーザインタフェース部108からユーザに対し再格納するか否かを問い合わせるようにしてもよい(S610、S611)。ユーザの嗜好が変わり、一定期間前の嗜好が現在と異なる場合にはユーザが再格納を拒否できるため、誤った嗜好情報を格納部から参照することがなくなる。

20

【0048】

図5は再格納を問い合わせる画面例を表す図である。集計嗜好情報の集計プロフィールに登録されている単語の一覧を示し、ユーザは、リモコンやマウスやキーボードからの操作により、再格納する嗜好情報の一覧504をチェックし確認502を指示しユーザ嗜好格納部104への再格納を行うこと(S612)、あるいはキャンセル501を指示し、再格納を行わないことを選択できる。再格納期間情報503は、再格納を確認する集計嗜好情報の期間情報である。

【0049】

一覧された単語が現在の嗜好と大幅に異なる場合には、ユーザは再格納を行わないことでより現在のユーザの状態に即したユーザ嗜好が抽出可能である。また、現在の嗜好と近い場合には、ユーザはその嗜好情報を再格納することにより、あまり番組を視聴できないことによる誤ったユーザ嗜好の生成を抑制することができる。一覧された単語の中で再格納したい単語をユーザが選択し、選択された単語のみを再格納してもよい。再格納される単語の取捨選択により、より正確な嗜好情報が得られる。

30

【0050】

以上のように、本発明のユーザ嗜好抽出装置は、嗜好情報を区分ごとに格納して利用することにより、嗜好変化への対応が容易である。また、メモリを効率よく使うことができ、嗜好情報の時期依存性をユーザの嗜好抽出に反映できるため、ユーザ嗜好を適切に抽出

40

ことができる。また、嗜好継続性を考慮した嗜好度の算出により、嗜好抽出精度を向上させることができる。最後に、ユーザ嗜好格納部104に格納された集計嗜好情報を用いた番組推薦の例について説明する。ユーザ嗜好格納部には、例えば一ヶ月などの一定の期間単位ごとに区分された集計嗜好情報が例えば数ヶ月などの一定期間分と、集計嗜好保存部105から再格納された、現在時期と一致した時期の集計嗜好情報が格納されている。現在時期と一致した時期の集計嗜好情報は数年程度とすると、使用保存領域を抑制しながら、周期的な嗜好情報を用いることができる。何年間の集計嗜好情報を保持するかはユーザが設定できる。

【0051】

これらの集計嗜好情報は、それぞれ単語とその嗜好度の集合として構成されている。格

50

納部に格納されている集計嗜好情報を統合し、統合嗜好情報を生成する。格納部に格納されている各集計嗜好情報に含まれる単語集合の和集合の全単語に対し、単語の統合的な嗜好度を求める。例えば、ある単語に対する統合的な嗜好度は、各集計嗜好情報におけるその単語の嗜好度の線形和などで求めることができる。計算においては、時期が一致している集計嗜好情報の重み係数を大きくしてもよいし、また、一定期間分格納されている集計嗜好情報については、新しい情報の重み係数を大きくしてもよい。これらにより、時期が一致する嗜好情報または現在に近い時期の嗜好情報を強く反映させることができる。

【0052】

番組を推薦する際には、上記で求めた統合嗜好情報を用いる。つまり、統合嗜好情報と番組情報を比較し、ユーザの嗜好と合致する番組を推薦する。具体的には、まず推薦するか否かを判定する番組の番組情報に対して形態素解析を行い、番組情報に含まれる単語を抽出する。推薦判定対象番組は番組情報が得られる範囲の番組とする。即ち、現在放映中もしくは放映予定の番組である。判定対象番組は上記範囲内の日付などにより、ユーザが指定することもできる。抽出された単語の嗜好度を統合嗜好情報を参照して求め、推薦判定対象番組の情報に含まれる全ての単語の嗜好度の総和が最大となる番組を推薦する。また、予め嗜好度の閾値を設定しておき、推薦判定対象番組情報から抽出される単語の嗜好度の総和が閾値を越える番組を推薦してもよい。さらに、推薦判定対象番組の情報に含まれる単語のうち最大の嗜好度を持つ単語の嗜好度に着目し、その嗜好度が最大となる場合もしくは閾値を越える番組を推薦してもよい。もちろん、推薦する番組数を指定し、閾値を越えて推薦候補となった複数の番組の中から、嗜好度が高い順に選択して推薦することも可能である。

【0053】

これらの構成により、高精度に抽出され一定の期間ごとに格納されている集計嗜好情報の中から、現在時期に関連した嗜好情報を選択的に利用し、ユーザの嗜好に合致した番組推薦を行うことができる。

【0054】

以上に示した発明の実施により、以下のような利点を得られる。

(1) 現在に近い時期もしくは時期が一致する嗜好情報を抽出し、その嗜好情報を用いて時期依存性の高い嗜好要素をもつ番組をユーザに推薦することができる。

(2) 断続的な視聴時間を考慮した、精度のよい嗜好度を抽出することができる。

【産業上の利用可能性】

【0055】

DVDレコーダ、TV等の家庭向け映像機器や、テレビ機能付きのパソコン、PDAや携帯電話等の情報機器、ネット配信番組を再生するセットトップボックス等に適用できる。

【図面の簡単な説明】

【0056】

【図1】本発明の一実施例の構成を説明する図である。

【図2】嗜好情報の構成の例を説明する図である。

【図3】集計嗜好情報の構成の例を説明する図である。

【図4】視聴履歴の例を説明する図である。

【図5】集計嗜好情報再格納の確認のためのユーザインタフェースの例を説明する図である。

【図6】集計視聴情報の保存処理の流れの例を説明する図である。

【符号の説明】

【0057】

101 番組情報、102 視聴履歴、103 ユーザ嗜好抽出処理部、104 ユーザ嗜好格納部、105 集計嗜好保存部、106 タイマー部、107 嗜好管理制御部、108 ユーザインタフェース部、201 嗜好情報、202 時刻情報、203 プロファイル、204 単語、205 嗜好度、301 集計嗜好情報、302 期間情報

10

20

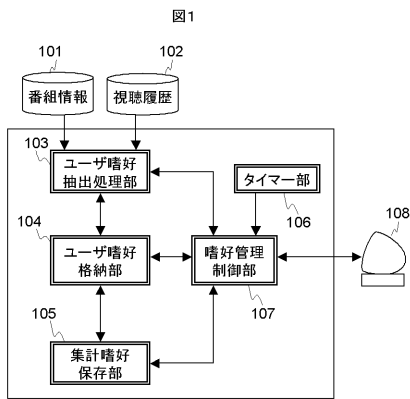
30

40

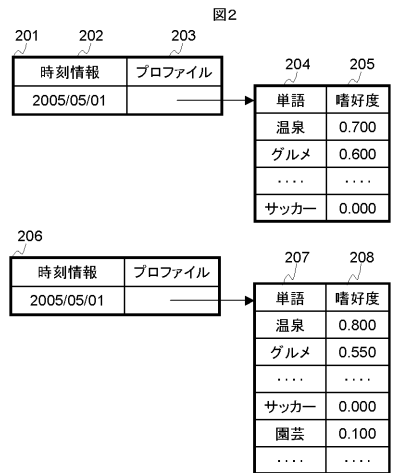
50

, 3 0 3 集計プロフィール, 3 0 4 単語, 3 0 5 嗜好度。

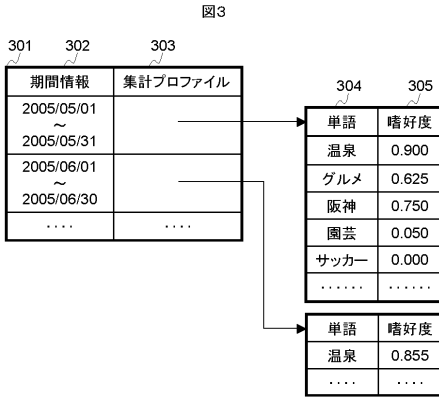
【 図 1 】



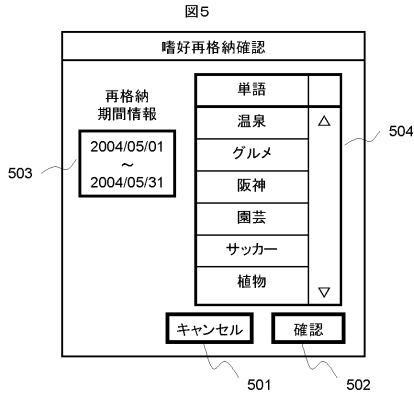
【 図 2 】



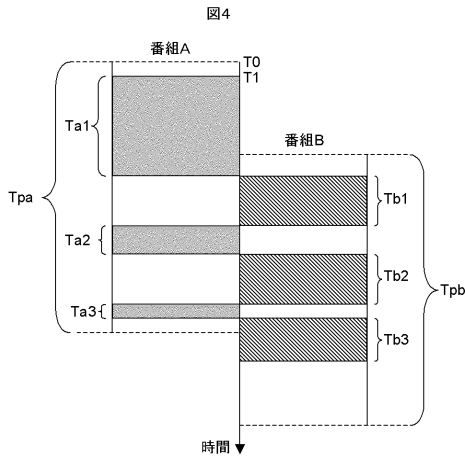
【 図 3 】



【 図 5 】



【 図 4 】



【 図 6 】

