(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2004 年6 月24 日 (24.06.2004)

PCT

(10) 国際公開番号
WO 2004/053736 A1

(51) 国際特許分類[7]: G06F 17/30

(21) 国際出願番号: PCT/JP2003/015927

(22) 国際出願日: 2003 年12 月12 日 (12.12.2003)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2002-361540
2002 年12 月12 日 (12.12.2002) JP
特願2003-285030 2003 年8 月1 日 (01.08.2003) JP

(71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都 品川区 北品川６丁目７番３５号 Tokyo (JP).

(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 宮嵜 充弘 (MIYAZAKI,Mitsuhiro) [JP/JP]; 〒141-0001 東京都 品川区 北品川６丁目７番３５号 ソニー株式会社内 Tokyo (JP). 山本 則行 (YAMAMOTO,Noriyuki) [JP/JP]; 〒141-0001 東京都 品川区 北品川６丁目７番３５号 ソニー株式会社内 Tokyo (JP). 斎藤 真里 (SAITO,Mari) [JP/JP]; 〒141-0001 東京都 品川区 北品川６丁目７番３５号 ソニー株式会社内 Tokyo (JP). 小池 宏幸 (KOIKE,Hiroyuki) [JP/JP]; 〒141-0001 東京都 品川区 北品川６丁目７番３５号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 稲本 義雄 (INAMOTO,Yoshio); 〒160-0023 東京都 新宿区 西新宿７丁目１１番１８号 ７１１ビルディング４階 Tokyo (JP).

[続葉有]

(54) Title: INFORMATION PROCESSING DEVICE AND METHOD, RECORDING MEDIUM, AND PROGRAM

(54) 発明の名称: 情報処理装置および方法、記録媒体、並びにプログラム

「ジャンル」,「出演者」 (a1)	グループID (a2)
バラエティ, タレントA (b)	××××
ニュース, タレントA (c)	××××
映画, タレントA (d)	××××
⋮	
バラエティ, タレントB (e)	××××
ニュース, タレントB (f)	××××
映画, タレントB (g)	××××
⋮	

a1..."GENRE", "PERFORMER"
a2...GROUP ID
b...VARIETY, TALENT A
c...NEWS, TALENT A
d...MOVIE, TALENT A
e...VARIETY, TALENT B
f...NEWS, TALENT B
g...MOVIE, TALENT B

(57) Abstract: An information processing device and method, a recording medium, and a program for appropriately generating taste information based on programs viewed continuously and programs viewed discontinuously and recommending a program corresponding to a user taste. A taste information extraction section (101) extracts program taste information based on meta data on the program and records a program viewing history. The taste information on the program viewed firstly is recorded as special program taste information. Moreover, user taste information is generated according to program taste information on a program whose viewing count exceeds a threshold value. A control section (103) sets reservation for automatic recording of a program whose viewing count exceeds the threshold value. When the reservation of automatic recording is released, a taste information update section (102) modifies the user taste information according to the meta data on the program of the time whose reservation has been released.

(57) 要約: 本発明は、連続して視聴される番組と、非連続的に視聴される番組に基づく嗜好情報を適正に生成し、よりユーザの嗜好に合う番組を推薦できるようにする情報処理装置および方法、記録媒体、並びにプログラムに関する。嗜好情報抽出部１０１で番組のメタデータに基づく番組嗜好情報を抽出し、番組の視聴履歴を記録する。初めて視聴された番組の嗜好情報は、特殊番組嗜好情報として記録される。また、視聴回数が閾値を超える番組の番組嗜好情報に基づいて、ユーザ嗜好情報が生成される。制御部１０３が、視聴回数が閾値を超える番組について自動録画の予約設定を行う。自動録画の予約が解除された場合、嗜好情報更新部１０２が、予約が解除された回の番組のメタデータに基づいて、ユーザ嗜好情報を変更する。

WO 2004/053736 A1

(81) 指定国 (国内): CN, KR, US.

(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

添付公開書類:
— 国際調査報告書

２文字コード及び他の略語については、定期発行される各*PCT*ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明細書

情報処理装置および方法、記録媒体、並びにプログラム

技術分野

本発明は、情報処理装置および方法、記録媒体、並びにプログラムに関し、特に、コンテンツの推薦を効率よく、かつ効果的に行うことができる情報処理装置および方法、記録媒体、並びにプログラムに関する。

背景技術

配信されるコンテンツから、ユーザの嗜好の合うコンテンツを検出し、そのコンテンツの情報を提供するコンテンツ推薦システムが存在する（例えば、特開２０００－２８７１８９号公報参照）。

このシステムでは、例えば、ユーザがよく利用するコンテンツの属性（例えば、ジャンル）を検出し、その属性毎にコンテンツが推薦される。

しかしながら、コンテンツの属性は、コンテンツの編成等の事情に基づいてコンテンツの提供元により設定されているので、その属性毎の推薦では、推薦すべきコンテンツを効率的または効果的に検出することができない場合があった。

発明の開示

本発明はこのような状況に鑑みてなされたものであり、コンテンツの推薦を行う側が、コンテンツの属性を利用してコンテンツをグループ化し、グループ毎にコンテンツの推薦を行うことができるようにするものである。

本発明の情報処理装置は、配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループ ID を付与し、コンテンツのグループ化を行うグループ化手段と、グループ ID 毎にコンテンツの利用頻度を算出する算出手段

と、算出手段により算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成手段と、生成手段により生成されたユーザ嗜好情報に基づいて、コンテンツを推薦する推薦手段とを備えることを特徴とする。

放送時間帯を表す属性項目と、少なくとも１つ以上の他の属性項目からなるグループ化項目が設定されており、グループ化手段は、そのグループ化項目に基づいてコンテンツのグループ化を行うことができる。

少なくとも放送時間帯を表す属性項目からなるグループ化項目と、他の属性項目からなるグループ化項目が設定されており、グループ化手段は、それらのグループ化項目に基づいてコンテンツのグループ化を行うことができる。

グループ化手段は、コンテンツの属性項目の内容を形態素解析し、その結果に基づいて、グループ化項目の内容の類似度を決定することができる。

生成手段は、グループに属するコンテンツの利用状態が所定の条件を満たしていないグループの利用頻度を、ユーザ嗜好情報の生成に利用しないようにすることができる。

推薦手段は、算出手段により算出された利用頻度が、予め設定された値より高いか否かを判定する判定手段と、判定手段により、利用頻度が、予め設定された値より高いと判定された場合、コンテンツの推薦情報に、頻繁に視聴されるコンテンツであることを表す定番フラグを設定する設定手段とを設けることができる。

生成手段は、算出手段により算出された利用頻度が予め設定された値より高いグループのコンテンツのメタデータを取得し、メタデータの特徴量を表すベクトルを抽出する抽出手段を備え、抽出手段により抽出されたベクトルに基づいて、嗜好情報を生成することができる。

生成手段は、算出手段により算出された利用頻度が予め設定された値より高いグループのコンテンツが、定番フラグが設定されたコンテンツ推薦情報に対応するコンテンツか否かを判定する定番判定手段を備え、定番判定手段により、コンテンツが、定番フラグが設定されたコンテンツ推薦情報に対応するコンテンツで

はないと判定された場合、抽出手段は、コンテンツのメタデータを取得し、メタデータの特徴量を表すベクトルを抽出することができる。

嗜好情報は、複数の属性とその属性の重要度を表す値により構成されるようにすることができる。

生成手段は、算出手段により算出された利用頻度に基づいて、コンテンツの熟知度を設定する熟知度設定手段を備え、熟知度に基づいて、嗜好情報の重要度を表す値に重み付けを行うことができる。

生成手段は、コンテンツの利用履歴に基づいて、利用回数が所定の値以下だけ利用されたコンテンツを検索する検索手段と、検索手段により検索されたコンテンツのメタデータに基づいて、特殊嗜好情報を生成する特殊嗜好情報生成手段とをさらに設けることができる。

嗜好情報または特殊嗜好情報の特徴量を表すベクトルを抽出する第１の抽出手段と、予め設定された期間に放送されるコンテンツのメタデータを取得し、メタデータの特徴量を表すベクトルを抽出する第２の抽出手段と、第１の抽出手段により抽出されたベクトルと第２の抽出手段により抽出されたベクトルの類似度を演算する演算手段とを備え、推薦手段は、類似度が高い順に、予め設定された数だけ第２の抽出手段により抽出されたベクトルを選択し、選択されたベクトルのメタデータに基づいて、コンテンツを推薦することができる。

本発明の情報処理方法は、配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループ ID を付与し、コンテンツのグループ化を行うグループ化ステップと、グループ ID 毎にコンテンツの利用頻度を算出する算出ステップと、算出ステップの処理で算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、生成ステップの処理で生成されたユーザ嗜好情報に基づいて、コンテンツを推薦する推薦ステップとを含むことを特徴とする。

本発明の記録媒体のプログラムは、配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDを付与することによっての、コンテンツのグループ化を制御するグループ化制御ステップと、グループID毎のコンテンツの利用頻度の算出を制御する算出制御ステップと、算出制御ステップの処理で算出された利用頻度に基づいての、ユーザの嗜好を表すユーザ嗜好情報の生成を制御する生成制御ステップと、生成制御ステップの処理で生成されたユーザ嗜好情報に基づいての、コンテンツの推薦を制御する推薦制御ステップとを含むことを特徴とする。

本発明のプログラムは、配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDを付与することによっての、コンテンツのグループ化を制御するグループ化制御ステップと、グループID毎のコンテンツの利用頻度の算出を制御する算出制御ステップと、算出制御ステップの処理で算出された利用頻度に基づいての、ユーザの嗜好を表すユーザ嗜好情報の生成を制御する生成制御ステップと、生成制御ステップの処理で生成されたユーザ嗜好情報に基づいての、コンテンツの推薦を制御する推薦制御ステップとを含むことを特徴とする。

本発明の情報処理装置および方法、並びにプログラムにおいては、配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDが付与され、コンテンツのグループ化が行われ、グループID毎にコンテンツの利用頻度が算出され、算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報が生成され、生成されたユーザ嗜好情報に基づいて、コンテンツが推薦される。

図面の簡単な説明

図１は、本発明を適用したコンテンツ推薦システムの構成例を示している。

図２は、メタデータの例を示す図である。

図３は、コンテンツのグループ化を説明する図である。

図４は、コンテンツのグループ化を説明する他の図である。

図５は、グループ ID が付されたメタデータの例を示す図である。

図６は、利用履歴の例を示す図である。

図７は、図１のコンテンツ推薦サーバの構成例を示すブロック図である。

図８は、図１のクライアント機器の構成例を示すブロック図である。

図９は、ユーザ嗜好情報生成処理を説明するフローチャートである。

図１０は、利用頻度の算出方法を説明する図である。

図１１Ａは、利用頻度の算出方法を説明する他の図である。

図１１Ｂは、利用頻度の算出方法を説明する他の図である。

図１２は、利用状態の確認処理を説明する図である。

図１３は、利用状態の確認処理を説明する他の図である。

図１４は、利用状態の確認処理を説明する他の図である。

図１５は、コンテンツ推薦情報生成処理を説明するフローチャートである。

図１６は、コンテンツ推薦情報の表示例を示す図である。

図１７は、他のコンテンツ推薦情報の表示例を示す図である。

図１８は、タイトルグループ化処理１について説明するフローチャートである。

図１９は、タイトルグループ化処理２について説明するフローチャートである。

図２０は、タイトルグループ化処理３について説明するフローチャートである。

図２１は、タイトルグループ化処理４について説明するフローチャートである。

図２２は、定番番組設定処理を説明するフローチャートである。

図２３は、嗜好情報抽出処理１を説明するフローチャートである。

図２４は、番組ベクトルの構成例を示す図である。

図２５は、嗜好情報の構成例を示す図である。

図２６は、嗜好情報抽出処理２を説明するフローチャートである。

図２７は、嗜好情報抽出処理３を説明するフローチャートである。

図２８は、嗜好情報変更処理を説明するフローチャートである。

図２９は、特殊嗜好情報生成処理を説明するフローチャートである。

図３０は、図７のＣＰＵの機能的構成例を示すブロック図である。

図３１は、推薦情報検索処理を説明するフローチャートである。

図３２は、特殊推薦情報検索処理を説明するフローチャートである。

発明を実施するための最良の形態

図１は、本発明を適用したコンテンツ推薦システムの構成例を示している。

配信サーバ３は、ストリーミングデータデータベース１から、ストリーミングデータを取得し、インターネットその他のネットワークを含むネットワーク６を介して、クライアント機器５に配信する。配信サーバ３はまた、メタデータデータベース２から、コンテンツのメタデータを取得し、ネットワーク６を介して、コンテンツ推薦サーバ４に供給する。

メタデータは、コンテンツ毎に、例えば、図２に示すような、「放送開始時刻」、「放送終了時刻」、「放送局」、「ジャンル」、「タイトル」、「出演者名」、「番組内容」、および「キーワード」等のコンテンツの属性を表す項目から構成されている。

コンテンツ推薦サーバ４は、１個以上の項目で設定されたグループ化項目について、その内容（グループ化項目を構成する各項目の要素）が一定以上類似する（各項目の要素が、全部一致、一部一致、または所定の類似度を表す値が所定以上となる）コンテンツに同一のグループＩＤを付与する（同一のグループにグループ化する）。

メタデータの項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目が設定されている場合、「８ch（放送局），００：００（放送開始時間）～０６：００（放送終了時刻）」のコンテンツ（８

chで、００：００～０６：００の間に放送される番組）には同一のグループIDが付与される。

すなわちこのグループ化項目の下では、図３に示すように、グループ化項目を構成する項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」の各要素の組み毎にコンテンツのグループ化がなされる。

また、メタデータの項目「ジャンル」および項目「出演者」からなるグループ化項目が設定されている場合、例えば、「バラエティ（ジャンル）、タレントＡ（出演者）」のコンテンツ（バラエティで、タレントＡが出演する番組）には同一のグループIDが付与される。

すなわちこのグループ化項目の下では、図４に示すように、グループ化項目を構成する項目「ジャンル」および項目「出演者」の各要素の組み毎にコンテンツのグループ化がなされる。

なおグループ化項目が複数設定されている場合、１個のコンテンツは、コンテンツの項目によっては、複数のグループに属することもあり得る。例えば、８chで、００：００～０６：００の間に放送され、バラエティで、タレントＡが出演する番組には、「８ch（放送局），００：００（放送開始時間）～０６：００（放送終了時刻）」のグループＩＤ（図３）と、「バラエティ（ジャンル）、タレントＡ（出演者）」のグループＩＤ（図４）が付与され、それぞれのグループに属することになる。

コンテンツ推薦サーバ４は、このようにグループＩＤが設定されたメタデータ（例えば、図５）を、適宜、クライアント機器５に送信する。

コンテンツ推薦サーバ４はまた、クライアント機器５から、コンテンツのグループＩＤを含む利用履歴を適宜取得し、その利用履歴に基づいて、グループ毎の利用頻度を算出する。そしてコンテンツ推薦サーバ４は、算出したその利用頻度をユーザの嗜好を表すものとして利用し、グループ毎にコンテンツの推薦を行う。例えば、高い利用頻度のグループに属するコンテンツに関する情報が、コンテンツ推薦情報としてクライアント機器５に送信される。

クライアント機器５は、配信サーバ３から配信されてきたコンテンツを利用するが、その利用履歴として、利用したコンテンツの、例えば図６に示すようなメタデータ（グループＩＤが設定されたもの）を、適宜、コンテンツ推薦サーバ４に供給する。

クライアント機器５は、コンテンツ推薦サーバ４から供給されたコンテンツ推薦情報を表示してユーザに提示する。ユーザは、それを参照することにより、自分の嗜好にあったコンテンツを選択することができる。

なお、ここでの配信サーバ３乃至クライアント機器５の通信は、ネットワーク６を介して行われているが、それぞれ直接通信する構成にすることもできる。

図７は、コンテンツ推薦サーバ４の構成例を示している。CPU（Central Processing Unit）１１は、ROM（Read Only Memory）１２に記憶されている、例えば、コンテンツ推薦用のプログラム等に従って所定の処理を実行する。RAM（Random Access Memory）１３には、CPU１１がその処理を実行する上において必要なデータなどが適宜記憶される。

CPU１１にはバス１４を介して入出力インタフェース１５が接続されている。入出力インタフェース１５には、キーボード、マウスなどよりなる入力部１６、LCD（Liquid Crystal Display）などよりなる出力部１７、メタデータ等を記憶する記憶部１８、およびネットワーク６を介して配信サーバ３またはクライアント機器５との通信を行う通信部１９が接続されている。

入出力インタフェース１５には、ドライブ２０が適宜接続され、CPU１１は、そこに装着される磁気ディスク３１、光ディスク３２、光磁気ディスク３３、または半導体メモリ３４との間でデータの授受を行う。

なお、ＣＰＵ１１の機能的構成例として、例えば、ユーザの嗜好情報を取得する嗜好情報取得部、配信サーバ３から番組のメタデータを取得するメタデータ取得部、およびコンテンツの推薦情報を生成する推薦情報生成部により構成されるようにすることも可能である。

図８は、クライアント機器５の構成例を示している。この構成は、コンテンツ推薦サーバ４の構成と基本的に同様であるので、その説明は省略する。

次に、ユーザ嗜好情報を生成する場合のコンテンツ推薦サーバ４の動作を、図９のフローチャートを参照して説明する。

ステップＳ１において、コンテンツ推薦サーバ４のCPU１１は、ユーザ嗜好情報を生成するタイミングであるか否かを判定し、そのタイミングであると判定した場合、ステップＳ２に進む。例えば、クライアント機器５からコンテンツ推薦情報（後述）の提供が要求されたとき、または予め決められた時期（例えば、毎週所定の時刻）が来たとき、ステップＳ２に進む。

ステップＳ２において、CPU１１は、通信部１９を介して、クライアント機器５から、所定の利用履歴を取得する。この例の場合、１週間前からの間に利用されたコンテンツのメタデータ（グループＩＤが設定されている）が取得される。CPU１１は、グループ毎のコンテンツの利用頻度（回数）を算出する。

項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目が設定されている場合、各メタデータには、そのグループ化項目の内容（グループ化項目を構成する各項目の要素の組み）に応じたグループＩＤが記述されているので、図１０に示すように、その内容（各項目の要素の組み）毎にコンテンツの利用頻度（回数）が算出される。

図１０が示すグループ毎の利用回数からは、８chで、２０：００～２１：００で放送される番組と、１０chで、１９：００～２０：００で放送される番組が最も多く視聴され（各７回）、それに続いて８chで、２２：００～２３：００の間に放送される番組が次に視聴（６回）されていることがわかる。

また項目「ジャンル」および項目「出演者」からなるグループ化項目が設定されている場合、各メタデータには、そのグループ化項目の内容（グループ化項目を構成する各項目の要素の組み）に応じたグループＩＤが記述されているので、図１１Ａに示すように、その内容（各項目の要素の組み）毎にコンテンツの利用頻度（回数）が算出される。

図１１Ａがグループ毎の利用回数からは、タレントＤが出演するバラエティ番組が最も多く視聴され（１０回）、それに続いてタレントＤが出演するニュース番組（８回）とタレントＣが出演するバラエティ番組（５回）が多く視聴されていることがわかる。

なお、利用回数は、コンテンツの配信数に応じて多くなる可能性があるので、そのままではユーザの嗜好を正確に対応しない。そのため、ステップＳ２で取得された利用履歴に対応する期間中に配信されたコンテンツ数で利用回数が正規化される。

例えば、例えば図１１の例の場合、タレントＤが出演するバラエティ番組が、その期間（この例の場合、１週間）の間に１０本配信され、タレントＤが出演するニュース番組が、その期間の間に１００本配信され、タレントＣが出演するバラエティ番組が、その期間の間に８０本配信された場合、図１１Ａの利用回数は、図１１Ｂに示すように、正規化される。このように利用回数を正規化することでユーザの嗜好を適切に対応した利用頻度を得ることができる。

図９に戻り、ステップＳ３において、コンテンツ編集サーバ４のCPU１１は、グループ化項目毎に、所定の閾値以上の利用回数（利用頻度）が得られたグループ（のグループのＩＤ）を検出する。

例えば、項目「放送局」、項目「放送開始時刻」、および項目「放送開始時刻」からなるグループ化項目に対する閾値が値７である場合、図１０の例では、「８ch，２０：００～２１：００」、および「１０ch，１９：００～２０：００」の２つのグループが検出される。

また項目「ジャンル」および項目「出演者」からなるグループ化項目に対する閾値が値０．０６である場合、図１１Ｂの例では、「バラエティ，タレントＤ」、「ニュース，タレントＤ」、および「バラエティ，タレントＣ」の３つのグループが検出される。

次に、ステップＳ４において、CPU１１は、ステップＳ３で検出した各グループのコンテンツがユーザの嗜好に合っているか否かを判定する。

例えば、そのグループに属するコンテンツの配信リストに基づいて、いまから遡って所定の回数（例えば、３回）連続して利用されなかったか否かが確認され、その回数連続して利用されなかった場合、そのグループのコンテンツはユーザの嗜好に合っていないと判定と判定される。

図１３に示すように、「バラエティ，タレントＤ」のグループの番組が３回連続して視聴されなかったとき、「バラエティ，タレントＤ」のグループのコンテンツは、ユーザの嗜好に合わないと判定される。

なお、図１２に示すように、グループ「８ch，２０：００～２１：００」のグループの、いまから遡って最も最近の番組Ａは視聴されなかったが、その前に配信された番組Ｂは視聴されているとき（視聴されないことが３回連続していないとき）、「８ch，２０：００～２１：００」のグループのコンテンツは、ユーザの嗜好に合わないとは判定されない（合っていると判定される）。

また図１４に示すように、過去に、所定の回数（例えば、３回）連続して利用されているとき、そのグループのコンテンツを、ユーザの嗜好にあっていると判定することもできる。

ステップＳ５において、CPU１１は、ステップＳ４での判定結果から、ユーザの嗜好に合うコンテンツのグループを検出する。

ステップＳ６において、CPU１１は、ステップＳ５で検出したグループのグループ ID を、ユーザ嗜好情報として、記憶部１８に記憶する。

この例の場合、項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目における「８ch，２０：００～２１：００」と「１０ch，１９：００～２０：００」のグループのグループ ID、および項目「ジャンル」および項目「出演者」からなるグループ化項目における「ニュース，タレントＤ」と「バラエティ，タレントＣ」のグループのグループ ID が、ユーザ嗜好情報として記憶部１８に記憶される。

次に、コンテンツ推薦情報を生成する場合のコンテンツ推薦サーバ４の動作を、図１５のフローチャートを参照して説明する。

ステップＳ２１において、コンテンツ推薦サーバ４のCPU１１は、クライアント機器５から、コンテンツ推薦情報の提供が要求されるまで待機し、その要求があったとき、ステップＳ２２に進み、記憶部１８から上述したようにして生成したユーザ嗜好情報を取得する。

ステップＳ２３において、CPU１１は、いまから配信されるコンテンツのメタデータ（グループIDが設定されているもの）から、ユーザ嗜好情報としてのグループIDと同じグループIDが設定されているメタデータを抽出する。CPU１１は、抽出したメタデータからコンテンツ推薦情報を生成する。

なお、ユーザ嗜好情報として記憶されているグループ ID が複数ある場合には、いずれのグループIDも付与されているコンテンツのメタデータが抽出されるようにすることができる。

ステップＳ２４において、CPU１１は、通信部１９を介して、ステップＳ２３で生成したコンテンツ推薦情報を、クライアント機器５に送信する。クライアント機器５は、コンテンツ推薦サーバ４から送信されてきたコンテンツ推薦情報を出力部５７に表示する。

図１６および図１７は、コンテンツ推薦情報の表示例を示している。

図１６の例では、項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目の「８ch，２０：００～２１：００」および「１０ch，１９：００～２０：００」のグループに属する番組の情報（タイトル等）がそれぞれ表示されている。

図１７の例では、項目「ジャンル」および項目「出演者」からなるグループ化項目の「ニュース，タレントＤ」および「バラエティ，タレントＣ」のグループに属する番組に関する情報（タイトル等）がそれぞれ表示されている。なお、各グループの番組の情報が表示されているウインドウは、表示画面の大きさによっては、図１７に示すように一部重なるように表示されるようにすることができる。

ユーザは、このように表示されたコンテンツ推薦情報を参照して、視聴する番組を選択することができる。

以上のように、ユーザの嗜好を把握するための利用頻度を、グループIDを利用してグループ毎に算出するようにしたので、メタデータの項目毎に利用頻度を算出する場合に比べ、その計算量を少なくすることができる。

また、コンテンツ推薦情報がグループ毎にまとまって表示されるようにしたので、表示スペースが小さいクライアント機器５においても適切にコンテンツ推薦情報を表示することができる。

また、以上においては、メタデータの項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」、並びに項目「ジャンル」および項目「出演者」を利用してグループ化を行ったが、項目「タイトル」や項目「内容」など他の項目を利用してグループ化を行うこともできる。その結果、例えば、再放送やスペシャル版の番組をオリジナルの番組と同じグループに属するコンテンツとして扱うことができるので、オリジナルであろうと、再放送されたものであろうと、その番組が視聴されれば、その利用履歴をユーザ嗜好情報生成に反映することができる。

ここで、図１８のフローチャートを参照して、項目「タイトル」を利用してグループ化を行う処理（項目「タイトル」からなるグループ化項目によるグループ化処理）（タイトルグループ化処理１）について説明する。

ステップＳ６１において、コンテンツ推薦サーバ４は、メタデータから、タイトルを抽出する。

ステップＳ６２において、コンテンツ推薦サーバ４は、タイトルを形態素解析し、単語に分解する。

例えばメタデータに含まれている映画の題名が「東海道三谷怪談」であった場合、これがタイトルとして形態素解析され、「東海道」、「三谷」、「怪談」の３つの単語が得られる。

ステップＳ６３において、コンテンツ推薦サーバ４は、解析された単語、もしくは、複数の単語から構成される単語群のうちのいずれかを抽出して、記憶部１８から、抽出された単語、または単語群に対応するグループＩＤを抽出する。

ここで、複数の単語から構成される単語群とは、形態素解析により得られた単語の組み合わせにより生成される単語群であり、例えば、形態素解析により得られた単語が「東海道」、「三谷」、「怪談」である場合、単語群は、「東海道・三谷」、「東海道・怪談」、「三谷・怪談」となる。

ステップＳ６４において、コンテンツ推薦サーバ４は、グループＩＤが抽出されたか否かを判断する。

ステップＳ６４において、対応するグループＩＤが抽出されなかったと判断された場合、抽出された単語、もしくは、複数の単語から構成される単語群には、まだグループＩＤが付けられていないので、ステップＳ６５において、抽出された単語、もしくは、複数の単語から構成される単語群に新たなグループＩＤを割り当てる。また、コンテンツ推薦サーバ４は、単語、もしくは、複数の単語から構成される単語群と、それに対応するグループＩＤを記憶する。

ステップＳ６４において、対応するグループＩＤが抽出されたと判断された場合、または、ステップＳ６５の処理の終了後、ステップＳ６６において、コンテンツ推薦サーバ４は、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループＩＤを抽出したか否かを判断する。

ステップＳ６６において、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループＩＤを抽出していないと判断された場合、処理は、ステップＳ６３に戻り、それ以降の処理が繰り返される。

ステップＳ６６において、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループＩＤが抽出されたと判断された場合、ステップＳ６７において、コンテンツ推薦サーバ４は、メタデータに、抽出されたまたは割り当てたグループＩＤを対応付けて、処理が終了される。

なお、類似したタイトルの番組を、同一のグループとするようにしても良い。例えば、タイトル「２年Ａ組銀八先生」の連続ドラマと、タイトル「２年Ａ組銀八先生スペシャル」の特別番組とを、同一のグループとしてグループ化することができるように、タイトルを構成する単語の形態素解析結果を基に、例えば、２

週間、１ヶ月、半年などの所定の期間の番組タイトルで、単語の一致率を総当りで算出し、単語の一致率が所定の値以上である場合、同一グループとするようにしてもよい。

次に、図１９のフローチャートを参照して、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理２（項目「タイトル」からなるグループ化項目による他のグループ化処理）について説明する。

ステップＳ４０１およびステップＳ４０２において、図１８を用いて説明した、ステップＳ６１およびステップＳ６２と同様の処理が実行される。すなわち、コンテンツ推薦サーバ４は、メタデータから、タイトルを抽出して形態素解析し、単語に分解する。

ステップＳ４０３において、コンテンツ推薦サーバ４は、解析された単語を基に、タイトル間の単語の一致度、すなわち、単語が一致している割合を示す一致率を算出する。

具体的には、タイトル「２年Ａ組銀八先生」と、タイトル「２年Ａ組銀八先生スペシャル」とが、それぞれ、「２」「年」「Ａ」「組」「銀八」「先生」と、「２」「年」「Ａ」「組」「銀八」「先生」「スペシャル」とに形態素分析された場合、この２つの番組のタイトルを構成する単語の一致率は、６／７で８５．７％となる。

ステップＳ４０４において、コンテンツ推薦サーバ４は、単語が、例えば、７０％などの所定の値以上一致しているか否かを判断する。この、一致率の閾値は、７０％以外のいかなる数値であっても良いことは言うまでもない。

ステップＳ４０４において、単語が、７０％などの所定の値以上一致していると判断された場合、ステップＳ４０５において、コンテンツ推薦サーバ４は、それらの番組に、同一のグループＩＤを対応付ける。また、コンテンツ推薦サーバ４は、一致した単語、または、単語群と、それに対応するグループＩＤを記憶する。

ステップＳ４０４において、７０％などの所定の値以下の一致率であると判断された場合、または、ステップＳ４０５の処理の終了後、ステップＳ４０６において、コンテンツ推薦サーバ４は、タイトルの総当りが終了したか否かを判断する。

ステップＳ４０６において、タイトルの総当りが終了していないと判断された場合、処理は、ステップＳ４０３に戻り、それ以降の処理が繰り返される。

ステップＳ４０６において、タイトルの総当りが終了したと判断された場合、処理が終了される。

このような処理により、タイトルを構成する単語の一致率を基にしたグループＩＤが対応付けられるので、例えば、連続ドラマとスペシャル番組などの類似したタイトルの番組を、同一のグループとして処理させるようにすることができる。

また、タイトルを構成する単語の一致率を基にグループを決定するようにすることにより、例えば、メタデータにおいて、数字の半角と全角、または、英字の半角と全角、もしくは、大文字と小文字などの表記ゆれがあった場合にも、同一タイトルの番組を、同一のグループとして検出することが可能となる。

また、単語の一致率に加えて、例えば、放送局や番組ジャンル、あるいは、放送開始時刻などを、グループ化の条件に加えるようにしても良い。例えば、ニュース番組などにおいては、タイトルが、「ニュース」を含む少ない単語によって構成されているので、図１９を用いて説明した処理では、異なる放送局の異なる形態のニュース番組であっても、同一のグループとして検出されてしまう恐れがあるので、単語の一致率に加えて、放送局も一致した場合、同一グループとするようにしても良い。

次に、図２０のフローチャートを参照して、放送局の一致を条件に加えて、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理３（項目「タイトル」と項目「放送局」からなるグループ化項目によるグループ化処理）について説明する。

ステップＳ４２１乃至ステップＳ４２４において、図１９を用いて説明した、ステップＳ４０１乃至ステップＳ４０４と同様の処理が実行される。すなわち、コンテンツ推薦サーバ４は、メタデータから、タイトルを抽出して形態素解析し、単語に分解する。そして、コンテンツ推薦サーバ４は、解析された単語を基に、タイトル間の単語の一致度を算出し、単語が、例えば、７０％などの所定の値以上一致しているか否かを判断する。

ステップＳ４２４において、単語が、７０％などの所定の値以上一致していると判断された場合、ステップＳ４２５において、コンテンツ推薦サーバ４は、それらの番組の放送局が一致しているか否かを判断する。

ステップＳ４２５において、これらの番組の放送局が一致していると判断された場合、ステップＳ４２６において、コンテンツ推薦サーバ４は、それらの番組に、同一のグループＩＤを対応付ける。また、コンテンツ推薦サーバ４は、一致した単語、または、単語群と、それに対応する放送局およびグループＩＤを記憶する。

ステップＳ４２４において、７０％などの所定の値以下の一致率であると判断された場合、ステップＳ４２５において、これらの番組の放送局が一致していないと判断された場合、または、ステップＳ４２６の処理の終了後、ステップＳ４２７において、コンテンツ推薦サーバ４は、タイトルの総当りが終了したか否かを判断する。

ステップＳ４２７において、タイトルの総当りが終了していないと判断された場合、処理は、ステップＳ４２３に戻り、それ以降の処理が繰り返される。

ステップＳ４２７において、タイトルの総当りが終了したと判断された場合、処理が終了される。

このような処理により、放送局の一致とタイトルを構成する単語の一致率を基にしたグループＩＤが対応付けられるので、例えば、類似したタイトルの番組を、同一のグループとする場合に、他局のニュース番組を同一のグループとするようなことを防ぐことができる。

なお、図２０においては、タイトルを構成する単語の一致率以外に、同一の放送局であるか否かを条件として、グループ化を行うものとして説明したが、放送局以外の、例えば、放送時間帯やジャンルなどを、タイトルを構成する単語の一致率以外の条件として、グループ化を実行するようにしても良いことは言うまでもない。

更に、例えば、連続ドラマや帯番組の放送開始時刻が、スポーツ中継や特別番組などのためにずれた場合においても、同一グループとして検出可能なように、タイトルを構成する単語の一致率以外の条件を、放送時刻が、例えば、１時間などの所定の時間範囲内のずれで一致しているか否かとして、グループ化を実行するようにしても良い。

図２１のフローチャートを参照して、放送時刻が、所定の時間範囲内のずれで一致しているか否かを条件に加えて、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理４（項目「タイトル」と項目「放送開始時刻」からなるグループ化項目によるグループ化処理）について説明する。

ステップＳ４４１乃至ステップＳ４４４において、図１９を用いて説明した、ステップＳ４０１乃至ステップＳ４０４と同様の処理が実行される。すなわち、コンテンツ推薦サーバ４は、メタデータから、タイトルを抽出して形態素解析し、単語に分解する。そして、コンテンツ推薦サーバ４は、解析された単語を基に、タイトル間の単語の一致度を算出し、単語が、例えば、７０％などの所定の値以上一致しているか否かを判断する。

ステップＳ４４４において、単語が、７０％などの所定の値以上一致していると判断された場合、ステップＳ４４５において、コンテンツ推薦サーバ４は、それらの番組の放送開始時刻が、例えば、１時間などの所定の範囲内のずれで一致しているか否かを判断する。

ステップＳ４４５において、それらの番組の放送開始時刻が所定の範囲内のずれで一致していると判断された場合、ステップＳ４４６において、コンテンツ推薦サーバ４は、それらの番組に、同一のグループＩＤを対応付ける。また、コン

テンツ推薦サーバ４は、一致した単語、または、単語群と、それに対応する放送開始時刻の範囲、およびグループＩＤを記憶する。

ステップＳ４４４において、７０％などの所定の値以下の一致率であると判断された場合、ステップＳ４４５において、それらの番組の放送開始時刻が所定の範囲以上にずれていると判断された場合、または、ステップＳ４４６の処理の終了後、ステップＳ４４７において、コンテンツ推薦サーバ４は、タイトルの総当りが終了したか否かを判断する。

ステップＳ４４７において、タイトルの総当りが終了していないと判断された場合、処理は、ステップＳ４４３に戻り、それ以降の処理が繰り返される。

ステップＳ４４７において、タイトルの総当りが終了したと判断された場合、処理が終了される。

このような処理により、放送開始時刻の所定の範囲内のずれを含む一致と、タイトルを構成する単語の一致率を基にしたグループＩＤが対応付けられるので、例えば、類似したタイトルの番組を同一のグループとする場合に、特別番組などによる放送時刻の変更のために、同一グループとして検出されるべき番組が、同一グループとして検出されないようなことを防ぐことができる。

なお、以上においては、コンテンツ推薦サーバ４が、ユーザ嗜好情報生成処理（図９）およびコンテンツ推薦情報処理（図１５）を行う場合を例として説明したが、クライアント機器５が、コンテンツ推薦サーバ４から供給されるグループＩＤが設定されたメタデータ（グループ化情報）を利用して、自分自身がグループ毎の利用頻度を算出してユーザ嗜好情報を生成し、それに基づいてコンテンツ推薦情報を生成することもできる。

また、頻繁に視聴される番組をいわゆる定番の番組として推薦し、推薦された番組が自動的に、視聴または録画されるようにすることもできる。図２２を参照して定番番組設定処理について説明する。この処理は、コンテンツ推薦サーバ４において、図１５を参照して上述したコンテンツ推薦情報生成処理を実行するのに先立って（事前に）実行される。

ステップＳ５０１において、CPU１１は、利用履歴を分析する。このとき、図９のステップＳ２の場合と同様に、クライアント機器５から、所定の期間に利用されたコンテンツのメタデータ（グループＩＤが設定されている）が取得され、グループ毎の利用回数（図１０）、または利用頻度（図１１Ｂ）が分析される。

ステップＳ５０２において、CPU１１は、利用回数（視聴回数）が所定の閾値を超えるグループがあるか否かを判定し、利用回数が閾値を超えるグループがあると判定された場合、ステップＳ５０３に進み、そのグループに属する番組（利用回数が閾値を超える番組）のコンテンツ推薦情報に、この番組が定番であることを表す定番フラグを設定する。

また、ステップＳ５０２において、視聴頻度が閾値を超えるグループがあるか否かが判定され、利用頻度が閾値を超える番組があると判定された場合、ステップＳ５０３において、そのグループに属する番組のコンテンツ推薦情報に定番フラグが設定されるようにしてもよい。

ステップＳ５０２において、視聴回数が閾値を超えるグループがないと判定された場合、処理は終了される。

このようにして、定番フラグが設定されたコンテンツ推薦情報が、図１５のコンテンツ推薦情報生成処理により、クライアント機器５に送信される。これにより、クライアント機器５において、例えば、定番フラグが設定されたコンテンツ推薦情報に対応する番組が自動録画されるようにすることができる。

上述した図９のユーザ嗜好情報生成処理においては、ユーザ嗜好情報としてグループＩＤが記憶されるようにしたが、番組のメタデータに含まれる複数の属性に基づいて、より詳細な嗜好情報が生成され、生成された嗜好情報に基づいて番組が推薦されるようにすることもできる。図２３を参照して、番組のメタデータに含まれる複数の属性に基づいて、より詳細な嗜好情報を生成する第１の例である嗜好情報抽出処理１について説明する。この処理は、例えば、予め決められた時期（毎週所定の時刻）に、コンテンツ推薦サーバ４において実行される。

ステップＳ５２１において、CPU１１は、利用履歴を分析する。このとき、図９のステップＳ２の場合と同様に、クライアント機器５から、所定の期間に利用されたコンテンツのメタデータ（グループＩＤが設定されている）が取得され、グループ毎の利用回数（図１０）、または利用頻度（図１１Ｂ）が分析される。

ステップＳ５２２において、CPU１１は、利用回数（利用頻度）が所定の閾値以上であるグループを検索する。なお、利用頻度が所定の閾値以上であるグループが検索されるようにしてもよい。

ステップＳ５２３において、CPU１１は、グループが検索されたか否かを判定し、グループが検索されたと判定された場合、ステップＳ５２４に進み、検索されたグループに属する番組のメタデータを分析する。このとき、番組が複数あった場合、複数の番組のメタデータが分析される。ステップＳ５２５において、CPU１１は、ステップＳ５２４で分析された番組のメタデータに基づいて、番組ベクトルを生成する。

図２４に、このとき生成される番組ベクトルＰＰの構成例を示す。この例では、番組ベクトルＰＰは、ステップＳ５２４で分析された番組のメタデータの属性「タイトル（番組名）」（Tm）、「ジャンル」（Gm）、「出演者」（Pm）、「放送局」（Sm）、「時間帯」（Hm）、・・・を要素とするベクトル PP＝（Ｔm，Ｇm，Ｐm，Ｓm，Ｈm，・・・）として構成されている。そして要素Ｔm，Ｇm，Ｐm，Ｓm，Ｈm，・・・も、複数の要素を持つベクトルとして構成される。

例えば属性「放送局」に対応するベクトルＳmは、ＭＨＫ総合、ＭＨＫ教育、亜細亜テレビ、ＴＡＳ、フシ、テレ日、東都、ＭＨＫ衛星第１、ＭＨＫ衛星第２、およびＷＯＷＯ（いずれも仮想的な放送局の名称）など、放送局の種類が限られているので、Ｓm＝｛ＭＨＫ総合，ＭＨＫ教育，亜細亜テレビ，ＴＴＳ，フジ，テレ日，東都，ＭＨＫ衛星第１，ＭＨＫ衛星第２，ＷＯＷＯ｝のように構成し、対応する放送局を値１、その他の放送局を値０とすることで得られる。すなわち対応する番組の放送局がＷＯＷＯであるとき、項目「放送局」のベクトルＳmは、Ｓm＝｛０，０，０，０，０，０，０，０，０，１｝とされる。

属性「ジャンル」に対応するベクトルＧｍも、ドラマ、バラエティ、スポーツ、映画、音楽、子供向け／教育、教養／ドキュメント、ニュース／報道、およびその他など、その種類が限られているので、Ｇｍ＝｛ドラマ，バラエティ，スポーツ，映画，音楽，子供向け／教育，教養／ドキュメント，ニュース／報道，その他｝のように構成し、対応するジャンルを値１、その他のジャンルを値０とすることで得られる。すなわち対応する番組のジャンルが教養／ドキュメントであるとき、項目「ジャンル」のベクトルＧｍは、Ｇｍ＝｛０，０，０，０，０，０，１，０，０｝とされる。

属性「時間帯」に対応するベクトルＨｍも、属性「放送局」のベクトルＳｍおよび「ジャンル」のベクトルＧｍと同様にして得ることができる。

一方、属性「タイトル」、「出演者」などのように、要素を限定することが容易でないものは、その属性を構成する単語とその頻度を表す数値の組を１つの要素とするベクトルがその属性のベクトルとなる。例えば、番組のメタデータの属性「出演者」が「personA，personB，・・・」である場合、属性「出演者」に対応するベクトルPmは、Ｐｍ＝｛（personA－１），（personB－３），・・・｝とされる。ここで、（personA－１）と（personB－３）は、メタデータの属性「出演者」を構成する単語として、personAとpersonBがそれぞれ１回または３回検出されたことを表す。

なお、ステップＳ５２２において複数の番組が検索された場合、ステップＳ５２５において、それぞれの番組毎に、番組ベクトルが生成される。

ステップＳ５２６において、CPU１１は、ステップＳ５２５において生成された番組ベクトルを統合して嗜好情報を生成する。このとき、例えば、複数の番組ベクトルのそれぞれの属性が足し合わされ、嗜好情報として生成される。

図２５に、このとき生成される嗜好情報の例を示す。この例では、嗜好情報は、ベクトルとして生成されており、属性「番組名（タイトル）」（Tup）、「ジャンル」（Gup）、「出演者」（Pup）、「放送局」（Sup）、「時間帯」（Hup）、・・・に対応するベクトルＵＰ＝（Ｔup，Ｇup，Ｐup，Ｓup，Ｈ

up，・・・）として構成されている。そして要素Tup，Gup，Pup，Sup，Hup，・・・も、複数の要素を持つベクトルとして構成される。

この例では、属性「番組名（タイトル）」に対応するベクトルTupは、Tup＝｛（title1－12），（title2—3），・・・｝とされている。これは、嗜好情報の属性「番組名」に要素「title1」と「title2」があり、それぞれの重要度が「12」と「3」に設定されていたことを表す。

重要度は、その要素に対するユーザの嗜好の度合いをあらわすもので、同一の要素が含む番組ベクトルが足し合わされるとき、重要度が1だけ加算される。例えば、番組ベクトルPP1乃至PP20の20個の番組ベクトルに基づいて、嗜好情報が生成される場合、番組ベクトルPP5、PP10、およびPP17の3つの番組ベクトルにおいて、属性Tmの中に「title2」が含まれていた場合、Tupの要素「title2」の重要度が「3」と設定される。

また、属性「ジャンル」に対応するベクトルGupは、Gup＝｛（ドラマー25），（バラエティー34），（スポーツー42），（映画－37），（音楽－73），（子供向け／教育－120），（教養／ドキュメント－3），（ニュース／報道－5），（その他－23）｝とされており、属性「ジャンル」に含まれる要素とその重要度により構成されている。

同様にして、嗜好情報の属性「放送局」に対応するベクトルSup、属性「出演者」に対応するベクトルPup、・・・が、各属性を構成する要素と重要度により構成されている。

ステップS523において、視聴回数が閾値以上のグループが検索されなかったと判定された場合、ステップS524乃至S526の処理はスキップされ、処理は終了する。

このようにして嗜好情報が生成される。嗜好情報は、所定の回数または頻度だけ利用された番組のメタデータに基づいて、生成されるのでよりユーザの嗜好を適確に反映したものとすることができる。

なお、嗜好情報は、ステップＳ５２１で、特定のユーザの利用履歴を分析することにより、ユーザ単位に生成されるものとしてもよいし、ステップＳ５２１で複数のユーザの利用履歴を分析することにより、一般的な（複数のユーザに共通の）嗜好情報が生成されるものとしてもよい。

ところで、図２３を参照して上述した嗜好情報抽出処理１によれば、同一の要素が含む番組ベクトルが足し合わされる都度、重要度が加算されるので、ユーザが頻繁に視聴する番組のメタデータに含まれる要素の重要度が極端に高くなり、偏った嗜好情報となってしまう場合もある。たとえば、毎日（月曜日から金曜日まで）放送される番組をユーザが欠かさず視聴している場合、その番組のメタデータに含まれる要素（例えば、タレントＡ）の重要度が、他の要素と比較して極端に高くなってしまう。このような、頻繁に視聴される番組（いわゆる定番の番組）のメタデータを嗜好情報に反映させないようにすることも可能である。図２６を参照して、番組のメタデータに含まれる複数の属性に基づいて、嗜好情報を生成する第２の例である嗜好情報抽出処理２について説明する。

ステップＳ５４１乃至５４３の処理は、図２３のステップＳ５２１乃至Ｓ５２３の処理と同様の処理なので、その説明は省略する。ステップＳ５４４において、CPU１１は、ステップＳ５４２で検索されたグループに属する番組が定番番組か否かを判定する。ここで、定番番組か否かの判定は、図２２を参照して上述した定番番組設定処理により設定された定番フラグに基づいて判定される。

ステップＳ５４４において、検索された番組が定番番組ではないと判定された場合、ステップＳ５４５に進む。そして、図２３のステップＳ５２４乃至Ｓ５２５の処理と同様にして、ステップＳ５４５において、番組のメタデータが分析され、ステップＳ５４６において番組ベクトルが生成され、ステップＳ５４７において嗜好情報が生成される。

一方、ステップＳ５４４において、検索された番組が定番番組であると判定された場合、ステップＳ５４５乃至Ｓ５４７の処理はスキップされる。

このようにすることで、定番番組に基づいて、嗜好情報が生成されることがなくなり、偏った嗜好情報が生成されることを防止することができる。

また、図２３を参照して上述した処理によれば、所定の回数（または頻度）以上に視聴されているグループに属する番組について、すべて同様に番組ベクトルが生成され、嗜好情報が生成されるようにしている。この例では、シリーズで放送される番組Ａ１，Ａ２，Ａ３，・・・（以下、個々に区別する必要がない場合、これらをまとめたシリーズ番組Ａと称する。他の場合も同様である）と番組Ｂ１，Ｂ２，Ｂ３，・・・（それぞれ１つのグループに属する番組）があり、例えば各グループの閾値が３回の場合、３回視聴した番組Ａ（正確には、シリーズ化された３個の番組が視聴されたシリーズ）についても、１０回視聴した番組Ｂ（シリーズ化された１０個の番組が視聴されたシリーズ）についても同様に番組ベクトルが生成される。

しかし、シリーズ番組Ａとシリーズ番組Ｂでは、ユーザの、その番組に対する知識が異なる場合がある。例えば、ユーザは、１０回視聴したシリーズ番組Ｂについては、番組の中にどのようなコーナーがあり、どのようなタレントが出演するかを熟知している可能性が高いが、３回しか視聴していないシリーズ番組Ａについては、番組のコーナー、タレントなどについて熟知していない可能性があり、場合によっては、シリーズ番組Ａを視聴したくないと感じる可能性もある。そこで、その番組に対する熟知度を反映して嗜好情報を生成する必要がある。図２６を参照して、番組のメタデータに含まれる複数の属性に基づいて、嗜好情報を生成する第３の例である嗜好情報抽出処理３について説明する。

ステップＳ５６１乃至Ｓ５６５の処理は、図２３のＳ５２１乃至Ｓ５２５の処理と同様の処理なので、その説明は省略する。ステップＳ５６６においてCPU１１は、番組の熟知度を特定する。熟知度は、ステップＳ５６１において分析された、シリーズ番組（すなわちグループ）の利用頻度に基づいて特定される。このとき、そのシリーズ番組の利用頻度に対応して３段階の熟知度が設定される。例えば、利用頻度が「０．１」以上のものは熟知度が「高」と設定され、利用頻度

が「０．０５」以上「０．１」未満のものは熟知度が「中」と設定され、利用頻度が「０．０５」未満のものは熟知度が「低」と設定される。

なお、熟知度の分類は、３段階に限られるものではない。また、熟知度が段階別に分類されずに、数値として設定されるようにしてもよい。あるいはまた、利用頻度ではなく、利用回数に基づいて、熟知度が設定されるようにしてもよい。

ステップＳ５６７において、CPU１１は、ステップＳ５６５で生成された番組ベクトルに対して、熟知度に基づいて、重み付けを行う。このとき、例えば、熟知度が「高」の番組ベクトルに含まれる要素に基づいて生成される、嗜好情報の重要度は３倍に設定され、熟知度が「中」の番組ベクトルに含まれる要素に基づいて生成される、嗜好情報の重要度は２倍に設定され、熟知度が「低」の番組ベクトルに含まれる要素に基づいて生成される、嗜好情報の重要度は１倍に設定される。

ステップＳ５６８において、CPU１１は、ステップＳ５６７において重み付けされた番組ベクトルに基づいて、嗜好情報を生成する。このとき、例えば、熟知度が「高」である番組ベクトルＰＰ１を構成するベクトルPm１が、Pm1＝（personA）であり、熟知度が「中」である番組ベクトルＰＰ２を構成するベクトルPm２が、Pm２＝（personB）であり、熟知度が「低」である番組ベクトルＰＰ３を構成するベクトルPm３が、Pm３＝（personC）である場合、嗜好情報の属性「出演者」に対応するベクトルPupは、Pup＝{（personA－３），（personB－２），（personC－１）}とされる。

このようにして、熟知度を反映した嗜好情報が生成される。なお、嗜好情報は、ステップＳ５６１で、特定のユーザの利用履歴を分析することにより、ユーザ単位に生成されるものとしてもよいし、ステップＳ５６１で複数のユーザの利用履歴を分析することにより、一般的な（複数のユーザに共通の）嗜好情報が生成されるものとしてもよい。例えば、まだ視聴履歴が蓄積されていないユーザに対しては、一般的な嗜好情報に基づいて番組（コンテンツ）を推薦することができる。嗜好情報は、番組の熟知度を反映して生成されているので、例えば、単に視聴率

の高い番組を推薦するより、信頼性の高い番組を、ユーザに推薦することができる。

以上においては、嗜好情報の重要度は、番組が視聴される毎に加算されるようにしているが、場合によっては重要度を減算する必要もある。例えば、ユーザはクライアント機器５において、自動録画が予約された定番番組について録画予約を解除することができる。ここで、録画予約が解除された番組は、それ以前は、頻繁に視聴されていたにもかかわらず、その回だけ録画予約をあえて解除したものであり、録画予約が解除された回は、ユーザの嗜好に合わない内容であったことが想定される。そこで、本発明においては、この録画予約が解除された番組のメタデータに基づいて、ユーザの嗜好情報の変更が行われる。

図２８を参照して、嗜好情報変更処理について説明する。この処理は、クライアント機器５のCPU５１により、自動録画の予約の解除が検知されたとき、コンテンツ推薦サーバ４に対して、自動録画予約が解除された番組の情報がネットワーク６を介して通知されたとき、コンテンツ推薦サーバ４により実行される。

ステップＳ５８１において、CPU１１は、自動録画予約が解除された回の番組（例えば、１０回放送されるシリーズ番組のうち、第３回目の番組）のメタデータを取得し、ステップＳ５８２において、取得したメタデータの属性を分析する。ステップＳ５８３において、CPU１１は、自動録画予約が設定された番組の嗜好情報の属性と、自動録画予約が解除された回の番組のメタデータの属性を比較し、ステップＳ５８４においてネガティブな要素を検出する。

例えば、番組Ｘの自動録画予約が設定され、ユーザによりその録画予約が解除された場合を考える。自動録画予約が設定された回の番組Ｘのメタデータに基づいて、生成された番組ベクトルＰＰ１において、属性「出演者」に対応するベクトルPm１が、Pm1＝（personA, personB）であり、自動録画予約が解除された回の番組Ｘのメタデータに基づいて、生成された番組ベクトルＰＰ２において、属性「出演者」に対応するベクトルPm２が、Pm２＝（personA, personB, personC）であった場合、自動録画予約が解除された回の番組Ｘには

「personC」が出演していたため、録画予約が解除されたと考えられ、ステップＳ５８４において、「personC」がネガティブな要素として検出される。

ステップＳ５８５において、CPU１１は、ステップＳ５８４で検出されたネガティブな要素に基づいて、ユーザの嗜好情報を変更する。このとき、ネガティブな要素の重要度が減算される。いまの場合、例えば、嗜好情報の属性「出演者」に対応するベクトルPupが、Pup＝{（personA－５），（personB－２），（personC－３）}であった場合、ステップＳ５８５において、Pup＝{（personA－５），（personB－２），（personC－２）}と変更され、「personC」の重要度が１だけ減算される。

このようにして、嗜好情報の変更が行われる。このようにすることで、ユーザが好まない属性の重要度は低く変更されるので、ユーザに番組（コンテンツ）の推薦を行うとき、よりユーザの嗜好に合った番組（コンテンツ）を推薦することができる。

また、以上においては、視聴回数が所定の回数以上のシリーズ番組のメタデータに基づいて嗜好情報が生成される例について説明したが、このようにして生成された嗜好情報に基づいた番組の推薦が続けて行われると、ユーザが飽きてしまう恐れがある。そこで、本発明においては、ユーザが初めて視聴した（過去に視聴していない）番組に注目する。初めて視聴した番組に対して、ユーザは特別な興味をもっていることが考えられるので、この番組のメタデータに基づいて、特殊嗜好情報を生成する。

図２９を参照して、コンテンツ推薦サーバ４による特殊嗜好情報生成処理について説明する。この処理は、例えば、ユーザにより、所定のコマンドを投入したとき実行されるようにしてもよいし、所定の周期（例えば、１週間）毎に自動的に実行されるようにしてもよい。

ステップＳ６０１において、CPU１１は利用履歴を検索する。このとき、クライアント機器５から、所定の期間（例えば、直近の６ヶ月間）に利用されたコン

テンツのメタデータ（グループＩＤが設定されている）が取得され、グループ毎の利用回数（図１０）が分析される。

ステップＳ６０２において、CPU１１は、視聴回数が１回のシリーズ番組（そのグループに属する番組の中の１番組だけが視聴されているグループ）を検出する。ステップＳ６０３において、CPU１１は、視聴回数が１回のシリーズ番組が検出されたか否かを判定し、シリーズ番組が検出されたと判定された場合、ステップＳ６０４に進み、検出されたシリーズ番組に属する番組のメタデータに基づいて、特殊嗜好情報を生成する。このとき、図２３のステップＳ５２４乃至Ｓ５２６と同様に、番組のメタデータに基づいて番組ベクトルが生成され、番組ベクトルに基づいて特殊嗜好情報が生成される。ステップＳ６０３において、視聴回数が１回の番組が検出されなかったと判定された場合、ステップＳ６０４の処理はスキップされる。

このようにして、ユーザが初めて視聴した番組のメタデータに基づいて特殊嗜好情報が生成される。

次に、図２３、図２６、および図２７を参照して上述した処理により生成した嗜好情報に基づいて、コンテンツが推薦される処理について説明する。

図３０は、図２３、図２６、および図２７を参照して上述した処理により生成した嗜好情報に基づいて、コンテンツを推薦する場合の、コンテンツ推薦サーバ４のCPU１１の機能的構成例を示すブロック図である。この例では、番組のメタデータを取得するメタデータ取得部１１１、特定のユーザの嗜好情報を取得する嗜好情報取得部１１２が設けられている。

メタデータ取得部１１１により取得された番組のメタデータは、番組ベクトル抽出部１１３に出力され、番組ベクトル抽出部１１３において、番組ベクトルが抽出される。また、嗜好情報取得部１１２により取得された嗜好情報は、嗜好ベクトル抽出部１１４に出力され、嗜好情報に基づく嗜好ベクトルが抽出される。番組ベクトル抽出部１１３により抽出された番組ベクトルと、嗜好ベクトル抽出

部１１４により抽出された嗜好ベクトルは、マッチング処理部１１５に出力され、マッチング処理部１１５は、番組ベクトルと嗜好ベクトルの類似度を算出する。

１つの嗜好ベクトルに対して複数の番組ベクトルとの類似度が算出され、マッチング処理部１１５は、類似度が高い順に所定の数の番組ベクトルを選択し、選択された番組ベクトルに対応する番組のメタデータを情報出力部１１６に出力する。

情報出力部１１６は、マッチング処理部１１５により選択された番組のメタデータを、例えば、記憶部１８に記憶させる。

次に、図３１のフローチャートを参照して、推薦情報検索処理について説明する。ステップＳ６２１において、メタデータ取得部１１１は、コンテンツ（番組）のメタデータを取得する。このとき、所定の基準に基づいて、複数の番組（例えば、今後１週間に放送される番組）のメタデータが取得される。ステップＳ６２２において、番組ベクトル抽出部１１３は、ステップＳ６２１で取得された番組のメタデータに基づいて、番組ベクトルを抽出する。このとき、図２４を参照して上述した番組ベクトルと同様に、複数の番組の番組ベクトルが抽出される。

ステップＳ６２３において、嗜好ベクトル抽出部１１４は、嗜好情報を取得する。このとき、特定のユーザの嗜好情報が取得される。ステップＳ６２４において、嗜好ベクトル抽出部１１４は、嗜好ベクトルを生成する。嗜好ベクトルは、図２５に示されるような嗜好情報が、そのまま嗜好ベクトルとして生成されるようにしてもよいし、嗜好情報を構成する特定の属性が抽出されて嗜好ベクトルとして生成されるようにしてもよい。

ステップＳ６２５において、マッチング処理部１１５は、ステップＳ６２２において生成された番組ベクトルＰＰと、ステップＳ６２４において生成された嗜好ベクトルＵＰのコサイン距離を算出する。いま、ベクトルＰＰとベクトルＵＰのなす角を θ とすると、$\cos\theta = PP \cdot UP / |PP||UP|$ となる。

例えば、嗜好ベクトルＵＰ（＝（Ｔup，Ｇup，Ｐup，Ｓup，Ｈup，・・・））におけるベクトルPupが、Ｐup＝｛（personA－１），（personＢ－１），（personＣ－１）｝であり、番組ベクトルＰＰ＝（Ｔm，Ｇm，Ｐm，Ｓm，Ｈm，・・・）におけるベクトルＰmが、Ｐm＝｛（personA－１），（personD－１），（personＥ－１）｝である場合、コサイン距離cosθpは、式（１）に従って算出される。

$\cos\theta p = (1 \cdot 1) / (\sqrt{3} \times \sqrt{3}) = 1/3$ ・・・（１）

なお、式中“・”は、内積を表し、“×”はスカラ演算を表す。

cosθpと同様にして、ベクトルTup，Gup，Sup，Hup，・・・と、ベクトルＴm，Ｇm，Ｓm，Ｈm，・・・とのコサイン距離cosθt，cosθg，cosθs，cosθh，・・・が算出される。

そして、算出されたコサイン距離を式（２）に従って合計し、類似度Simを計算する。

$Sim = \cos\theta t + \cos\theta g + \cos\theta p + \cos\theta s + \cos\theta h$ ・・・（２）

このようにして、嗜好ベクトルＵＰと番組ベクトルＰＰの類似度が算出される。なお、１つの嗜好ベクトルＵＰに対して複数の番組ベクトルＰＰとの類似度が算出される。これにより、それぞれの番組のメタデータについて、ユーザの嗜好情報との類似度が算出される。

ステップＳ６２６において、マッチング処理部１１５は、類似度の高い番組のメタデータを選択し、情報出力部１１６に出力する。このとき、ステップＳ６２５で算出された類似度に基づいて、類似度の高い順、すなわちSimの値が大きい順に所定の数（例えば、１０）だけ番組ベクトルＰＰが選択され、選択された番組ベクトルＰＰに対応する番組のメタデータが出力される。なお、類似度が所定の値より大きい番組ベクトルＰＰが全て選択され、選択された番組ベクトルＰＰに対応する番組のメタデータが出力されるようにしてもよい。

ステップＳ６２７において、情報出力部１１６は、ステップＳ６２６で抽出された番組のコンテンツ推薦情報をクライアント機器５に送信する。

このようにして、嗜好情報に基づく番組の推薦が行われる。

ところで、番組の推薦は、図２９を参照して説明した処理で生成された特殊嗜好情報に基づいて行うこともできる。図３２を参照して、コンテンツ推薦サーバ４による特殊推薦情報検索処理について説明する。この処理は、例えば、ユーザにより、所定のコマンドを投入したとき実行されるようにしてもよいし、所定の周期（例えば、１週間）毎に自動的に実行されるようにしてもよい。

ステップＳ６４１とＳ６４２の処理は、図３１のステップＳ６２１とＳ６２２の処理と同様の処理であるので、その説明は省略する。

ステップＳ６４３において、嗜好ベクトル抽出部１１４は、特殊嗜好情報を取得する。このとき、図２９を参照して上述した特殊嗜好情報生成処理により生成された特殊嗜好情報が取得される。そして、ステップＳ６４４において、嗜好ベクトル抽出部１１４は、ステップＳ６４３で取得された特殊嗜好情報に基づいて、嗜好ベクトルを生成する。

ステップＳ６４５とＳＳ６４６の処理は、図２３のステップＳ６２５とＳ６２６の処理と同様の処理であるので、その説明は省略する。

ステップＳ６２７において、情報出力部１１６は、ステップＳ６４６において抽出された番組のコンテンツ推薦情報をクライアント機器５に送信する。

このようにして、特殊嗜好情報に基づいて、コンテンツの推薦が行われる。上述したように、特殊嗜好情報は、ユーザが初めて視聴した番組のメタデータに基づいて生成されたものであり、特殊嗜好情報に基づいて、ユーザにコンテンツを推薦することにより、より新鮮な印象を与えることができる。

上述した一連の処理は、ソフトウェアにより実行することもできる。そのソフトウェアは、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

この記録媒体は、図７，８に示すように、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク３１または７１（フレキシブルディスクを含む）、光ディスク３２または７２（CD-ROM（Compact Disk-Read Only Memory），DVD（Digital Versatile Disk）を含む）、光磁気ディスク３３または７３（MD（Mini-Disk）（商標）を含む）、もしくは半導体メモリ３４または７４などよりなるパッケージメディアなどにより構成される。

また、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

なお、本明細書において、システムとは、複数の装置により構成される装置全体を表すものである。

産業上の利用可能性

本発明によれば、コンテンツの属性を表す項目から生成されたグループ化項目におけるグループ毎の利用頻度から、コンテンツ推薦を行うことができる。

請求の範囲

1． 配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDを付与し、コンテンツのグループ化を行うグループ化手段と、

前記グループID毎にコンテンツの利用頻度を算出する算出手段と、

前記算出手段により算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成手段と、

前記生成手段により生成された前記ユーザ嗜好情報に基づいて、コンテンツを推薦する推薦手段と

を備えることを特徴とする情報処理装置。

2．放送時間帯を表す属性項目と、少なくとも１つ以上の他の属性項目からなるグループ化項目が設定されており、

前記グループ化手段は、そのグループ化項目に基づいてコンテンツのグループ化を行う

ことを特徴とする請求の範囲第１項に記載の情報処理装置。

3． 少なくとも放送時間帯を表す属性項目からなる前記グループ化項目と、他の属性項目からなる前記グループ化項目が設定されており、

前記グループ化手段は、それらのグループ化項目に基づいてコンテンツのグループ化を行う

ことを特徴とする請求の範囲第１項に記載の情報処理装置。

4． 前記グループ化手段は、コンテンツの前記属性項目の内容を形態素解析し、その結果に基づいて、前記グループ化項目の内容の類似度を決定する

ことを特徴とする請求の範囲第１項に記載の情報処理装置。

5． 前記生成手段は、グループに属するコンテンツの利用状態が所定の条件を満たしていないグループの利用頻度を、前記ユーザ嗜好情報の生成に利用しない

ことを特徴とする請求の範囲第１項に記載の情報処理装置。

6． 前記推薦手段は、

前記算出手段により算出された前記利用頻度が、予め設定された値より高いか否かを判定する判定手段と、

前記判定手段により、前記利用頻度が、予め設定された値より高いと判定された場合、前記コンテンツの推薦情報に、頻繁に視聴されるコンテンツであることを表す定番フラグを設定する設定手段と

を備えることを特徴とする請求の範囲第１項に記載の情報処理装置。

７．　前記生成手段は、

前記算出手段により算出された前記利用頻度が予め設定された値より高いグループのコンテンツのメタデータを取得し、前記メタデータの特徴量を表すベクトルを抽出する抽出手段を備え、

前記抽出手段により抽出されたベクトルに基づいて、前記嗜好情報を生成することを特徴とする請求の範囲第１項に記載の情報処理装置。

８．　前記生成手段は、

前記算出手段により算出された前記利用頻度が予め設定された値より高いグループのコンテンツが、前記定番フラグが設定されたコンテンツ推薦情報に対応するコンテンツか否かを判定する定番判定手段を備え、

前記定番判定手段により、前記コンテンツが、前記定番フラグが設定されたコンテンツ推薦情報に対応するコンテンツではないと判定された場合、前記抽出手段は、コンテンツのメタデータを取得し、前記メタデータの特徴量を表すベクトルを抽出する

ことを特徴とする請求の範囲第７項に記載の情報処理装置。

９．　前記嗜好情報は、複数の属性とその属性の重要度を表す値により構成される

ことを特徴とする請求の範囲第７項に記載の情報処理装置。

１０．　前記生成手段は、

前記算出手段により算出された前記利用頻度に基づいて、前記コンテンツの熟知度を設定する熟知度設定手段を備え、

前記熟知度に基づいて、前記嗜好情報の重要度を表す値に重み付けを行う

ことを特徴とする請求の範囲第７項に記載の情報処理装置。

１１．　前記生成手段は、

前記コンテンツの利用履歴に基づいて、利用回数が所定の値以下だけ利用されたコンテンツを検索する検索手段と、

前記検索手段により検索されたコンテンツのメタデータに基づいて、特殊嗜好情報を生成する特殊嗜好情報生成手段とをさらに備える

ことを特徴とする請求の範囲第７項に記載の情報処理装置。

１２．　前記嗜好情報または前記特殊嗜好情報の特徴量を表すベクトルを抽出する第１の抽出手段と、

予め設定された期間に放送されるコンテンツのメタデータを取得し、前記メタデータの特徴量を表すベクトルを抽出する第２の抽出手段と、

前記第１の抽出手段により抽出されたベクトルと第２の抽出手段により抽出されたベクトルの類似度を演算する演算手段とを備え、

前記推薦手段は、前記類似度が高い順に、予め設定された数だけ前記第２の抽出手段により抽出されたベクトルを選択し、選択されたベクトルのメタデータに基づいて、コンテンツを推薦する

ことを特徴とする請求の範囲第１１項に記載の情報処理装置。

１３．　配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループ ID を付与し、コンテンツのグループ化を行うグループ化ステップと、

前記グループ ID 毎にコンテンツの利用頻度を算出する算出ステップと、

前記算出ステップの処理で算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、

前記生成ステップの処理で生成された前記ユーザ嗜好情報に基づいて、コンテンツを推薦する推薦ステップと

を含むことを特徴とする情報処理方法。

１４．　配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDを付与することによっての、コンテンツのグループ化を制御するグループ化制御ステップと、

前記グループID毎のコンテンツの利用頻度の算出を制御する算出制御ステップと、

前記算出制御ステップの処理で算出された前記利用頻度に基づいての、ユーザの嗜好を表すユーザ嗜好情報の生成を制御する生成制御ステップと、

前記生成制御ステップの処理で生成された前記ユーザ嗜好情報に基づいての、コンテンツの推薦を制御する推薦制御ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

１５．　配信されるコンテンツの属性を表す属性項目の中の１個以上の属性項目からなるグループ化項目が一定以上の類似度をもって類似するコンテンツに、同一のグループIDを付与することによっての、コンテンツのグループ化を制御するグループ化制御ステップと、

前記グループID毎のコンテンツの利用頻度の算出を制御する算出制御ステップと、

前記算出制御ステップの処理で算出された前記利用頻度に基づいての、ユーザの嗜好を表すユーザ嗜好情報の生成を制御する生成制御ステップと、

前記生成制御ステップの処理で生成された前記ユーザ嗜好情報に基づいての、コンテンツの推薦を制御する推薦制御ステップと

を含む処理をコンピュータに実行させることを特徴とするプログラム。

図１

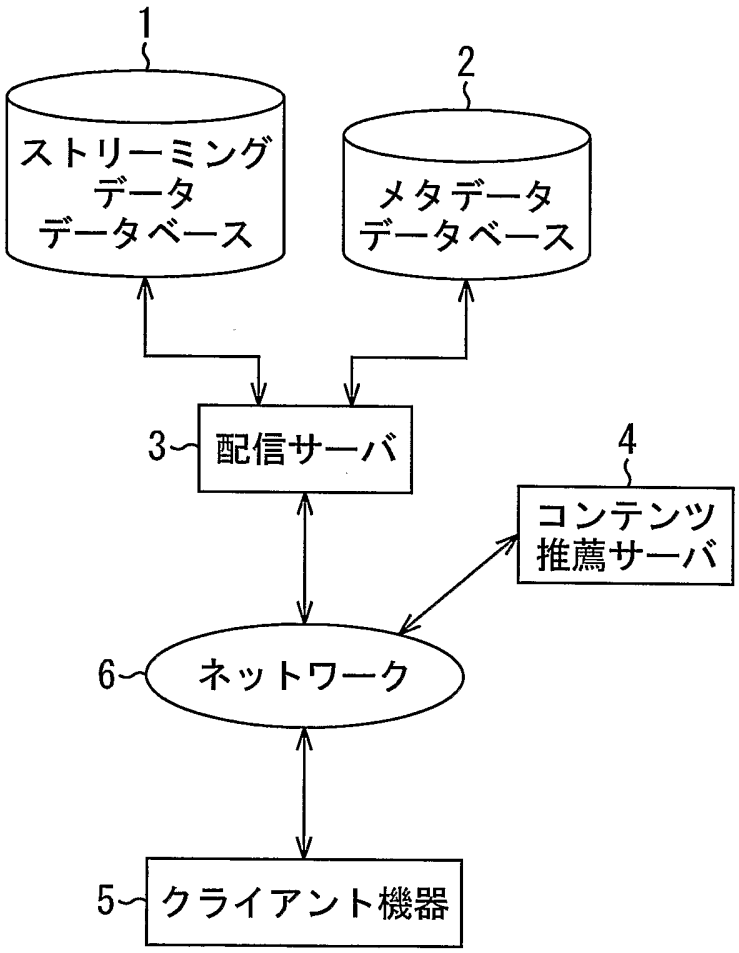

1
ストリーミング
データ
データベース
2
メタデータ
データベース
3
配信サーバ
4
コンテンツ
推薦サーバ
6
ネットワーク
5
クライアント機器

図２

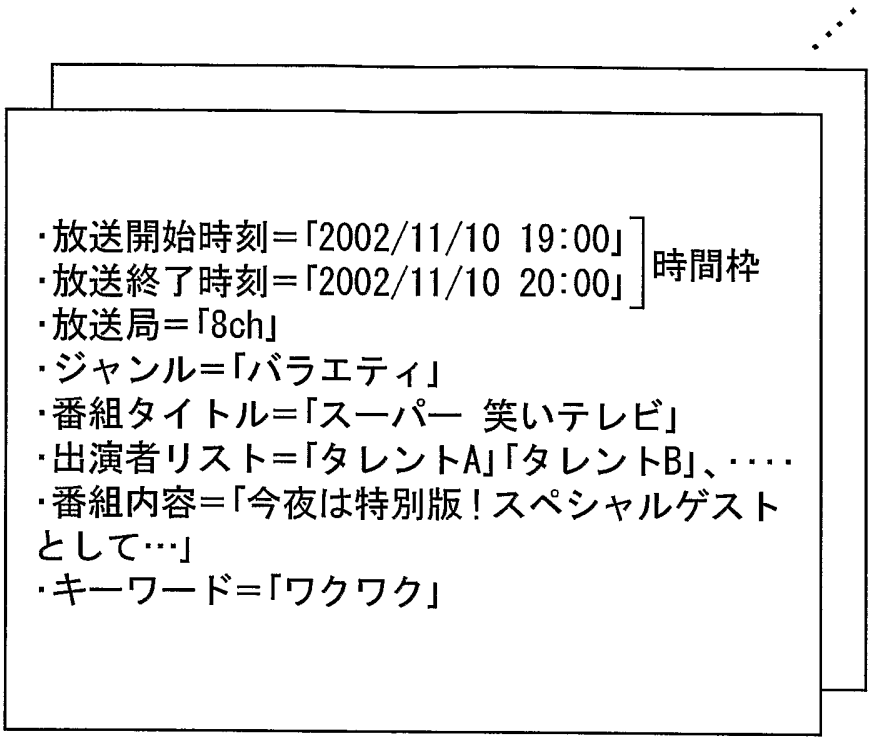

・放送開始時刻＝「2002/11/10 19:00」
・放送終了時刻＝「2002/11/10 20:00」
時間枠
・放送局＝「8ch」
・ジャンル＝「バラエティ」
・番組タイトル＝「スーパー 笑いテレビ」
・出演者リスト＝「タレントA」「タレントB」、・・・・
・番組内容＝「今夜は特別版！スペシャルゲストとして…」
・キーワード＝「ワクワク」

図３

「放送局」,「開始時刻」,「終了時刻」	グループID
8ch, 00:00～06:00	××××
8ch, 06:00～07:00	××××
8ch, 07:00～08:00	××××
⋮	
10ch, 00:00～06:00	××××
10ch, 06:00～07:00	××××
10ch, 07:00～08:00	××××
⋮	

図４

「ジャンル」,「出演者」	グループID
バラエティ, タレントA	××××
ニュース, タレントA	××××
映画, タレントA	××××
⋮	
バラエティ, タレントB	××××
ニュース, タレントB	××××
映画, タレントB	××××
⋮	

図５

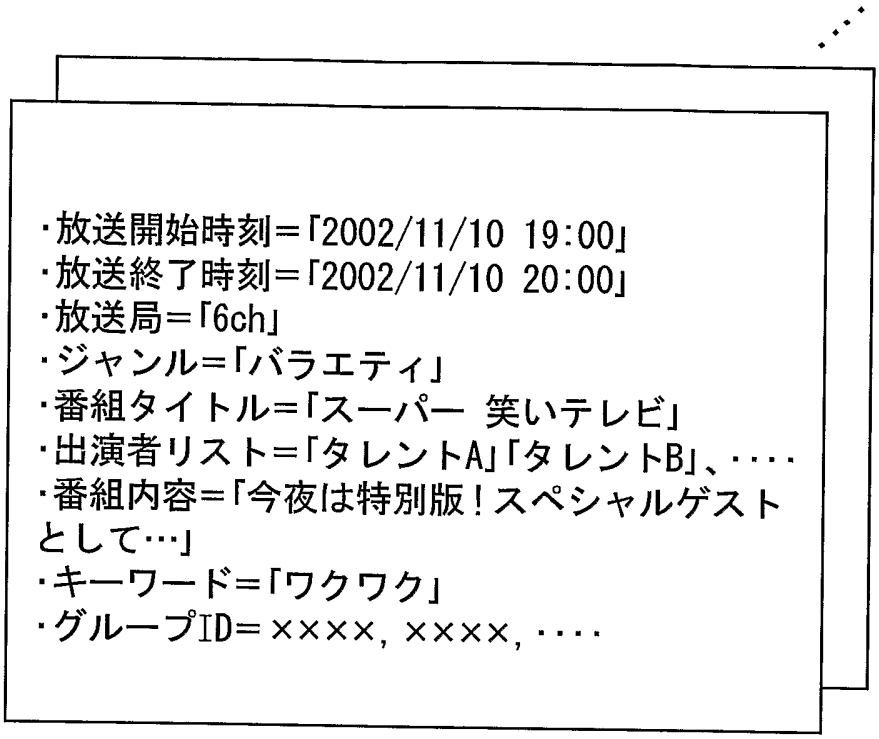

・放送開始時刻＝「2002/11/10 19:00」
・放送終了時刻＝「2002/11/10 20:00」
・放送局＝「6ch」
・ジャンル＝「バラエティ」
・番組タイトル＝「スーパー 笑いテレビ」
・出演者リスト＝「タレントA」「タレントB」、・・・・
・番組内容＝「今夜は特別版！スペシャルゲストとして…」
・キーワード＝「ワクワク」
・グループID＝××××, ××××, ・・・・

図６

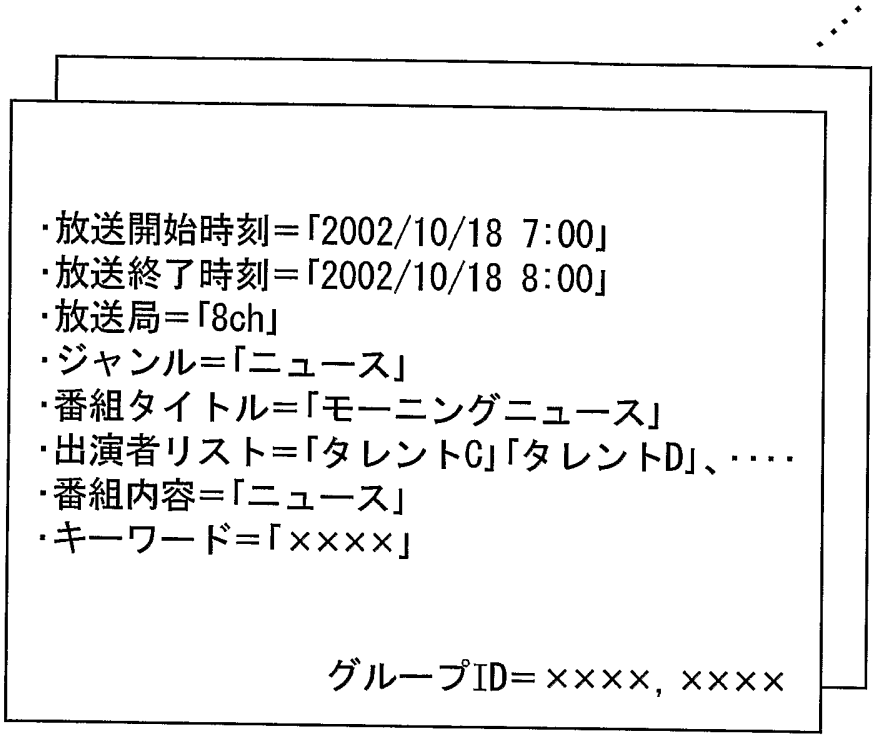

・放送開始時刻＝「2002/10/18 7:00」
・放送終了時刻＝「2002/10/18 8:00」
・放送局＝「8ch」
・ジャンル＝「ニュース」
・番組タイトル＝「モーニングニュース」
・出演者リスト＝「タレントC」「タレントD」、・・・・
・番組内容＝「ニュース」
・キーワード＝「××××」
グループID＝××××, ××××

図7

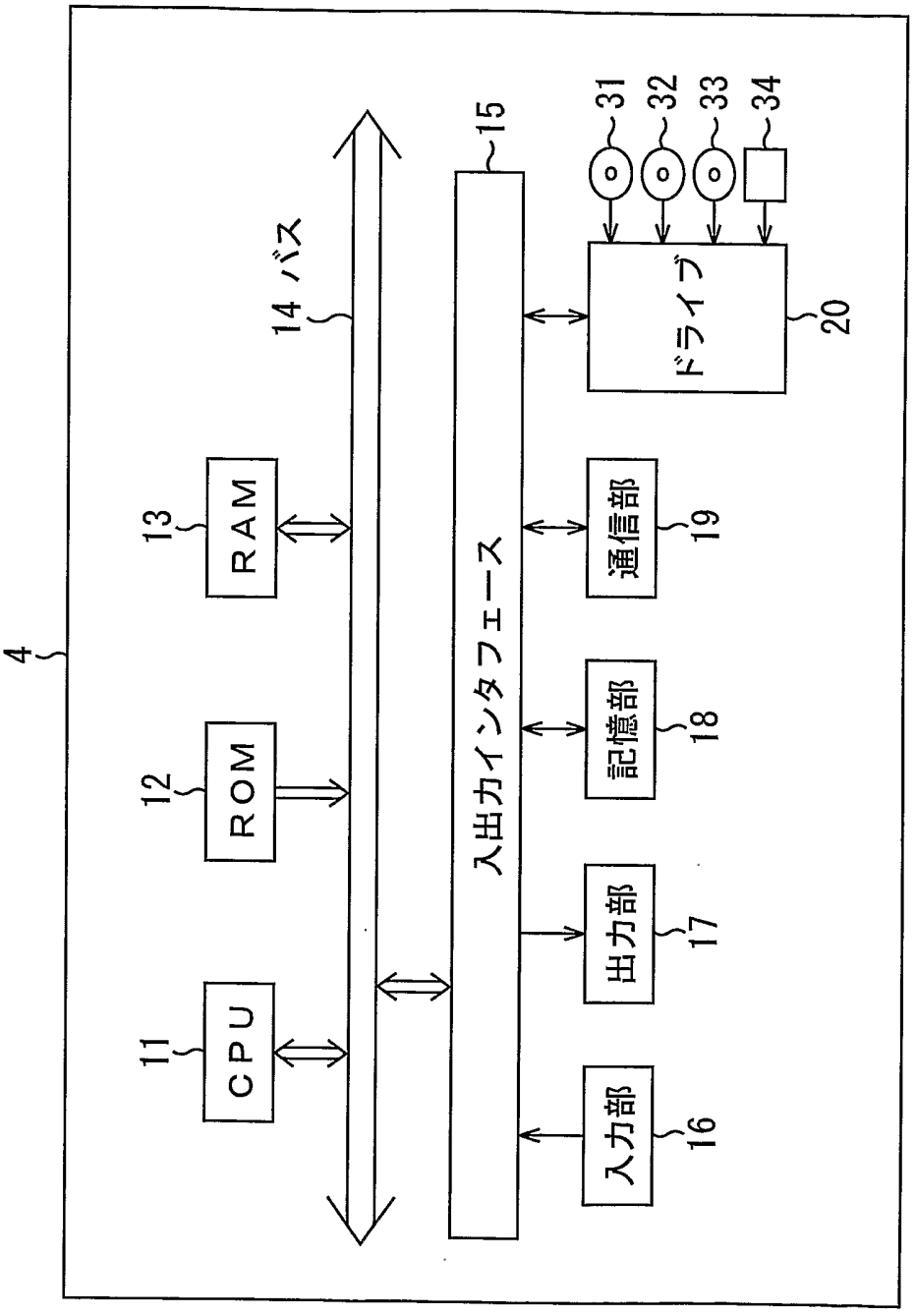

4
11
CPU
12
ROM
13
RAM
14 バス
15
入出力インタフェース
16
入力部
17
出力部
18
記憶部
19
通信部
20
ドライブ
31
32
33
34

図8

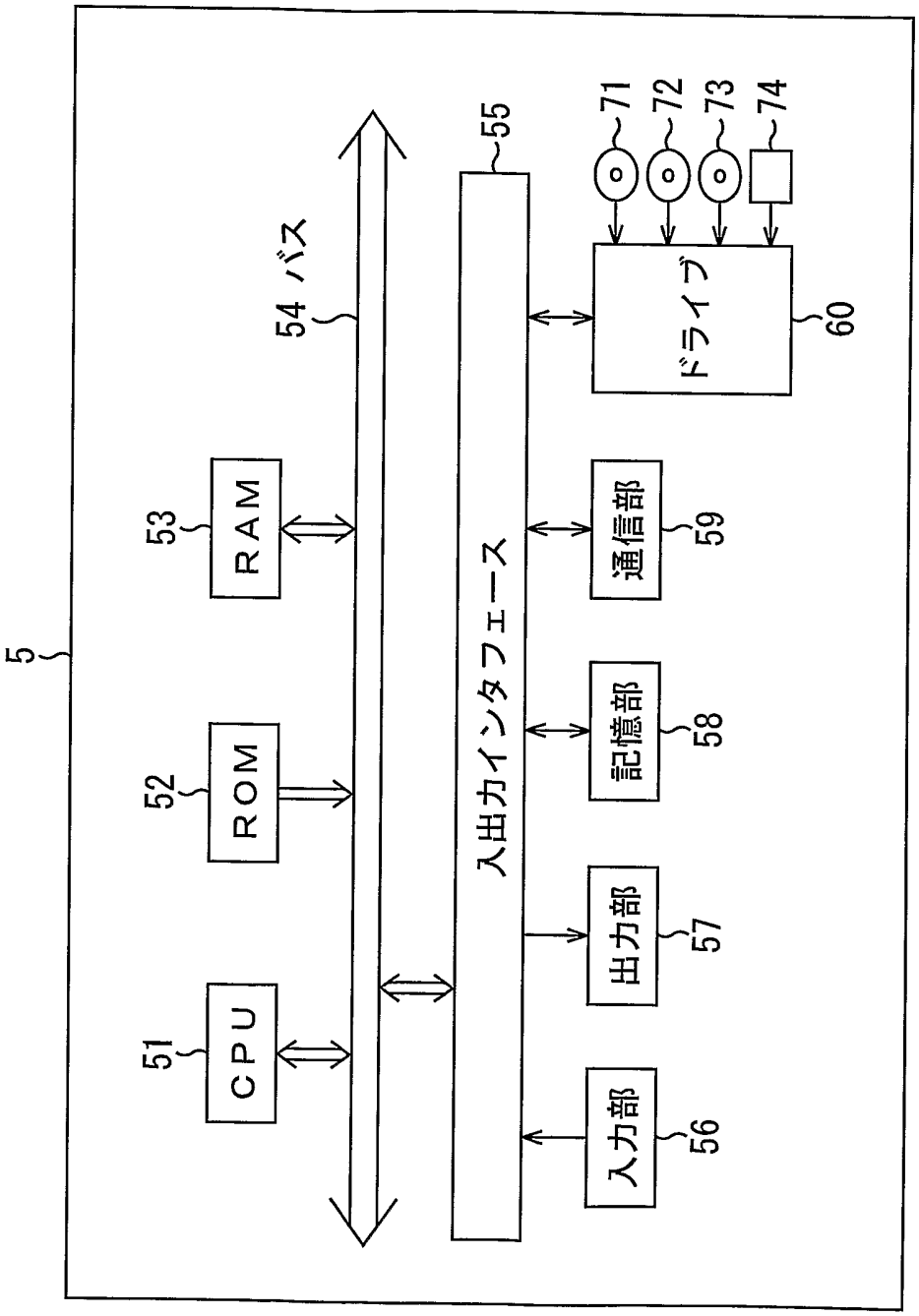

5
CPU
51
ROM
52
RAM
53
54 バス
入出力インタフェース
55
入力部
56
出力部
57
記憶部
58
通信部
59
ドライブ
60
71
72
73
74

図９

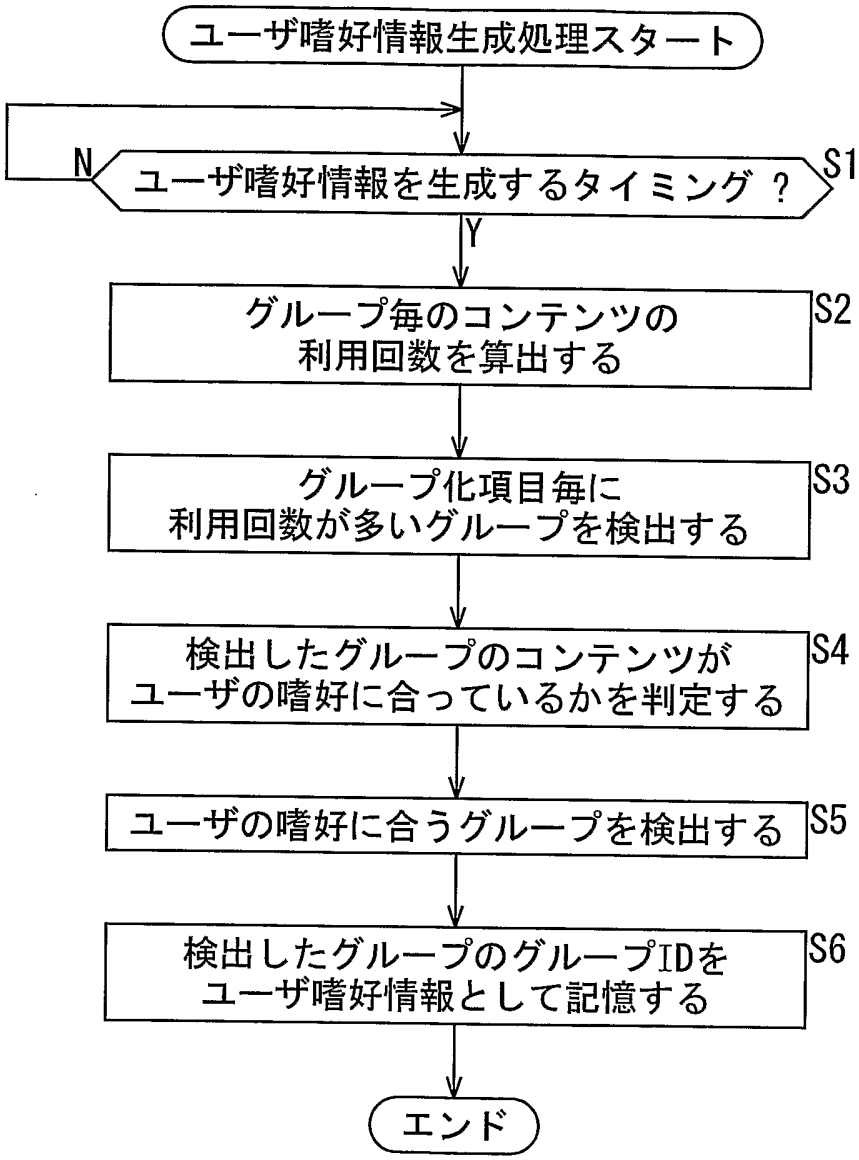

ユーザ嗜好情報生成処理スタート
N
ユーザ嗜好情報を生成するタイミング ？
S1
Y
グループ毎のコンテンツの
利用回数を算出する
S2
グループ化項目毎に
利用回数が多いグループを検出する
S3
検出したグループのコンテンツが
ユーザの嗜好に合っているかを判定する
S4
ユーザの嗜好に合うグループを検出する
S5
検出したグループのグループIDを
ユーザ嗜好情報として記憶する
S6
エンド

図10

「放送局」,「開始時刻」,「終了時刻」	回数
8ch, 20:00～21:00	7
10ch, 19:00～20:00	7
8ch, 22:00～23:00	6
⋮	

図11A

「ジャンル」,「出演者」	回数
バラエティ, タレントD	10
ニュース, タレントD	8
バラエティ, タレントC	5
⋮	

図11B

「ジャンル」,「出演者」	正規化回数
バラエティ, タレントD	0.25
ニュース, タレントD	0.08
バラエティ, タレントC	0.0625
⋮	

図12

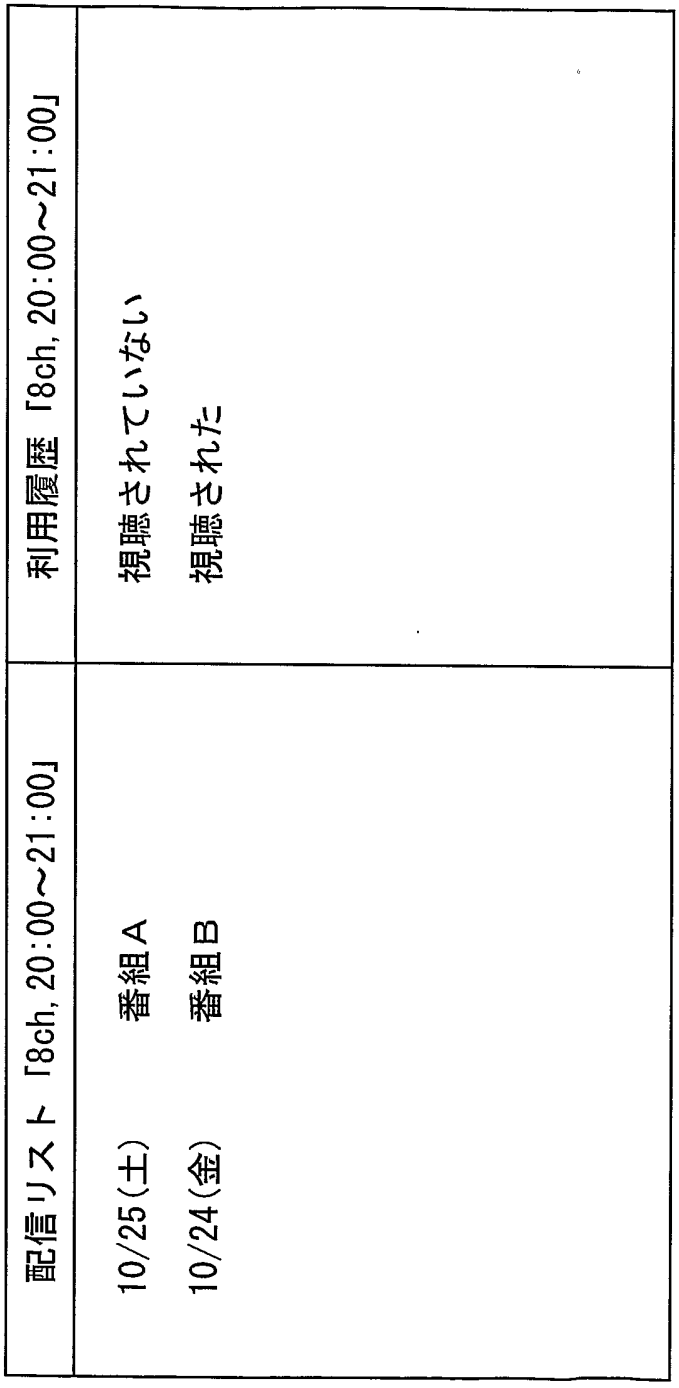

配信リスト「8ch, 20:00～21:00」
10/25(土) 番組A
10/24(金) 番組B
利用履歴「8ch, 20:00～21:00」
視聴されていない
視聴された

図13

配信リスト「バラエティ，タレントD」		利用履歴「バラエティ，タレントD」
10/25(土)	番組ＡＡ	視聴されていない
	番組ＡＢ	視聴されていない
10/24(金)	番組ＢＡ	視聴されていない

図14

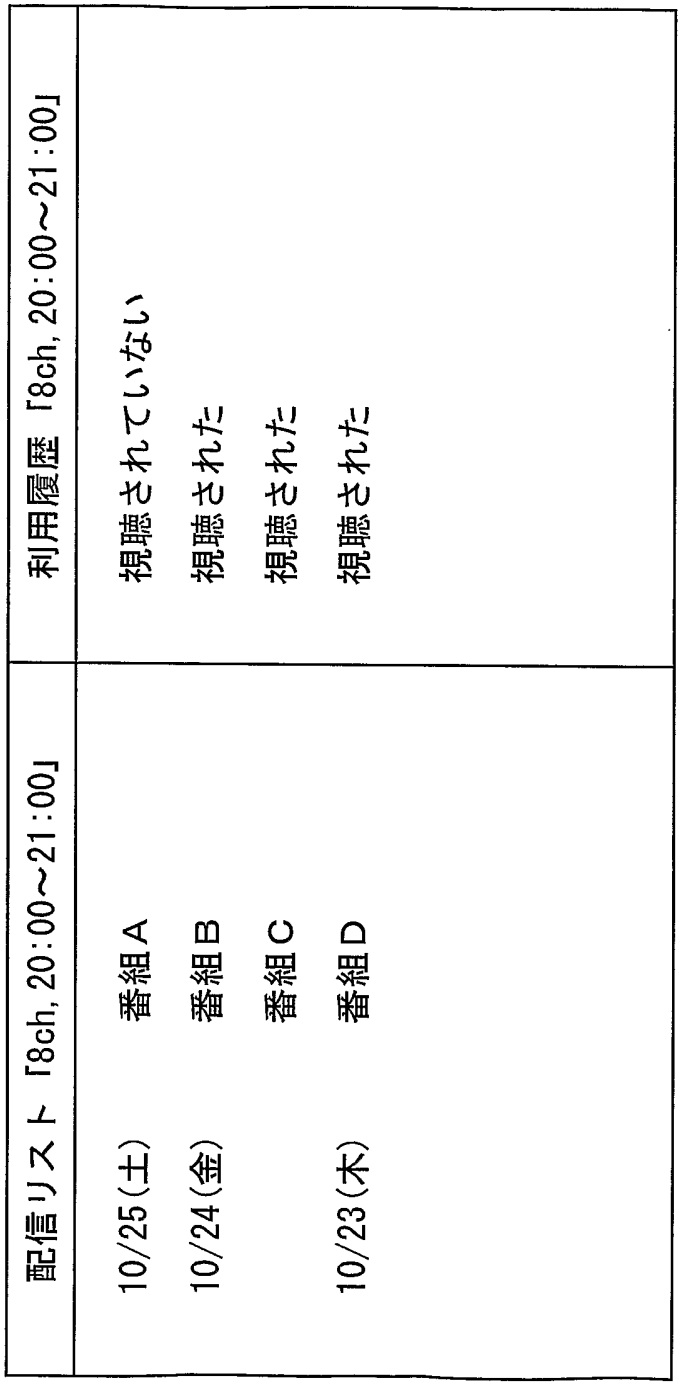

配信リスト「8ch, 20:00～21:00」
利用履歴「8ch, 20:00～21:00」
10/25(土) 番組A
視聴されていない
10/24(金) 番組B
視聴された
番組C
視聴された
10/23(木) 番組D
視聴された

図15

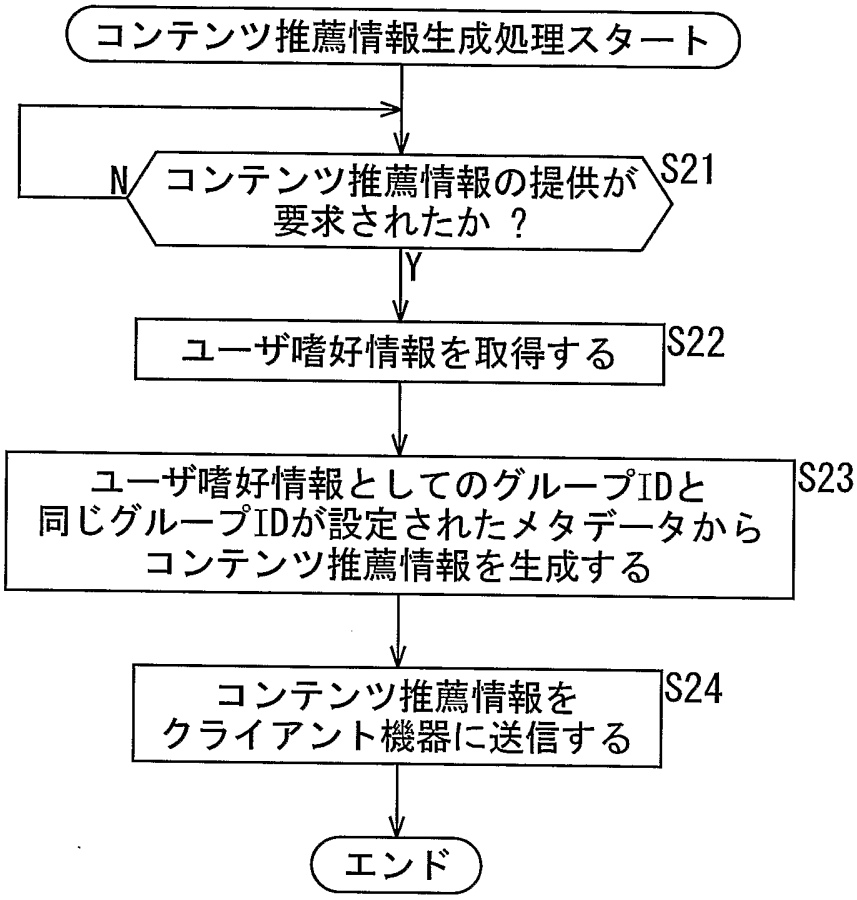

コンテンツ推薦情報生成処理スタート
N
コンテンツ推薦情報の提供が要求されたか ？
S21
Y
ユーザ嗜好情報を取得する
S22
ユーザ嗜好情報としてのグループIDと同じグループIDが設定されたメタデータからコンテンツ推薦情報を生成する
S23
コンテンツ推薦情報をクライアント機器に送信する
S24
エンド

図16

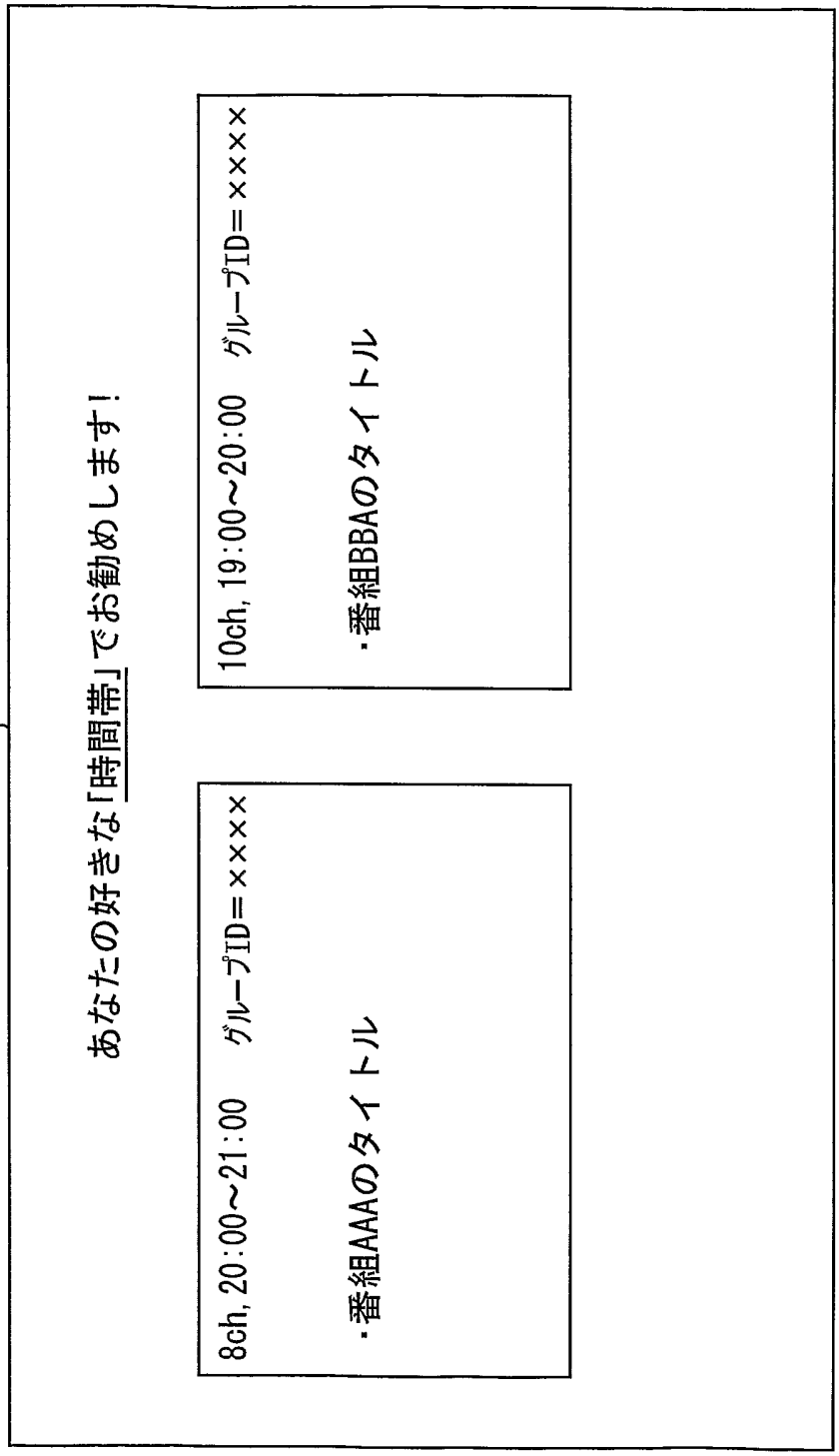

57
あなたの好きな「時間帯」でお勧めします！
8ch, 20:00～21:00 グループID＝××××
・番組AAAのタイトル
10ch, 19:00～20:00 グループID＝××××
・番組BBAのタイトル

図17

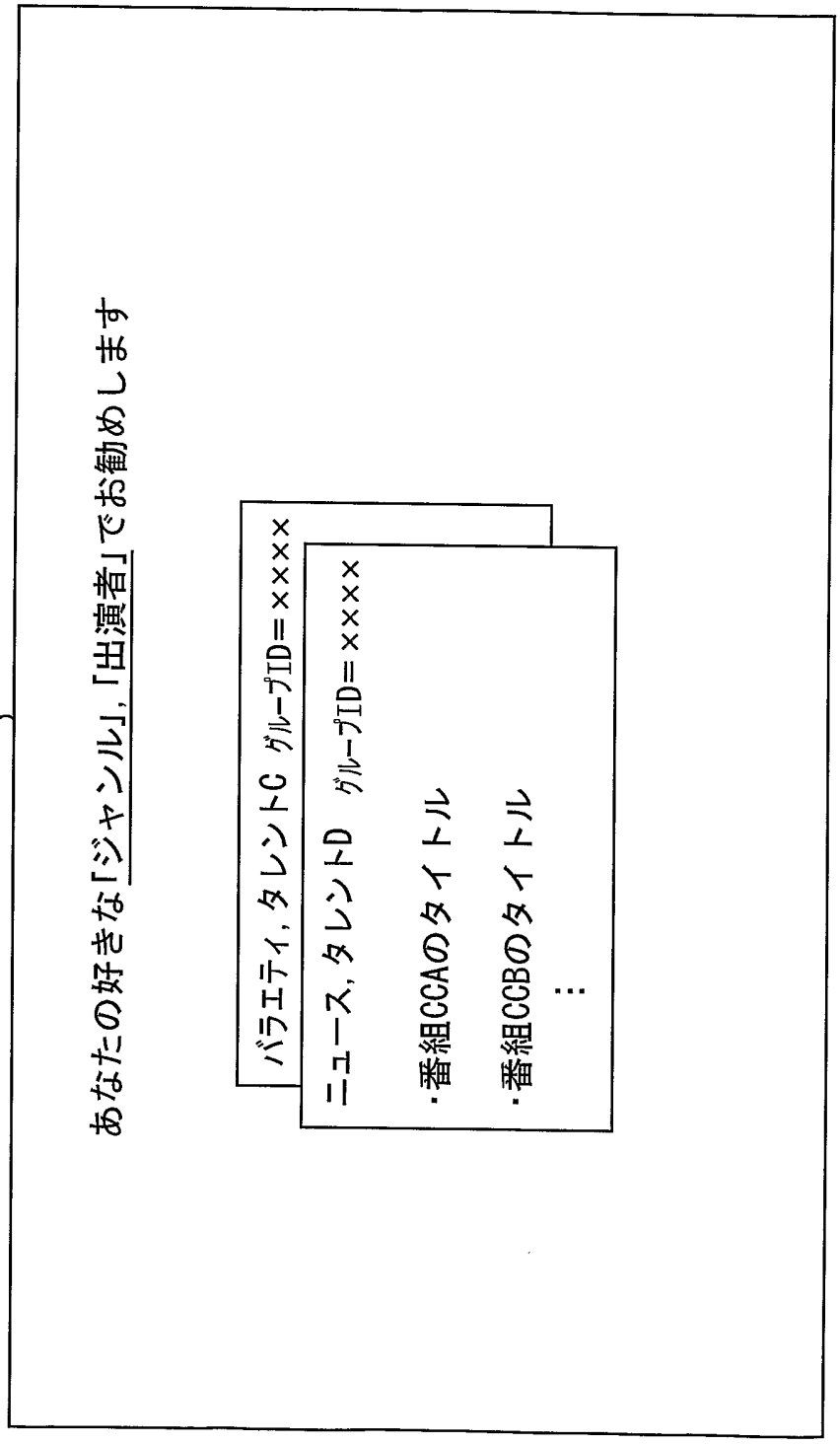

57
あなたの好きな「ジャンル」,「出演者」でお勧めします
バラエティ,タレントC グループID=××××
ニュース,タレントD グループID=××××
・番組CCAのタイトル
・番組CCBのタイトル
・・・

図18

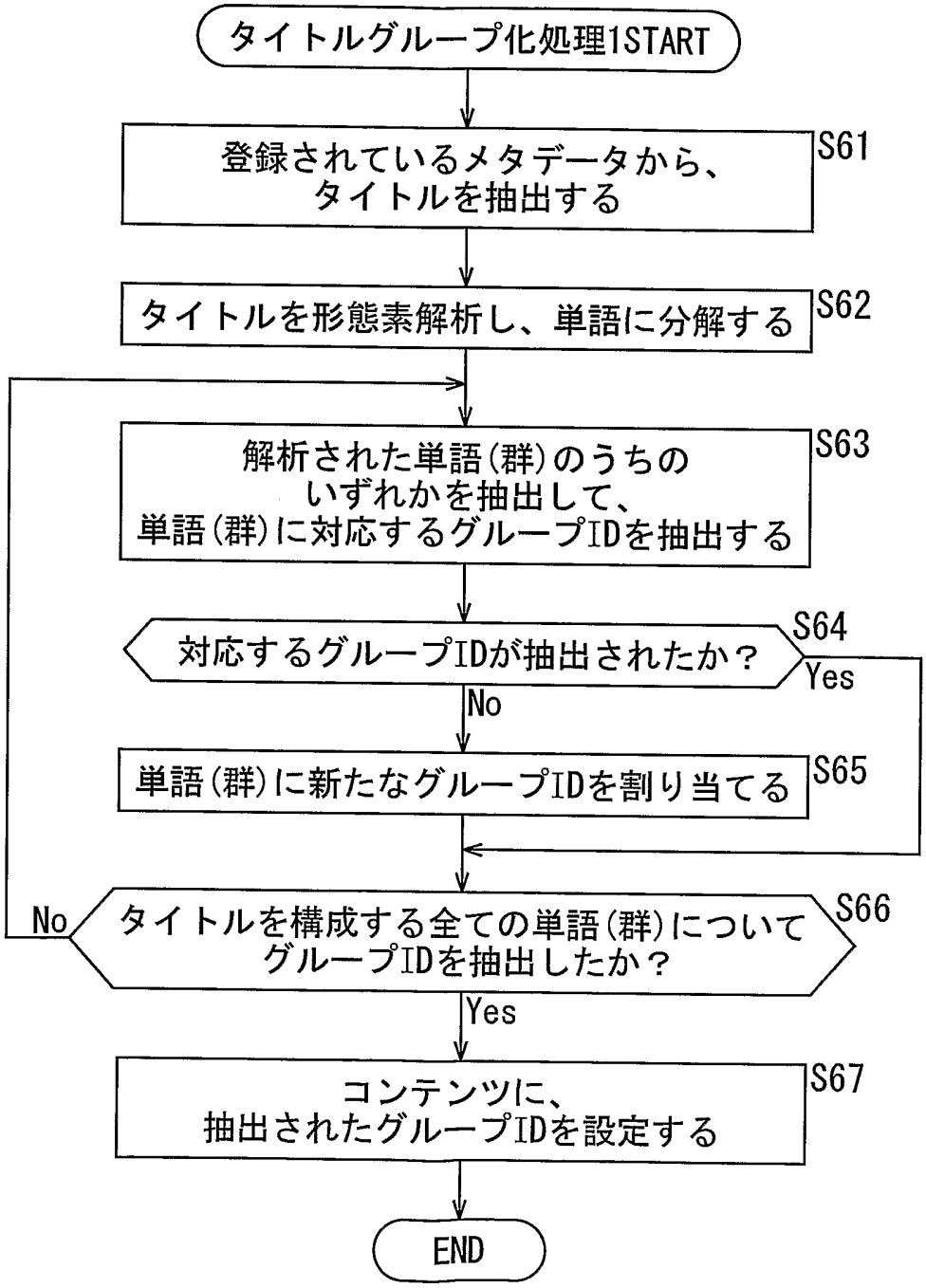

タイトルグループ化処理1START
登録されているメタデータから、
タイトルを抽出する
S61
タイトルを形態素解析し、単語に分解する
S62
解析された単語(群)のうちの
いずれかを抽出して、
単語(群)に対応するグループIDを抽出する
S63
対応するグループIDが抽出されたか？
S64
Yes
No
単語(群)に新たなグループIDを割り当てる
S65
No
タイトルを構成する全ての単語(群)について
グループIDを抽出したか？
S66
Yes
コンテンツに、
抽出されたグループIDを設定する
S67
END

図19

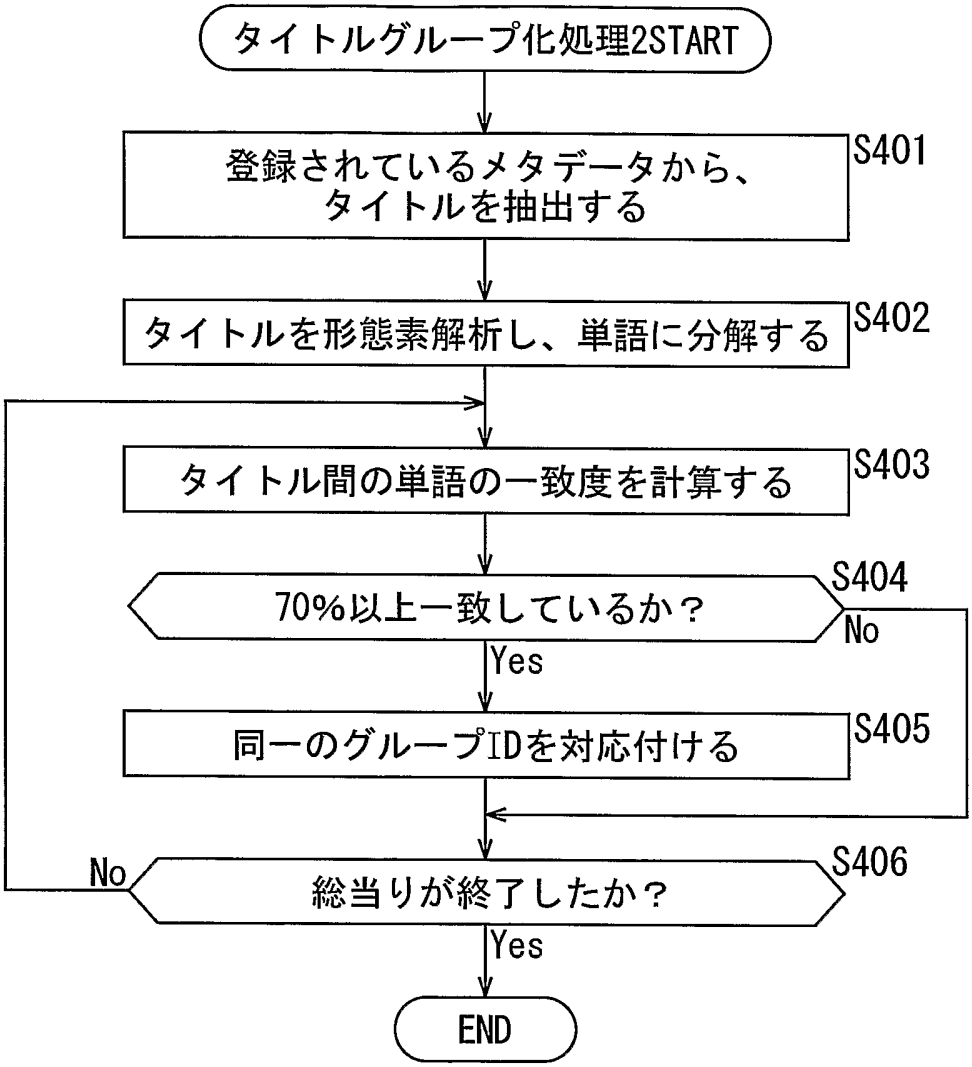

タイトルグループ化処理2START
登録されているメタデータから、
タイトルを抽出する
S401
タイトルを形態素解析し、単語に分解する
S402
タイトル間の単語の一致度を計算する
S403
70%以上一致しているか？
S404
No
Yes
同一のグループIDを対応付ける
S405
No
総当りが終了したか？
S406
Yes
END

図20

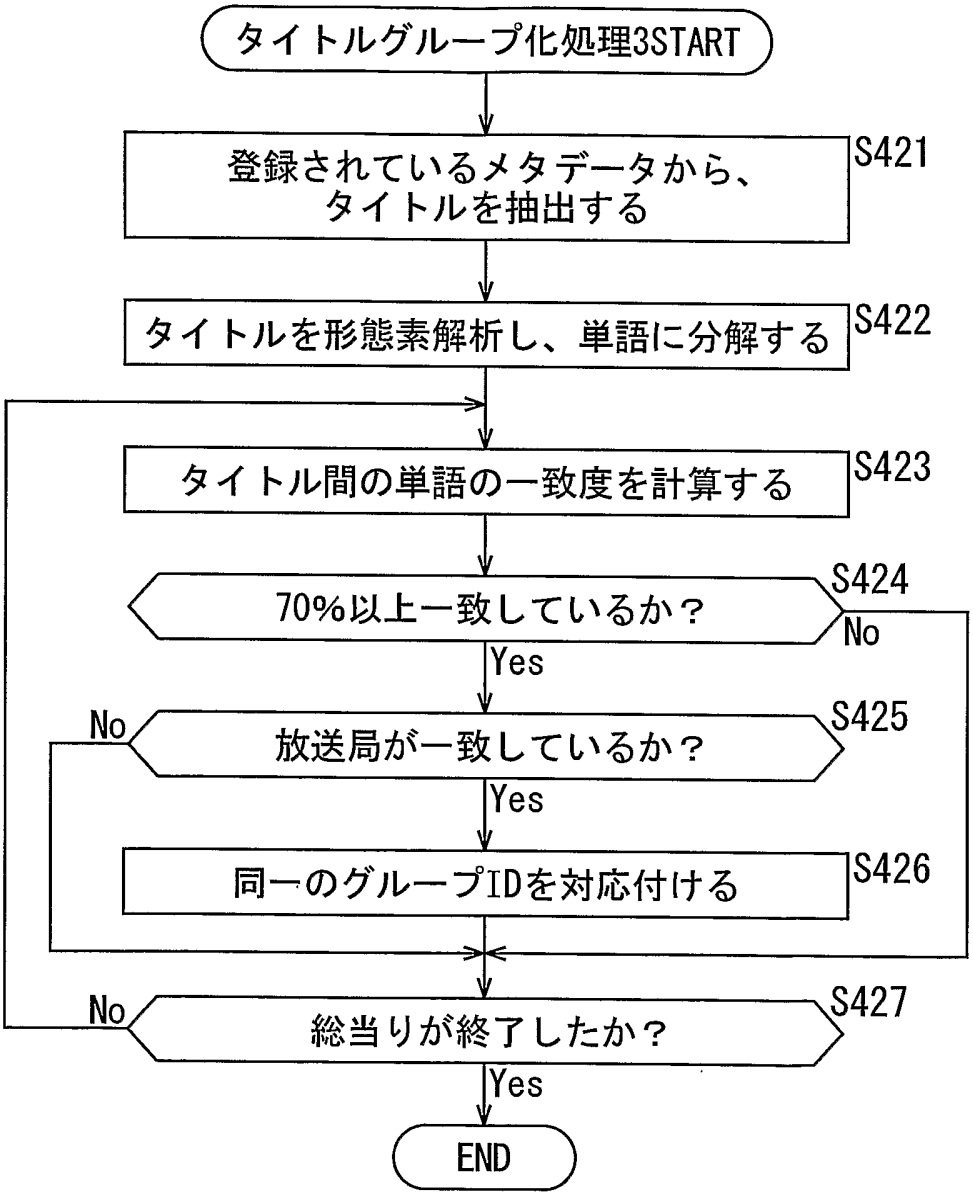

タイトルグループ化処理3START
登録されているメタデータから、
タイトルを抽出する
S421
タイトルを形態素解析し、単語に分解する
S422
タイトル間の単語の一致度を計算する
S423
70%以上一致しているか？
S424
No
Yes
No
放送局が一致しているか？
S425
Yes
同一のグループIDを対応付ける
S426
No
総当りが終了したか？
S427
Yes
END

図21

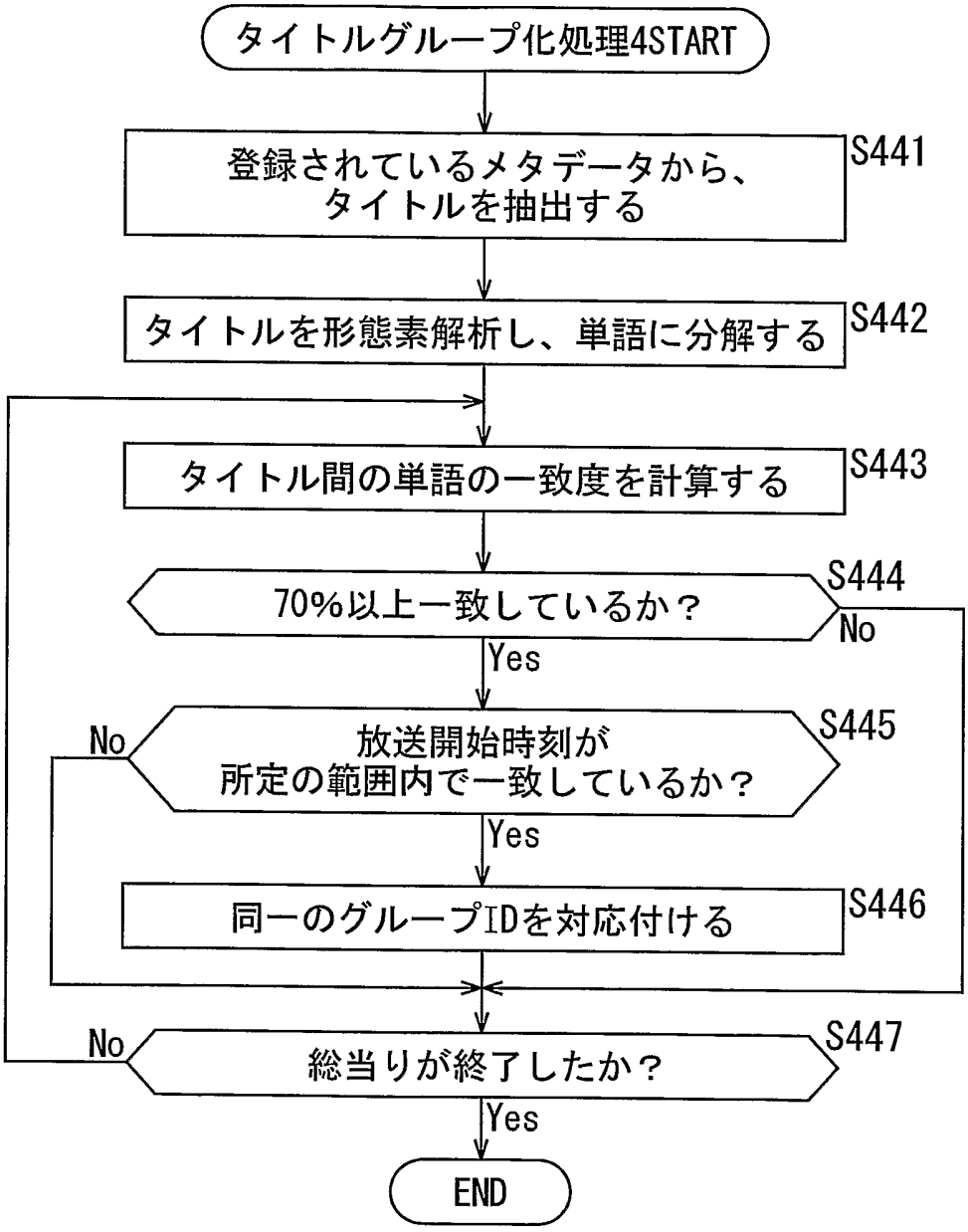

タイトルグループ化処理4START
登録されているメタデータから、
タイトルを抽出する
S441
タイトルを形態素解析し、単語に分解する
S442
タイトル間の単語の一致度を計算する
S443
70%以上一致しているか？
S444
No
Yes
No
放送開始時刻が
所定の範囲内で一致しているか？
S445
Yes
同一のグループIDを対応付ける
S446
No
総当りが終了したか？
S447
Yes
END

図22

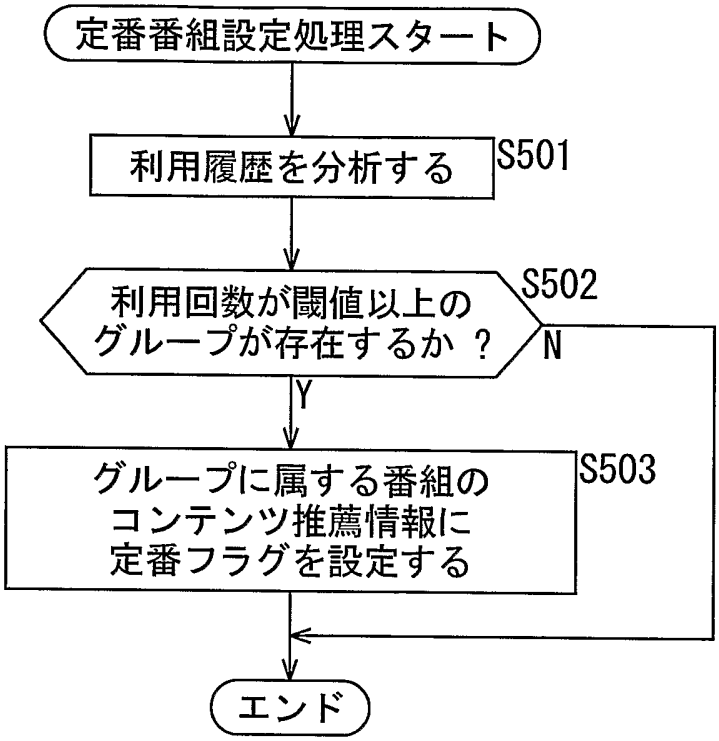

定番番組設定処理スタート
利用履歴を分析する
S501
利用回数が閾値以上の
グループが存在するか ？
S502
N
Y
グループに属する番組の
コンテンツ推薦情報に
定番フラグを設定する
S503
エンド

図23

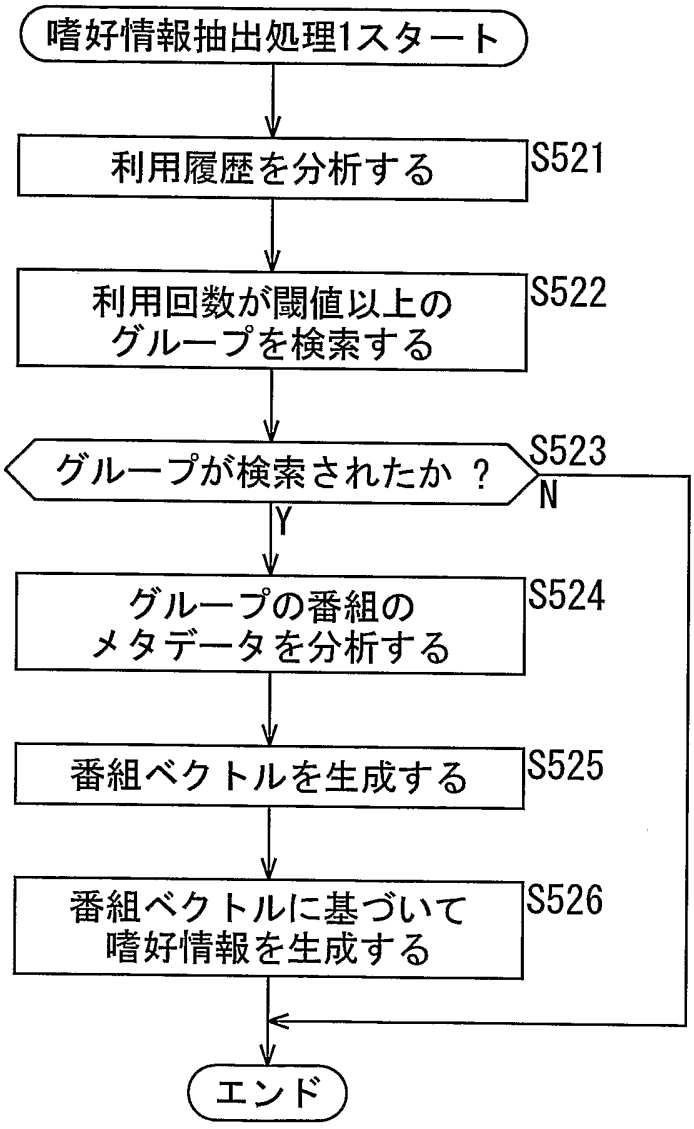

嗜好情報抽出処理1スタート
利用履歴を分析する
S521
利用回数が閾値以上の
グループを検索する
S522
グループが検索されたか ？
S523
N
Y
グループの番組の
メタデータを分析する
S524
番組ベクトルを生成する
S525
番組ベクトルに基づいて
嗜好情報を生成する
S526
エンド

図24

番組ベクトル	PP＝(Tm, Gm, Pm, Sm, Hm, …)
T:タイトル(Title)	Tm＝{Title}
G:ジャンル(Genre)	Gm＝{ドラマ、バラエティ、スポーツ、映画、音楽、子供向け/教育、教養/ドキュメント、ニュース/報道、その他}
P:出演者(Person)	Pm＝{person A-1, personB-3, …, }
S:放送局(TV Station)	Sm＝{MHK総合、MHK教育、亜細亜テレビ、TAS、フシ、テレ日、東都、MHK衛星第1、MHK衛星第2, WOWO}
H:時間帯(Hour)	Hm＝{朝、昼、夕方、ゴールデン、深夜}
⋮	⋮

図25

嗜好情報	UP＝(Tup, Gup, Pup, Sup, …)
タイトル	Tup＝{(Title1-12), (Title2-3), , }
ジャンル	Gup＝{(ドラマ-25), (バラエティ-34), (スポーツ-42), (映画-37), (音楽-73), (子供向け/教育)-120, (教養/ドキュメント-3), (ニュース/報道-5), (その他-23)}
出演者	Pup＝{(person A-10), (personB-43), …, }
放送局	Sup＝{MHK総合-7、MHK教育-4、亜細亜テレビ-3、TAS-2、フシ-5、テレ日-6、東都-0、MHK衛星第1-9、MHK衛星第2-1, WOWO-1}
⋮	⋮

図26

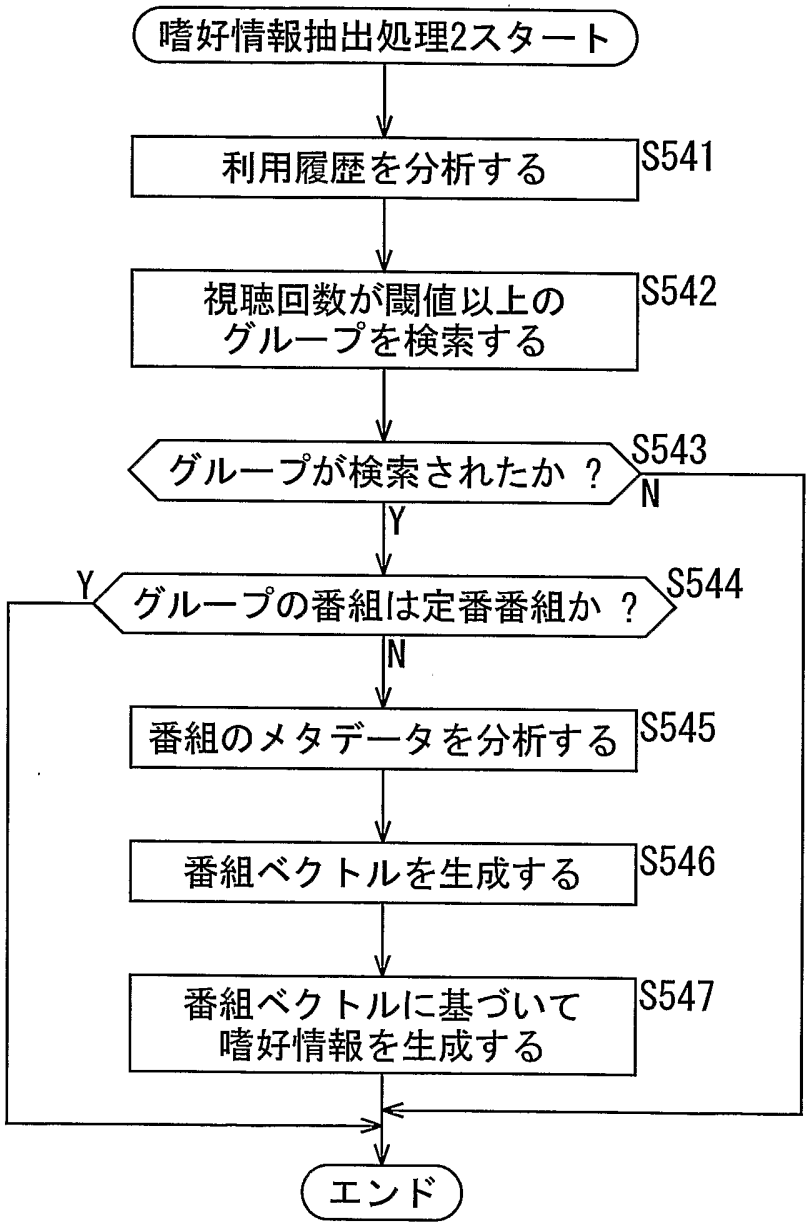

嗜好情報抽出処理2スタート
利用履歴を分析する
S541
視聴回数が閾値以上の
グループを検索する
S542
グループが検索されたか ？
S543
N
Y
Y
グループの番組は定番番組か ？
S544
N
番組のメタデータを分析する
S545
番組ベクトルを生成する
S546
番組ベクトルに基づいて
嗜好情報を生成する
S547
エンド

図27

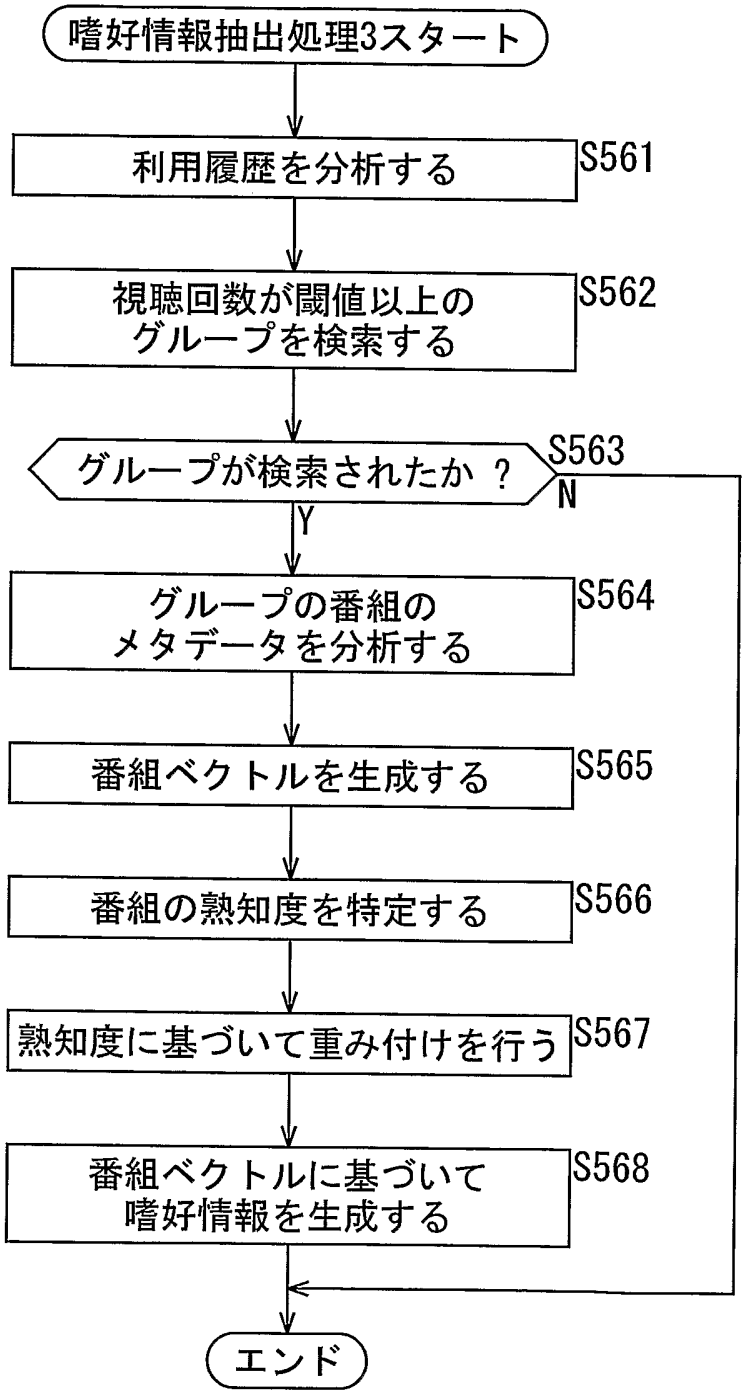

嗜好情報抽出処理3スタート
利用履歴を分析する
S561
視聴回数が閾値以上の
グループを検索する
S562
グループが検索されたか ？
S563
N
Y
グループの番組の
メタデータを分析する
S564
番組ベクトルを生成する
S565
番組の熟知度を特定する
S566
熟知度に基づいて重み付けを行う
S567
番組ベクトルに基づいて
嗜好情報を生成する
S568
エンド

図28

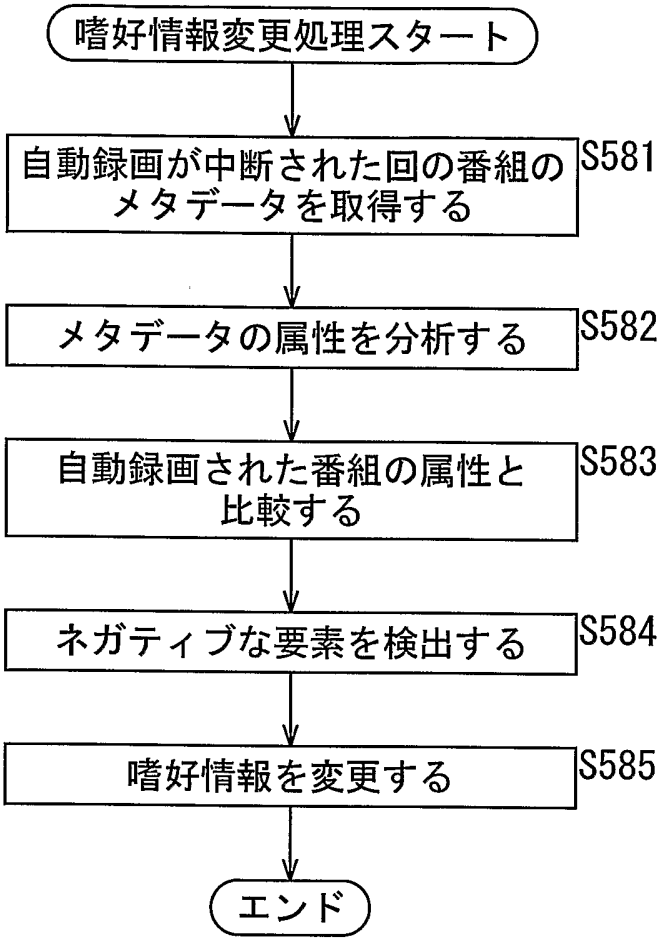

嗜好情報変更処理スタート
自動録画が中断された回の番組のメタデータを取得する
S581
メタデータの属性を分析する
S582
自動録画された番組の属性と比較する
S583
ネガティブな要素を検出する
S584
嗜好情報を変更する
S585
エンド

図29

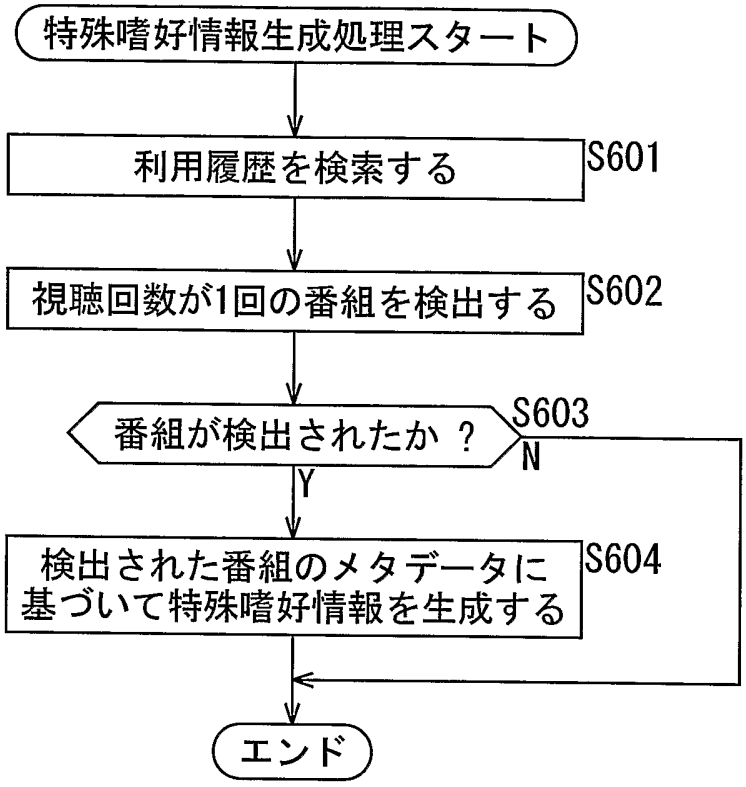

特殊嗜好情報生成処理スタート
利用履歴を検索する
S601
視聴回数が1回の番組を検出する
S602
番組が検出されたか ？
S603
N
Y
検出された番組のメタデータに
基づいて特殊嗜好情報を生成する
S604
エンド

図30

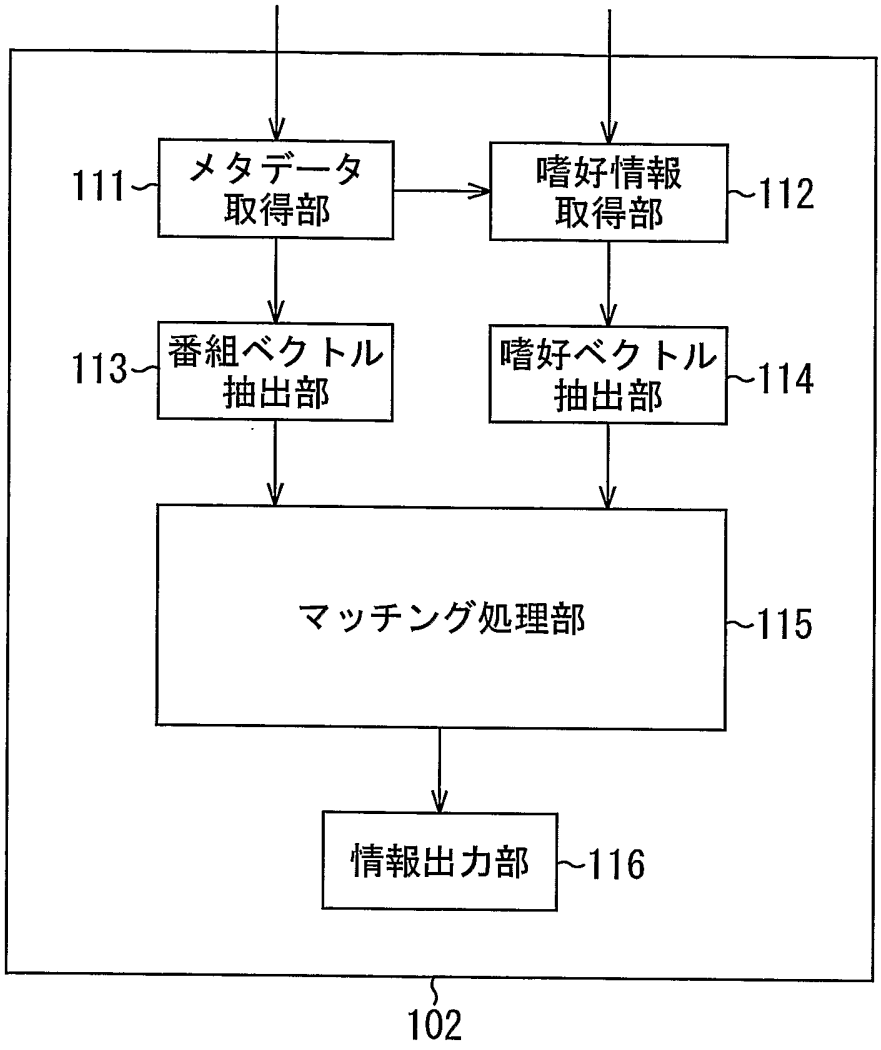

111
メタデータ
取得部
嗜好情報
取得部
112
113
番組ベクトル
抽出部
嗜好ベクトル
抽出部
114
マッチング処理部
115
情報出力部
116
102

図31

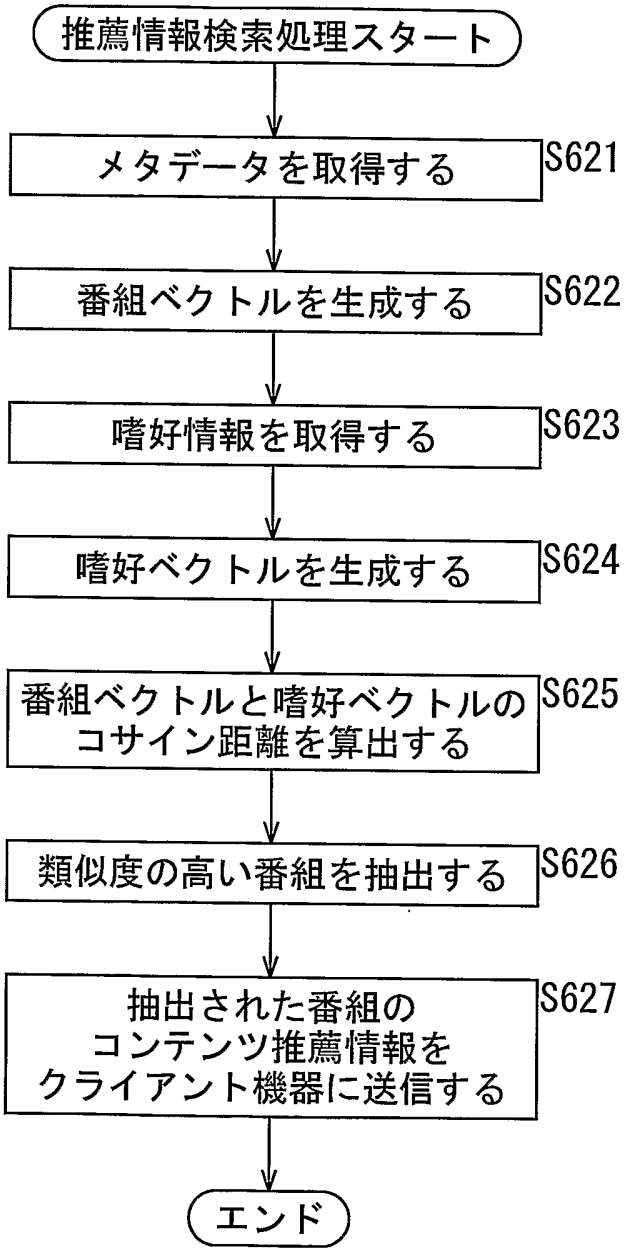
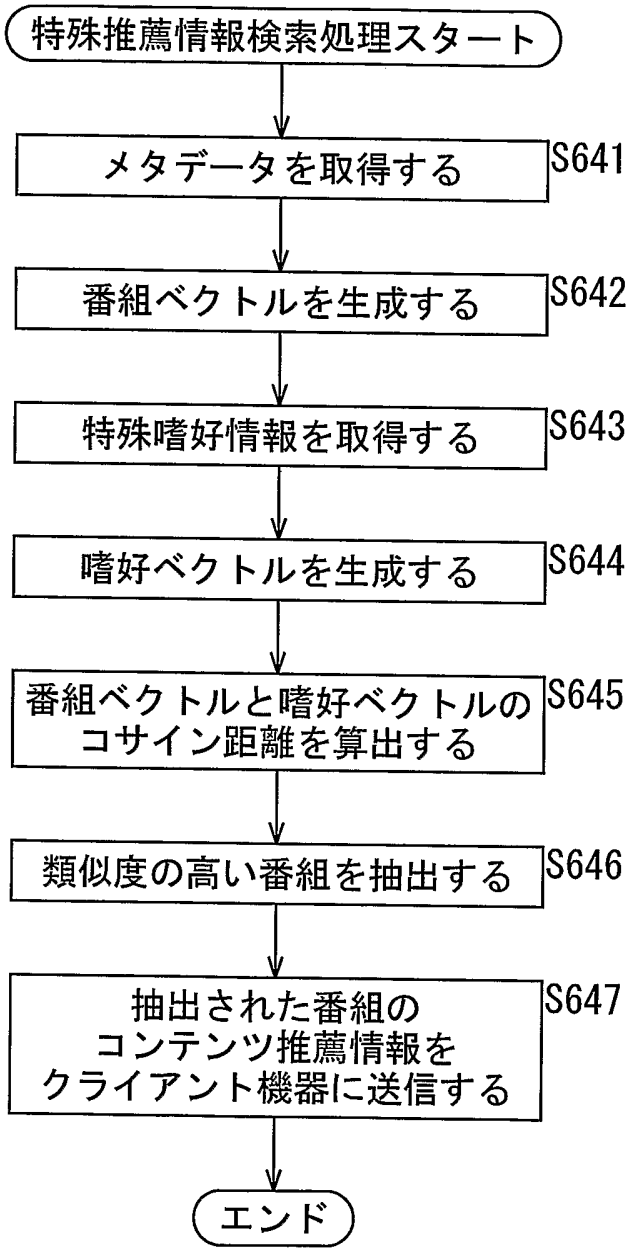

推薦情報検索処理スタート
メタデータを取得する
S621
番組ベクトルを生成する
S622
嗜好情報を取得する
S623
嗜好ベクトルを生成する
S624
番組ベクトルと嗜好ベクトルの
コサイン距離を算出する
S625
類似度の高い番組を抽出する
S626
抽出された番組の
コンテンツ推薦情報を
クライアント機器に送信する
S627
エンド

図32


特殊推薦情報検索処理スタート
メタデータを取得する
S641
番組ベクトルを生成する
S642
特殊嗜好情報を取得する
S643
嗜好ベクトルを生成する
S644
番組ベクトルと嗜好ベクトルの
コサイン距離を算出する
S645
類似度の高い番組を抽出する
S646
抽出された番組の
コンテンツ推薦情報を
クライアント機器に送信する
S647
エンド

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/15927

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[7] G06F17/30

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[7] G06F17/30, H04N17/00

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

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST FILE(JOIS), WPI, INSPEC(DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 2002-320159 A (NEC Corp.), 31 October, 2002 (31.10.02), Full text & CN 1383328 A	1,4,7,9,10, 13-15 2,3,5,6,8, 11,12
Y	JP 06-124309 A (Hitachi, Ltd.), 06 May, 1994 (06.05.94), Full text (Family: none)	2,3
Y	JP 11-196389 A (Kabushiki Kaisha Jisedai Joho Hoso System Kenkyusho), 21 July, 1999 (21.07.99), Full text (Family: none)	5

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 January, 2004 (20.01.04)	03 February, 2004 (03.02.04)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/15927

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2001-92832 A (Matsushita Electric Industrial Co., Ltd.), 06 April, 2001 (06.04.01), Full text (Family: none)	6,8,11,12

国際調査報告

国際出願番号　PCT／JP03／15927

A.　発明の属する分野の分類（国際特許分類（IPC））

Int.Cl[7]　G06F17／30

B.　調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl[7]　G06F17／30　H04N17／00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922－1996年
日本国公開実用新案公報	1971－2004年
日本国登録実用新案公報	1994－2004年
日本国実用新案登録公報	1996－2004年

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

JICSTファイル（JOIS），WPI，INSPEC（DIALOG）

C.　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
X Y	JP　2002－320159　A（日本電気株式会社），2002．10．31，　全文&CN　1383328　A	1, 4, 7, 9, 10 13-15 2, 3, 5, 6, 8 11, 12
Y	JP　06－124309　A（株式会社日立製作所），1994．05．06，　全文（ファミリーなし）	2, 3
Y	JP　11－196389　A（株式会社次世代情報放送システム研究所），　1999．07．21，　全文（ファミリーなし）	5

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

＊　引用文献のカテゴリー
「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日 20．01．2004	国際調査報告の発送日 03. 2. 2004
国際調査機関の名称及びあて先 日本国特許庁（ISA／JP） 郵便番号100－8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 野崎　大進　　5M　3042 電話番号　03－3581－1101　内線　3597

国際調査報告	国際出願番号　ＰＣＴ／ＪＰ０３／１５９２７

C（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Ｙ	ＪＰ　２００１－９２８３２　Ａ（松下電器産業株式会社），２００１．０４．０６　全文（ファミリーなし）	6, 8, 11, 12