

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

W02010/053160

発行日 平成24年4月5日 (2012.4.5)

(43) 国際公開日 平成22年5月14日 (2010.5.14)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30 380F	5E501
G06F 3/048 (2006.01)	G06F 17/30 350C	
G10L 11/00 (2006.01)	G06F 17/30 170G	
	G06F 3/048 656A	
	G10L 11/00 402B	

審査請求 未請求 予備審査請求 未請求 (全 57 頁) 最終頁に続く

出願番号 特願2010-536802 (P2010-536802)
 (21) 国際出願番号 PCT/JP2009/069005
 (22) 国際出願日 平成21年11月6日 (2009.11.6)
 (31) 優先権主張番号 61/193, 232
 (32) 優先日 平成20年11月7日 (2008.11.7)
 (33) 優先権主張国 米国 (US)

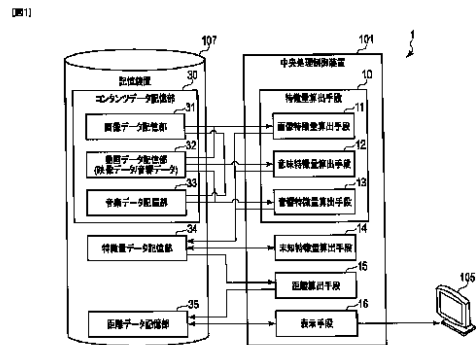
(71) 出願人 504173471
 国立大学法人北海道大学
 北海道札幌市北区北8条西5丁目
 (74) 代理人 100083806
 弁理士 三好 秀和
 (74) 代理人 100101247
 弁理士 高橋 俊一
 (72) 発明者 長谷山 美紀
 北海道札幌市北区北14条西9丁目 国立
 大学法人北海道大学 大学院情報科学研究
 科内
 Fターム(参考) 5E501 AC33 BA05 CA01 EA09 FA47

最終頁に続く

(54) 【発明の名称】 コンテンツ検索装置およびコンテンツ検索プログラム

(57) 【要約】

コンテンツ検索装置 1 は、複数の種別の各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、特徴量データ 3 4 a を記憶する特徴量算出手段 1 0 と、特徴量データにおいて 3 4 a、コンテンツ識別子に対応づけられた特徴種別の特徴量から、コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として特徴量データ 3 4 a に記憶する未知特徴量算出手段 1 4 と、特徴量データ 3 4 a に記憶されている特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する距離算出手段 1 5 と、距離算出手段 1 5 によって算出された距離に基づいて、コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置に表示する表示手段 1 6 とを備える。



- 107 STORAGE DEVICE
- 30 CONTENT DATA STORAGE UNIT
- 31 IMAGE DATA STORAGE SECTION
- 32 MOVING IMAGE DATA STORAGE SECTION (VIDEO DATA/ACOUSTIC DATA)
- 33 MUSIC DATA STORAGE SECTION
- 34 FEATURE QUANTITY DATA STORAGE UNIT
- 35 FEATURE QUANTITY DATA STORAGE UNIT
- 36 DISTANCE DATA STORAGE UNIT
- 101 CENTRAL PROCESSING CONTROLLER
- 10 FEATURE QUANTITY COMPUTING MEANS
- 11 IMAGE FEATURE QUANTITY COMPUTING MEANS
- 12 SEMANTIC FEATURE QUANTITY COMPUTING MEANS
- 13 ACOUSTIC FEATURE QUANTITY COMPUTING MEANS
- 14 UNKNOWN FEATURE QUANTITY COMPUTING MEANS
- 15 DISTANCE COMPUTING MEANS
- 16 DISPLAY MEANS

【特許請求の範囲】**【請求項 1】**

異なる種別のコンテンツデータについて、類似するコンテンツデータを検索するコンテンツ検索装置であって、

画像特徴、音響特徴および意味特徴のうちいずれか一つ以上を有する複数のコンテンツデータが、コンテンツ識別子に対応づけて記憶装置に記憶されたコンテンツデータ記憶部と、

前記コンテンツデータ記憶部に記憶された各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、前記コンテンツ識別子と、画像特徴、音響特徴および意味特徴の特徴種別と、その特徴量とを関連づけた特徴量データを、前記記憶装置に記憶する特徴量算出手段と、

前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、前記コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として前記特徴量データに記憶する未知特徴量算出手段と、

前記特徴量データに記憶されている前記特徴量および前記特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する距離算出手段と、

前記距離算出手段によって算出された距離に基づいて、前記コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置に表示する表示手段

とを備えることを特徴とするコンテンツ検索装置。

【請求項 2】

前記未知特徴量算出手段は、前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、主成分分析を用いて前記特徴推定値を算出することを特徴とする請求項 1 に記載のコンテンツ検索装置。

【請求項 3】

前記未知特徴量算出手段は、前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、正準相関分析を用いて前記特徴推定値を算出することを特徴とする請求項 1 に記載のコンテンツ検索装置。

【請求項 4】

前記特徴量算出手段は、音響特徴の特徴量を算出する際、音響信号のうち繰り返し現れるメロディーを特定し、その繰り返しメロディーの特徴量を、前記音響特徴の特徴量とする

ことを特徴とする請求項 1 に記載のコンテンツ検索装置。

【請求項 5】

ユーザ識別子と、当該ユーザの嗜好に合うコンテンツ識別子とが関連づけられたユーザ嗜好データが、前記記憶装置に記憶されたユーザ嗜好データ記憶部と、

各前記コンテンツ識別子について、前記ユーザ嗜好データにおいて当該コンテンツに関連づけられたユーザ識別子の集合を作成するとともに、各集合において、ユーザ間に有向辺を作成し、各集合を代表する重要ユーザを決定する重要ユーザ決定手段と、をさらに備え、

前記距離算出手段は、前記ユーザ嗜好データにおいて前記重要ユーザのユーザ識別子に関連づけられた前記コンテンツ識別子を抽出し、抽出した前記コンテンツ識別子に対応する前記特徴量および前記特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する

ことを特徴とする請求項 1 に記載のコンテンツ検索装置。

【請求項 6】

前記表示手段はさらに、前記ユーザ嗜好データにおいて、所定のコンテンツに関連づけられた複数のユーザの識別子を抽出し、前記複数のユーザの識別子のそれぞれをノードと、前記ユーザ間の有向辺を考慮した前記ノード間のリンクを含むユーザネットワークを表示して、表示装置に表示する

ことを特徴とする請求項 5 に記載のコンテンツ検索装置。

10

20

30

40

50

【請求項 7】

異なる種別のコンテンツデータについて、類似するコンテンツデータを検索するコンテンツ検索プログラムであって、

コンピュータに、

画像特徴、音響特徴および意味特徴のうちいずれか一つ以上を有する複数のコンテンツデータが、コンテンツ識別子に対応づけて記憶装置に記憶された各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、前記コンテンツ識別子と、画像特徴、音響特徴および意味特徴の特徴種別と、その特徴量とを関連づけた特徴量データを、前記記憶装置に記憶する特徴量算出手段と、

10

前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、前記コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として前記特徴量データに記憶する未知特徴量算出手段と、

前記特徴量データに記憶されている前記特徴量および前記特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する距離算出手段と、

前記距離算出手段によって算出された距離に基づいて、前記コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置に表示する表示手段

とを備えることを特徴とするコンテンツ検索プログラム。

【請求項 8】

前記未知特徴量算出手段は、前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、主成分分析を用いて前記特徴推定値を算出する

20

ことを特徴とする請求項 7 に記載のコンテンツ検索プログラム。

【請求項 9】

前記未知特徴量算出手段は、前記特徴量データにおいて、前記コンテンツ識別子に対応づけられた特徴種別の特徴量から、正準相関分析を用いて前記特徴推定値を算出する

ことを特徴とする請求項 7 に記載のコンテンツ検索プログラム。

【請求項 10】

前記特徴量算出手段は、音響特徴の特徴量を算出する際、音響信号のうち繰り返し現れるメロディーを特定し、その繰り返しメロディーの特徴量を、前記音響特徴の特徴量とする

30

ことを特徴とする請求項 7 に記載のコンテンツ検索プログラム。

【請求項 11】

ユーザ識別子と、当該ユーザの嗜好に合うコンテンツ識別子とが関連づけられたユーザ嗜好データが、前記記憶装置に記憶されたユーザ嗜好データにおいて当該コンテンツに関連づけられたユーザ識別子の集合を、各前記コンテンツ識別子について作成するとともに、各集合において、ユーザ間に有向辺を作成し、各集合を代表する重要ユーザを決定する重要ユーザ決定手段と、をさらに前記コンピュータに実行させ、

前記距離算出手段は、前記ユーザ嗜好データにおいて前記重要ユーザのユーザ識別子に関連づけられた前記コンテンツ識別子を抽出し、抽出した前記コンテンツ識別子に対応する前記特徴量および前記特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する

40

ことを特徴とする請求項 7 に記載のコンテンツ検索プログラム。

【請求項 12】

前記表示手段はさらに、前記ユーザ嗜好データにおいて、所定のコンテンツに関連づけられた複数のユーザの識別子を抽出し、前記複数のユーザの識別子のそれぞれをノードと、前記ユーザ間の有向辺を考慮した前記ノード間のリンクを含むユーザネットワークを表示して、表示装置に表示する

ことを特徴とする請求項 11 に記載のコンテンツ検索プログラム。

【発明の詳細な説明】**【技術分野】**

50

【 0 0 0 1 】

本発明は、類似するコンテンツデータを検索するコンテンツ検索装置およびコンテンツ検索プログラムに関する。

【 背景技術 】

【 0 0 0 2 】

近年の情報通信ネットワークや記憶媒体の発達に伴い、ユーザは大量のコンテンツデータを取得できるようになった。ユーザは、自身が作成したコンテンツデータを取得できるのみならず、配信サイトなどに接続してコンテンツデータを取得することができる。このコンテンツデータは、画像データ、動画データ、音楽データ等の各種のコンテンツデータが含まれる。画像データは、静止画等のデータである。動画データは、テレビ録画物、ビデオ録画物、映画やアニメーションなどのデータである。音楽データは、クラシック音楽、歌謡曲、BGMなどのデータである。

10

【 0 0 0 3 】

これらコンテンツデータの数は膨大である。従ってユーザは、膨大な数のコンテンツデータから所望のコンテンツデータを検索するために、コンピュータにより実現される検索システムを利用することが一般的である。

【 0 0 0 4 】

一般的には、これらの各コンテンツ種別について、所望のコンテンツを検索する装置が開示されている。例えば動画データを検索する検索装置（例えば、特許文献1参照。）、音楽データを検索する検索装置（例えば、特許文献2参照。）などがある。また、楽曲の繰り返し区間を検出する方法もある（例えば、非特許文献1参照。）。

20

【 先行技術文献 】

【 特許文献 】

【 0 0 0 5 】

【 特許文献 1 】 国際公開第 2 0 0 2 / 0 3 3 5 8 9 号

【 特許文献 2 】 特開 2 0 0 5 - 1 0 7 7 1 号 公 報

【 非特許文献 】

【 0 0 0 6 】

【 非特許文献 1 】 M. A. Bartsch and G. H. Wake_eld, "To chatch a chorus: using chroma-based representations for audio thumbnailing," Proc. WAS-PAA'01, pp. 15-18, 2001.

30

【 発明の概要 】

【 0 0 0 7 】

しかしながら、上記特許文献1または特許文献2に記載の検索システムにおいては、各種コンテンツデータのうち、特定の1種類のコンテンツデータのみを検索対象としている。従って、従来の検索システムにおいては、動画像、画像、音楽のそれぞれのコンテンツの検索において個別の検索システムを用いなければならなかった。また、動画像と画像、画像と音楽、および音楽と動画像など、異なるコンテンツ種別のコンテンツデータから、類似するコンテンツを検索することはできなかった。

【 0 0 0 8 】

また、従来、各コンテンツデータについて予め、作成者、コンテンツ内容等のメタデータが付与されており、これらのメタデータを用いて類似するコンテンツデータを検索する検索システムが多い。このような検索システムを利用する場合、予め各コンテンツデータにメタデータを付与する必要があり、膨大な数のコンテンツデータの処理をするに際し、弊害となるおそれがある。

40

【 0 0 0 9 】

そこで、ユーザが、コンテンツ種別を意識することなく、またメタデータがなくとも所望のコンテンツデータを容易に検索することができる技術の開発が期待されている。

【 0 0 1 0 】

従って本発明の目的は、異なる種別のコンテンツデータについて、類似するコンテンツ

50

データを検索するコンテンツ検索装置およびコンテンツ検索プログラムを提供することである。

【0011】

上記課題を解決するために、本発明の第1の特徴は、異なる種別のコンテンツデータについて、類似するコンテンツデータを検索するコンテンツ検索装置に関する。即ち本発明の第1の特徴に係るコンテンツ検索装置は、画像特徴、音響特徴および意味特徴のうちいずれか一つ以上を有する複数のコンテンツデータが、コンテンツ識別子に対応づけて記憶装置に記憶されたコンテンツデータ記憶部と、コンテンツデータ記憶部に記憶された各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、コンテンツ識別子と、画像特徴、音響特徴および意味特徴の特徴種別と、その特徴量とを関連づけた特徴量データを、記憶装置に記憶する特徴量算出手段と、特徴量データにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として特徴量データに記憶する未知特徴量算出手段と、特徴量データに記憶されている特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する距離算出手段と、距離算出手段によって算出された距離に基づいて、コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置に表示する表示手段とを備える。

10

ここで、未知特徴量算出手段は、例えば、特徴量データにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、主成分分析を用いて特徴推定値を算出する。また他の例としては、未知特徴量算出手段は、特徴量データにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、正準相関分析を用いて特徴推定値を算出する。

20

【0012】

特徴量算出手段は、音響特徴の特徴量を算出する際、音響信号のうち繰り返し現れるメロディーを特定し、その繰り返しメロディーの特徴量を、音響特徴の特徴量としても良い。

【0013】

第1の特徴に係るコンテンツ検索装置は、さらに、ユーザ識別子と、当該ユーザの嗜好に合うコンテンツ識別子とが関連づけられたユーザ嗜好データが、記憶装置に記憶されたユーザ嗜好データ記憶部と、各コンテンツ識別子について、ユーザ嗜好データにおいて当該コンテンツに関連づけられたユーザ識別子の集合を作成するとともに、各集合において、ユーザ間に有向辺を作成し、各集合を代表する重要ユーザを決定する重要ユーザ決定手段と、をさらに備えても良い。この場合、距離算出手段は、ユーザ嗜好データにおいて重要ユーザのユーザ識別子に関連づけられたコンテンツ識別子を抽出し、抽出したコンテンツ識別子に対応する特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出することが好ましい。

30

【0014】

表示手段はさらに、ユーザ嗜好データにおいて、所定のコンテンツに関連づけられた複数のユーザの識別子を抽出し、複数のユーザの識別子のそれぞれをノードと、ユーザ間の有向辺を考慮したノード間のリンクを含むユーザネットワークを表示して、表示装置に表示しても良い。

40

【0015】

本発明の第2の特徴は、異なる種別のコンテンツデータについて、類似するコンテンツデータを検索するコンテンツ検索プログラムに関する。即ち本発明の第2の特徴に係るコンテンツ検索プログラムは、コンピュータに、画像特徴、音響特徴および意味特徴のうちいずれか一つ以上を有する複数のコンテンツデータが、コンテンツ識別子に対応づけて記憶装置に記憶された各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、コンテンツ識別子と、画像特徴、音響特徴および意味特徴の特徴種別と、その特徴量とを関連づけた特徴量データを、記憶装置に記憶する特徴量算出手段と、特徴量データにおいて、コンテンツ識

50

別子に対応づけられた特徴種別の特徴量から、コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として特徴量データに記憶する未知特徴量算出手段と、特徴量データに記憶されている特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する距離算出手段と、距離算出手段によって算出された距離に基づいて、コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置に表示する表示手段として実現させる。

【0016】

ここで、未知特徴量算出手段は、例えば、特徴量データにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、主成分分析を用いて特徴推定値を算出する。また他の例としては、未知特徴量算出手段は、特徴量データにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、正準相関分析を用いて特徴推定値を算出する。

10

【0017】

特徴量算出手段は、音響特徴の特徴量を算出する際、音響信号のうち繰り返し現れるメロディーを特定し、その繰り返しメロディーの特徴量を、音響特徴の特徴量としても良い。

【0018】

第2の特徴に係るコンテンツ検索プログラムは、ユーザ識別子と、当該ユーザの嗜好に合うコンテンツ識別子とが関連づけられたユーザ嗜好データが、記憶装置に記憶されたユーザ嗜好データにおいて当該コンテンツに関連づけられたユーザ識別子の集合を、各コンテンツ識別子について作成するとともに、各集合において、ユーザ間に有向辺を作成し、各集合を代表する重要ユーザを決定する重要ユーザ決定手段と、をさらにコンピュータに実行させても良い。この場合、距離算出手段は、ユーザ嗜好データにおいて重要ユーザのユーザ識別子に関連づけられたコンテンツ識別子を抽出し、抽出したコンテンツ識別子に対応する特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出することが好ましい。

20

【0019】

表示手段はさらに、ユーザ嗜好データにおいて、所定のコンテンツに関連づけられた複数のユーザの識別子を抽出し、複数のユーザの識別子のそれぞれをノードと、ユーザ間の有向辺を考慮したノード間のリンクを含むユーザネットワークを表示して、表示装置に表示しても良い。

30

【0020】

本発明によれば、異なる種別のコンテンツデータについて、類似するコンテンツデータを検索するコンテンツ検索装置およびコンテンツ検索プログラムを提供することができる。

【図面の簡単な説明】

【0021】

【図1】図1は、本発明の実施の形態に係るコンテンツ検索装置の機能を説明する図である。

【図2】図2は、本発明の実施の形態に係るコンテンツ検索装置で算出する特徴量を説明する図である。

40

【図3】図3は、本発明の実施の形態に係るコンテンツ検索装置で算出する未知の特徴量を説明する図である。

【図4】図4は、本発明の実施の形態に係るコンテンツ検索装置のハードウェア構成を説明する図である。

【図5】図5は、本発明の実施の形態に係るコンテンツ検索装置の特徴量データのデータ構造とデータの一例を説明する図である。

【図6】図6は、本発明の実施の形態に係るコンテンツ検索装置の画像特徴量算出処理を説明するフローチャートである。

【図7】図7は、本発明の実施の形態に係るコンテンツ検索装置の意味特徴量算出処理を説明するフローチャートである。

50

【図 8】図 8 は、本発明の実施の形態に係るコンテンツ検索装置の音響特徴量算出処理を説明するフローチャートである。

【図 9】図 9 は、本発明の実施の形態に係るコンテンツ検索装置において、主成分分析による未知特徴量算出処理を説明するフローチャートである。

【図 10】図 10 は、本発明の実施の形態に係るコンテンツ検索装置において、正準相関分析による未知特徴量算出処理を説明するフローチャートである。

【図 11】図 11 は、本発明の実施の形態に係るコンテンツ検索装置において、重みなし距離を算出する距離算出処理を説明するフローチャートである。

【図 12】図 12 は、本発明の実施の形態に係るコンテンツ検索装置において、重み付き距離を算出する距離算出処理を説明するフローチャートである。

【図 13】図 13 は、本発明の実施の形態に係るコンテンツ検索装置において、画面中央にクエリコンテンツデータのサムネイルを表示し、その周囲に検索結果のコンテンツデータのサムネイルを表示する画面構成を説明する図である。

【図 14】図 14 は、本発明の実施の形態に係るコンテンツ検索装置において、クエリコンテンツデータを指定する画面の一例を説明する図である。

【図 15】図 15 は、本発明の実施の形態に係るコンテンツ検索装置において、画面中央にクエリコンテンツデータのサムネイルを表示し、その周囲に検索結果のコンテンツデータのサムネイルを表示する画面の一例を説明する図である。

【図 16】図 16 は、本発明の実施の形態に係るコンテンツ検索装置において、コンテンツデータの種別毎に表示エリアを分けて表示する画面構成を説明する図である。

【図 17】図 17 は、本発明の実施の形態に係るコンテンツ検索装置において、コンテンツデータの種別毎に表示エリアを分けて表示する画面の一例を説明する図である。

【図 18】図 18 は、本発明の実施の形態に係るコンテンツ検索装置において、画面左上にクエリコンテンツデータのサムネイルを表示し、クエリコンテンツデータのサムネイルから放射状に検索結果のコンテンツデータのサムネイルを表示する画面構成を説明する図である。

【図 19】図 19 は、本発明の実施の形態に係るコンテンツ検索装置において、画面左上にクエリコンテンツデータのサムネイルを表示し、クエリコンテンツデータのサムネイルから放射状に、検索結果の一つの動画データにフォーカスをあてて表示する画面の一例を説明する図である。

【図 20】図 20 は、本発明の実施の形態に係るコンテンツ検索装置において、画面左上にクエリコンテンツデータのサムネイルを表示し、クエリコンテンツデータのサムネイルから放射状に、検索結果の一つの音楽データにフォーカスをあてて表示する画面の一例を説明する図である。

【図 21】図 21 は、本発明の実施の形態に係るコンテンツ検索装置において、画面左上にクエリコンテンツデータのサムネイルを表示し、クエリコンテンツデータのサムネイルから放射状に、検索結果の一つの画像データにフォーカスをあてて表示する画面の一例を説明する図である。

【図 22】図 22 は、本発明の実施の形態に係るコンテンツ検索装置において、クロマベクトル算出の概要を説明する図である。

【図 23】図 23 は、本発明の実施の形態に係るコンテンツ検索装置において、遷移行列を説明する図である。

【図 24】図 24 は、本発明の実施の形態に係るコンテンツ検索装置において、正準相関分析で得られた相関行列を説明する図である。

【図 25】図 25 は、本発明の第 2 の変形例に係るコンテンツ検索装置の機能を説明する図である。

【図 26】図 26 は、本発明の第 2 の変形例に係るコンテンツ検索装置において、重要ユーザを決定する重要ユーザ決定処理を説明するフローチャートである。

【図 27】図 27 は、本発明の第 2 の変形例に係るコンテンツ検索装置において、ユーザネットワークを表示する画面の一例を説明する図である。

10

20

30

40

50

【発明を実施するための形態】

【0022】

次に、図面を参照して、本発明の実施の形態を説明する。以下の図面の記載において、同一又は類似の部分には同一又は類似の符号を付している。

【0023】

(実施の形態)

図1に示すように、本発明の実施の形態に係るコンテンツ検索装置1は、異なる種類のコンテンツデータについて、類似するコンテンツデータを検索する。具体的にはコンテンツ検索装置1は、コンテンツデータ記憶部30に記憶された各種コンテンツデータについて、各コンテンツデータから特徴量を算出する。この際、コンテンツ検索装置1は、各種コンテンツデータにメタデータが付与されているか否かにかかわらず、特徴量を算出する。

10

さらにコンテンツ検索装置1は、各コンテンツデータの未知の特徴量を、算出可能な特徴量から推定する。コンテンツ検索装置1は、算出された特徴量および推定された特徴量から、各コンテンツデータの距離を算出し、類似度を決定し、表示装置105に検索結果を3次元の空間として表現して表示する。

【0024】

本発明の実施の形態において「コンテンツデータ」は、画像データ、動画データおよび音楽データを示す。またコンテンツデータとして、テキストデータが含まれていても良い。また本発明の実施の形態において「特徴量」は、画像特徴量、意味特徴量および音響特徴量を示す。画像特徴量は、静止画の画像データや動画の各フレームのデータの色ヒストグラムに基づいて算出される。意味特徴量は、単語が出現する頻度などに基づいて算出される。音響特徴量は、音階の時間変化に基づいて算出される。

20

【0025】

図2および図3を参照して、各コンテンツデータについて、算出される特徴量と、推定される未知の特徴量とを説明する。

【0026】

図2に示すように、画像データは画像特徴を有しているので、コンテンツ検索装置1は、画像データから画像特徴量を算出する。また、画像データにメタデータが付与されている場合、コンテンツ検索装置1は、画像データのメタデータから意味特徴量も算出する。さらにコンテンツ検索装置1は、未知特徴量として音響特徴量を算出するとともに、メタデータがない場合、さらに意味特徴量も算出する。これにより、図3に示すように画像データは、画像特徴量、意味特徴量および音響特徴量の全ての特徴量を有する。

30

【0027】

図2に示すように、音楽データは音響特徴を有しているので、コンテンツ検索装置1は、音楽データから音響特徴量を算出する。また、音楽データにメタデータが付与されている場合、コンテンツ検索装置1は、音楽データのメタデータから意味特徴量も算出する。さらにコンテンツ検索装置1は、未知特徴量として画像特徴量を算出するとともに、メタデータがない場合、さらに意味特徴量も算出する。これにより、図3に示すように音楽データは、画像特徴量、意味特徴量および音響特徴量の全ての特徴量を有する。

40

【0028】

図2に示すように、動画データは画像特徴および音響特徴を有しているので、コンテンツ検索装置1は、動画データから画像特徴量および音響特徴量を算出する。また、動画データにメタデータが付与されている場合、コンテンツ検索装置1は、動画データのメタデータから意味特徴量も算出する。また、動画データに人の声やキャプションなどが含まれている場合、これらのデータをテキストデータとして抽出することにより、コンテンツ検索装置1は、動画データの意味特徴量も算出する。意味特徴量が算出されない場合、さらにコンテンツ検索装置1は、未知特徴量として意味特徴量を算出する。これにより、図3に示すように動画データは、画像特徴量、意味特徴量および音響特徴量の全ての特徴量を有する。

50

【 0 0 2 9 】

このように、本発明の実施の形態に係るコンテンツ検索装置 1 は、コンテンツの種別にかかわらず、各コンテンツデータについて、画像特徴量、意味特徴量および音響特徴量を算出することができる。さらに、コンテンツ検索装置 1 は、各コンテンツデータに対応付けられた画像特徴量、意味特徴量および音響特徴量に基づいてコンテンツ間の距離を算出し、コンテンツの類似度を算出することができる。従って、コンテンツの種別にかかわらず、類似するコンテンツを検索することができる。

【 0 0 3 0 】

このようなコンテンツ検索装置 1 は、例えば、画像データに類似する音楽データを検索することにより、画像に合う音楽を抽出することができる。また、コンテンツ検索装置 1 は、音楽データに類似する動画データを検索することにより、音楽に合う動画を抽出することができる。

10

【 0 0 3 1 】

(コンテンツ検索装置のハードウェア構成)

図 4 に示すように、本発明の実施の形態に係るコンテンツ検索装置 1 は、中央処理制御装置 1 0 1、R O M (Read Only Memory) 1 0 2、R A M (Random Access Memory) 1 0 3 及び入出力インタフェース 1 0 9 が、バス 1 1 0 を介して接続されている。入出力インタフェース 1 0 9 には、入力装置 1 0 4、表示装置 1 0 5、通信制御装置 1 0 6、記憶装置 1 0 7 及びリムーバブルディスク 1 0 8 が接続されている。

20

【 0 0 3 2 】

中央処理制御装置 1 0 1 は、入力装置 1 0 4 からの入力信号に基づいて R O M 1 0 2 からコンテンツ検索装置 1 を起動するためのブートプログラムを読み出して実行し、更に記憶装置 1 0 7 に記憶されたオペレーティングシステムを読み出す。更に中央処理制御装置 1 0 1 は、入力装置 1 0 4 や通信制御装置 1 0 6 などの入力信号に基づいて、各種装置の制御を行ったり、R A M 1 0 3 や記憶装置 1 0 7 などに記憶されたプログラム及びデータを読み出して R A M 1 0 3 にロードするとともに、R A M 1 0 3 から読み出されたプログラムのコマンドに基づいて、データの計算又は加工など、後述する一連の処理を実現する処理装置である。

【 0 0 3 3 】

入力装置 1 0 4 は、操作者が各種の操作を入力するキーボード、マウスなどの入力デバイスにより構成されており、操作者の操作に基づいて入力信号を作成し、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送信される。表示装置 1 0 5 は、C R T (Cathode Ray Tube) ディスプレイや液晶ディスプレイなどであり、中央処理制御装置 1 0 1 からバス 1 1 0 及び入出力インタフェース 1 0 9 を介して表示装置 1 0 5 において表示させる出力信号を受信し、例えば中央処理制御装置 1 0 1 の処理結果などを表示する装置である。通信制御装置 1 0 6 は、L A N カードやモデムなどの装置であり、コンテンツ検索装置 1 をインターネットや L A N などの通信ネットワークに接続する装置である。通信制御装置 1 0 6 を介して通信ネットワークと送受信したデータは入力信号又は出力信号として、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送受信される。

30

40

【 0 0 3 4 】

記憶装置 1 0 7 は半導体記憶装置や磁気ディスク装置であって、中央処理制御装置 1 0 1 で実行されるプログラムやデータが記憶されている。リムーバブルディスク 1 0 8 は、光ディスクやフレキシブルディスクのことであり、ディスクドライブによって読み書きされた信号は、入出力インタフェース 1 0 9 及びバス 1 1 0 を介して中央処理制御装置 1 0 1 に送受信される。

【 0 0 3 5 】

本発明の実施の形態に係るコンテンツ検索装置 1 の記憶装置 1 0 7 には、コンテンツ検索プログラムが記憶されるとともに、図 1 に示すように、記憶装置 1 0 7 は、コンテンツデータ記憶部 3 0、特徴量データ記憶部 3 4 および距離データ記憶部 3 5 を備える。また

50

、コンテンツ検索プログラムがコンテンツ検索装置 1 の中央処理制御装置 101 に読み込まれ実行されることによって、特徴量算出手段 10、未知特徴量算出手段 14、距離算出手段 15 および表示手段 16 が、コンテンツ検索装置 1 に実装される。

【0036】

次に図 1 を参照して、本発明の実施の形態に係るコンテンツ検索装置 1 を説明する。

記憶装置 107 は、コンテンツデータ記憶部 30、特徴量データ記憶部 34 および距離データ記憶部 35 を備える。

【0037】

コンテンツデータ記憶部 30 は、コンテンツデータが記憶された記憶領域である。コンテンツデータは、画像特徴、音響特徴および意味特徴のうちいずれか一つ以上を有している。コンテンツデータ記憶部 30 において、各コンテンツデータは、コンテンツ識別子に対応づけられて記憶されている。コンテンツデータ記憶部 30 は、画像データ記憶部 31、動画データ記憶部 32 および記憶された音楽データ記憶部 31 を備える。画像データ記憶部 31 には、画像データ 31a がコンテンツ識別子と関連づけられて記憶されている。画像データ記憶部 31 に、複数の画像データが記憶されていても良い。動画データ記憶部 32 には、動画データ 32a がコンテンツ識別子と関連づけられて記憶されている。動画データ記憶部 32 に、複数の動画データが記憶されていても良い。音楽データ記憶部 33 には、音楽データ 33a がコンテンツ識別子と関連づけられて記憶されている。音楽データ記憶部 33 に、複数の音楽データが記憶されていても良い。

10

【0038】

特徴量データ記憶部 34 には、特徴量データ 34a が記憶されている。特徴量データは、各コンテンツデータの画像特徴量、音響特徴量および意味特徴量を関連づけたデータである。特徴量データ 34a は、各コンテンツ識別子と、画像特徴量、音響特徴量および意味特徴量を関連づけたデータである。各特徴量は、後述する特徴量算出手段 10 によって算出された特徴量、または未知特徴量算出手段 14 によって算出された特徴量である。

20

【0039】

図 5 を参照して、本発明の実施の形態に係る特徴量データ 34a を説明する。図 5 (a) に示す特徴量データ 34a は、コンテンツ識別子、コンテンツ種別、画像特徴量、音響特徴量および意味特徴量の各項目を有する。コンテンツ識別子は、各コンテンツデータ記憶部 30 に記憶されたコンテンツデータに関連づけられた識別子である。コンテンツ種別は、コンテンツデータ記憶部 30 に記憶されたコンテンツの種別であり、図 5 に示す例では、画像、音楽および動画のいずれかの種別を有する。画像特徴量、音響特徴量および意味特徴量の各項目には、コンテンツデータの各特徴の特徴量が関連づけられている。

30

【0040】

特徴量算出手段 10 によって各コンテンツデータの特徴量が算出されると、図 5 (a) に示すように、特徴量算出手段 10 によって算出された特徴量が特徴量データ 34a に登録される。図 5 (a) に示すように、コンテンツ種別が「画像」の場合、画像特徴量が、関連づけられる。コンテンツ種別が「音楽」の場合、音響特徴量が関連づけられる。コンテンツ種別が「動画」の場合、画像特徴量および音響特徴量がそれぞれ関連づけられる。メタデータ、音声データまたはキャプチャデータ等からテキストデータを取得できる場合、各コンテンツ識別子に、意味特徴量が関連づけられる場合もある。

40

【0041】

さらに未知特徴量算出手段 14 によって各コンテンツデータの未知の特徴量が算出されると、図 5 (b) に示すように、特徴量算出手段 10 によって算出された特徴量が特徴量データ 34a に登録され、各コンテンツデータについて、画像特徴量、音響特徴量および意味特徴量が関連づけられる。図 5 (b) に示すように、コンテンツ種別が「画像」の場合、音響特徴量が、コンテンツ種別が「音楽」の場合、画像特徴量が、それぞれ新たに関連づけられる。各コンテンツデータについて特徴量算出手段 10 によって意味特徴量が関連づけられていない場合、未知特徴量算出手段 14 によって算出される意味特徴量が関連づけられる。

50

【0042】

距離データ記憶部35には、距離データ35aが記憶されている。距離データ35aは、コンテンツデータ記憶部30に記憶される各コンテンツデータの任意の2つのコンテンツデータのコンテンツ識別子と、画像特徴量、意味特徴量および音響特徴量のそれぞれの類似度を示す距離を関連づけたデータである。後述する特徴量算出手段10および未知特徴量算出手段14によって、コンテンツデータ記憶部30に記憶される全てのコンテンツデータについて、画像特徴量、意味特徴量および音響特徴量が算出され、コンテンツ識別子に対応付けられた特徴量データ34aが生成される。さらに距離算出手段15は、特徴量データ34aから任意の2つのコンテンツデータの画像特徴量、意味特徴量および音響特徴量を抽出し、画像特徴、意味特徴および音響特徴のそれぞれについて、この2つのコンテンツデータの類似度となる距離を算出する。距離算出手段15は、この2つのコンテンツデータの識別子と、画像特徴の距離、意味特徴の距離および音響特徴の距離を関連づけて、距離データ35aを生成する。

10

【0043】

中央処理制御装置101には、コンテンツ検索プログラムがインストールされることにより、特徴量算出手段10、未知特徴量算出手段14、距離算出手段15および表示手段16が実装される。

【0044】

特徴量算出手段10は、コンテンツデータ記憶部30に記憶された各コンテンツデータについて、各コンテンツデータが有する画像特徴、音響特徴および意味特徴のうちいずれか一つ以上の特徴量を算出し、コンテンツ識別子と、画像特徴、音響特徴および意味特徴の特徴種別と、その特徴量とを関連づけた特徴量データ34aを、記憶装置107の特徴量データ記憶部34に記憶する。特徴量算出手段10は、画像特徴量算出手段11、意味特徴量算出手段12および音響特徴量算出手段13を備える。

20

【0045】

画像特徴量算出手段11は、画像データ記憶部31に記憶された画像データ31aについて、画像特徴量として、色ヒストグラムを算出する。さらに画像特徴量算出手段11は、動画データ記憶部32に記憶された動画データ32aのうち映像データについて、全フレームからそれぞれ色ヒストグラムを算出し、そのベクトルメジアンを画像特徴量とする。

30

【0046】

図6(a)を参照して、本発明の実施の形態に係る画像特徴量算出手段11が、画像データ31aについて画像特徴量を算出する処理を説明する。各画像データについて、ステップS101ないしステップS103の処理が繰り返される。

【0047】

具体的には、ステップS101において画像特徴量算出手段11は、一つのコンテンツ識別子に対応する画像データについて、RGB表示系における色ヒストグラムを算出する。このときのRGBの各ピン数は、例えば4、4および4とする。次にステップS102において画像特徴量算出手段11は、ステップS101で得られたベクトルを、当該画像データの画像特徴量として出力する。ステップS101においてRGBのピン数を4、4および4にした場合、このベクトルは64次元である。ステップS103において画像特徴量算出手段11は、ステップS102で出力した画像特徴量を、当該画像データのコンテンツ識別子に対応する画像特徴量として、特徴量データ34aに記録する。

40

【0048】

図6(b)を参照して、本発明の実施の形態に係る画像特徴量算出手段11が、動画データ32aについて画像特徴量を算出する処理を説明する。各動画データの映像データについて、ステップS151ないしステップS155の処理が繰り返される。

【0049】

具体的には、ステップS151において画像特徴量算出手段11は、一つのコンテンツ識別子に対応する動画データの映像データ部分を抽出し、各フレームについて、64次元

50

ベクトルを算出する。まずステップ S 1 5 1 において画像特徴量算出手段 1 1 は、一つのフレームの画像データについて、RGB 表示系における色ヒストグラムを算出する。このときの RGB の各ビン数は、例えば 4、4 および 4 とする。次にステップ S 1 5 2 において画像特徴量算出手段 1 1 は、ステップ S 1 5 1 で得られたベクトルを、当該フレームの特徴量として出力する。ステップ S 1 5 1 において RGB のビン数を 4、4 および 4 にした場合、ステップ S 1 5 2 で取得されるベクトルは 6 4 次元である。

【 0 0 5 0 】

一つの動画データの映像データ部分の各フレームについて 6 4 次元ベクトルが取得されると、ステップ S 1 5 3 において、各フレームの 6 4 次元ベクトルからベクトルメジアンを算出し、ステップ S 1 5 4 として当該映像データの画像特徴量として出力する。ステップ S 1 5 5 において画像特徴量算出手段 1 1 は、ステップ S 1 5 4 で出力した画像特徴量を、当該動画データのコンテンツ識別子に対応する画像特徴量として、特徴量データ 3 4 a に記録する。

10

【 0 0 5 1 】

意味特徴量算出手段 1 2 は、意味特徴量を算出する。画像データ 3 1 a、動画データ 3 2 a および音楽データ 3 3 a にテキスト形式のメタデータが付与されている場合、意味特徴量算出手段 1 2 は、これらのメタデータから意味特徴量を算出する。意味特徴量算出手段 1 2 は、さらに、動画データ 3 2 a に人の声やキャプチャのデータが含まれている場合、これらのデータをテキスト形式に変換して意味特徴量を算出しても良い。また意味特徴量算出手段 1 2 は、音楽データ 3 3 a に歌声が含まれている場合、この歌声のデータをテキスト形式に変換して歌詞データを取得し、意味特徴量を算出しても良い。

20

【 0 0 5 2 】

図 7 を参照して、本発明の実施の形態に係る意味特徴量算出手段 1 2 が、各コンテンツデータについて意味特徴量を算出する処理を説明する。各コンテンツデータのテキストデータについて、ステップ S 2 0 1 ないしステップ S 2 0 8 の処理が繰り返される。

【 0 0 5 3 】

まず意味特徴量算出手段 1 2 は、各コンテンツデータのテキストデータに含まれる各単語について、ステップ S 2 0 1 ないしステップ S 2 0 5 の処理を繰り返し、TFIDF を取得する。TFIDF は、テキスト中の特徴的な単語について、TF (Term Frequency : 出現頻度) および IDF (Inverse Document Frequency : 逆出現頻度) の二つの指標に基づいて算出される。

30

【 0 0 5 4 】

ステップ S 2 0 1 において意味特徴量算出手段 1 2 は、各単語がコンテンツ中に出現する回数 TF を算出する。次にステップ S 2 0 2 において意味特徴量算出手段 1 2 は、各単語が出現するコンテンツ数 DF を算出し、ステップ S 2 0 3 において、ステップ S 2 0 2 で算出された DF を用いて IDF を算出する。ステップ S 2 0 4 において意味特徴量算出手段 1 2 は、ステップ S 2 0 1 で算出された TF と、ステップ S 2 0 3 で算出された IDF とを用いて、各単語の TFIDF を算出する。ステップ S 2 0 5 において意味特徴量算出手段 1 2 は、ステップ S 2 0 4 で算出した TFIDF を正規化する。

【 0 0 5 5 】

所定のコンテンツに含まれる各単語について、TFIDF が正規化されると、ステップ S 2 0 6 において意味特徴量算出手段 1 2 は、特徴ベクトルとして算出する。特徴ベクトルは、全コンテンツに含まれる単語数を次元数とするベクトルである。ステップ S 2 0 7 において意味特徴量算出手段 1 2 は、ステップ S 2 0 6 において算出された特徴ベクトルを、当該コンテンツデータのコンテンツ識別子に対応する意味特徴量として、特徴量データ 3 4 a に記録する。

40

【 0 0 5 6 】

音響特徴量算出手段 1 3 は、音楽データ記憶部 3 3 に記憶された音楽データ 3 3 a について、音響特徴量を算出する。さらに音響特徴量算出手段 1 3 は、動画データ記憶部 3 2 に記憶された動画データ 3 2 a のうち音響データについて、音響特徴量を算出する。

50

【0057】

図8を参照して、本発明の実施の形態に係る音響特徴量算出手段13が、音楽データ33aと動画データ32aのうちの音響データについて画像特徴量を算出する処理を説明する。各音楽データまたは音響データの音響信号について、ステップS301ないしステップS306の処理が繰り返される。

【0058】

ステップS301ないしステップS304において音響特徴量算出手段13は、各時刻の音響信号についてクロマベクトルを算出する。具体的には音響特徴量算出手段13は、ステップS301において、所定時刻の音響信号について帯域通過フィルタを施し、ステップS302において、各音階に対するパワーを抽出する。これにより、各時刻における音階のパワーの分布を得ることができる。ステップS303において音響特徴量算出手段13は、ステップS302で抽出した各音階に対応する周波数のパワーを、オクターブ毎に換算し、ステップS304において各周波数パワーを要素とするクロマベクトルを算出する。

10

【0059】

所定の音楽データまたは音響データの各時刻の音響信号についてクロマベクトルが算出されると、ステップS305において音響特徴量算出手段13は、各時刻のクロマベクトルから、その時間変化を表す音階の遷移行列を算出する。このとき音響特徴量算出手段13は、所定時刻の1単位時間前のクロマベクトルから、所定時刻のクロマベクトルを推定した際の平均二乗誤差が最小となるように、遷移行列を算出する。ステップS306において音響特徴量算出手段13は、ステップS305において算出された音階の遷移行列の各要素を、当該コンテンツデータのコンテンツ識別子に対応する音響特徴量として、特徴量データ34aに記録する。

20

【0060】

未知特徴量算出手段14は、特徴量データ34aにおいて、コンテンツ識別子に対応づけられた特徴種別の特徴量から、コンテンツ識別子に対応づけられていない特徴種別の未知の特徴量を算出し、特徴推定値として特徴量データ34aに記憶する。

例えば、図2に示すように、画像データ31aについて、画像特徴量算出手段11によって画像特徴量が算出される。しかし画像データ31aは音響信号を有していないので、音響特徴量は算出されない。また、画像データ31aにメタデータが付与されていない場合、意味特徴量は算出されない。そこで、未知特徴量算出手段14は、画像データ31aについて、図3に示すように音響特徴量および意味特徴量を算出する。

30

動画データ32aについて、画像特徴量算出手段11によって画像特徴量が算出されるとともに、画像特徴量算出手段13によって音響特徴量が算出される。動画データ32aにメタデータが付与されていない場合や、人の声やキャプションなどが含まれていない場合、意味特徴量は算出されない。そこで、未知特徴量算出手段14は、動画データ32aについて、図3に示すように意味特徴量を算出する。

音楽データ33aについて、音響特徴量算出手段13によって音響特徴量が算出される。しかし音楽データ31aは画像信号を有していないので、画像特徴量は算出されない。また、音楽データ33aにメタデータが付与されていない場合、意味特徴量は算出されない。そこで、未知特徴量算出手段14は、音楽データ33aについて、図3に示すように画像特徴量および意味特徴量を算出する。

40

【0061】

未知の特徴量を算出するためには、異なる特徴間の相関が求める必要がある。未知の特徴量を算出する方法としては、主成分分析による方法、正準相関分析による方法などが考えられる。

正準相関分析が異なる特徴間の相関を求めるのに最適な方法であるのに対して、主成分分析による方法では、同一特徴間での相関も求める。その一方、主成分分析では、全ての特徴間で相関が求まっていることから、どの特徴が既知である、未知であることの拘束は存在しない。従って、主成分分析では、計算量を削減することができる。

50

これに対し正準相関分析による方法では、既知のデータ、未知のデータがはっきりとしている場合に、最適な相関を算出することが可能である。

未知特徴量算出手段14は、データの状況や、要求される処理能力等に応じて、いずれかの計算手法を選択することが好ましい。

【0062】

図9を参照して、未知特徴量算出手段14の主成分分析による未知特徴量算出処理を説明する。未知特徴量算出手段14は、特徴量算出手段10によって得られた各特徴量に対して主成分分析を施し、その固有空間を得る。このように算出される固有空間を用いて、未知特徴量算出手段14は、各コンテンツデータにおける未知の特徴量を算出することができる。

10

【0063】

まずステップS401におよびステップS402において未知特徴量算出手段14は、各コンテンツデータについて、その特徴値のベクトル x を算出する。具体的にはステップS401において未知特徴量算出手段14は、特徴量データ34aから、所定のコンテンツデータについて、画像特徴ベクトル(画像特徴量)、音響特徴ベクトル(音響特徴量)および意味特徴ベクトル(意味特徴量)を抽出する。ステップS402において未知特徴量算出手段14は、ステップS401で抽出した各ベクトルを並べて、ベクトル x を求める。このとき、画像データの音響特徴量などの未知の特徴量については、デフォルト値を設定する。

【0064】

全てのコンテンツデータについてベクトル x が算出されると、ステップS403において未知特徴量算出手段14は、各コンテンツのベクトル x について主成分分析を施し、固有ベクトルを並べた行列 U を取得する。

20

【0065】

ステップS404ないしステップS409において未知特徴量算出手段14は、各コンテンツデータについて特徴値が関連づけられていない特徴種別について、特徴値を算出する。具体的には、ステップS404において未知特徴量算出手段14は、所定のコンテンツデータのベクトル x について、全ての特徴量が関連づけられているか判定する。例えば、コンテンツデータが動画データの場合で、かつ、動画データにメタデータが関連づけられている場合や、動画データ中に人の声やキャプションが含まれている場合である。全ての特徴量が関連づけられている場合、未知特徴量算出手段14は、次のコンテンツデータについて、ステップS404を処理する。

30

【0066】

一方、いずれかの特徴量が関連づけられていない場合、ステップS405において未知特徴量算出手段14は、コンテンツ種別に応じて処理を振り分ける。コンテンツデータが画像データの場合、ステップS406において未知特徴量算出手段14は、ステップS403で取得した行列 U を利用して、当該コンテンツデータの音響特徴を推定したベクトルを算出する。当該コンテンツデータに意味特徴量が関連づけられていない場合、未知特徴量算出手段14はさらに、当該コンテンツデータの意味特徴を推定したベクトルを算出する。

40

【0067】

コンテンツデータが音楽データの場合、ステップS407において未知特徴量算出手段14は、ステップS403で取得した行列 U を利用して、当該コンテンツデータの画像特徴を推定したベクトルを算出する。当該コンテンツデータに意味特徴量が関連づけられていない場合、未知特徴量算出手段14はさらに、当該コンテンツデータの意味特徴を推定したベクトルを算出する。

【0068】

コンテンツデータが動画データの場合、ステップS408において未知特徴量算出手段14は、ステップS403で取得した行列 U を利用して、当該コンテンツデータの意味特徴を推定したベクトルを算出する。

50

【 0 0 6 9 】

ステップ S 4 0 6 ないしステップ S 4 0 8 で、行列 U を利用して未知の特徴量が算出されると、ステップ S 4 0 9 において、算出された特徴量を、当該コンテンツデータの識別子に関連づけて特徴量データ 3 4 a に登録する。

全てのコンテンツデータについてステップ S 4 0 4 ないしステップ S 4 0 9 が実行されると、未知特徴量算出処理は終了する。

【 0 0 7 0 】

図 1 0 を参照して、未知特徴量算出手段 1 4 の正準相関分析による未知特徴量算出処理を説明する。未知特徴量算出手段 1 4 は、特徴量算出手段 1 0 によって得られた各特徴量に対して正準相関分析を施し、その特徴の相関が最大となる正準相関係数およびその相関を得る。

10

【 0 0 7 1 】

まずステップ S 5 0 1 において未知特徴量算出手段 1 4 は、特徴量データ 3 4 a を読み出し、各コンテンツデータの特徴量を抽出する。ステップ S 5 0 2 において未知特徴量算出手段 1 4 は、ステップ S 5 0 1 で抽出した特徴量に対して、正準相関分析を施す。ステップ S 5 0 3 において未知特徴量算出手段 1 4 は、特徴量間の相関が最大となる正準相関係数およびその相関を算出する。

【 0 0 7 2 】

ステップ S 5 0 4 ないしステップ S 5 0 6 において未知特徴量算出手段 1 4 は、各コンテンツデータについて特徴値が関連づけられていない特徴種別について、特徴値を算出する。具体的には、ステップ S 6 0 4 において未知特徴量算出手段 1 4 は、所定のコンテンツデータのベクトル x について、全ての特徴量が関連づけられているか判定する。例えば、コンテンツデータが動画データの場合で、かつ、動画データにメタデータが関連づけられている場合や、動画データ中に人の声やキャプションが含まれている場合である。全ての特徴量が関連づけられている場合、未知特徴量算出手段 1 4 は、次のコンテンツデータについて、ステップ S 5 0 4 を処理する。

20

【 0 0 7 3 】

一方、いずれかの特徴量が関連づけられていない場合、ステップ S 5 0 5 において未知特徴量算出手段 1 4 は、ステップ S 5 0 3 で算出した正準相関係数およびその相関を用いて、未知の特徴量を算出する。ステップ S 5 0 6 において未知特徴量算出手段 1 4 は、ステップ S 5 0 5 で算出した各特徴量を、当該コンテンツデータの識別子に関連づけて特徴量データ 3 4 a に登録する。

30

全てのコンテンツデータについてステップ S 5 0 4 ないしステップ S 5 0 6 が実行されると、未知特徴量算出処理は終了する。

【 0 0 7 4 】

距離算出手段 1 5 は、特徴量データ 3 4 a に記憶されている特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する。各コンテンツデータが類似するほど距離は小さくなり、類似しないほど距離は大きくなる。算出された距離は、特徴データ記憶部 3 4 に記憶される。

【 0 0 7 5 】

本発明の実施の形態においてコンテンツ検索装置 1 は、全てのコンテンツデータについて、コンテンツ種別を問うことなく、画像特徴量、意味特徴量および音響特徴量が関連づけることができる。従って、画像データと音楽データの類似性、動画データと音楽データの類似性など、異なるコンテンツ間の類似度を算出することができる。

40

従って、ユーザは特定のコンテンツの種類を指定することなく、異なるコンテンツ種別のコンテンツについて、類似するコンテンツを検索することができる。

類似性の高いコンテンツを関連づけることにより、例えば「画像に合う音楽」や「動画に合う音楽」などを抽出することができる。従って、従来は知っているコンテンツでなければ関連づけができなかったところ、本発明の実施の形態によれば、操作者にとって未知のコンテンツを結びつけて新たなコンテンツを制作したり、コンテンツ制作の発想を支援

50

することができる。

【0076】

また、本発明の実施の形態においては、画像特徴、意味特徴および音響特徴のそれぞれについて距離を算出する。従って距離算出手段14は、「動画データの映像データ(画像特徴)に類似する音楽データ」や、「動画データの意味特徴に類似する画像データ」など、一つの特徴に着目して類似するコンテンツを検索することもできる。

【0077】

本発明の最良の実施の形態において距離算出手段15は、各特徴について、重みなしの距離、および重み付き距離のいずれかを算出する。図11および図12を参照して、距離算出手段15による距離算出処理を説明する。

10

【0078】

図11を参照して、重みなしの距離を算出する処理を説明する。図11では、コンテンツデータ記憶部31に記憶された所定のクエリコンテンツデータと、それ以外の検索対象のコンテンツデータとの距離を算出する場合について説明する。図11に示す例では、特徴量のL2距離を、コンテンツ間の距離として出力する。

【0079】

距離算出手段15は、特徴量データ34から、ステップS601においてクエリコンテンツデータのコンテンツ識別子に関連づけられた特徴量を抽出するとともに、ステップS602において検索対象コンテンツデータのコンテンツ識別子に関連づけられた特徴量を抽出する。ステップS601およびステップS602で抽出される特徴量は、特徴量算出手段10および未知特徴量算出手段14によって算出された画像特徴量、意味特徴量および音響特徴量である。

20

【0080】

ステップS603において距離算出手段15は、クエリコンテンツと検索対象コンテンツ間の距離として、ステップS601およびステップS602で抽出した各特徴量のベクトルのL2距離を算出する。距離算出手段15は、画像特徴の重みなし距離、意味特徴の重みなし距離、および音響特徴の重みなし距離をそれぞれ算出する。ステップS604において、ステップS603で算出した各特徴のL2距離を、コンテンツ間の距離として出力する。

【0081】

さらに、距離算出手段15は、他の検索対象コンテンツについてもステップS601ないしステップS604の処理を繰り返し、クエリコンテンツと、各々の検索対象コンテンツについて、画像特徴、意味特徴および音響特徴のそれぞれの距離を出力する。

30

【0082】

図12を参照して、重み付き距離を算出する処理を説明する。図12に示す例では、共分散行列を用いて算出されるマハラノビス汎距離を、コンテンツ間の距離として出力する。

【0083】

まずステップS651において距離算出手段15は、各コンテンツデータについて、画像特徴、意味特徴および音響特徴の共分散行列を求める。ステップS652において距離算出手段15は、ステップS651で算出した各特徴の共分散行列から、マハラノビス汎距離に基づいた各特徴の重み付き距離を算出する。

40

【0084】

距離算出手段15によって算出されたコンテンツ間の距離は、表示手段16によって各コンテンツが表示される位置の算出に用いられる。

【0085】

表示手段16は、クエリコンテンツが与えられた場合に、そのクエリコンテンツに類似するコンテンツの検索や、その結果を表示する。表示手段16は、距離算出手段15によって算出された距離に基づいて、各コンテンツデータに対応するサムネイルの表示位置を決定して、表示装置105に表示する。表示手段16は例えば、クエリコンテンツを画面

50

の中央に表示する。さらに表示手段 16 は、距離が小さく類似するコンテンツをクエリコンテンツの近くに表示するとともに、距離が大きく類似しないコンテンツをクエリコンテンツの遠くに表示する。

【0086】

表示手段 16 は、例えば、多次元尺度構成法 (MDS: MultiDimensional Scaling) を用いて、コンテンツ間の距離に基づく表示位置を決定し、表示装置 105 に表示する。「多次元尺度構成法」は、多変量解析の一手法である。「多次元尺度構成法」は、分類対象物の関係を低次元空間における点の位置関係によって表現する。本発明の実施の形態においては、各コンテンツにおいて算出された特徴ベクトル間の距離に対して適用することで、コンテンツ間の類似関係を保持した検索結果を可視化する。

10

【0087】

表示手段 16 は、さらに、画像データ、映像データおよび音楽データの各種コンテンツを再生するユーザインタフェースである。表示手段 16 は、映像データの再生、画像データの再生、音楽データの再生などを実現する。

【0088】

表示手段 16 による画像データの再生を説明する。表示手段 16 は、画像データ記憶部 31 の画像データ 31a のサムネイルを、表示装置 105 の表示画面の任意の位置に配置するとともに、画像データ 31a を再生する。このとき、表示手段 16 は、画像データ 31a に含まれる色に光る装飾を画像データ 31a の縦横のサイズに合わせて生成する。表示手段 16 は、画像データ 31a の周囲にこの装飾を配置して表示する。このように、画像データ 31a の周囲に装飾を行うことにより、動画データ 32a と同時に画面に表示する場合、その動画データ 32a とのコンテンツ種別の差異を明確にする視覚効果が得られる。

20

【0089】

表示手段 16 による映像データの再生を説明する。表示手段 16 は、動画データ記憶部 32 の動画データ 32a のサムネイルを、表示装置 105 の表示画面の任意の位置に配置するとともに、動画データ 32a を再生する。このとき表示手段 16 は、再生中の動画データ 32a の現在のフレームを表示するとともに、現在再生中のフレームより後方のフレームの画像を、3次元空間上で再生中のフレームの後方に配置して表示する。表示手段 16 は、動画データ 32a の再生に伴って、現在再生しているフレームの画像と、その後に表示されるフレームの画像を常に更新する。これにより、動画データ 32a の再生中に、画像が奥から手前に向かって流れているような視覚効果が得られる。

30

【0090】

表示手段 16 による音楽データの再生を説明する。表示手段 16 は、音楽データ記憶部 33 の音楽データ 33a のサムネイルを、表示装置 105 の表示画面の任意の位置に配置するとともに、音楽データ 33a を再生する。このとき表示手段 16 は、音楽データ 33a のサムネイルの配置された位置で、サムネイルを回転させるなどのアニメーションを行うことにより、音楽データ 33a が再生中であることを表す視覚効果が得られる。表示手段 16 は、音楽データ 33a の信号長や、音楽データ 33a をフーリエ変換し、この結果に基づいて、アニメーションデータを生成し、表示装置 105 に表示する。

40

【0091】

クエリコンテンツが与えられた場合に、表示手段 16 が、そのクエリコンテンツに類似するコンテンツの検索結果を表示する場合を説明する。本発明の実施の形態においては、動画データ等のコンテンツデータの再生中に、ユーザによって検索指示が入力されると、コンテンツ検索装置 1 の表示手段 16 は、再生中のコンテンツデータを、クエリコンテンツデータとして、コンテンツデータ記憶部 30 から類似するコンテンツを検索し、表示装置 105 に検索結果を表示する。

【0092】

コンテンツの検索においては、距離算出手段 15 によって算出されたコンテンツ間の距離に基づいて、コンテンツデータ記憶部 30 に記憶された各種コンテンツデータから、ク

50

エリコンテンツデータとの距離の小さいコンテンツデータが抽出される。具体的には、表示手段16は、距離データ35aから、クエリコンテンツデータと距離の小さいコンテンツ識別子を抽出し、抽出されたコンテンツ識別子に対応するコンテンツデータを、類似するコンテンツデータとして出力する。ここで、距離データ35aは画像特徴、意味特徴および音響特徴のそれぞれについて距離が算出されているので、画像特徴、意味特徴および音響特徴の各パラメータに重みを設定して類似するコンテンツデータを出力してもよい。これによれば、表示手段16は、画像特徴のみに着目して類似するコンテンツデータを検索することができる。また表示手段16は、音響特徴を3、意味特徴を7とする重みを設定して、類似するコンテンツデータを検索することができる。

【0093】

表示手段16は、類似するコンテンツデータを検索すると、多次元尺度構成法によって算出された位置関係に基づいて、そのコンテンツデータのサムネイルを表示装置105に表示する。表示装置105の表示画面において、クエリコンテンツデータのサムネイルの表示位置を原点とした2次元空間で表示される。表示手段16は、多次元尺度構成法を用い算出された座標をそれぞれの検索結果に与え、各検索結果の表示位置および大きさを決定する。さらにユーザによってそのコンテンツデータのサムネイルが選択されると、表示手段16は、選択されたサムネイルのコンテンツデータをクエリコンテンツデータとした類似コンテンツを検索することができる。

【0094】

表示手段16が、クエリコンテンツデータと、クエリコンテンツデータに類似するコンテンツデータを表示する際、例えば、図13、図16および図18に示すような表示パターンが考えられる。図13、図16および図18の表示パターンで表示した具体例を、それぞれ図14、図15、図17および図19ないし図21に示す。図15、図17および図19ないし図21において、二重枠で囲まれている画像のサムネイルは、画像データのサムネイルである。また一重枠で囲まれている画像のサムネイルは、動画データのサムネイルである。この動画データのサムネイルでは、動画データが再生されていても良い。また、音譜の画像のサムネイルは、音楽データのサムネイルである。この音楽データが再生されている間、その音楽データのサムネイルは、回転するなどして再生中であることを示すアニメーションが実行される。

【0095】

図13に示す例では、画面中央にクエリコンテンツデータのサムネイルを表示し、その周囲に検索結果のコンテンツデータのサムネイルを表示している。クエリコンテンツデータのサムネイルに近いほど、類似するコンテンツデータである。図13に示す例では、画像データ、動画データ、音楽データ等のコンテンツの種類に関係なく、検索結果のコンテンツデータのサムネイルが表示される。

この場合、動画データ32a等のコンテンツデータの再生中に、ユーザによって検索指示が入力された場合、クエリコンテンツデータのサムネイルを中心として、多次元尺度構成法を用いて算出された2次元空間の座標に基づいて、検索結果のコンテンツデータのサムネイルの座標を再計算し、表示装置105の表示画面に配置する。図13に示す例において、表示されている複数のコンテンツの位置関係によって、ユーザは、クエリコンテンツデータと検索されたコンテンツデータの類似する度合い、および検索されたコンテンツデータ同士が類似する度合いを、視覚的に把握することができる。

【0096】

図14および図15を参照して、図13に示す例に従ってサムネイルを表示した画面例を説明する。まず図14において、本発明の実施の形態に係るコンテンツ検索装置1が、最初に表示する画面である。図14の中央に、クエリコンテンツデータを指定する枠が設けられている。ユーザによって、この枠にクエリコンテンツデータが設定されると、本発明の実施の形態に係るコンテンツ検索装置1は、コンテンツデータ記憶部30に記憶された各種コンテンツデータのうち、クエリコンテンツデータに類似するコンテンツデータを検索し、図15に示す画面の様に、その結果を表示する。図15においては、中心にクエ

10

20

30

40

50

リコンテンツデータのサムネイルが表示される。またクエリコンテンツデータのサムネイルの周辺には、類似するコンテンツデータのサムネイルが表示され、類似するほど近くに表示される。

【0097】

図16に示す例では、画面中央にクエリコンテンツデータのサムネイルを表示し、その周囲に検索結果のコンテンツデータのサムネイルを表示する点で図13に示す例と類似する。図13に示す例では、検索結果のコンテンツデータのサムネイルを表示する際、コンテンツ種別に関係なく表示していたが、図16に示す例では、コンテンツデータの種別毎に表示エリアを分けている点異なる。図16に示す例では、検索結果のコンテンツデータのサムネイルのうち、画面左に画像データのサムネイルを、画面右上に音楽データのサムネイルを、画面右下に動画データのサムネイルを表示している。

10

この場合、動画データ32a等のコンテンツ再生中に、ユーザによって検索指示が入力された場合、表示手段16は、クエリコンテンツデータのサムネイルを中心として、多次元尺度構成法を用いて2次元空間の座標を算出する。さらに表示手段16は、算出された2次元空間の座標に基づいて、検索結果のコンテンツデータのコンテンツ種別ごとに、そのサムネイルの座標を再計算し、表示装置105の表示画面に配置する。図16に示す例において、表示されている複数のコンテンツのエリアおよび位置関係によって、ユーザは、検索されたコンテンツの種別、クエリコンテンツデータと検索されたコンテンツデータの類似する度合い、および検索されたコンテンツデータ同士が類似する度合いを、視覚的に把握することができる。

20

【0098】

図17を参照して、図16に示す例に従ってサムネイルを表示した画面例を説明する。図17に示す例では、中央にクエリコンテンツデータのサムネイルが表示されている。また、クエリコンテンツデータのサムネイルの左側には画像データのサムネイルが、右上には音楽データのサムネイルが、右下には動画データのサムネイルが、それぞれ表示されている。

【0099】

図18に示す例では、画面左上にクエリコンテンツデータのサムネイルを表示し、クエリコンテンツデータのサムネイルから放射状に検索結果のコンテンツデータのサムネイルを表示する。図18に示す例では、画像データ、動画データ、音楽データ等の、クエリコンテンツデータと同じ種類のコンテンツデータについて、検索結果のコンテンツデータのサムネイルが表示される。

30

【0100】

この場合、動画データ32a等のコンテンツ再生中に、ユーザによって検索指示が入力された場合、表示手段16は、クエリを画面端に配置し、クエリコンテンツデータのサムネイルを中心として多次元尺度構成法を用いて算出された2次元空間の座標に基づいて、検索結果のコンテンツデータのサムネイルの座標を、メディア別に再計算し表示装置105の表示画面に配置する。

一つのメディアにフォーカスを当てた配置で表示される検索結果は、表示されているコンテンツの位置関係によって、クエリコンテンツデータと検索結果が類似する度合いを視覚的に把握することができる。図18に示す例では、画像データ、動画データおよび音楽データのいずれか一つのメディアにのみフォーカスを当てて、検索結果を表示する。この場合、フォーカスが当てられていないコンテンツ種別のサムネイルについては、座標を再計算する必要はない。

40

【0101】

図19ないし図21を参照して、図18に示す例に従ってサムネイルを表示した画面例を説明する。図19ないし図21に示す画面例は、例えば図15または図17において示されたサムネイルのいずれか一つが選択された場合に、その選択されたサムネイルのコンテンツ種別にフォーカスをあてて検索結果を表示した画面例である。ここで、「フォーカスをあてる」とは、「動画像」、「音響」または「画像」のいずれかのコンテンツ種別を

50

特定して検索および表示することを言う。フォーカスを当てない場合は、「動画像」、「音響」および「画像」の全てのコンテンツ種別について検索および表示することを言う。

【0102】

図19では、動画データにフォーカスをあてて、クエリの動画データに類似する動画データのサムネイルが表示されている。各サムネイルでは、動画が再生されている。また、画面の中央から各サムネイルに向かって、これから再生されるフレームの複数の画像が3次元的に表示されている。図20では、音楽データにフォーカスをあてて、クエリコンテンツデータに類似する音楽データのサムネイルが表示されている。図21では、画像データにフォーカスをあてて、動画データに類似する画像データのサムネイルが表示されている。

10

【0103】

(効果)

このように本発明の実施の形態に係るコンテンツ検索装置1によれば、各種コンテンツについて、画像特徴量、意味特徴量および音響特徴量を算出する。従って、コンテンツ検索装置1は、画像データと音楽データ、動画データと画像データ、動画データと音楽データなど、コンテンツ種別にかかわらず、類似するコンテンツデータを検索することができる。

【0104】

さらに、検索されたコンテンツデータについて、表示手段16は、コンテンツ種別ごとに表示することもできる。また、所定のコンテンツに着目して、そのコンテンツをクエリコンテンツデータとして、そのコンテンツに類似するコンテンツを検索することもできる。

20

【0105】

なお、コンテンツデータとして、文字や音声のみからなるテキストデータを用いても良い。テキストデータから、特徴量算出手段10の意味特徴量算出手段12が意味特徴量を算出することができる。さらに、未知特徴量算出手段14が画像特徴量および音響特徴量を算出することができる。これによりテキストデータについても、画像特徴量、意味特徴量および音響特徴量を算出することができるので、距離算出手段15は、テキストデータと音楽データ、テキストデータと画像データ、テキストデータと動画データのそれぞれについても距離を算出できる。従って、テキストデータについても、コンテンツ種別にかかわらず類似するコンテンツデータを検索することができる。

30

【0106】

以下で、各処理手段の詳細なアルゴリズムについて説明する。

【0107】

(特徴量算出処理のアルゴリズム)

以下で、特徴量算出手段10による特徴量算出処理のアルゴリズムを説明する。本発明の実施の形態においては、画像データ31aと、動画データ32aの映像データの画像特徴量として、色ヒストグラムが用いられる。画像データ31a、動画データ32aおよび音楽データ33aの意味特徴量として、TFIDFが用いられる。動画データ32aの音響データと、音楽データ33aの音響特徴量として、音階の遷移に基づく特徴量が用いられる。

40

【0108】

まず、画像特徴量算出手段11による画像特徴量算出のアルゴリズムを説明する。本発明の実施の形態において画像特徴量算出手段11は、静止画像 f^{still} からHSV色ヒストグラムを求め、その要素を順に並べたベクトル v^{still} を静止画像 f^{still} の特徴ベクトルとして用いる。このとき、得られる特徴ベクトルはp次元である。またその値は、色相、彩度、明度のヒストグラムのビン数の積となる。また、画像特徴量算出手段11は、動画像信号 f^{video} の各フレームについて、HSV色ヒストグラムを求め、その要素を順に並べたベクトルを算出する。さらに、画像特徴量算出手段11は、全フレームの特徴ベクトルの全体からベクトルメジアンを求め、それを画像特徴ベクトル

50

v f i d e o とする。

【 0 1 0 9 】

つぎに、意味特徴量算出手段 1 2 による意味特徴量算出のアルゴリズムを説明する。本発明の実施の形態において意味特徴量算出手段 1 2 は、動画像信号 f f i d e o のコンテンツに含まれるテキストに対して、以下の処理により T F - I D F を算出する。T F - I D F 法は、あるコンテンツ中に出現するそれぞれの単語に対して、そのコンテンツを特徴づける度合いを算出する手法である。T F - I D F 法では、単語が特徴的である度合いを以下の式により算出する。

【 数 1 】

$$TFIDF(t_i, C_j) = TF(t_i, C_j)IDF(t_i)$$

$$IDF(t_i) = \log \frac{N}{DF(t_i)}$$

10

t_i : $TFIDF(t_i, C_j)$ を算出する対象となる単語。 i は単語の識別子を表す添字。

C_j : 単語 t_i が含まれるコンテンツ。 $j(j = 1, 2, \dots, N_c; N_c$ はコンテンツ数) はコンテンツの識別子を表す添字。

N : コンテンツの総数。

$TF(t_i, C_j)$: Term Frequency の略。単語 t_i がコンテンツ C_j に出現する頻度。

$DF(t_i)$: Document Frequency の略。単語 t_i が 1 回以上出現するコンテンツの数。

$IDF(t_i)$: Inverse Document Frequency の略。コンテンツの総数 N に対する $DF(t_i)$ の比の逆数。

20

【 0 1 1 0 】

上式により算出される $TFIDF(t_i, C_j)$ の値は、 $TF(t_i, C_j)$ が大きく $DF(t_i)$ が小さいほど大きくなる。つまり、単語 t_i がコンテンツ C_j 中に多く出現し、他のコンテンツ中にあまり出現しなければ、 $TFIDF(t_i, C_j)$ の値は大きくなる。ここで、コンテンツ中の単語の総数が多い場合に $TFIDF$ が大きくなりやすいため、コンテンツ間で $TFIDF$ の値を比較する際、正規化した値を用いることが望ましい。

30

【 数 2 】

$$w(t_i, C_j) = \frac{TFIDF(t_i, C_j)}{\sqrt{\sum_{c=1}^T (TFIDF(t_c, C_j))^2}}$$

T : 全コンテンツ中に出現する単語の種類の数。

【 0 1 1 1 】

本発明の最良の実施の形態において、意味特徴量算出手段 1 2 は、T F - I D F 法を動画像信号 f f i d e o に適用することで、コンテンツ C_j におけるテキストの特徴ベクトル

40

【 数 3 】

$$w_{C_j} = (w(t_1, C_j), w(t_2, C_j), \dots, w(t_T, C_j))$$

を得る。なお、T F - I D F 法以外の特徴量として、各コンテンツ C_j がテキスト t_i を持つ場合に

【数 4】

$$\tilde{w}(t_i, C_j) = 1$$

そうでない場合を

【数 5】

$$\tilde{w}(t_i, C_j) = 0$$

10

とし、

【数 6】

$$w(t_i, C_j) = \frac{\tilde{w}(t_i, C_j)}{\sqrt{\sum_{c=1}^T \tilde{w}(t_c, C_j)^2}}$$

20

とする特徴量も利用できる。

【0112】

さらに、音響特徴量算出手段 13 による音響特徴量算出のアルゴリズムを説明する。本発明の実施の形態において音響特徴量算出手段 13 は、楽曲中の音階の遷移に着眼したメロディーを表す指標を算出する。本発明の実施の形態においては、音響信号 *f^{music}* および動画像信号 *f^{video}* に含まれる音響信号から、音階の遷移に基づく特徴量を算出する。前処理としてハニング窓である窓関数 *h(t)* を用いた短時間フーリエ変換 (STFT) を、高速フーリエ変換 (FFT) によって計算する。FFT のフレームは時間軸方向にシフトすることで、楽曲全体に処理が施される。本発明の実施の形態において音響特徴量算出手段 13 は、全ての処理対象区間からクロマベクトルを算出する。さらに、音響特徴量算出手段 13 は、得られるクロマベクトルを用いて、音階の遷移行列を、メロディーを表す指標として算出する。

30

以下で、クロマベクトルの算出および音階の遷移行列の算出について説明する。

【0113】

まず、クロマベクトルの算出について説明する。一般に楽音が 12 音階によって進行していることから、本発明の実施の形態ではクロマベクトルを 12 次元で表現し、ベクトルの各次元が平均律の異なる音階のパワーを表すものとする。図 22 に、クロマベクトル算出の概要を示す。音響特徴量算出手段 13 は、時刻 *t* の入力楽曲信号に対する STFT を計算した後に、周波数軸を対数スケールの周波数 *g* に変換して、パワースペクトル *p(g, t)* を求める。対数スケールの周波数は *cent* 単位で表し、Hz で表された周波数 *g_{Hz}* を、次のように *cent* で表された周波数 *g_{cent}* に変換する。

40

【数 7】

$$g_{cent} = 1200 \log_2 \frac{g_{Hz}}{440 \times 2^{\frac{3}{12} - 5}}$$

平均律の半音は 100 *cent* に、1 オクターブは 1200 *cent* に相当するため、音名 *c* (*c* は 1 *c* 12 の整数で音階 C, C_{1s}, . . . , B に対応)、オクターブ位置 *h* の周波数

50

【数 8】

$$G_{c,h}^{cent}$$

は

【数 9】

$$G_{c,h} = 1200h + 100(c - 1) \quad 10$$

と表すことができる。

【0114】

この対数スケール軸のパワースペクトル $p(g, t)$ から、音階 c の位置のパワーを Oct_L から Oct_H のオクターブ範囲で加算して、12次元のベクトル

【数 10】

$$\tilde{\mathbf{x}}(t) \quad 20$$

の各次元

【数 11】

$$\tilde{x}_c(t)$$

【数 12】

$$\tilde{x}_c(t) = \sum_{h=Oct_L}^{Oct_H} \int_{-\infty}^{\infty} BPG_{c,g} \Psi_p(g, t) dg \quad 30$$

を求める。ここで、 $BPG_{c,g}$ は、音階 c 、オクターブ位置 h のパワーを通過させるバンドパスフィルタで、

【数 13】

$$BPG_{c,g}(g) = \frac{1}{2} \left(1 - \cos \frac{2\pi(g - (G_{c,h} - 100))}{200} \right) \quad 40$$

ハニング窓の形状で定義する。本発明の実施の形態では、算出されるベクトル

【数 14】

$$\tilde{\mathbf{x}}(t)$$

を以下の式により正規化を行うことで、クロマベクトル $x(t)$ を算出する。

【数 15】

$$x(t) = \frac{\tilde{x}(t)}{\|\tilde{x}(t)\|}$$

本発明の実施の形態では、このようにクロマベクトルを算出することで、多重音を含む楽曲信号においても音階のパワーの分布を適切に得ることができる。

10

【0115】

つぎに、音階の遷移行列の算出について説明する。本発明の最良の実施の形態において音響特徴量算出手段 13 は、クロマベクトル $x(t)$ を用いて音階の遷移行列を算出する。本発明の実施の形態において、時刻 t におけるクロマベクトルの集合を $X_t = [x(1), x(2), \dots, x(T-1)]$ としたとき、時刻 t から $t+1$ への音階の遷移を遷移行列 A を用いて次式によって表す。

【数 16】

$$X_{t+1} = AX_t$$

20

・・・ 式 (1)

このとき、遷移行列 A は、次式のように得られる。

【数 17】

$$A = X_{t+1} X_t' (X_t X_t')^{-1}$$

【0116】

30

上式により得られる遷移行列 A は 12×12 となり、その各要素は時刻 t から $t+1$ へ音階が遷移する度合いを表す。例えば楽曲のメロディーが C D A D と演奏される場合、図 23 に示されるような遷移行列 A の要素が高くなる。このとき、遷移行列 A は、音階のパワーの分布であるクロマベクトル X に基づいて算出されるために、多重音を含む楽曲信号においても主要な音階の遷移が行列 A の要素に反映される。したがって、楽曲の特徴的な音階の遷移を、遷移行列 A は表現可能であると考えられる。本発明の実施の形態では、

【数 18】

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{12} \end{pmatrix}$$

10

$$a_k = (a_{k1}, a_{k2}, \dots, a_{k12}) \quad (k = 1, 2, \dots, 12)$$

と表すと、音響信号 f^{music} および動画像信号 f^{video} に含まれる音響信号の特徴ベクトルとして、それぞれ

【数 19】

$$a^{\text{music}} = (a_1^{\text{music}}, \dots, a_{12}^{\text{music}}) \text{ および } a^{\text{video}} = (a_1^{\text{video}}, \dots, a_{12}^{\text{video}})$$

20

を用いる。

【0117】

本発明の実施の形態では、式(1)により音階の遷移を表現する。従って、従来では推定困難とされる特定の楽器音の示す基本周波数を推定することなく、音階の遷移であるメロディーを遷移行列 A として得ることができる。これにより、行列 A の要素を特徴量として用いることができる。

【0118】

(未知特徴量算出処理のアルゴリズム)

以下で、未知特徴量算出手段 14 による未知特徴量算出処理のアルゴリズムを説明する。未知特徴量算出手段 14 は、主成分分析あるいは正準相関分析に基づいて、未知の特徴量を分析する。

30

【0119】

まず主成分分析による未知特徴量算出処理を説明する。主成分分析は、多変量解析の分野において多変数データからの説明変数を抽出する手法として用いられる。以下では、主成分分析に基づく特徴量推定の方法について具体的に説明を行う。

【0120】

本発明の実施の形態において未知特徴量算出手段 14 は、データベース中に存在する全てのコンテンツ C_j ($j = 1, 2, \dots, N$) に対してその画像特徴ベクトル v_j 、音響特徴ベクトル a_j 、および意味に基づく特徴ベクトル w_j を求める。未知特徴量算出手段 14 は、それらを順に並べたベクトル

40

【数 20】

$$x = [v'_j, a'_j, w'_j]'$$

を求める。尚、 v_j 、 a_j 、 w_j のいずれかが未知の場合、未知特徴量算出手段 14 は、そのベクトルの要素は全て同じ値とする。さらに未知特徴量算出手段 14 は、得られるベクトル x_1 、 x_2 、 \dots 、 x_N に対して主成分分析を施し、その固有ベクトル u_1 、 u_2 、 \dots 、 u_D (D 次元) を順に並べた行列

50

【数 2 1】

$$U = [u_1, u_2, \dots, u_D]$$

を得る。この行列を用いることで、本発明の実施の形態においては、新しいコンテンツに対して未知の特徴量を以下のように推定する。

【0 1 2 1】

まず画像データについて説明する。

10

未知特徴量算出手段 1 4 は、画像特徴 v 、音響特徴 a 、意味特徴 w のうち、音響特徴および意味特徴が未知である場合、ベクトル x を

【数 2 2】

$$x = [v', 0', 0']'$$

とする。未知特徴量算出手段 1 4 は、その音響特徴および意味特徴を推定したベクトル

20

【数 2 3】

$$x^{pca} = [\tilde{v}', \tilde{a}', \tilde{w}']'$$

を

【数 2 4】

$$x^{pca} = UU'(x - \bar{x}) + \bar{x}$$

30

・・・式(2)

として求める。ただし、

【数 2 5】

$$\bar{x}$$

は、 x_j ($j = 1, 2, \dots, N$) の平均ベクトルである。

【0 1 2 2】

次に音楽データについて説明する。

40

未知特徴量算出手段 1 4 は、画像特徴 v 、音響特徴 a 、意味特徴 w のうち、画像特徴および意味特徴が未知である場合、ベクトル x を、

【数 2 6】

$$x = [0', a', 0']'$$

とする。未知特徴量算出手段 1 4 は、その画像特徴および意味特徴を推定したベクトル x^{pca} を式(2)で得る。

50

【 0 1 2 3 】

さらに動画データについて説明する。

未知特徴量算出手段 1 4 は、画像特徴 v 、音響特徴 a 、意味特徴 w のうち、意味特徴が未知である場合、ベクトル x を

【数 2 7】

$$\mathbf{x} = [\mathbf{v}', \mathbf{a}', \mathbf{0}']'$$

10

とする。未知特徴量算出手段 1 4 は、その意味特徴を推定したベクトル x^{pca} を式 (2) で得る。

【 0 1 2 4 】

本発明の実施の形態において特徴量算出手段 1 0 は、あらかじめ対象の内容を表すテキストが付与された画像データ、音楽データ、および動画データの各コンテンツデータから画像、音響、および意味の各々について特徴量を算出する。その後、未知特徴量算出手段 1 4 は、得られた各特徴量に対して主成分分析を施し、その固有空間を得る。このように算出される固有空間を用いて、本発明の実施の形態に係るコンテンツ検索装置 1 は、各データにおける未知の特徴量の推定を可能とする。ただし、主成分分析を施す画像データ、音楽データ、および動画データの各コンテンツデータから得られる各特徴量は対象データの特性を十分に反映するように広く選択されていることが好ましい。

20

【 0 1 2 5 】

つぎに正準相関分析による未知特徴量算出処理を説明する。まず、 R 組の変数群

【数 2 8】

$$X_i^1, X_i^2, \dots, X_i^{p_i} \quad (i = 1, \dots, R; p_i \text{ は } i \text{ 番目の変数群の次元数を表す})$$

が与えられ、これら R 組の変数群に関する n 個のデータ

【数 2 9】

$$\mathbf{x}_i^{p_i} (= [X_i^{p_i}(1), \dots, X_i^{p_i}(n)]')$$

30

を R 個の行列

【数 3 0】

$$X_i = [x_i^1, x_i^2, \dots, x_i^{p_i}] \quad (i = 1, 2, \dots, R)$$

によって表現する。

40

【 0 1 2 6 】

未知特徴量算出手段 1 4 は、正準相関分析では線形変換

【数 3 1】

$$\mathbf{g}_i = X_i \mathbf{w}_i$$

で得られる重みベクトル w_i を求める。ただし、未知特徴量算出手段 1 4 は、

【数 3 2】

$$\mathbf{X}_i^{\mathbf{p}_i}$$

に含まれる各成分の平均は 0 とする。

【0 1 2 7】

ここで、未知数からなる n 次元のベクトル y を導入し、未知特徴量算出手段 1 4 は、 y と $\mathbf{X}_i \mathbf{w}_i$ の各成分について次式のように、差の平方和を最小とする \mathbf{w}_i を推定する。

10

【数 3 3】

$$Q(y, \mathbf{w}_i) = \sum_{i=1}^R \|\mathbf{y} - \mathbf{X}_i \mathbf{w}_i\|^2 \quad (\mathbf{y}'\mathbf{y} = 1)$$

未知特徴量算出手段 1 4 は、まず、 y を既知として、上式をベクトル \mathbf{w}_i の各成分について最小とすると、最小二乗法の性質より

20

【数 3 4】

$$\begin{aligned} Q(y, \mathbf{w}_i) &\geq \sum_{i=1}^R \|\mathbf{y} - \mathbf{X}_i (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' \mathbf{y}\|^2 \\ &= R \|\mathbf{y}\|^2 - \mathbf{y}' \sum_{i=1}^R (\mathbf{X}_i (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i') \mathbf{y} \\ &= Q(\mathbf{y}) \end{aligned}$$

30

となる。なお、等号は

【数 3 5】

$$\mathbf{w}_i = (\mathbf{X}_i' \mathbf{X}_i)^{-1} \mathbf{X}_i' \mathbf{y}$$

40

のときに成立する。

【0 1 2 8】

次に $\mathbf{y}'\mathbf{y} = 1$ の条件で $Q(\mathbf{y})$ を最小にするためには、同一の条件下で $Q(\mathbf{y})$ の第 2 項の符号を変えた部分を \mathbf{y} について最大にすればよい。したがって \mathbf{y} は

【数 3 6】

$$\sum_{i=1}^R (\mathbf{X}_i (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i) \mathbf{y} = \lambda \mathbf{y}$$

の固有値問題の解として求まる。また、未知特徴量算出手段 1 4 は、上式を満たす

【数 3 7】

$$N_e (= \text{rank}(\mathbf{X}'_i \mathbf{X}_j))$$

個の正の固有値 λ_k ($k = 1, 2, \dots, N_e$) を得る。これらに対応する固有ベクトル \mathbf{y}_k ($k = 1, 2, \dots, N_e$) から得られる重みベクトルを順に並べた行列を

【数 3 8】

$$\mathbf{W}_i = [\mathbf{w}_i^1, \mathbf{w}_i^2, \dots, \mathbf{w}_i^{N_e}] \quad (i = 1, 2, \dots, R)$$

とおくと、次式が成立する。

【数 3 9】

$$\mathbf{W}'_i \mathbf{X}'_i \mathbf{X}_j \mathbf{W}_j = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & & \\ \vdots & & \ddots & \\ 0 & & & \mu_{N_e} \end{bmatrix} = \Lambda_{ij}$$

ここで、 Λ_{ij} は正準相関係数 μ_k ($k = 1, 2, \dots, N_e$) を対角要素とする相関行列である。

【0 1 2 9】

以上に示す正準相関分析を用いることにより未知特徴量算出手段 1 4 は、異なる R 組の変数群間の相関が最大となる線形変換を求めることができる。本発明の実施の形態に係る未知特徴量算出手段 1 4 は、図 2 4 に示すように、メディア数を $R = 3$ 、コンテンツ数を n 個として正準相関分析を行い、相関行列を得る。これらの行列を用いることで本発明の実施の形態において未知特徴量算出手段 1 4 は、新しいコンテンツに対して既知の特徴から未知の特徴量を以下のいずれかの式により推定する。

【0 1 3 0】

未知特徴量算出手段 1 4 は、新しいコンテンツに対して既知の特徴から、未知の画像特徴量を以下のいずれかの式により推定する。

10

20

30

40

【数 4 0】

$$\begin{aligned}\tilde{\mathbf{v}} &= \mathbf{W}'_v \mathbf{v} \\ \tilde{\mathbf{v}} &= \Lambda_{va} \mathbf{W}'_a \mathbf{a} \\ \tilde{\mathbf{v}} &= \Lambda_{wv} \mathbf{W}'_w \mathbf{w}\end{aligned}$$

10

【0 1 3 1】

未知特徴量算出手段 1 4 は、新しいコンテンツに対して既知の特徴から、未知の音響特徴量を以下のいずれかの式により推定する。

【数 4 1】

$$\begin{aligned}\tilde{\mathbf{a}} &= \Lambda_{va} \mathbf{W}'_v \mathbf{v} \\ \tilde{\mathbf{a}} &= \mathbf{W}'_a \mathbf{a} \\ \tilde{\mathbf{a}} &= \Lambda_{wa} \mathbf{W}'_w \mathbf{w}\end{aligned}$$

20

【0 1 3 2】

未知特徴量算出手段 1 4 は、新しいコンテンツに対して既知の特徴から、未知の意味特徴量を以下のいずれかの式により推定する。

【数 4 2】

$$\begin{aligned}\tilde{\mathbf{w}} &= \Lambda_{wv} \mathbf{W}'_v \mathbf{v} \\ \tilde{\mathbf{w}} &= \Lambda_{wa} \mathbf{W}'_a \mathbf{a} \\ \tilde{\mathbf{w}} &= \mathbf{W}'_w \mathbf{w}\end{aligned}$$

30

40

【0 1 3 3】

また未知特徴量算出手段 1 4 は、未知の特徴が推定されたベクトル

【数 4 3】

$$\mathbf{X}^{cca} = [\tilde{\mathbf{v}}', \tilde{\mathbf{a}}', \tilde{\mathbf{w}}']'$$

を得る。

【0 1 3 4】

50

本発明の実施の形態において特徴量算出手段 10 は、あらかじめ対象の内容を表すテキストが付与された画像データ、音楽データ、および動画データの各コンテンツデータから画像、音響、および意味の各々について特徴量を算出する。その後、未知特徴量算出手段 14 は、得られた各特徴量に対して正準相関分析を施し、各特徴の相関が最大となる正準相関係数およびその相関を得る。これにより本発明の実施の形態において、そのままでは比較できなかった各特徴量を比較することが可能となる。ただし、正準相関分析を施す画像データ、音楽データ、および動画データの各コンテンツデータから得られる各特徴量は対象データの特性を十分に反映するように広く選択されていることが好ましい。

【0135】

(距離算出処理のアルゴリズム)

10

以下で、距離算出手段 15 による距離算出処理のアルゴリズムを説明する。距離算出手段 15 は、重みなし距離または重み付き距離を算出する。

【0136】

まず重みなし距離を算出する場合の距離算出処理を説明する。

今、クエリコンテンツが C^q 、対象コンテンツが C^f であるとする。このとき距離算出手段 15 は、それらコンテンツ間における画像特徴の重みなし距離

【数 4 4】

$$d_v^{L2}$$

20

、音響特徴の重みなし距離

【数 4 5】

$$d_a^{L2}$$

30

、および意味特徴の重みなし距離

【数 4 6】

$$d_w^{L2}$$

をそれぞれ以下の式で算出する。

【数 4 7】

40

$$d_v^{L2} = \|\tilde{\mathbf{v}}^q - \tilde{\mathbf{v}}^f\|^2$$

$$d_a^{L2} = \|\tilde{\mathbf{a}}^q - \tilde{\mathbf{a}}^f\|^2$$

$$d_w^{L2} = \|\tilde{\mathbf{w}}^q - \tilde{\mathbf{w}}^f\|^2$$

50

ただし、ベクトル

【数 4 8】

$$\tilde{v}^q \text{ および } \tilde{v}^f$$

はそれぞれ、未知特徴量算出手段 1 4 による未知特徴量算出処理で得られる C^q および C^f のベクトル x における画像特徴のベクトルである。同様に

【数 4 9】

$$\tilde{a} \text{ および } \tilde{w}$$

10

は、音響特徴および意味特徴のベクトルである。

【0 1 3 7】

つぎに重み付き距離を算出する場合の距離算出処理を説明する。距離算出手段 1 5 は、マハラノビス汎距離に基づいたそれらコンテンツ間における画像特徴の重み付き距離

【数 5 0】

20

$$d_v^m$$

、音響特徴の重み付き距離

【数 5 1】

$$d_a^m$$

30

、および意味特徴の重み付き距離

【数 5 2】

$$d_w^m$$

40

をそれぞれ以下の式で算出する。

【数 5 3】

$$\begin{aligned}
 d_v^m &= (\tilde{v}^q - \tilde{v}^f)' \mathbf{W}_v (\tilde{v}^q - \tilde{v}^f) \\
 d_a^m &= (\tilde{a}^q - \tilde{a}^f)' \mathbf{W}_a (\tilde{a}^q - \tilde{a}^f) \\
 d_w^m &= (\tilde{w}^q - \tilde{w}^f)' \mathbf{W}_w (\tilde{w}^q - \tilde{w}^f)
 \end{aligned}
 \tag{10}$$

【0 1 3 8】

ここで、行列 \mathbf{W}^v 、 \mathbf{W}^a 、 \mathbf{W}^w はそれぞれ、検索・分類対象の全コンテンツにおける画像、音響、意味特徴の共分散行列であり、次式により求められる。

【数 5 4】

$$\begin{aligned}
 \mathbf{W}^v &= \frac{1}{N} \mathbf{X}_v \mathbf{X}_v' \\
 \mathbf{W}^a &= \frac{1}{N} \mathbf{X}_a \mathbf{X}_a' \\
 \mathbf{W}^w &= \frac{1}{N} \mathbf{X}_w \mathbf{X}_w'
 \end{aligned}
 \tag{20}$$

【0 1 3 9】

ただし、

【数 5 5】

$$\begin{aligned}
 \mathbf{X}_v &= [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_N]', \\
 \mathbf{X}_a &= [\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_N]', \\
 \mathbf{X}_w &= [\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_N]'
 \end{aligned}
 \tag{30}$$

であり、 N は全コンテンツ数である。

【0 1 4 0】

(表示処理のアルゴリズム)

50

以下で、表示手段 16 による表示処理のアルゴリズムを説明する。表示手段 16 は、距離算出手段 15 によって出力された距離データ 35 に基づいて、検索結果を表示装置 105 の表示画面に表示する。所定のコンテンツとクエリコンテンツの距離が小さいほど、具体的にはコンテンツが類似するほど、そのコンテンツのサムネイルをクエリコンテンツのサムネイル近くに表示する。一方、所定のコンテンツとクエリコンテンツの距離が大きいほど、具体的にはコンテンツが類似しないほど、そのコンテンツのサムネイルをクエリコンテンツのサムネイル遠くに表示する。

本発明の実施の形態において、多次元尺度構成法 (MDS) を、各コンテンツにおいて算出された特徴ベクトル間の距離に対し適用することで、コンテンツ間の類似関係を保持した検索結果が可視化される。

10

【0141】

n 個のコンテンツにおいて算出された特徴ベクトルを x_1, \dots, x_n とし、 x_i, x_j 間の距離を 2 乗した値を要素とする $n \times n$ の行列を D とする。また、単位行列から、全要素が $1/n$ の行列を引いた結果、得られる $n \times n$ 行列を J とする。さらに、

【数 56】

$$P = (-1/2)JDJ^T$$

20

とし、 i 番目のコンテンツに対する k 次元空間の推定位置を

【数 57】

$$\hat{x}_i$$

、それを n 行集めた $n \times k$ 行列を求める X とすると、

【数 58】

30

$$XX^T$$

は次式を最小化する。

【数 59】

$$\phi = \text{trace}[(P - XX^T)^2]$$

40

・・・ 式 (3)

【0142】

P の固有値のうち、最大の k 個の固有値を対角要素とする行列を Λ_k 、対応する固有ベクトルの行列を Q_k とすると、多次元尺度構成法によるコンテンツの配置

【数 6 0】

$$\hat{X}$$

は以下の式により算出される。

【数 6 1】

$$\hat{X} = Q_r \Lambda_k^{1/2}$$

10

ただし、Pは最小でもk個の正の固有値が存在するものとする。式(3)により得られるコンテンツの配置

【数 6 2】

$$\hat{X}$$

20

は、式(3)を最小化し、コンテンツ間の距離を最小二乗の基準によりk次元で近似する解となる。

【0143】

このように本発明の実施の形態に係るコンテンツ検索装置1は、画像、動画または音楽のいずれかのコンテンツデータから、画像、意味、音響のいずれか一つ以上の特徴量を算出する。さらにコンテンツ検索装置1は、それぞれのコンテンツデータについて算出された特徴量から、コンテンツデータに未知な特徴量を算出し、各コンテンツデータに、画像、意味および音響の全ての特徴量を関連づけることができる。

30

【0144】

このようなコンテンツ検索装置1によれば、画像データに類似する音楽データ、動画データに類似する画像データなど、コンテンツの種別に関わらず、横断的に検索することができる。従って、ユーザは、コンテンツ検索装置1を使用してクエリコンテンツデータから類似するコンテンツデータを次々と検索することができる。ユーザは大量のコンテンツデータから所望のコンテンツデータを検索することができる。また、横断的に検索結果を出力することにより、画像データに似合う音楽データが検索されるなど、感覚的な検索結果を取得することができる。

40

【0145】

(第1の変形例)

本発明の第1の変形例として、音響信号の繰り返し部分に着目した特徴量を算出する場合について説明する。ここで、繰り返し部分とは、例えば曲の「サビ」の部分である。

本発明の実施の形態においては、楽曲全体から特徴量を算出している。従って、算出される特徴量はその時間長により平滑化されてしまい、抽出精度が劣化する可能性がある。そこで、第1の変形例においては、楽曲の繰り返し区間を検出する手法を導入することで、楽曲中に最も多く現れる区間を検出し、その区間より得られる遷移行列を用いて楽曲間の類似度を算出する。このように、楽曲の繰り返し区間から算出される類似度を用いることで、第1の変形例においては、高精度な類似楽曲を抽出することができる。

【0146】

50

本発明の第1の変形例に係るコンテンツ検索装置1aは、図1を参照して説明した本発明の実施の形態に係るコンテンツ検索装置1と比べて、音響特徴量算出手段13の処理が異なる。

音響特徴量算出手段13は、音響信号のうち繰り返し現れるメロディーを特定し、その繰り返しメロディーの特徴量を、音響特徴の特徴量とする。従って特徴量データ記憶部34に記憶される音響特徴は、音響信号の繰り返し区間についての特徴量であって、その楽曲の特徴にフォーカスをあてた特徴量であると考えられる。本発明の第1の変形例に異なるコンテンツ検索装置1aにおいて、音響特徴量算出手段13が処理する楽曲は、コンテンツデータ記憶部30の動画データ記憶部32に記憶された動画データのうち音響データ部分と、音楽データ記憶部33に記憶された音楽データである。

10

【0147】

つぎに、本発明の第1の変形例に係る音響特徴量算出手段13の処理を説明する。特徴量算出手段13は、前処理として楽曲信号を固定長のフレームに分割し、時刻tにおけるフレームから特徴量としてクロマベクトル X_t を算出する。さらに音響特徴量算出手段13は、全てのフレームの組み合わせに対して算出される類似度を用いることで、繰り返し区間を検出する。

【0148】

具体的に音響特徴量算出手段13は、 X_t と X_{t-1} ($0 \leq t$)間の相関係数を類似度 $r(t, 1)$ ($-1 \leq r(t, 1) \leq 1$)として求める。類似度 $r(t, 1)$ を、横軸がt、縦軸が1のt-1平面に描画した場合、楽曲中の繰り返し区間に対応して、類似度が連続して高くなる領域が時間軸に並行な線分として現れる。したがって、このような線分を検出することで、音響特徴量算出手段13は、楽曲中の繰り返し区間を得ることができる。

20

ただし、各要素がほぼ等しいクロマベクトルは、他のクロマベクトルとの類似度が高くなってしまふ傾向があり、t-1平面中に類似度の高い直線として現れる場合がある。このような直線を除去するために、本発明の第1の変形例に係る音響特徴量算出手段13は、移動平均フィルタを用いて時間軸に並行な線分を強調することで、高精度に繰り返し区間を検出する。

【0149】

本発明の第1の変形例では、このように得られる楽曲の繰り返し区間の中から、最も多く繰り返されるメロディーを含む区間に注目する。一般に、楽曲中で最も代表的な主題が提示される部分は繰り返されることが多いことから、音響特徴量算出手段13は、このような区間から抽出される特徴量を用いることで、楽曲間の適切な類似度を算出する。

30

【0150】

次に、音響特徴量算出手段13が、繰り返し区間から、楽曲の音階の遷移を表す特徴量を算出する処理を説明する。本発明の第1の変形例では、クロマベクトル X_t から X_{t+1} への時間変化が、遷移行列Aを用いて次式によって表される。

【数63】

$$X_{t+1} = AX_t$$

40

このとき、繰り返し区間中のtに関する

【数64】

$$E(t) = \|X_{t+1} - AX_t\|^2$$

の平均を最小とする行列Aは、次式のように得られる。

50

【数 6 5】

$$A = X_{t+1} X'_t (X_t X'_t)^{-1}$$

・・・ 式 (4)

ただし、 X_t 、 X_{t+1} は、それぞれ、

【数 6 6】

$$X_t = [x(1), x(2), \dots, x(T-1)]$$

$$X_{t+1} = [x(2), x(3), \dots, x(T)]$$

で表される。ここで、 T は、繰り返し区間から算出されるクロマベクトルの総数を表す。
このように得られる遷移行列のサイズは、 $L \times L$ ($L = 12$) である。

【0 1 5 1】

本発明の第 1 の変形例においては、式 (4) により遷移行列 A を算出することにより、
従来では困難であった、音階の遷移を表す特徴量を得ることができる。

【0 1 5 2】

次に、音響特徴量算出手段 13 は、音階の遷移を表す特徴量を用いて、その遷移行列の
相関係数を、楽曲間の類似度として算出する。

ここで、クエリ楽曲の繰り返し区間 n ($n = 1, 2, \dots, N$; N は繰り返し区間の総
数) より算出される音階の遷移行列を

【数 6 7】

$$A_Q^n$$

、コンテンツデータ記憶部 30 に記憶された楽曲の繰り返し区間 m ($m = 1, 2, \dots, M$; M は繰り返し区間の総数) より算出される音階の遷移行列を

【数 6 8】

$$A_R^m$$

とする。

【0 1 5 3】

このとき音響特徴量算出手段 13 は、クエリ楽曲の繰り返し区間 n とコンテンツデータ
記憶部 30 に記憶された楽曲の繰り返し区間 m の相関係数を、次式により算出する。

10

20

30

40

【数 6 9】

$$R_{n,m} = \frac{S_{A_Q^n} A_R^m}{S_{A_Q^n} S_{A_R^m}}$$

ただし、

【数 7 0】

10

$$S_{A_Q^n} A_R^m, S_{A_Q^n}, S_{A_R^m}$$

は、次式により得られる。

【数 7 1】

$$S_{A_Q^n} A_R^m = \sum_{i=1}^L \sum_{j=1}^L (A_Q^n(i, j) - \mu_Q^n)(A_R^m(i, j) - \mu_R^m)$$

20

$$S_{A_Q^n} = \sqrt{\sum_{i=1}^L \sum_{j=1}^L (A_Q^n(i, j) - \mu_Q^n)^2}$$

$$S_{A_R^m} = \sqrt{\sum_{i=1}^L \sum_{j=1}^L (A_R^m(i, j) - \mu_R^m)^2}$$

30

【数 7 2】

$$\mu_Q^n = \frac{1}{L^2} \sum_{i=1}^L \sum_{j=1}^L A_Q^n(i, j)$$

10

$$\mu_R^m = \frac{1}{L^2} \sum_{i=1}^L \sum_{j=1}^L A_R^m(i, j)$$

20

【0154】

(第2の変形例)

図25を参照して、本発明の第2の変形例に係るコンテンツ検索装置1bを説明する。図25に示すコンテンツ検索装置1bは、図1に示す本発明の実施の形態に係るコンテンツ検索装置1と比較して、記憶装置107がユーザ実行データ記憶部36および重要ユーザデータ記憶部37を備えている点と、中央処理制御装置101が、重要ユーザ決定手段17を備えている点が異なる。また、図25において距離算出手段15aおよび表示手段16aの処理は、図1に示した距離算出手段15および表示手段16のそれぞれの処理と一部の処理が異なる。

【0155】

30

ユーザ嗜好データ記憶部36は、ユーザ嗜好データ36aが記憶された記憶領域である。ユーザ嗜好データ36aは、ユーザ識別子と、当該ユーザの嗜好に合うコンテンツ識別子とが関連づけられたデータである。

ユーザ嗜好データ36aは例えば、ユーザによって使用されるユーザ端末で保存されているプレイリストから生成されても良い。コンテンツ検索装置1bは、ユーザ端末から、ユーザの識別子とプレイリストで記録されているコンテンツの識別子とを受信して、ユーザ嗜好データ36aを生成しても良い。

また、ユーザ嗜好データ36aは、例えば、ユーザ端末に嗜好にあうコンテンツを入力する画面を表示し、ユーザによって個別に特定されたコンテンツをコンテンツ検索装置1bが受信して、ユーザの識別子とその受信したコンテンツ識別子とを関連づけることにより、ユーザ嗜好データ36aを生成しても良い。

40

【0156】

重要ユーザデータ記憶部37には、重要ユーザデータ37aが記憶されている。重要ユーザデータ37aは、そのコンテンツを好むユーザの集合のうち、その集合を代表する重要ユーザの識別子が記憶されたデータである。従って、重要ユーザデータ37aは、コンテンツの識別子と、そのコンテンツの重要ユーザの識別子が対応づけられたデータである。

後述する重要ユーザ決定手段17が、重要ユーザを決定するとともに、重要ユーザデータ37aを生成して、重要ユーザデータ記憶部37に記憶する。

重要ユーザは、一つのコンテンツについて一人だけ選ばれても良い。また、重要ユーザ

50

は、下記の式(6)で算出されるスコアが一定値以上のユーザであって、各コンテンツについて、複数のユーザが選ばれても良い。

【0157】

重要ユーザ決定手段17は、各コンテンツ識別子について、ユーザ嗜好データ36aにおいて当該コンテンツに関連づけられたユーザ識別子の集合を作成するとともに、各集合において、ユーザ間に有向辺を作成し、各集合を代表する重要ユーザを決定する。重要ユーザは、例えば、その重要ユーザに関連づけられたコンテンツ識別子によって、他の多数のユーザの特徴を表現できるユーザである。重要ユーザは、入力されたキーワードに関して適格な情報を持つとされるユーザである。理想的には、重要ユーザには、比較的少ない数のコンテンツ識別子が関連づけられており、それらコンテンツ識別子が、他のユーザにも関連づけられていることが好ましい。

10

【0158】

重要ユーザ決定手段17は、ユーザの嗜好に基づいてユーザ間に有向辺を作成し、ユーザのネットワークを形成する。ユーザが共通のコンテンツを有している場合に、有向辺(リンク)が設けられ、共通のコンテンツを有していない場合は、有向辺が設けられない。この有向辺は、所持しているコンテンツが多いユーザがリンク元となり、少ないユーザがリンク先となる。

重要ユーザ決定手段17は、このネットワークに対し、HITS(Hypertext Induced Topic Selection)アルゴリズムを適用し、オーソリティを決定する。本発明の実施の形態において重要ユーザ決定手段17は、このオーソリティとなるユーザを重要ユーザとして出力する。

20

【0159】

図26を参照して、本発明の実施の形態に係る重要ユーザ決定手段17による重要ユーザ決定処理を説明する。

【0160】

まずステップS701において重要ユーザ決定手段17は、ユーザ嗜好データ36aを読み出し、ステップS702において、同一コンテンツを好むユーザの集合を作成する。このユーザの集合は、コンテンツ識別子1つに対して、1つ作成される。さらに、ステップS702において重要ユーザ決定手段17は、それぞれの集合において、ユーザの好むコンテンツに基づき、各ユーザ間に有向辺を作成する。

30

【0161】

ステップS704において重要ユーザ決定手段17は、ステップS703において作成されたユーザ間が有向辺で関連づけられたコンテンツごとの集合を、ユーザのネットワークと定義する。ステップS705において重要ユーザ決定手段17は、ステップS704で定義されたネットワークに対し、HITSアルゴリズムを適用する。ステップS706において重要ユーザ決定手段17は、オーソリティを重要ユーザとして決定し、出力する。

【0162】

距離算出手段15aは、ユーザ嗜好データ36aにおいて重要ユーザのユーザ識別子に関連づけられたコンテンツ識別子を抽出する。さらに距離算出手段15aは、特徴量データ34aから、抽出したコンテンツ識別子に対応する特徴量および特徴推定値を抽出し、その特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する。

40

特徴量データ34aに記憶されている特徴量および特徴推定値に基づいて、各コンテンツデータの類似度を示す距離を算出する際、距離算出手段15aは、ユーザ嗜好データ36aにおいて重要ユーザのユーザ識別子に関連づけられたコンテンツ識別子を考慮して、各コンテンツデータの類似度を示す距離を算出する。

【0163】

クエリコンテンツを好みとするユーザのネットワーク内における重要ユーザは、ネットワーク内の全てのユーザが好みとするコンテンツを適切に表現していると考えられる。従

50

って、距離算出手段 15 は、ユーザ嗜好データ 36 a において重要ユーザの識別子に関連づけられたコンテンツ識別子の集合に対して、重要ユーザでないユーザの嗜好データを考慮することなく、主成分分析や正準相関分析を施しても良い。さらに距離算出手段 15 は、主成分分析や正準相関分析した後、市街地距離やマハラノビス距離に基づいて距離を算出し、コンテンツ間の類似度を算出しても良い。特定の重要ユーザに関連づけられたコンテンツのみを用いて類似度を算出することにより、少ない計算量で、精度の高い類似度を算出することができる。

【0164】

表示手段 16 a は、距離算出手段 19 によって算出された距離に基づいて、コンテンツデータに対応するサムネイルの表示位置を決定して、例えば図 27 に示すような画面を、表示装置 105 に表示する。さらに表示手段 16 a は、ユーザ嗜好データ 36 a において、所定のコンテンツに関連づけられた複数のユーザの識別子を抽出し、複数のユーザの識別子のそれぞれをノードと、ユーザ間の有向辺を考慮したノード間のリンクを含むユーザネットワークを表示して、表示装置 105 に表示する。

10

【0165】

このユーザネットワークが表示装置 105 に表示されると、ノードが選択されると、表示手段 16 a は、ユーザ嗜好データ 36 a において、そのノードに対応するユーザの識別子に対応づけられたコンテンツの識別子を抽出し、そのコンテンツの識別子の一覧を表示手段 16 a に表示しても良い。例えば、ユーザは、自分の類似度の高いユーザ、すなわちユーザネットワークで近くに表示され、太いリンクで接続されたユーザのノードを選択して、そのユーザが興味を示しているコンテンツの一覧を参照することができる。これによりユーザは、自身と嗜好の近いユーザのコンテンツリストの中から、自分の嗜好に合ったコンテンツを把握することができる。

20

【0166】

ここで、ユーザネットワークのリンクの距離は、特徴量の類似度によって決定される。また、ユーザネットワークのリンクの太さは、ユーザ嗜好データ 36 a において、これらのユーザの識別子に関連づけられたコンテンツの識別子のうち、共通するコンテンツの識別子の数に対応する。例えば、ユーザの嗜好コンテンツの特徴量の類似度が高く、また共通するコンテンツを多く所有している場合、ノード間のリンクの長さは短くなり、そのリンクは太くなる。

30

【0167】

このような本発明の第 2 の変形例に係るコンテンツ検索装置 1 b は、同一のコンテンツを好むユーザの集合を、コンテンツ毎に作成する。そして、各ユーザが好むコンテンツの包含関係に基づき、各ユーザ間に有向辺を作成する。これにより得られる有向グラフに対してリンク解析等を適用することで、それぞれのネットワークにおける重要ユーザを抽出することができる。

さらに、その重要ユーザの嗜好にあったコンテンツリストに基づいて、コンテンツの類似度を算出することにより、類似度の精度を維持しつつ、計算量を削減することができる。

また、このユーザネットワークを表示し、そのユーザの嗜好にあったコンテンツリストを取得することで、ユーザは、新たなコンテンツを知ることができる。

40

【0168】

(詳細なアルゴリズム)

まず、重要ユーザ決定手段 17 による、HITS を用いたユーザの好みに基づくネットワークの作成処理を説明する。

【0169】

本発明の第 2 の変形例において、各ユーザが同一のコンテンツを好む場合に、各々のユーザは同一のネットワークに属するものとする。また、各々のコンテンツに対応するネットワークは、各ユーザ間のリンク関係を要素とする次式で表される行列 L_k ($k \in \{1, 2, \dots, N\}$; N はコンテンツの総数) を用いて定義される。ここで、行列 L_k の i

50

行 j 列の要素は $L_k(i, j)$ で表される。ユーザ i の好むコンテンツ総数は、 F_i で表される。ユーザ i とユーザ j が好む共通のコンテンツ数は、 $C_{i, j}$ で表される。

【数 7 3】

$$L_k(i, j) = \begin{cases} \frac{C_{i,j}}{F_i} & (F_i \geq F_j)(i \neq j) \\ 0 & (F_i < F_j) \end{cases}$$

... 式 (5)

10

【0170】

以上により、本発明の第 2 の変形例は、得られる行列 L_k に対してリンク解析等の処理を施すことにより、ユーザが好みとするコンテンツを考慮したネットワーク内における重要ユーザ（以降、オーソリティ（authority）とも呼ぶ）を抽出することができる。したがって、抽出されるオーソリティが好みとするコンテンツに注目し検索することにより、個々のユーザの嗜好を考慮した類似コンテンツを提示することができる。

以降では、ユーザの嗜好を考慮した類似コンテンツの検索を実現するため、オーソリティを抽出する具体的な方法の一例として HITS アルゴリズムを用いた方法について説明する。

20

【0171】

HITS は、入力されたキーワードに関する確かな情報をもつとされるユーザ（オーソリティ）と、これらのオーソリティに複数のリンクをもつユーザ（ハブ）を、リンク解析により求める手法である。このようなオーソリティとハブは、「良いオーソリティは複数のハブからリンクを張られ、また良いハブは複数のオーソリティへリンクを張る」と再帰的に定義される。この定義に基づいてリンク解析によりスコア付けを行うことでオーソリティおよびハブが求められる。

ここで、本発明においてリンクは、ユーザ嗜好データにおいて各ユーザに関連づけられた好みのコンテンツの特徴量によって算出される。例えば、画像データについてユーザが青の画像が多い場合、同じく青の画像が多いユーザと、太いリンクで結ばれることになる。

30

以下において、HITS の処理手順が示される。

【0172】

手順 (1)

全てのユーザをベース集合とし、ベース集合に属するユーザのリンクパターンを全て調べ、隣接行列 $L = [l_{i,j}]$ が作成される。ここで隣接行列は、ユーザ i からユーザ j へのリンクが存在する場合には $l_{i,j} = 1$ とし、それ以外は 0 とした行列である。ただし、HITS では、同一ホスト l に属するユーザ間のリンクの存在は無視して、隣接行列が作成される。

40

【0173】

手順 (2)

ユーザ i に対し、オーソリティスコア a_i と、ハブスコア h_i の 2 種類のスコアが、式 (6) および式 (7) のようにそれぞれ定義される。さらに、式 (6)、式 (7) により両スコアが算出される。

【数 7 4】

$$a_i = \alpha \sum_{j=1}^n l_{ji} h_j$$

10

・・・ 式 (6)

【数 7 5】

$$h_i = \beta \sum_{j=1}^n l_{ij} a_j$$

20

・・・ 式 (7)

ここで、 n は、ベース集合に属するユーザの総数を表す。 α および β は、正規化定数を表す。「式 (6) で得られる a_i を、式 (7) の a_j に代入」する処理と、「式 (7) で得られる h_i を式 (6) の h_j に代入」する処理とが、 a_i と h_i が変化しなくなるまで繰り返されて更新される。この更新処理が、ベース集合に含まれる全てのユーザに対して行われる。

30

【0174】

手順 (3)

オーソリティスコアが大きいユーザ、ハブスコアが大きいユーザをそれぞれオーソリティ、ハブとする。この手順 (3) を、各ユーザのオーソリティスコアとハブスコアが収束するまで反復する。

【0175】

式 (6) では、ユーザ i にリンクを張るユーザのハブスコアの総和から、ユーザ i のオーソリティスコアを算出する。また式 (7) では、ユーザ i からリンクを張られているユーザのオーソリティスコアの総和からユーザ i のハブスコアを算出している。この算出処理を収束するまで繰り返すことにより、重要ユーザ決定手段 17 は、上で述べた再帰的な定義に適合するオーソリティとハブを得る。

40

【0176】

次に、重要ユーザ決定手段 17 によるオーソリティに基づいた、距離算出手段 15 a の距離算出処理を説明する。

【0177】

本発明の第 2 の変形例では、本発明の実施の形態で説明した距離算出法を用いることで、上述した HITS により決定されるオーソリティが好みとするコンテンツの特徴を考慮して、距離を算出する。距離の算出方法について以下で、市街地距離に基づく距離算出と、マハラノビス汎距離に基づく距離算出とについて説明する。

【0178】

50

まず、市街地距離に基づく距離算出処理を説明する。

距離算出手段15aは、HITSにより決定されるオーソリティが好みとするコンテンツの集合に対して、本発明の実施の形態の距離算出手段15で説明した主成分分析あるいは正準相関分析を施す。これにより距離算出手段15aは、行列Uおよび W_v 、 W_a 、 W_w 、 $v_a = (a_v)$ 、 $a_w = (w_a)$ 、 $v_w = (w_v)$ を求め、ベクトル x^{pca} およびベクトル x^{cca} を得る。さらに距離算出手段15aは、本発明の実施の形態の距離算出手段15で説明した算出方法に従い、 x^{pca} およびベクトル x^{cca} を用いて距離を算出する。

【0179】

次に、マハラノビス汎距離に基づく距離算出処理を説明する。

距離算出手段15aは、HITSにより決定されるクエリコンテンツqに対するオーソリティが好みとする全コンテンツを用いて、重み行列を算出する。ここで、オーソリティが好みとする全コンテンツとは、ユーザ嗜好データ36aにおいて、オーソリティであるユーザの識別子に関連づけられたコンテンツ識別子に対応するコンテンツである。重み行列は、下記の式によって算出される。

【数76】

$$W_q^v = \frac{1}{N_{auth}} X_v X_v'$$

$$W_q^a = \frac{1}{N_{auth}} X_a X_a'$$

$$W_q^w = \frac{1}{N_{auth}} X_w X_w'$$

ただし、 $X_v = [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_{N_{auth}}]'$,

$X_a = [\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_{N_{auth}}]'$, $X_w = [\tilde{w}_1, \tilde{w}_2, \dots, \tilde{w}_{N_{auth}}]'$

N_{auth} は authority が好みとする全コンテンツ数

【0180】

距離算出手段15aは、以上の重み行列を用いて、本発明の実施の形態の距離算出手段15で説明した算出方法に従い、距離を算出する。

【0181】

ここで、クエリコンテンツを好みとするユーザのネットワーク内におけるオーソリティはネットワーク内の全てのユーザが好みとするコンテンツを適切に表現している。したがって、距離算出手段15aは、以上のように距離を算出することで、オーソリティが好みとするコンテンツを考慮した距離に対する重み付けが可能となる。これにより、コンテンツ検索装置1bは、ユーザの嗜好を反映した類似コンテンツの検索または分類を実現することができる。

10

20

30

40

50

【 0 1 8 2 】

次に、表示手段 1 6 a が表示するユーザネットワークについて詳述する。

本発明の第 2 の変形例において、動画像信号等のコンテンツの再生中にユーザネットワーク表示が、ユーザによって操作されることにより、再生中のコンテンツを好みとするユーザのネットワークが構築され、表示手段 1 6 a は、その構築結果を表示装置 1 0 5 に表示する。ユーザネットワークは、多次元尺度構成法によって算出された 3 次元空間の座標位置関係に基づいて表示される。ユーザネットワークにおいて、ネットワーク内に存在する各ユーザはノードして表示される。ネットワーク内の任意のユーザが選択されると、表示手段 1 6 a は、そのユーザが好みとするコンテンツの一覧を表示する。

【 0 1 8 3 】

ここで表示手段 1 6 a は、選択されたユーザと、ネットワーク内におけるオーソリティとなるユーザと、ハブとなるユーザを、それぞれ異なる色や模様で表示する。また、ネットワークにおけるユーザ同士のリンク関係を、線の太さとボールの流れで表現する。本発明の第 2 の変形例では、リンク元からリンク先にボールが流れるように表示をすることにより、リンクの方向を表示する。このとき、式 (5) で算出されるネットワークを表現する行列の要素の値に従って、その値が大きいほど線を太く、そのリンク方向をボールの流れで表現する。

以上により可視化されるユーザネットワークは、各ユーザの嗜好を反映したユーザ同士のリンク関係やその度合いを視覚的に表現することができる。さらに、ユーザが好みとするコンテンツをクエリとして検索操作が入力されると、表示手段 1 6 a は、表示される検索結果をクエリを中心として多次元尺度構成法を用いて算出された 2 次元空間の座標に基づいて、座標を再計算し配置して表示する。これにより、ネットワーク内のユーザの嗜好を反映した検索結果が表示され、表示されている複数のコンテンツの位置関係によって、ユーザに、コンテンツ同士が類似する度合いを視覚的に把握させることができる。

(その他の実施の形態)

上記のように、本発明の実施の形態、第 1 の変形例および第 2 の変形例によって記載したが、この開示の一部をなす論述及び図面はこの発明を限定するものであると理解すべきではない。この開示から当業者には様々な代替実施の形態、実施例及び運用技術が明らかとなる。

例えば、本発明の実施の形態に記載したコンテンツ検索装置は、図 1 に示すように一つのハードウェア上に構成されても良いし、その機能や処理数に応じて複数のハードウェア上に構成されても良い。又、既存の情報システム上に実現されても良い。

【 0 1 8 4 】

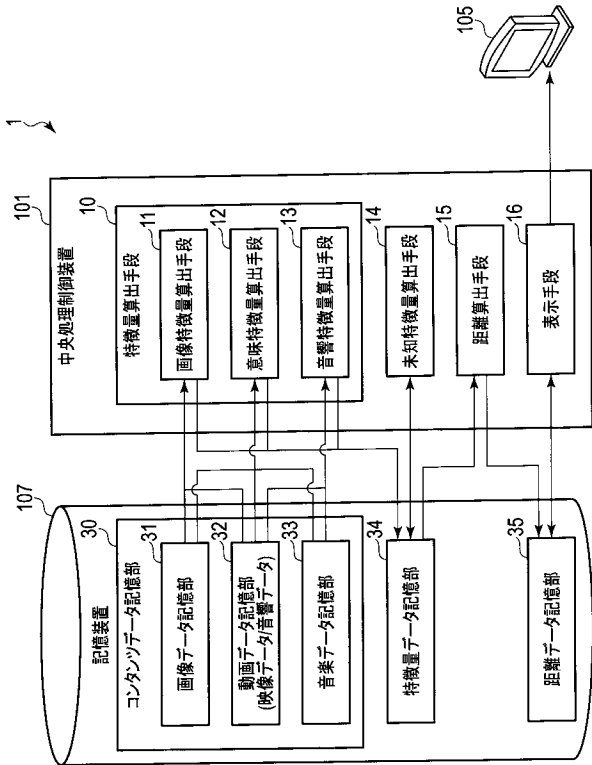
本発明はここでは記載していない様々な実施の形態等を含むことは勿論である。従って、本発明の技術的範囲は上記の説明から妥当な特許請求の範囲に係る発明特定事項によってのみ定められるものである。

10

20

30

【 図 1 】



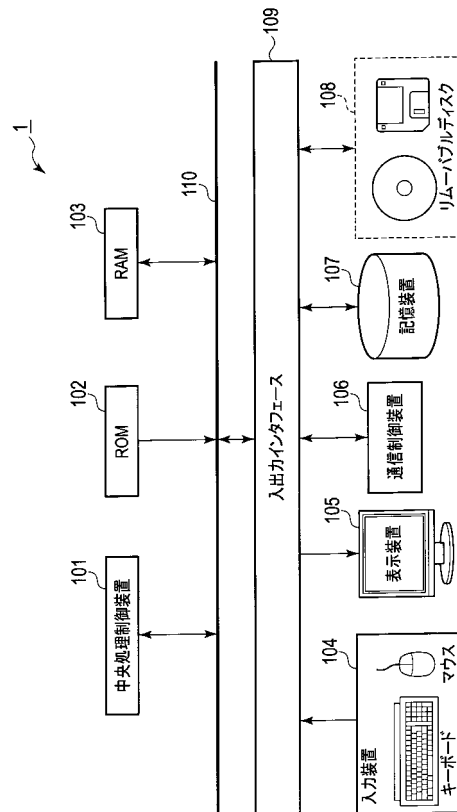
【 図 2 】

	画像データ	音楽データ	動画データ
画像特徴量	あり	なし	あり
暗黙特徴量	なし	あり	あり
意味特徴量	(あり)	(あり)	(あり)

【 図 3 】

	画像データ	音楽データ	動画データ
画像特徴量	あり	あり	あり
暗黙特徴量	あり	あり	あり
意味特徴量	あり	あり	あり

【 図 4 】



【 図 5 】

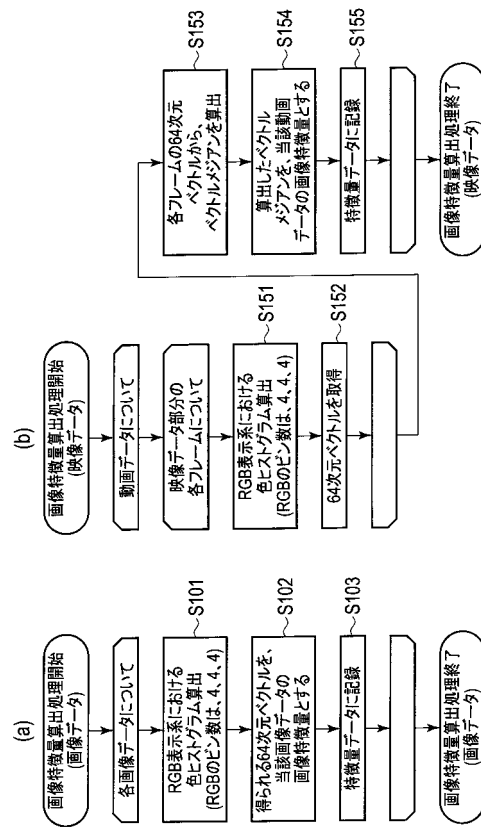
(a)

コンテンツ識別子	コンテンツ種別	画像特徴量	音響特徴量	意味特徴量
xx	画像	xx	-	xx
xx	音楽	-	xx	xx
xx	動画	xx	xx	xx
xx	音楽	-	xx	xx
...

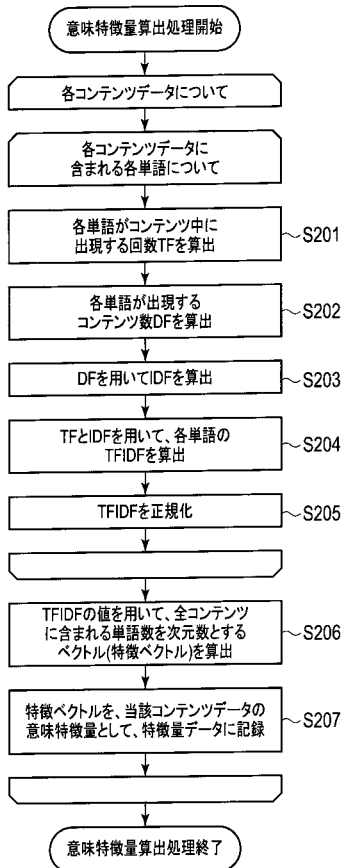
(b)

コンテンツ識別子	コンテンツ種別	画像特徴量	音響特徴量	意味特徴量
xx	画像	xx	xx	xx
xx	音楽	xx	xx	xx
xx	動画	xx	xx	xx
xx	音楽	xx	xx	xx
...

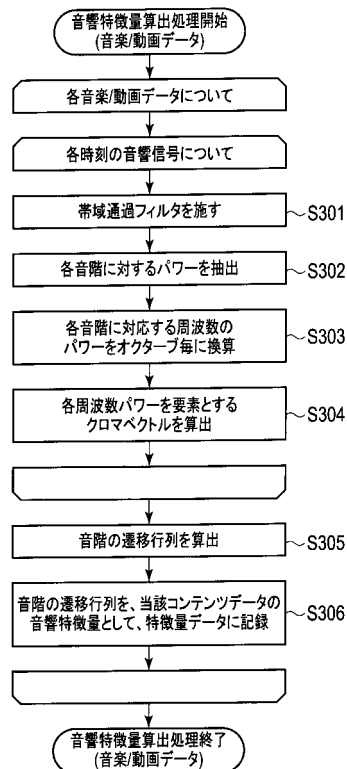
【 図 6 】



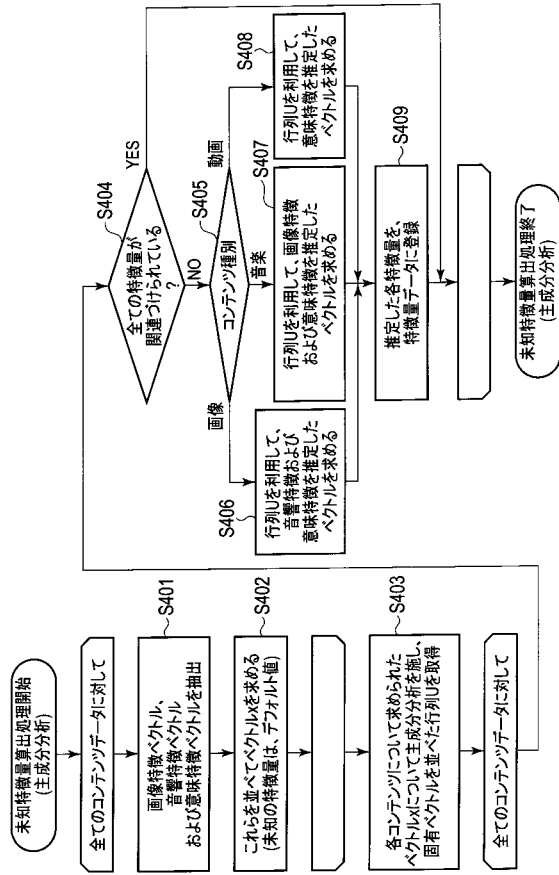
【 図 7 】



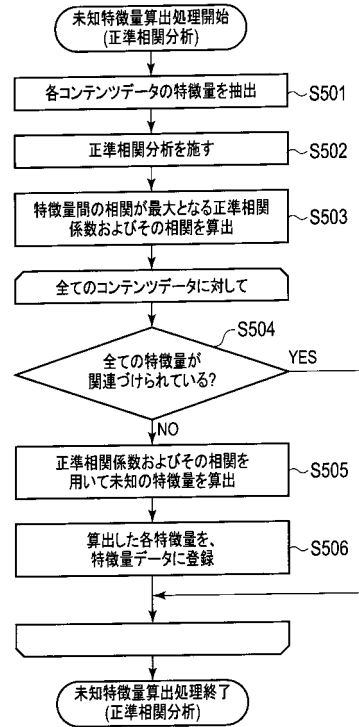
【 図 8 】



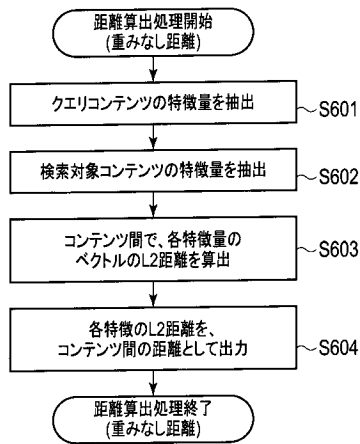
【 図 9 】



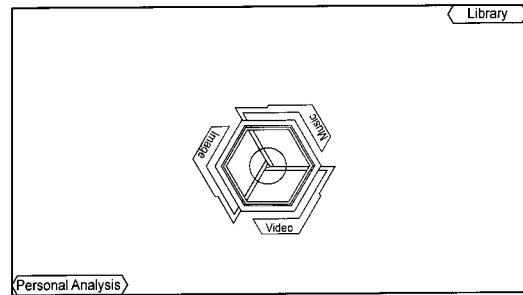
【 図 10 】



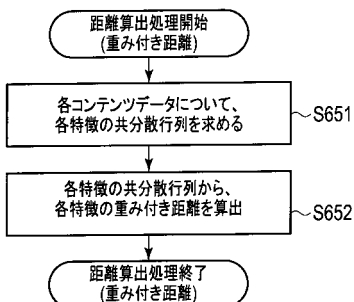
【 図 11 】



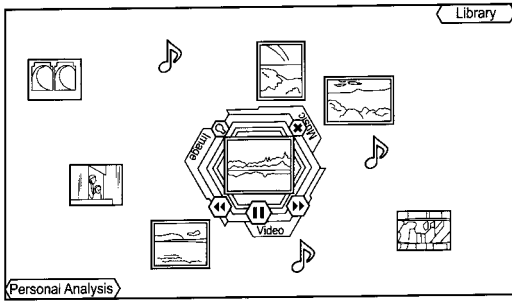
【 図 14 】



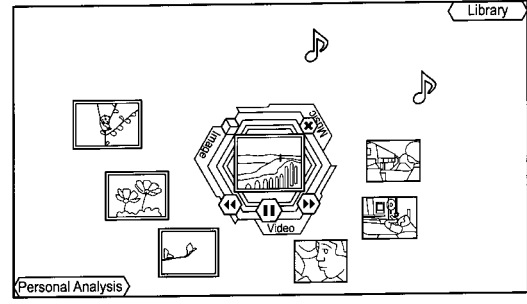
【 図 12 】



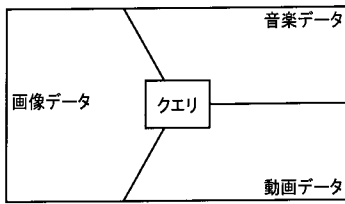
【 図 1 5 】



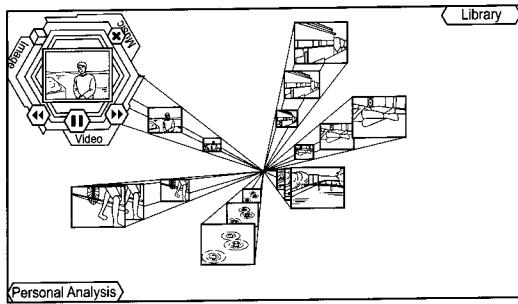
【 図 1 7 】



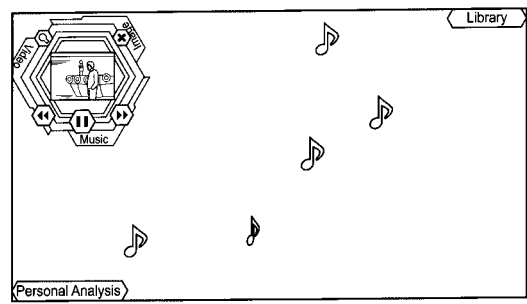
【 図 1 6 】



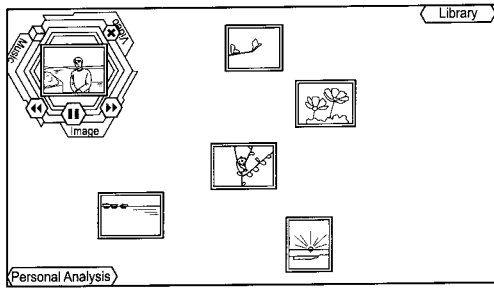
【 図 1 9 】



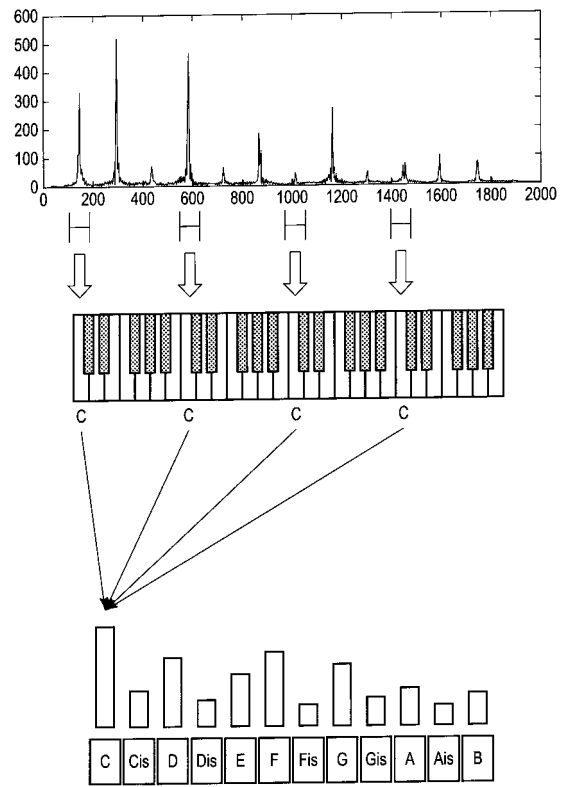
【 図 2 0 】



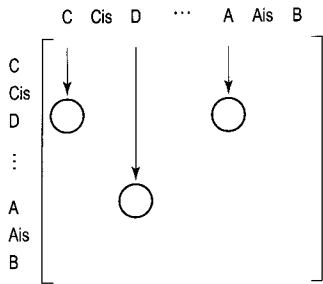
【 図 2 1 】



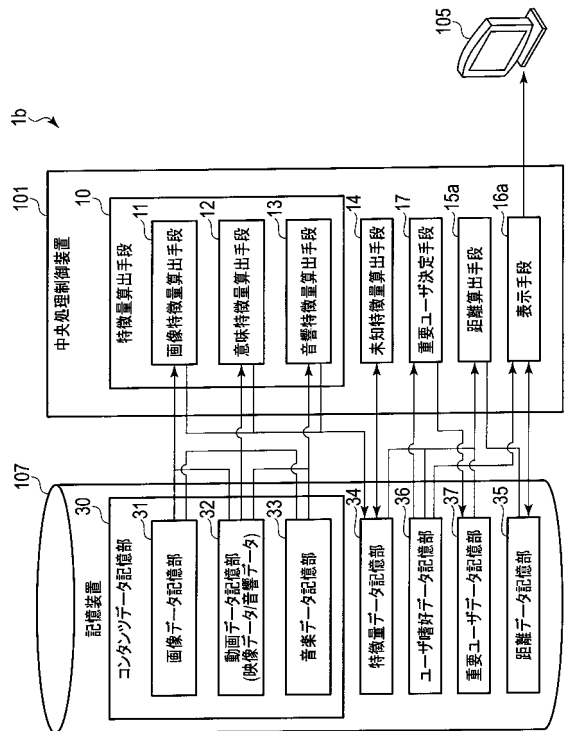
【 図 2 2 】



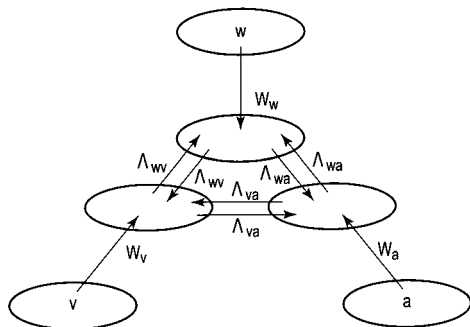
【 図 2 3 】



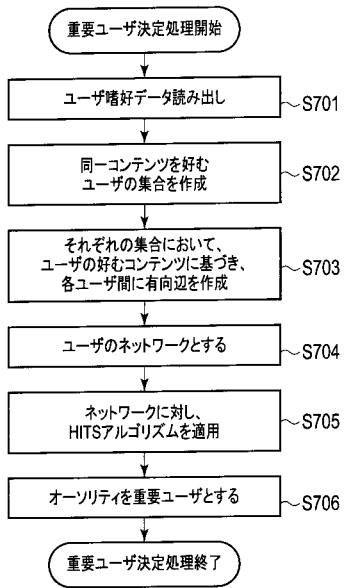
【 図 2 5 】



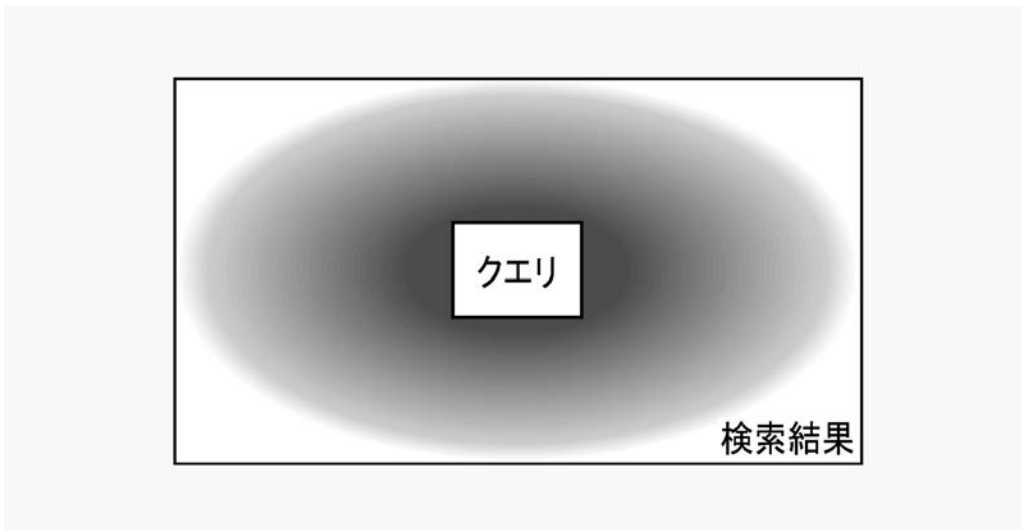
【 図 2 4 】



【 図 2 6 】



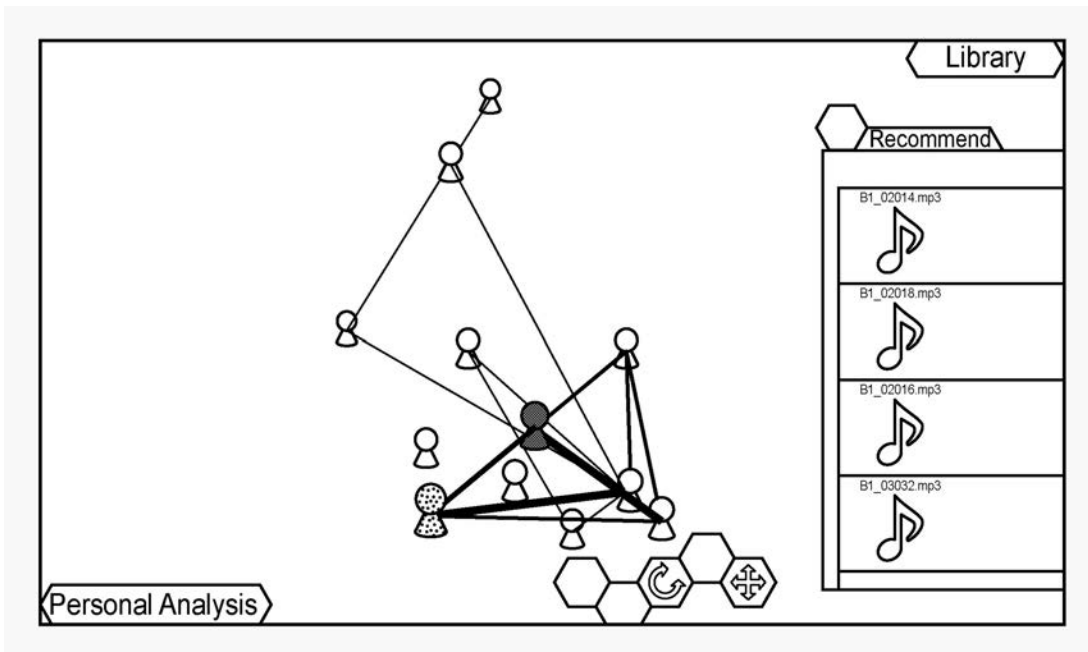
【 図 1 3 】



【 図 1 8 】



【 図 2 7 】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP2009/069005
A. CLASSIFICATION OF SUBJECT MATTER G06F17/30(2006.01)i, G06T1/00(2006.01)i, G10H1/18(2006.01)i, G10L11/00(2006.01)i, H04N5/76(2006.01)i 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) G06F17/30, G06T1/00, G10H1/18, G10L11/00, H04N5/76 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2009 Kokai Jitsuyo Shinan Koho 1971-2009 Toroku Jitsuyo Shinan Koho 1994-2009 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2006-344155 A (National University Corporation Hokkaido University), 21 December 2006 (21.12.2006), entire text; all drawings (Family: none)	1-12
A	Daigo MICHİYAMA, "Eizo Kensaku no Tameno Gazo Oyobi Onkyo Shingo ni yoru Scene-kan no Soido Teigi ni Kansuru Kento", ITE Technical Report, 21 February 2008 (21.02.2008), vol.32, no.9, pages 55 to 59	1-12
A	Kazuya KOBAYASHI, "Melody ni Chakugan shita Ruijido ni Motozuku Ruiji Gakkyoku Chushutsu no Seido ni Kansuru Ichikosatsu", ITE Technical Report, 18 February 2008 (18.02.2008), vol.32, no.9, pages 49 to 53	1-12
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> 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 application or patent 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 27 November, 2009 (27.11.09)		Date of mailing of the international search report 08 December, 2009 (08.12.09)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer Telephone No.
Facsimile No.		

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2009/069005

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Teruhisa HOCHIN, "Kansei ni Motozuku Cross Media Kensaku System ni Tsuite", IPSJ SIG Notes, 17 July 2002 (17.07.2002), vol.2002, no.67, pages 523 to 528	1-12
A	JP 2007-193222 A (Casio Computer Co., Ltd.), 02 August 2007 (02.08.2007), entire text; all drawings (Family: none)	1-12
A	WO 2007/066450 A1 (NEC Corp.), 14 June 2007 (14.06.2007), entire text; all drawings & US 2009/0132074 A	1-12

国際調査報告		国際出願番号 PCT/JP2009/069005									
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. G06F17/30(2006.01)i, G06T1/00(2006.01)i, G10H1/18(2006.01)i, G10L11/00(2006.01)i, H04N5/76(2006.01)i											
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. G06F17/30, G06T1/00, G10H1/18, G10L11/00, H04N5/76											
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 <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2009年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2009年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2009年</td> </tr> </table>				日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2009年	日本国実用新案登録公報	1996-2009年	日本国登録実用新案公報	1994-2009年
日本国実用新案公報	1922-1996年										
日本国公開実用新案公報	1971-2009年										
日本国実用新案登録公報	1996-2009年										
日本国登録実用新案公報	1994-2009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)											
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 請求項の番号									
A	JP 2006-344155 A (国立大学法人北海道大学) 2006.12.21, 全文、全図 (ファミリーなし)	1-12									
A	道山 大悟, 映像検索のための画像及び音響信号によるシーン間の相違度定義に関する検討, 映像情報メディア学会技術報告, 2008.02.21, Vol. 32 No. 9, 第55頁乃至第59頁	1-12									
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。		<input type="checkbox"/> 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		の日の後に公表された文献									
「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		「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									
「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		「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									
「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)		「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									
「O」口頭による開示、使用、展示等に言及する文献		「&」同一パテントファミリー文献									
「P」国際出願日前で、かつ優先権の主張の基礎となる出願											
国際調査を完了した日 27.11.2009		国際調査報告の発送日 08.12.2009									
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) 紀田 馨	5M 3042								
		電話番号 03-3581-1101 内線 3599									

国際調査報告		国際出願番号 PCT/JP2009/069005
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 請求項の番号
A	小林 和也, メロディーに着眼した類似度に基づく類似楽曲抽出の精度に関する一考察, 映像情報メディア学会技術報告, 2008.02.18, Vol. 32 No. 9, 第49頁乃至第53頁	1-12
A	宝珍 輝尚, 感性に基づくクロスメディア検索システムについて, 情報処理学会研究報告, 2002.07.17, Vol. 2002 No. 67, 第523頁乃至第528頁	1-12
A	JP 2007-193222 A (カシオ株式会社) 2007.08.02, 全文、全図 (ファミリーなし)	1-12
A	WO 2007/066450 A1 (日本電気株式会社) 2007.06.14, 全文、全図 & US 2009/0132074 A	1-12

フロントページの続き

(51) Int.Cl.	F I	テーマコード (参考)
	G 1 0 L 11/00	4 0 2 K
	G 0 6 F 3/048	6 5 2 Z

(81) 指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(出願人による申告) 平成 20 年度、総務省、研究開発委託事業、産業技術力強化法 19 条の適用を受ける特許出願

(注) この公表は、国際事務局 (W I P O) により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願 (日本語実用新案登録出願) の国際公開の効果は、特許法第 184 条の 10 第 1 項 (実用新案法第 48 条の 13 第 2 項) により生ずるものであり、本掲載とは関係ありません。