

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6783443号  
(P6783443)

(45) 発行日 令和2年11月11日(2020.11.11)

(24) 登録日 令和2年10月26日(2020.10.26)

(51) Int.Cl. F I  
G O 6 F 16/906 (2019.01) G O 6 F 16/906

請求項の数 23 (全 33 頁)

<p>(21) 出願番号 特願2016-46084 (P2016-46084)                  (22) 出願日 平成28年3月9日(2016.3.9)                  (65) 公開番号 特開2016-197406 (P2016-197406A)                  (43) 公開日 平成28年11月24日(2016.11.24)                          審査請求日 平成31年1月24日(2019.1.24)                  (31) 優先権主張番号 特願2015-78034 (P2015-78034)                  (32) 優先日 平成27年4月6日(2015.4.6)                  (33) 優先権主張国・地域又は機関                          日本国(JP)</p> <p>(出願人による申告)平成25年度、独立行政法人新エネルギー・産業技術総合開発機構委託研究「センシング技術を用いた情報提供システムの研究開発」、産業技術力強化法第19条の適用を受ける特許出願</p>	<p>(73) 特許権者 301021533                  国立研究開発法人産業技術総合研究所                  東京都千代田区霞が関1-3-1                  (74) 代理人 100079108                          弁理士 稲葉 良幸                  (74) 代理人 100109346                          弁理士 大貫 敏史                  (74) 代理人 100144325                          弁理士 小澁 高弘                  (72) 発明者 河本 満                  東京都江東区青海2-3-26 国立研究                  開発法人産業技術総合研究所臨海副都心セ                  ンター内</p>
---	--

最終頁に続く

(54) 【発明の名称】 情報処理装置、情報処理システム、情報処理方法、プログラム、及び記録媒体

(57) 【特許請求の範囲】

【請求項1】

環境音を表す複数の音データを取得する取得部と、  
 各音データの特徴量を抽出する抽出部と、  
 抽出された特徴量を用いて前記複数の音データを各カテゴリに分類する分類部と、  
 前記分類部により分類された各カテゴリに基づく前記音データを、前記音データの識別情報と、前記音データの方向及び/又は前記音データの発生時間とを用いて解析する解析部と、

前記解析部により解析された解析結果を可視化する可視化部と、

を備え、

前記分類部は、

抽出された特徴量を用いて生成された各音データ間の類似度行列に基づく第1評価値及び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成して各カテゴリに分類し、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、情報処理装置。

【請求項2】

前記各カテゴリに基づく前記音データの第1特徴量、第2特徴量及び第3特徴量を抽出

する音特徴抽出部と、

前記第 1 特徴量を、第 1 色により表現し、前記第 2 特徴量を、第 2 色により表現し、かつ、前記第 3 特徴量を、第 3 色により表現する表現部と、をさらに備える、請求項 1 に記載の情報処理装置。

【請求項 3】

前記音データに対して、所定領域内の音源位置を推定する推定部をさらに備え、

前記表現部は、

前記音源位置の音データを、前記第 1 色、前記第 2 色、及び前記第 3 色で表現する、請求項 2 に記載の情報処理装置。

【請求項 4】

前記音特徴抽出部は、

前記カテゴリの中心周波数に基づく前記音データの高低を示す前記第 1 特徴量と、前記音データの大小を示す前記第 2 特徴量と、前記音データの継続性を示す前記第 3 特徴量とを抽出する、請求項 2 に記載の情報処理装置。

【請求項 5】

前記分類部は、

前記類似度行列の各列ベクトルにおける最大値の標準偏差に基づいて、該列ベクトル内の類似度を選択し、選択した類似度を前記類似度行列の自己相関の初期値に設定する、請求項 1 に記載の情報処理装置。

【請求項 6】

前記分類部は、

前記第 1 評価値及び前記第 2 評価値、並びに前記自己相関の値それぞれの更新を繰り返すことにより、前記最終的なクラスタを生成する、請求項 1 乃至 5 いずれか一項に記載の情報処理装置。

【請求項 7】

前記分類部は、

前記特徴量を用いて前記類似度行列を生成する生成部と、

前記第 1 評価値及び前記第 2 評価値を算出する算出部と、

前記第 1 評価値及び前記第 2 評価値の算出過程において生成されたクラスタ内の外れデータを求め、該外れデータの前記自己相関の値を、他の値に更新する更新部と、

更新された値を用いて前記算出部により新たに算出された第 1 評価値及び第 2 評価値に基づく値が収束するか否かを判定する判定部と、

を含む、請求項 1 乃至 6 いずれか一項に記載の情報処理装置。

【請求項 8】

前記更新部は、

前記外れデータに関する前記類似度行列の列ベクトル内の類似度に基づく値を、前記他の値に設定する、請求項 7 に記載の情報処理装置。

【請求項 9】

前記更新部は、

前記列ベクトル内の類似度を所定順に並べ替え、所定番目の類似度を、前記他の値に設定する、請求項 8 に記載の情報処理装置。

【請求項 10】

前記音データは、マイクにより集音された音データであり、

前記特徴量は、周波数領域におけるパワースペクトルに関する特徴量である、請求項 1 乃至 9 いずれか一項に記載の情報処理装置。

【請求項 11】

前記特徴量は、前記パワースペクトルに非負値行列因子分解を適用して得られた基底ベクトル、前記パワースペクトルの次元を揃えた分散行列、前記パワースペクトルにメルフィルタバンク分析を行って得られたベクトル、及び前記非負値行列因子分割により得られた係数ベクトルと前記基底ベクトルを用いて再構成されたパワースペクトルの第 1 主成分

10

20

30

40

50

ベクトルのうちの少なくとも1つを含む、請求項10に記載の情報処理装置。

【請求項12】

前記分類部は、

前記類似度行列に含まれる類似度を、前記特徴量に含まれる各ベクトル又は前記分散行列の類似度に重み付けを行うことで算出する、請求項11に記載の情報処理装置。

【請求項13】

種類が異なる各音データと、該種類を示す識別情報とを関連付けて記憶する記憶部と、前記分類部により生成されたクラスタ中心の音データにマッチングする、前記記憶部に記憶された音データを検索する検索部と、

検索された音データに関連付けられた識別情報を含むラベルを、前記クラスタ中心の音データに付与する付与部と、

をさらに備える、請求項1乃至12のいずれか一項に記載の情報処理装置。

10

【請求項14】

前記解析部は、

前記ラベルに含まれる識別情報を用いて前記複数の音データを解析する、請求項13に記載の情報処理装置。

【請求項15】

前記取得部により取得された音データは、マイクロフォンアレイにより取得された音データであり、

前記ラベルは、前記音データの方向及び/又は前記音データの発生時間をさらに含む、請求項14に記載の情報処理装置。

20

【請求項16】

前記解析部は、

前記複数の音データにおける各識別情報の割合を算出し、該割合を解析結果にする、請求項15に記載の情報処理装置。

【請求項17】

前記可視化部は、

前記音データの方向及び/又は前記音データの発生時間を用いて、前記解析結果をグラフ化する、請求項16に記載の情報処理装置。

【請求項18】

30

前記解析部は、

前記ラベル付けされた音データと、前記特徴量とを用いてアンサンブル学習を行う学習部と、

前記学習部による学習により、前記音データに関するモデルを構築する構築部と、を含む、請求項15に記載の情報処理装置。

【請求項19】

前記可視化部は、

前記モデルを用いて解析された解析結果が表示部において可視化されるように処理する、請求項18に記載の情報処理装置。

【請求項20】

40

環境音を表す複数の音データを取得する取得部と、

各音データの特徴量を抽出する抽出部と、

抽出された特徴量を用いて前記複数の音データを各カテゴリに分類する分類部と、

前記分類部により分類された各カテゴリに基づく前記音データを、前記音データの識別情報と、前記音データの方向及び/又は前記音データの発生時間とを用いて解析する解析部と、

前記解析部により解析された解析結果を可視化する可視化部と、

を備え、

前記分類部は、

抽出された特徴量を用いて生成された各音データ間の類似度行列に基づく第1評価値及

50

び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成して各カテゴリに分類し、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、情報処理システム。

【請求項21】

コンピュータが、  
 環境音を表す複数の音データを取得することと、  
 各音データの特徴量を抽出することと、  
 抽出された特徴量を用いて前記複数の音データを各カテゴリに分類することと、  
 分類された各カテゴリに基づく前記音データを、前記音データの識別情報と、前記音データの方向及びノ又は前記音データの発生時間とを用いて解析することと、  
 解析された解析結果を可視化することと、  
 を実行し、  
 前記分類することは、

抽出された特徴量を用いて生成された各音データ間の類似度行列に基づく第1評価値及び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成して各カテゴリに分類し、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、情報処理方法。

【請求項22】

コンピュータに、  
 環境音を表す複数の音データを取得することと、  
 各音データの特徴量を抽出することと、  
 抽出された特徴量を用いて前記複数の音データを各カテゴリに分類することと、  
 分類された各カテゴリに基づく前記音データを、前記音データの識別情報と、前記音データの方向及びノ又は前記音データの発生時間とを用いて解析することと、  
 解析された解析結果を可視化することと、  
 を実行させ、  
 前記分類することは、

抽出された特徴量を用いて生成された各音データ間の類似度行列に基づく第1評価値及び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成して各カテゴリに分類し、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、プログラム。

【請求項23】

コンピュータに、  
 環境音を表す複数の音データを取得することと、  
 各音データの特徴量を抽出することと、  
 抽出された特徴量を用いて前記複数の音データを各カテゴリに分類することと、  
 分類された各カテゴリに基づく前記音データを、前記音データの識別情報と、前記音データの方向及びノ又は前記音データの発生時間とを用いて解析することと、  
 解析された解析結果を可視化することと、  
 を実行させ、  
 前記分類することは、

抽出された特徴量を用いて生成された各音データ間の類似度行列に基づく第1評価値及び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成して各カテゴリに分類し、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、プログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、情報処理装置、情報処理システム、情報処理方法、プログラム、及び記録媒体に関する。

【背景技術】

【0002】

従来、多数のデータを特徴的なカテゴリに分類するクラスタリング手法が知られている。最近のクラスタリング手法として、Affinity Propagation (AP) 法が知られている（例えば非特許文献1、2参照）。

【0003】

AP法では、パラメータ $r(i,k)$ は、クラスタ内のデータポイント $i$ がそのクラスタの中心点(exemplar)候補 $k$ に割り当てられる適切さを示すスコアであり、データポイント $i$ からexemplar候補 $k$ に送られるメッセージであると考えられている。また、パラメータ $a(i,k)$ は、exemplar候補 $k$ のクラスタ内に、データポイント $i$ がメンバーとして属する適切さを示すスコアであり、exemplar候補 $k$ からそのクラスタメンバーとなる見込みのデータポイント $i$ に対して送られるメッセージであると考えられている。つまり、AP法はメッセージ交換型クラスタリング手法である。

20

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】Brendan J. Frey and Delbert Dueck, "Clustering by Passing Messages Between Data Points", University of Toronto Science 315, 972-976, February 2007

30

【非特許文献2】Inmar E. Givoni and Brendan J. Frey, "A Binary Variable Model for Affinity Propagation", Neural Computation, Vol 21, issue 6, pp 1589-1600, June 2009

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、上記の従来技術では、クラスタリングされたクラスタ内に、クラス内の代表値とは異なる外れ値を含む可能性があり、分類精度に改善の余地がある。ここで、データの一例として環境音を用いてAP法が適用されると、クラスタ内の代表音と、同じクラスタ内の音とで、異なる音が同じクラスタに分類されることがある。このように、AP法に関して、分類精度を向上させることが求められている。

40

【0006】

そこで、本発明は、AP法に関して、クラスタへの分類精度を向上させることができる情報処理装置、情報処理システム、情報処理方法、プログラム、及び記録媒体を提供する。

【課題を解決するための手段】

【0007】

本発明の一態様における情報処理装置は、複数のデータを取得する取得部と、各データの特徴量を抽出する抽出部と、

50

抽出された特徴量を用いて生成された各データ間の類似度行列に基づく第1評価値及び第2評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成する分類部であって、前記類似度行列の自己相関は、前記第1評価値及び前記第2評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて設定され、前記第1評価値は、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す値であり、前記第2評価値は、前記メンバーデータが、前記候補データのクラスタに属することの適切さを示す値である、前記分類部と、を備える。

【図面の簡単な説明】

【0008】

【図1】第1実施形態における情報処理システムの一例を示す概念図である。 10

【図2】第1実施形態における情報処理装置のハードウェア構成の一例を示すブロック図である。

【図3】第1実施形態における情報処理装置の機能構成の一例を示すブロック図である。

【図4】環境音を計測した場所を示す図である。

【図5】マイクロフォンアレイの一例を示す図である。

【図6】マイクロフォンの位置を示す図である。

【図7】パワースペクトルの一例を示す図である。

【図8】基底ベクトルの一例を示す図である。

【図9】係数ベクトルの一例を示す図である。

【図10】基底ベクトル及び係数ベクトルを用いて再構築したパワースペクトルの一例を示す図である。 20

【図11】改良AP法を用いて音データをクラスタリングした結果の一例を示す図である。

【図12】改良AP法の精度結果(その1)の一例を示す図である。

【図13】改良AP法の精度結果(その2)の一例を示す図である。

【図14】一日の環境音の発生割合の一例を示す図である。

【図15】あるイベント日の10時から13時までのイベントAにおける環境音発生割合の一例を示す図である。

【図16】あるイベント日の14時から18時までのイベントAにおける環境音発生割合の一例を示す図である。

【図17】あるイベント日の14時から18時までのイベントBにおける環境音発生割合の一例を示す図である。 30

【図18】あるイベント日の10時から13時までのイベントCにおける環境音発生割合の一例を示す図である。

【図19】あるイベント日の14時から18時までのイベントCにおける環境音発生割合の一例を示す図である。

【図20】第1実施形態におけるデータ解析処理の一例を示すフローチャートである。

【図21】第1実施形態における分類処理の一例を示すフローチャートである。

【図22】第1実施形態におけるタグ付け処理の一例を示すフローチャートである。

【図23】第1実施形態における解析処理の一例を示すフローチャートである。

【図24】第2実施形態における情報処理装置の機能構成の一例を示すブロック図である。 40

【図25】パワースペクトルの一例を示す図である。

【図26】エッジ検出されたパワースペクトルの一例を示す図である。

【図27】分類された各カテゴリの一例を示す図である。

【図28】音の大小度の計算を説明するための図である。

【図29】音の継続度の計算を説明するための図である。

【図30】実験地における音模様を表現する所定領域を示す図である。

【図31】平常時の音模様を示す図である。

【図32】イベント時の音模様を示す図である。

【図33】第2実施形態における音模様表現処理の一例を示すフローチャートである。 50

## 【発明を実施するための形態】

## 【0009】

以下、図面を参照して本発明の実施の形態を説明する。ただし、以下に説明する実施形態は、あくまでも例示であり、以下に明示しない種々の変形や技術の適用を排除する意図はない。即ち、本発明は、その趣旨を逸脱しない範囲で種々変形して実施することができる。また、以下の図面の記載において、同一または類似の部分には同一または類似の符号を付して表している。図面は模式的なものであり、必ずしも実際の寸法や比率等とは一致しない。図面相互間においても互いの寸法の関係や比率が異なる部分が含まれていることがある。

## 【0010】

10

## 〔第1実施形態〕

以下、本発明の第1実施形態における情報処理システムについて説明する。第1実施形態における情報処理システムは、取得された複数のデータに対し、適切な分類を行うことができる。また、情報処理システムは、適切に分類されたクラスタやデータを用いて、そのデータに適したモデルを構築し、データの解析などを行うことができる。

## 【0011】

## &lt;システム概要&gt;

図1は、第1実施形態における情報処理システムの一例を示す概念図である。図1に示すように、例えば情報処理システム1は、データを処理する情報処理装置(サーバ装置)10A、10Bと、データを取得する取得装置20A、20B、20CとがネットワークNを介して接続される。情報処理装置10A、10Bは、処理を分散してもよいし、いずれか一つの装置で処理するようにしてもよい。

20

## 【0012】

各情報処理装置は、個別に区別して説明する場合には符号10A、10Bを用い、個別に区別する必要がなく、まとめて説明する場合には符号10を用いる。なお、取得装置の符号の使い方は、情報処理装置と同様である。

## 【0013】

取得装置20は、分類対象のデータを取得する。例えば、データが音データであれば、取得装置20は、集音可能なマイクやマイクロフォンアレイなどであり、データが画像データであれば、取得装置20は、カメラなどの撮像装置であり、データが文字データであれば、文字データを取得する装置である。

30

## 【0014】

情報処理装置10は、取得装置20から複数のデータを取得し、複数のデータに対して分類処理、所謂クラスタリングを行う。また、情報処理装置10は、クラスタリングの結果を用いて、ラベル付けを行ったり、対象データのモデルを構築したりしてもよい。また、情報処理装置10は、このモデルを用いて対象データの各種解析を行ってもよい。なお、情報処理装置10は、自身の記憶部に記憶されているデータに対してクラスタリングや、モデル構築や、各種解析を行ってもよい。

## 【0015】

ネットワークNは、インターネット等であり、無線LANのアクセスポイントや携帯電話の基地局などを含む。

40

## 【0016】

## &lt;ハードウェア構成&gt;

次に、情報処理装置10のハードウェア構成について説明する。図2は、第1実施形態における情報処理装置10のハードウェア構成の一例を示すブロック図である。

## 【0017】

図2に示すように、情報処理装置10は、制御部102と、通信インタフェース106と、記憶部108と、表示部114と、入力部116と、を有し、各部はバスラインを介して接続される。

## 【0018】

50

制御部 102 は、CPU、ROM、RAM 104 等を含む。制御部 102 は、記憶部 108 に記憶される処理プログラム 110 等を実行することにより、一般的なサーバ装置としての機能に加え、例えば分類機能や、ラベル付け機能や、モデル構築機能などを実現するように構成される。

【0019】

また、RAM 104 は、各種情報を一時的に保持したり、CPU が各種処理を実行する際のワークエリアとして使用されたりする。

【0020】

通信インタフェース 106 は、ネットワーク N を介した取得装置 20 との通信を制御する。

10

【0021】

記憶部 108 は、例えば HDD 等からなり、一般的な情報処理装置としての機能を実現するためのアプリケーション及びデータ（図示省略）を記憶することに加え、処理プログラム 110 を記憶する。また、記憶部 108 は、情報記憶部 112 を有している。

【0022】

処理プログラム 110 は、分類などの処理を行うためのプログラムであり、取得装置 20 から各データを取得し、各データを処理するプログラムである。処理プログラム 110 は、コンピュータに読み取り可能な記録媒体に保存され、この記録媒体から読み出されて、記憶部 108 に記憶されてもよい。

【0023】

情報記憶部 112 は、例えば各取得装置 20 から取得されたデータや、分類に関する情報、ラベルに関する情報、モデルに関する情報などを関連付けて記憶する。

20

【0024】

表示部 114 は、情報処理装置 10 の管理者等に情報を表示する。入力部 116 は、管理者等からの入力を受け付けたり、管理者等からの指示を受け付けたりする。また、情報処理装置 10 は、表示部 114 と入力部 116 とを必ずしも設ける必要はなく、表示部 114 及び入力部 116 は、外部から情報処理装置 10 に接続されるようにしてもよい。

【0025】

< 機能構成 >

図 3 は、第 1 実施形態における情報処理装置 10 の機能構成の一例を示すブロック図である。図 3 に示す情報処理装置 10 は、送信部 202、受信部 204、処理制御部 206、及び記憶部 240 を含む。

30

【0026】

送信部 202 は、例えば通信インタフェース 106 等により実現されうる。送信部 202 は、分類結果や、ラベル付け結果や、構築したモデルなどを、ネットワークを介して接続された各装置に送信してもよい。

【0027】

受信部 204 は、例えば通信インタフェース 106 等により実現されうる。受信部 204 は、例えば、各取得装置 20 から複数のデータを受信する。受信されたデータは、例えばマイクなどにより集音された環境音の音データとする。処理制御部 206 に出力される。

40

【0028】

記憶部 240 は、分類に用いるデータや、タグ付けなどに用いるデータや、モデル構築に用いるデータなどの各種データを記憶する。例えば、記憶部 240 は、ラベル付けのための特徴的なデータや、ラベル付けする識別情報などを記憶する。

【0029】

処理制御部 206 は、例えば制御部 102 等により実現されうる。処理制御部 206 は、受信部 204 により受信された複数のデータ、例えば音データに対して、分類処理や、ラベル付け処理や、モデル構築処理などを実行する。

【0030】

50

処理制御部 206 は、上述した処理を実行するため、取得部 210、抽出部 212、分類部 214、検索部 224、付与部 226、解析部 228、可視化部 234 を備える。

【0031】

取得部 210 は、受信部 204 から、複数のデータを取得する。取得部 210 は、受信部 204 以外にも自装置内の記憶部などから分類等の対象データを取得してもよい。

【0032】

抽出部 212 は、各データの特徴量を抽出する。データが音データである場合、抽出部 212 は、音データを周波数領域に変換し、この周波数領域におけるパワースペクトルに関する特徴量を 1 つ又は複数抽出する。抽出部 212 は、データに適した特徴量を算出すればよい。

10

【0033】

より具体的には、抽出部 212 は、特徴量として、パワースペクトルに非負値行列因子分解を適用して得られた基底ベクトル、パワースペクトルの次元を揃えた分散行列、パワースペクトルにメルフィルタバンク分析を行って得られたベクトル、及び非負値行列因子分割により得られた係数ベクトルと基底ベクトルを用いて再構成されたパワースペクトルの第 1 主成分ベクトルのうちの少なくとも 1 つを含む。抽出された各データの特徴量は、分類部 214 に出力される。

【0034】

分類部 214 は、例えば、抽出部 212 により抽出された特徴量を用いて、AP法を改良した手順によりデータの分類を行う。分類部 214 は、後述する第 1 評価値、第 2 評価値それぞれの更新を繰り返すことにより、最終的なクラスタを生成する。例えば、分類部 214 は、類似度行列の自己相関の更新値を用いて更新された第 1 評価値及び第 2 評価値に基づく値が収束するまで、クラスタ中心の候補データの決定及びクラスタの生成を行う。分類部 214 は、この第 1 評価値及び第 2 評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて、類似度行列の自己相関の値を設定する。分類部 214 は、この分類法を用いるために、生成部 216、算出部 218、更新部 220、判定部 222 を含む。

20

【0035】

生成部 216 は、抽出部 212 により抽出された特徴量を用いて、各データ間の類似度行列を算出する。生成部 216 は、所定の手法により 2 つのデータ間の類似度を算出することで類似度行列を生成する。

30

【0036】

例えば、音データに対して、上述したベクトルや分散行列が抽出された場合、生成部 216 は、類似度行列に含まれる類似度を、特徴量に含まれる各ベクトル又は分散行列の類似度に重み付けを行うことで算出してもよい。生成された類似度行列は、算出部 218 に出力される。

【0037】

算出部 218 は、例えば、分類手法の AP法で用いられる responsibility と availability とを算出する。ここでは、responsibility を  $r(i,k)$  で表し、第 1 評価値と称し、availability を  $a(i,k)$  で表し、第 2 評価値と称す。

40

【0038】

算出部 218 は、生成部 216 により生成された類似度行列に基づく、クラスタ中心の候補データが、クラスタに属するメンバーデータのクラスタ中心になることの適切さを示す第 1 評価値を算出する。また、算出部 218 は、類似度行列の自己相関に基づく、メンバーデータが、候補データのクラスタに属することの適切さを示す第 2 評価値を算出する。また、算出部 218 は、更新部 220 から類似度行列の自己相関の更新値を取得した場合、自己相関の更新値や新たなクラスタに基づいて第 1 評価値及び第 2 評価値を新たに算出し、更新する。

【0039】

更新部 220 は、従来の AP法とは異なり、類似度行列の自己相関を示す  $s(k,k)$  の値の設

50

定について、新規な手法を用いる。例えば、更新部 220 は、第 1 評価値及び第 2 評価値の算出過程において生成されたクラスタ内の外れデータに関する類似度を用いて、自己相関を設定する。

【0040】

これは、クラスタリングの性質上、類似度が低いデータであっても、いずれかのクラスタに分類する必要があるため、クラスタ内には、類似度の低いデータも含まれてしまう。そのため、類似度の低い外れ値に対して適切な処理を施すことで、第 1 実施形態では、適切なクラスタ数を生成し、分類精度の向上を図る。

【0041】

従来の AP 法では、クラスタ数を決定するのに影響を与えるパラメータである類似度行列の自己相関の値  $s(k,k)$  は更新されない。また、第 1 評価値  $r(i,k)$ 、第 2 評価値  $a(i,k)$  の更新時においても、 $s(k,k)$  の値は、更新されない。したがって、従来の AP 法においては、自己相関  $s(k,k)$  の値に基づくクラスタ数が求まるのみであり、求めたクラスタには外れ値が存在する可能性が大いにある。

【0042】

他方、提案手法では、第 1 評価値  $r(i,k)$ 、第 2 評価値  $a(i,k)$  を更新していく過程で求められるクラスタの外れ値を検出し、それらに対応する類似度行列の自己相関を更新する手法を用いる。つまり、第 1 評価値  $r(i,k)$ 、第 2 評価値  $a(i,k)$  の更新過程でクラスタ数の決定に影響を与える外れデータの  $s(k,k)$  の値も更新されるので、外れデータを中心としてクラスタが生成される可能性があり、最終的に更新が収束したときのクラスタリング結果では、それぞれのクラスタ内に外れ値が存在する可能性は従来手法より小さくなる。したがって、第 1 実施形態では、分類精度の向上を図ることができる。

【0043】

また、更新部 220 は、クラスタ内の外れデータに対する類似度行列の自己相関を、他の値に更新してもよい。また、更新部 220 は、類似度行列の外れデータに関する列ベクトルに基づく値を、他の値に設定してもよい。また、更新部 220 は、この列ベクトルを所定順に並べ替え、所定番目の値を、他の値に設定してもよい。これにより、一旦クラスタに分類された外れ値を、他のクラスタに分類することを可能にし、外れデータを更新しながら、より適切なデータから構成されるクラスタが生成されやすくなる。

【0044】

また、更新部 220 は、類似度行列の各列ベクトルにおける最大値の標準偏差に基づいて、この列ベクトル内の類似度を選択し、選択した類似度を類似度行列の自己相関の初期値に設定してもよい。これにより、列ベクトルの最大値の標準偏差を用いて外れデータを事前に識別し、この外れデータに対する自己相関の初期値をより適切な値に設定することができる。初期値が適切に設定されるだけでも、クラスタ内の外れデータが少なくなり、分類精度が向上する。

【0045】

判定部 222 は、算出部 218 により算出された第 1 評価値と、第 2 評価値とに基づく値が収束するかどうかを判定する。例えば分類部 214 は、第 1 評価値と第 2 評価値との加算値が収束すれば、そのときのクラスタ中心データを代表データにする。分類部 214 は、加算値が収束していなければ、候補データの再決定やクラスタを再編し、もう一度分類を行う。

【0046】

以上の分類処理を行うことにより、クラスタ内の外れ値に対して、クラスタの分類数などに影響を与える自己相関を更新することで、外れ値がクラスタから除外されやすくなり、分類精度を向上させることができる。

【0047】

以下、上記処理で適切に分類されたデータを用いて、新たなモデルを構築することを考える。このとき、分類されたデータに、そのデータの特徴をラベル付けしておくことで、モデル化を行うことができる。

10

20

30

40

50

## 【0048】

例えば、第1実施形態では、環境音の音データに対して分類されたクラスタに対し、代表音にラベル付けを行い、ラベル付けしたデータの特徴量を用いた学習をすることで、新たなモデルを構築する。この新たなモデルは、環境音全体を表すモデルであり、従来のような特定の音の推定や、音の方向の推定を表すモデルなどではない。

## 【0049】

次に、ラベル付け処理や、モデル化に関する処理について説明する。記憶部240は、種類が異なる各データと、この種類を示す識別情報とを関連付けて記憶する。例えば、音データの場合、電車音、拍手音、イベント音、風音などの特徴的な各種音データと、その音データの種類を示す識別情報とが関連付けられている。識別情報は、電車音を示す情報、拍手音を示す情報などであり、ラベルの一種である。

10

## 【0050】

検索部224は、分類部214により生成されたクラスタ中心のデータにマッチングする、記憶部240に記憶されたデータを検索する。例えば、クラスタ中心の代表音が、記憶部240に記憶されたどの音データにマッチングするかを、類似度などを用いて検索する。なお、クラスタ中心のデータについては、クラスタの特徴を示す代表データであれば他のデータでもよい。

## 【0051】

付与部226は、検索部224により検索されたデータに関連付けられた識別情報を含むラベルを、クラスタ中心のデータに付与する。例えば、取得部210により取得されたデータが、マイクロフォンアレイにより取得された音データであれば、音の方向及び/又は音の発生時間が関連付けられているので、付与部226は、これらの情報をラベルに含めてもよい。

20

## 【0052】

なお、検索部224及び付与部226を用いれば自動で代表データにラベルを付与することができるが、ユーザがクラスタ内のデータを実際に調べることで、適切なラベルを付与してもよい。例えば、ユーザが、クラスタ内の代表音を実際に聞くことで、その代表音が何かを判別し、代表音にラベルを付与してもよい。

## 【0053】

解析部228は、ラベルに含まれる識別情報を用いて複数のデータを解析する。例えば、解析部228は、複数のデータにおける各識別情報の割合を算出し、この割合を解析結果にしてもよい。また、解析部228は、音の方向及び/又は音の発生時間を用いて、解析結果を細分化してもよい。

30

## 【0054】

また、解析部228は、分類されたデータを用いて解析モデルを構築するため、学習部230と、構築部232とを含む。

## 【0055】

学習部230は、ラベル付けされたデータと、特徴量とを用いてアンサンブル学習を行う。アンサンブル学習とは、個々に学習した識別器を複数個用意し、それらを、単純には出力の平均を用いる等して、まとめ合わせて一つの識別器を構築する学習法である。アンサンブル学習には、ランダムサブスペース法(参考 T.K.Ho, "The Random Subspace Method for Constructing Decision Forests," IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 20, No. 8, pp. 832-844, 1998)を用いたアンサンブル学習(参考 "D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," Journal of Artificial Intelligence Research 11, pp. 169-196, 1999)を適用する。

40

## 【0056】

構築部232は、学習部230による学習により、データに関するモデルを構築する。モデル構築には、例えば、MATLAB関数(参考 <http://jp.mathworks.com/help/stats/random-subspace.html>)を用いる。

## 【0057】

50

可視化部 234 は、構築されたモデルを用いて解析された解析結果が表示部において可視化されるように処理する。例えば、可視化部 234 は、解析結果をグラフ化し、ユーザーに把握しやすくすることができる。

【0058】

以上により、このモデルを用いれば、例えば環境音に対して発生割合などの環境の雰囲気(活性度)を調べることができる。

【0059】

<具体例>

ここでは、第1実施形態の具体例として、データは、環境計測から得られる環境音の音データとする。解析手法においては、様々な環境音をどのように整理して、分かり易い情報に変換するのかを主な課題とし、その課題を解決する手法を以下に説明する。

【0060】

環境音の計測場所及び取得装置

図4は、環境音を計測した場所を示す図である。図4に示すように、駅STの右側の建物BL1の所定の高さに、取得装置20(マイクロフォンアレイ)が設置される。また、この建物BL1と建物BL2との間に広場ARがあり、この広場ARでは、イベント等が頻繁に開催される。また、建物BL1及び建物BL2は、ショッピングセンターである。

【0061】

図5は、マイクロフォンアレイの一例を示す図である。図6は、マイクロフォンの位置を示す図である。環境音計測には、図5、6に示すマイクロフォンアレイ20A~D(以下、単にマイクロフォンアレイ20と称す。)を用いる。

【0062】

ここで、マイクロフォンアレイ20は、複数のマイクロフォンMCを使用した音計測機器であるため、音源方向や音源位置の推定が可能となり、それらの情報もデータとして利活用が可能となる。マイクロフォンアレイ20により集音される環境音は、そのショッピングセンター内で発生した音である。より具体的には、環境音は、電車音、広場整備音、イベント音、ハシャギ声、風音などである。マイクロフォンアレイ20により計測される音の長さは、例えば最長で2secとする。計測された環境音は、周波数領域に変換され、パワースペクトルが算出される。算出されたパワースペクトルから特徴量が抽出され、抽出された特徴量は、環境音識別の解析モデルの構築に用いられる。

【0063】

概要

まず、環境音の解析手法の概要について説明する。処理制御部206は、ある程度振幅を持った発生音データを対象にし、短時間フーリエ変換を適用し、得られたパワースペクトルPsに非負値行列因子分解(参考 澤田宏:“非負値行列因子分解NMFの基礎とデータ/信号解析への応用,”電子情報通信学会誌, Vol.95, No.9, pp.829-833, 2012)を施し、基底ベクトルPvと係数ベクトルPcを求める。

【0064】

ここで、処理制御部206は、求めた値P(PはPs、Pv、Pcを含む)を環境音の特徴量とみなす。処理制御部206は、これらの特徴量を使って類似度行列を算出して生成する。次に、処理制御部206は、改良したAffinity Propagation(AP)法を用いてデータをカテゴリ(クラス)に分類する。これにより、本具体例では、適切なカテゴリを生成し、分類精度を向上させることができる。

【0065】

また、処理制御部206は、分類されたそれぞれのカテゴリにラベル付けを行う。処理制御部206は、ラベル付けされたデータを利用してランダムサブスペース法を用いたアンサンブル学習を行い、環境音を識別するための解析モデルを構築する。ここで、アンサンブル学習において、例えば、弱学習器にはk最近傍法(k-NN法)を用いる。

【0066】

さらに、本具体例では、環境音識別の解析モデルにより、計測環境がどのような音環境

10

20

30

40

50

を形成しているのかを把握することができる。また、本解析手法の応用例として、例えば、イベント時などで、イベント音、歓声、拍手などを区別することにより、イベントの盛り上がりなどの活性度が音によって計測できるサービスを提供することができる。

【0067】

解析モデル

次に、解析モデルを構築するための分類処理、ラベル付け処理、モデル構築処理を順に説明する。

【0068】

(1) 分類処理

(1.1) 特徴量抽出

抽出部212は、環境音の音データから特徴量を抽出する。例えば、抽出部212は、ある環境音に短時間フーリエ変換を施し、パワースペクトルを求める。

【0069】

図7は、パワースペクトルの一例を示す図である。図7に示す例では、縦軸は、周波数(N個)を表し、横軸は、短時間フーリエ変換のシフト数で決まるフレーム数(Mフレーム)を表す。このとき、このパワースペクトルの値は、行列の要素と考えられ、N×M行列Psが考えられる。ここで、Nは固定で、256次元である。Mは、切り出す音の長さに応じて変化する。

【0070】

抽出部212は、行列Psに、非負値行列因子分解を適用し、基底ベクトルと係数ベクトルとを求める。

【0071】

図8は、基底ベクトルの一例を示す図である。図9は、係数ベクトルの一例を示す図である。非負値行列因子分解には、例えばGamma Process非負値行列因子分解(Gap-NMF) (参考 M. Hoffman, D. Blei, and P. Cook, "Bayesian Nonparametric Matrix Factorization for Recorded Music," Proceedings of the 27<sup>th</sup> International Conference on Machine Learning, Haifa, 2010)を用いる。

【0072】

Gap-NMFは、特徴的な周波数が選択されるように基底ベクトルを決定する因子分解手法である。このとき、基底ベクトルの数は、パワースペクトルの形状によって異なる。また、係数ベクトルは、それぞれの基底ベクトルの時間的変化を表す値を要素を持ったベクトルとなる。図8、9に示す例では、濃い部分が大きな値を表しており、段階的に薄い色に変化していくに従って値が小さくなっていく。

【0073】

図10は、基底ベクトル及び係数ベクトルを用いて再構築したパワースペクトルの一例を示す図である。図10に示すパワースペクトルは、図8に示す基底ベクトルと、図9に示す係数ベクトルとを乗算することで生成される。つまり、図10に示すパワースペクトルは、図7に示すパワースペクトルを再構築したものである。

【0074】

Gap-NMFの特徴から、再構築されたパワースペクトル(図10参照)は、元のパワースペクトル(図7参照)の形状の中で特徴的な周波数が強調されたパワースペクトルの形状になることが分かる。以下、再構築されたパワースペクトルをPr<sub>s</sub>と表す。

【0075】

抽出部212は、環境音の特徴量として、例えば以下の4つの値を算出する。

・元のパワースペクトルを以下の計算で正方行列にした行列 P<sub>1</sub>、

【数1】

$$P_1 = P_s P_s^T \dots\dots(1)$$

10

20

30

40

50

- ・基底ベクトルの平均値、ベクトルf、
  - ・以下の計算で得られる正方行列  $\Pi_2$  に主成分分析が適用され、得られた第1主成分ベクトルe、
- 【数2】

$$\Pi_2 = Pr_s Pr_s^T \quad \dots\dots (2)$$

- ・メルフィルタバンク分析によって  $P_s$  を20次元に圧縮した振幅スペクトルfa
- ここで、faは、Mが異なるパワースペクトルに対して、パワースペクトルの時間変化を同次元（20次元）で比較できるようにする特徴量として取り扱われる。

【0076】

(1.2) 分類手法

分類部214は、上述した4つの特徴量を用いて、一日に発生した環境音の分類を行う。ここで、マイクロフォンアレイ20により計測された一日の音データ数をLとする。分類部214は、分類手法としてAP法と同様の方法を用いる。AP法は、音データそれぞれの類似度に応じて、以下の2つの式を使って、代表音(exemplar)と、それに類似する環境音とを求めることができるクラスタリング手法である。算出部218は、以下の式(3)(4)を用いて第1評価値及び第2評価値を算出する。

【数3】

$$r(i,k) \leftarrow s(i,k) - \max_{k's.t.k' \neq k} \{a(i,k') + s(i,k')\} \quad \dots\dots (3)$$

【数4】

$$a(i,k) \leftarrow \min\{0, r(k,k) + \sum_{i's.t.i' \neq i,k} \max\{0, r(i',k)\}\} \quad \dots\dots (4)$$

【0077】

ここで、第1評価値  $r(i,k)$  は、AP法では、環境音データiからexemplar候補kに送られるメッセージを表す。また、第2評価値  $a(i,k)$  は、exemplar候補kからそのクラスタに含められる可能性がある環境音データiに送られるメッセージを表す。このとき、パラメータ  $s(i,k)$  は、各要素間の関係（例えば類似度）を表したものである。パラメータ  $a(k,k)$  は、以下の式を用いて更新する。

【数5】

$$a(k,k) \leftarrow \sum_{i's.t.i' \neq i,k} \max\{0, r(i',k)\} \quad \dots\dots (5)$$

このように、AP法は、exemplar候補kと音データi間でメッセージをやり取りする。つまり、分類部214は、式(3)と式(4)とを交互に求めることによって、exemplarが求まり、そのexemplarに分類される音データが決定するメッセージ交換型クラスタリング手法である。このとき、 $a(i,k)$  の初期値は0に設定される。

【0078】

ここで、AP法は、従来のクラスタリング手法、例えば、k-means法のように事前にクラスタ数の情報を必要とせず、データ依存でexemplarの数が決まる手法である。このexemplarの数は、 $s(k,k)$ 、つまり、環境音データの自己相関の大きさで決まることが知られている。また、従来手法では、全類似度  $s(i,k)$  の中央値を  $s(k,k)$  に設定してAP法を動作させて

10

20

30

40

50

いる例が多い。

【0079】

しかしながら、例えば分類対象データが音データの場合、従来の $s(k,k)$ の設定方法では、分類がうまくいかないことが少なくない。つまり、クラスタリングの性質上、データはどこかのクラスタに分類されてしまうので、クラスタ内のデータの中には、多くのデータに対して類似度が低いデータ（以下、外れデータとも称す。）が含まれてしまう。

【0080】

この欠点を解決するために、第1実施形態では、改良した $s(k,k)$ の設定方法が用いられる。ここで、 $s(i,k)$ を要素として持つ類似度行列を $S$ と表記する。また、その類似度行列の縦ベクトルを $s_k$  ( $k=1,2,\dots,L$ )と表記する。このとき、それぞれの $s_k$ に対する最大値を $s_{\max k}$  ( $k=1,2,\dots,L$ )と表す。

10

【0081】

ここでは、 $s(k,k)$ の設定方法（更新方法）として、以下の2つを例に挙げて説明する。1つ目の更新方法は、式(6)が用いられる。更新部220は、式(6)により、外れデータを求める。

【数6】

$$abs(Xi-mean(X)) > \alpha std(X) \quad \dots\dots(6)$$

20

【0082】

ここで、 $abs(x)$ は、 $x$ の絶対値を表し、 $mean(x)$ は、 $x$ の平均値を表し、 $std(x)$ は、 $x$ の標準偏差を表し、パラメータ $\alpha$ は、正数を表す。つまり、クラスタ内データ $X$ の外れデータは、式(6)により検出することができる。更新部220は、検出された外れデータのデータポイント $k$ の $s(k,k)$ に対して、以下のルールを適用して更新する。

【数7】

$$s(k,k) \leftarrow sort(s_k)_{nth} \quad \dots\dots(7)$$

30

【0083】

ここで、 $s_k$ は、類似度行列の $k$ 列ベクトルを表し、 $sort(s_k)_{nth}$ は、降順の $k$ 列ベクトル $s_k$ の上位から $n$ 番目の値を表す。本具体例では、例えば4番目の値を用いるが、どの値を用いるかは与えられたデータの分類形態で変更する。

【0084】

また、式(6)のパラメータ $\alpha$ は、どの程度の外れデータが想定されて分類が行われるかによって変更される。 $s(k,k)$ の初期値は、従来のAP法で用いられている類似度行列のすべての要素の中央値が設定される。

【0085】

次に、2つ目の更新方法について説明する。更新部220は、最大値 $s_{\max k}$  ( $k=1,2,\dots,L$ )の標準偏差 $\sigma_s$ を基準に、 $s(k,k)$ の初期値を決める方法を用いる。すなわち、更新部220は、以下の式(8)を閾値 $Th$ として用いて $s(k,k)$ を決める。

40

$$Th = \sigma_s / a \quad (8)$$

【0086】

ここで、定数 $a$ は、任意の正の整数とする。閾値 $Th$ は、 $s_k$ の要素の中で、できるだけ最大値に近い値を $s(k,k)$ に設定しようとして、発明者によって考えられた閾値である。

【0087】

更新部220は、各 $s_k$ の要素において、 $\sigma_s / a$ よりも大きい要素で、 $\sigma_s / a$ に最も近い要素の値を $s(k,k)$ の初期値に設定する。もし、 $s_k$ が該当する要素を持たなければ、更新部220は、 $s_k$ の2番目の値を $s(k,k)$ の初期値に設定する。

50

## 【0088】

2つ目の更新方法の場合、分類部214は、改良されたAP法でクラスタリングした事後処理として、exemplarの値に応じて、クラスタ内の外れデータを求め、この外れデータをexemplar候補としてさらにクラスタリングすればよい。また、更新部220は、上述した2つの更新方法のいずれかを用いてもよいし、自己相関の初期値を2つ目の更新方法を用いて設定し、さらに、1つ目の更新方法を用いて、その初期値を更新してもよい。

## 【0089】

分類部214は、 $r(i,k) + a(i,k)$ の加算値が収束するまで分類処理を繰り返す。収束とは、例えば、再分類を所定回数以上行っても、加算値が所定範囲内にあることをいう。これにより、本具体例では、クラスタ内の外れデータを処理する改良AP法におけるクラスタリングを用いることで、分類精度を向上させることができる。

10

## 【0090】

## (2) ラベル付け

検索部224は、分類部214により分類された各カテゴリの代表音(例えばexemplarの音)に類似する音データを、記憶部240に記憶された音データから検索する。例えば、検索部224は、代表音との類似度が一番大きい音データを検索する。

## 【0091】

付与部226は、検索された音データに関連する識別情報を含むラベルを、代表音に付与する。なお、ラベルは、識別情報以外にも、音の発生位置や発生時間帯などを含んでもよい。また、付与部226は、代表音と同じラベルを、クラスタ内の音データに付与してもよい。これにより、全データに対して検索処理を行う必要がなくなり、処理負荷を軽減させることができる。

20

## 【0092】

また、ラベル付けとして、ユーザが、代表音を聞き、代表音に合ったラベルを選択し、ラベル付けを行ってもよい。この場合、付与部226は、ユーザが付与した代表音のラベルを、代表音と同じカテゴリの音データに付与してもよい。

## 【0093】

## (3) 学習処理

解析部228は、ラベル付けされた各音データを教師信号とし、そのそれぞれに対する環境音の上述した特徴量( $f', e, fa$ )を学習データとして用いて、一日の環境音を識別するための解析モデルを構築する。ここで、 $f'$ は、 $Pr_s$ の周波数平均値である。また、 $f'$ は、基底ベクトルの平均値、ベクトル $f$ を用いてもよい。

30

## 【0094】

本具体例では、解析部228は、アンサンブル学習を用いて環境音を識別するための解析モデルを構築する。学習部230は、アンサンブル学習として、上述したランダムサブスペース法を用いたアンサンブル学習を実行する。また、弱学習器には、k-最近傍法を用いてモデルを構築する。ここで、モデル構築には、例えばMATLAB関数(`fitensemble`)を用いることにする。

## 【0095】

## 実験

ここで、発明者らが行った実験について説明する。発明者らは、図4に示す位置に、図5, 6に示すマイクロフォンアレイ20を設置し、マイクロフォンアレイ20は、公共空間内に発生した環境音を計測する。

40

## 【0096】

環境音のモデル構築には、平成26年7月7日(月)から平成26年7月11日(金)において計測された5日分の音データが用いられる。まず、分類部214による分類結果について説明する。この実験では、更新方法として2つ目を適用し、式(8)のパラメータ $a$ は3とする。

## 【0097】

また、類似度行列の要素 $s(i,k)$ は、それぞれの環境音データの特徴量を用いて、Itakur

50

a spectral距離 (Di) とCOSH spectral距離 (Dcosh) を用いて算出する。

【数 8】

$$s(i,k)=c_1D_{cosh}(ei,ek)+c_2D_{cosh}(fi,fk)+c_3Di(\Pi_i,\Pi_k)+\dots\dots(9) \\ c_4D_{cosh}(fmi,fmk)$$

パラメータCi (i=1,2,3,4)は、それぞれの距離の重み係数を表し、正数値が与えられる。

図 1 1 は、改良AP法を用いて音データをクラスタリングした結果の一例を示す図である。

図 1 1 に示す例では、データ数に対するクラスタ数が示されている。図 1 1 に示す例では、圧縮率 (=クラスタ数 / データ数) が46.2%となっている。

10

【0098】

図示しないが、従来手法で用いられている全s(i,k)の中央値をs(k,k)に設定する方法を用いたAP法では、圧縮率が91.7%となる。つまり、クラスタ数が少ないため、外れデータがクラスタ内に含まれている可能性が高い。なお、自己相関の更新において、1つ目の更新方法を用いた場合には、その更新方法に合わせた圧縮率になるように、事前実験等により適宜パラメータを調整すればよい。

【0099】

次に、上述した処理で構築された解析モデルの精度の検証について説明する。検証に用いた音データは、平成26年7月14日から平成26年7月18日に計測された環境音である。それぞれの音データに対し、平成26年7月7日(月)から平成26年7月11日(金)の音データを用いて構築された解析モデルが適用される。このとき、曜日毎に解析モデルが適用され、間違っ

20

たラベルが出力された回数(代表音とは異なる音が、その代表音のクラスタに分類された数)がカウントされる。

【0100】

図 1 2 は、改良AP法の精度結果(その1)の一例を示す図である。図 1 2 に示すように、各曜日の正解率の平均は91%である。なお、誤解答数とは、クラスタに分類された音データを人が聞いて、その音データがクラスタの代表音とは異なる数を示す。

【0101】

なお、図示しないが、従来のAP法によりクラスタリングされた代表音を使ってラベリングされたときの解析モデルを用いた場合では、正解率が約78%という結果が得られている。このことから、改良AP法により、そのクラスタリング結果は、適切にカテゴリが生成され、そのカテゴリの分類精度が増加したと考えることができる。

30

【0102】

さらに、曜日毎ではなく、7月7日から7月11日の全ての音データを使って、解析モデルが構築された場合を検討する。7月14日から7月18日のそれぞれの環境音データに、この解析モデルを適用し、モデル構築に使用するデータ数が多くなると、分類精度が向上するの

かどうかを発明者らは検証した。

【0103】

図 1 3 は、改良AP法の精度結果(その2)の一例を示す図である。図 1 3 に示すように、各曜日の正解率の平均は93%である。図 1 3 に示す例では、各曜日において、誤解答数が減り、平均正解率が向上している。このことから、多くのデータを使って解析モデルを構築した方が、モデルの信頼性が向上することが分かる。

40

【0104】

次に、上述した解析モデルを用いて、可視化部 2 3 4 により分類された環境音の発生割合を示す。

【0105】

図 1 4 は、一日の環境音の発生割合の一例を示す図である。図 1 4 に示すように、本実験において、マイクロフォンアレイ 2 0 を設置した場所は、すぐ近くに駅STがあることから(図 4 参照)、電車音の発生割合が一番高い。

【0106】

50

2番目は、広場ARのザワツキを表したような雑踏音の発生割合が高く、3番目は、広場の整備音の発生割合が高い。その他に、ハシャギ音、風音、イベント音などが発生している。

【0107】

付与部226は、このような音にラベル付けを行い、可視化部234は、その発生割合を可視化することができる。図14に示すグラフから、その日の音環境の雰囲気を理解することが可能になる。さらに、改良AP法を用いて生成された解析モデルは、図12や図13に示すように正答率が90%を超えているので、音環境の可視化は、信頼でき得るレベルに達していると言える。

【0108】

上述したように、本具体例における解析モデルは、環境音の発生割合の可視化から、その音環境の雰囲気を把握することができる。そこで、この解析モデルの応用例として、広場ARで開催されたイベントに関して、イベント中に発生する音から、そのイベントの雰囲気を把握することができるのかを発明者らは検証した。

【0109】

解析モデルの作成に関しては、イベントが開催された日に発生した一日の環境音の音データを使って、イベント音（歌、音楽等）、拍手音に新たにラベルが付与され、以前作成した解析モデルに、新たなラベルが付与されたクラスが追加されるという手順でモデルが構築された。

【0110】

図15は、あるイベント日の10時から13時までのイベントAにおける環境音発生割合の一例を示す図である。ここで、以下に示す全ての例は、モデル構築に使用されていない音データから得られた結果である。以降の図では、環境音に、ハシャギ音、イベント音、拍手音などが含まれる。

【0111】

図16は、あるイベント日の14時から18時までのイベントAにおける環境音発生割合の一例を示す図である。図16に示すイベントAと図15に示すイベントAは同じイベントである。図15及び図16に示すグラフによれば、イベント中盤から後半にかけて、ハシャギ音、イベント音、拍手音がほぼ同じ割合で発生していることが分かる。

【0112】

これは、実際のイベントによるアンケート結果などと照らし合わせて、このイベントの雰囲気や盛り上がりがどうだったかを、このグラフと照らし合わせて検討することが可能になる。

【0113】

図17は、あるイベント日の14時から18時までのイベントBにおける環境音発生割合の一例を示す図である。図17に示す例では、イベント音の発生割合が高く、これは、音楽や歌が中心のコンサート型のイベントであることが分かる。また、ハシャギ音や拍手音が発生しているが、イベント音の割合が圧倒的に多いことから、音楽や歌で盛り上がったイベントであったと推察することができる。

【0114】

また、上述した環境音とは雰囲気が異なる例を図18及び図19を用いて説明する。図18は、あるイベント日の10時から13時までのイベントCにおける環境音発生割合の一例を示す図である。

【0115】

図19は、あるイベント日の14時から18時までのイベントCにおける環境音発生割合の一例を示す図である。図18に示すイベントCと図19に示すイベントCは同じイベントである。図18及び図19に示すグラフによれば、イベント開始時には、拍手音などが発生しており、盛り上りの雰囲気があったが、イベントの中盤、後半に向けて、徐々に盛り上りの雰囲気が減少しているような傾向を把握することができる。

【0116】

10

20

30

40

50

以上より、環境音を識別するための解析モデルを使うことによって、イベントの雰囲気  
を把握することができる。よって、この解析モデルのさらなる応用例として、イベント評  
価への活用にも期待できる。

【 0 1 1 7 】

本具体例では、環境音データの分類方法、その分類結果から環境音を識別するための解  
析モデルの構築手法を提供することができる。本具体例では、環境音の識別を約90%以上  
の確率で実現することができ、環境音の発生場所や、音環境の雰囲気などを理解すること  
ができる。

【 0 1 1 8 】

また、本具体例の応用例として、イベント時の発生音からイベントの雰囲気の把握、さ  
らにはイベントの盛り上がりなどの評価にも利用することが期待できる。また、イベント時  
の環境音発生割合から推察する結果と、実際の雰囲気とが一致しているかどうか、イベン  
ト参加者にアンケートなどで検証することができる。

10

【 0 1 1 9 】

< 動作 >

次に、第1実施形態における情報処理システム1の動作について説明する。図20は、  
第1実施形態におけるデータ解析処理の一例を示すフローチャートである。図20に示す  
ステップS102で、分類部214は、複数のデータに対して分類処理を行う。分類処理  
の詳細は、図21を用いて説明する。

【 0 1 2 0 】

20

ステップS104で、検索部224及び付与部226は、ラベル付け処理を行う。ラベ  
ル付け処理の詳細は、図22を用いて説明する。

【 0 1 2 1 】

ステップS106で、解析部228は、分類されたデータを用いて、データ解析を行う  
。データ解析の詳細は、図23を用いて説明する。

【 0 1 2 2 】

図21は、第1実施形態における分類処理の一例を示すフローチャートである。例えば  
図21に示す処理は、定期的に必要なデータが集まったときに実行される。例えば音デー  
タの場合、一日ごとに、データの取得処理が終わった後に分類処理が開始されればよい。

【 0 1 2 3 】

30

図21に示すステップS202で、取得部210は、取得装置20から受信した複数の  
データを取得する。例えば、取得装置20が外に設置されたマイクロフォンの場合、取得  
部210は、環境音の音データを取得する。また、複数のデータは、自装置の記憶部24  
0から取得されてもよい。

【 0 1 2 4 】

ステップS204で、抽出部212は、各データから特徴量を抽出する。抽出される特  
徴量は、データに応じて変更されればよい。例えば、音データの場合、周波数領域のパワ  
ースペクトルに関する特徴量が抽出される。

【 0 1 2 5 】

ステップS206で、生成部216は、抽出された特徴量を用いて、各データ間の類似  
度で表される類似度行列を生成する。

40

【 0 1 2 6 】

ステップS208で、算出部218は、分類の収束判定に用いる評価値を算出する。例  
えば、算出部218は、改良AP法が用いられる場合、第1評価値 $r(i,k)$ と、第2評価値 $a(i,k)$   
とを算出する。さらに算出部218は、分類処理の終了判定パラメータとして、第1  
評価値と第2評価値との加算値を算出する。

【 0 1 2 7 】

ステップS210で、判定部222は、終了判定パラメータが収束したか否かを判定す  
る。例えば、判定部222は、加算値が収束したか否かを判定する。終了判定パラメータ  
が収束していれば(ステップS210 - YES)処理はステップS214に進み、終了判

50

定パラメータが収束していなければ(ステップS210-NO)処理はステップS212に進む。

【0128】

ステップS212で、更新部220は、クラスタ内の外れデータを検出し、検出された外れデータの自己相関を、他の値に更新する。自己相関は、類似度行列の中に含まれている。更新処理が終わると、処理はステップS208に戻り、更新後の自己相関を用いて第1評価値や第2評価値が更新される

【0129】

ステップS214で、分類部214は、判定対象パラメータが終了したときのクラスタを最終の結果として決定する。これにより、従来のAP法よりも、外れデータを考慮してクラスタリングを行うため、クラスタ内の外れデータの数を減らすことができ、分類精度を向上させることができる。

10

【0130】

図22は、第1実施形態におけるタグ付け処理の一例を示すフローチャートである。図22に示すステップS302で、検索部224は、各クラスタから1つのクラスタを選択し、そのクラスタ内の代表データを選択する。代表データは、例えばクラスタ中心のデータである。

【0131】

ステップS304で、検索部224は、選択した代表データに対応するデータを、記憶部240の中から検索する。例えば、検索部224は、代表データに一番類似するデータを検索する。

20

【0132】

ステップS306で、付与部226は、検索されたデータに関連付けられている識別情報をラベルに含めて、代表データに付与する。ラベルには、識別情報以外にも、そのデータに関する他の情報が含まれてもよい。

【0133】

ステップS308で、付与部226は、全ての代表データにラベルを付与したか否かを判定する。代表データは、例えば各クラスタに1つずつ含まれる。全ての代表データにラベルが付与されていれば(ステップS308-YES)処理は終了し、全ての代表データにラベルが付与されていないければ(ステップS308-NO)処理はステップS302に戻り、他の代表データが選択される。

30

【0134】

図23は、第1実施形態における解析処理の一例を示すフローチャートである。図23に示すステップS402で、学習部230及び構築部232は、ラベル付けされた各データを教師信号とし、各教師信号の特徴量を学習データとして用いてアンサンブル学習を行い、解析モデルを構築する。

【0135】

ステップS404で、分類部214は、この解析モデルを適用して、解析対象のデータを分類する。さらに、解析部228は、分類の正答率などを算出する場合には、事前に割り当てられたデータのラベルと、分類後に割り当てられたラベルとを比較し、ラベルが異なるデータをカウントする。また、人が分類後のデータを分析して、誤って分類されたデータの数をカウントしてもよい。

40

【0136】

ステップS406で、可視化部234は、解析結果を表示部において可視化されるように処理する。例えば、可視化部234は、解析結果をグラフ化し、把握しやすくする。

【0137】

以上、取得されたデータ全体に対して解析モデルを構築することができ、さらに、解析結果を一目で把握しやすいように可視化することができる。

【0138】

以上、第1実施形態では、多くのデータを処理するのに悩めるサービス産業の手助けを

50

するため、まずはデータの分類において分類精度を向上させることができる。さらに、データ利活用を実行するためのデータ解析手法を提案することができる。また、その解析から得られた結果をどのようなサービスに活かすことができるのかを提示することができる。

#### 【0139】

例えば、第1実施形態によれば、音データを解析する場合、音の方向や位置を推定するだけでなく、その音の発生状況から、その音環境がどのような雰囲気になっているか、音環境模様を把握することができる。

#### 【0140】

より具体的には、イベント時に、その音環境を計測し、どのような雰囲気イベントであったのかを、盛り上がり、賑わいなどの音による客観的指標を提示することができる。また、共同住宅において、その周りの音環境を計測し、どのような音環境であるのかを、入居予定者に提示することができる。また、クラスタ内の外れデータを検出することで、この外れデータを分析し、何らかの特徴を把握してもよい。

#### 【0141】

また、第1実施形態によれば、大量にあるデータから代表データを検出し、その代表データを用いて解析を行うことができるため、データ解析に時間がかからない。また、データのモデル構築については、適切な分類手法によりデータがカテゴリに分類されていれば、そのカテゴリに対してラベル付けを行い、モデルを構築することが可能である。

#### 【0142】

##### [第2実施形態]

次に、第2実施形態について説明する。第2実施形態では、音環境を色模様としてモニタリングする技術に関する。第2実施形態で取得されるデータは、音データである。第2実施形態において、音データから1又は複数の特徴量を抽出し、抽出した特徴量を色で表現することにより、環境音などの音データを色模様（以下、音模様とも称す。）で表現することができる。

#### 【0143】

例えば、第2実施形態における情報処理装置は、環境音の音データから、音の高低、大小、及び継続の3つの特徴を抽出し、それぞれに所定の色を割り当て、音模様を表現する。さらに、音の音源位置を推定することで、所定領域内における音模様を表現することができる。この音模様は、画面上に表示されることで、ユーザにとって視認可能となる。

#### 【0144】

なお、第2実施形態におけるシステム構成、及びハードウェア構成は、第1実施形態におけるシステム構成、及びハードウェア構成とそれぞれ同様であるため、その説明を省略する。また、第2実施形態におけるシステム構成及びハード構成については、図1及び図2と同じ符号を用いる。

#### 【0145】

##### <機能構成>

図24は、第2実施形態における情報処理装置10の機能構成の一例を示すブロック図である。図24に示す情報処理装置10は、送信部202、受信部204、処理制御部300、及び記憶部310を含む。送信部202、及び受信部204は、第1実施形態と同様であるため、同じ符号を付し、その説明を省略する。記憶部310は、分類に用いるデータや、特徴量抽出などに用いるデータや、モデル構築に用いるデータなどの各種データを記憶する。例えば、記憶部310は、色表現に用いるデータや、色表現された画像データなどを記憶する。

#### 【0146】

処理制御部300は、取得部210、抽出部212、分類部214、音特徴抽出部302、表現部304、及び推定部306を含む。なお、取得部210、抽出部212、分類部214は、第1実施形態と同様であるため、同じ符号を付し、その説明を省略する。

#### 【0147】

10

20

30

40

50

音特徴抽出部 302 は、分類部 214 により分類された各カテゴリに基づく音データの第 1 特徴量を抽出する。例えば、音特徴抽出部 302 は、分類部 214 により改良 AP 法を用いて所定期間内の音データが特徴ごとに分類された各カテゴリに対し、各カテゴリの代表となる周波数（例えばカテゴリの重心の周波数）を用いて第 1 特徴量を抽出する。

【0148】

より具体的な例として、音特徴抽出部 302 は、代表周波数を大、中、小の 3 つに分類し、ヒストグラムを生成する。次に、音特徴抽出部 302 は、ヒストグラム内の度数に応じて、大、中、小ごとに 0 又は 1 の値を与える。音特徴抽出部 302 は、（大、中、小）の（0, 0, 1）から（1, 1, 1）までの 7 つのスコアを第 1 特徴量とする。

【0149】

表現部 304 は、音データから、第 1 特徴量を、第 1 色により表現する。第 1 色は、例えば G（緑）とする。これにより、音データの特徴を色で表現することができ、音という視認しにくいものを視認しやすくすることができる。

【0150】

また、音特徴抽出部 302 は、音データから第 2 特徴量及び第 3 特徴量を抽出してもよい。例えば、音特徴抽出部 302 は、音データの大小を示す第 2 特徴量と、音データの継続性を示す第 3 特徴量を求めてもよい。音特徴抽出部 302 は、取得された音データの大小を示す第 2 特徴量と、音データの継続性を示す第 3 特徴量とを生成する。

【0151】

第 2 特徴量として、音特徴抽出部 302 は、例えば、公知のラウドネス（参考 ITU-R BS.17701-1 勧告 <http://www.itu.int/rec/R-REC-BS.1770/>）の計算によって求められる音量（単位：LKFS）を基準とし、音の強弱を 7 段落のスコアで表す。なお、音特徴抽出部 302 は、このスコアが大きくなるほど、音量が大きくなるようにスコアを設定する。また、音特徴抽出部 302 は、単純に音データの波形のエンベロープを、所定の時間単位で求めて、その平均値を音の大きさにしてもよい。また、音特徴抽出部 302 は、パワースペクトルを周波数軸方向に足して、時間軸方向の平均値を用いて音の大きさの代わりとしてもよい。

【0152】

第 3 特徴量として、音特徴抽出部 302 は、例えば、音データのパワースペクトルにおけるパワー最大の時刻を基準にして相互情報量からヒストグラムを求める（参考 角田拓也，中川匡弘，「相互情報量と相互相関関数を用いた筋電と筋音」の時間差の推定，信学技法，MBE2015-60，pp.37-42，2015.）。また、音特徴抽出部 302 は、求めたヒストグラムの尖度（Kurtosis）を計算する。このとき、音特徴抽出部 302 は、尖度の値が大きいくほど、突発的な音を表し、尖度の値が小さくなるほど、変動が少なく持続性が高い音を表すことになる。よって、音特徴抽出部 302 は、尖度値に応じた度合を音の継続度とする。例えば、音の継続度は、7 段階のスコアで表される。

【0153】

上述したように、3 つの特徴量が抽出される場合、表現部 304 は、第 2 特徴量を、第 2 色により表現し、かつ、第 3 特徴量を、第 3 色により表現する。第 2 色は、R（赤）、第 3 色は、B（青）として表す。例えば、それぞれの色は、0 ~ 255 の値で表される。これにより、環境音などの音に対して色を用いて表現することができる。なお、色の組み合わせは、光の三原色を用いたが、これに限られない。

【0154】

推定部 306 は、音データに対して、所定領域内の音源位置を推定する。例えば、推定部 306 は、所定領域をブロック分割し、分割ごとに周波数選択を行って、その周波数を用いる MUSIC 法（参考 R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," IEEE Trans. on Antennas and Propagation, Vol. 34, No. 3, pp. 276-280, 1986.）により、音源位置を推定することができる。また、推定部 306 は、主に音の到来方向が推定できる方法として、ビームフォーマ法、線形予測法や最小分散法などを用いてもよい（参考 大賀寿郎、山崎芳男、金田豊著、音響システムとデジタル

10

20

30

40

50

処理、電子情報通信学会出版)。

【0155】

このとき、表現部304は、推定部306により推定された音源位置の音データに対して、第1特徴量を第1色で表現し、第2特徴量を第2色で表現し、第3特徴量を第3色で表現する。これにより、所定領域内の音を音源位置ごとにRGBで表すことができ、色を用いて音模様を視覚的に表現することができる。

【0156】

<音特徴量の抽出>

次に、音特徴量の抽出について説明する。まず、音特徴抽出部302は、第1音特徴量を求めるため、音データのパワースペクトルを求める。ここで、音特徴抽出部302は、所定期間内において、時間と共に変化するパワースペクトルを1つの画像として捉える。

10

【0157】

図25は、パワースペクトルの一例を示す図である。図25に示す例では、所定値以上のパワースペクトルが表示されている。所定値には、音が発生していることを示すための適切な閾値が設定される。音特徴抽出部302は、所定値以上のパワースペクトルが表現された図に対し、エッジ検出を行う。

【0158】

図26は、エッジ検出されたパワースペクトルの一例を示す図である。図26に示す例では、エッジ検出で囲まれた領域のそれぞれの重心において、分類部214により、距離が近いものが1つのカテゴリ内に改良AP法により分類される。

20

【0159】

ここで、改良AP法を用いる理由としては、分類が詳細になり、周波数の分布把握がより精度よくなるからである。

【0160】

図27は、分類された各カテゴリの一例を示す図である。図27に示す例では、カテゴリCT1~7までの7つのカテゴリに分類されている。音特徴抽出部302は、例えば、各カテゴリ内の重心の周波数(代表周波数)を大(4kHz以上)、中(1kHz以上4kHz未満)、小(1kHz未満)の各領域に分類する。各領域の閾値は一例であって、これに限定されない。

30

【0161】

次に、音特徴抽出部302は、それぞれの周波数領域のカテゴリを、時間軸上において、いくつあるかをカウントする。例えば、大の領域において横軸のカテゴリの数をカウントすると2などである。音特徴抽出部302は、各領域のカウント数の最大値を用いて、カウント値を正規化する。例えば、音特徴抽出部302は、各カウント値を最大値で除算する。このとき、或る閾値(例えば0.4)以上の値は1に設定され、閾値未満の値は0に設定される。例えば、(大、中、小)が(1、0.2、0.5)の場合、(1、0、1)となる。

【0162】

また、音特徴抽出部302は、(0,0,1)~(1,1,1)からの各パターンについてスコア(第1特徴量)を付ける。例えば、音特徴抽出について、以下の処理が行われる。

40

パターン：スコア値

(0,0,1)：1

(0,1,0)：2

(0,1,1)：3

(1,0,0)：4

(1,0,1)：5

(1,1,0)：6

(1,1,1)：7

50

このように、音特徴抽出部 302 は、第 1 特徴量として、1 ~ 7 の数字を用いる。

【0163】

図 28 は、音の大小度の計算を説明するための図である。図 28 に示す例では、横軸が時間で、縦軸が音の大きさである。音の大きさはラウドネスを用いるが、音の大きさが計算できれば、いずれの指標が用いられてもよい。

【0164】

音特徴抽出部 302 は、音のレベルによって、例えば 7 段落のスコアを付ける。音特徴抽出部 302 は、平常時の音のレベル（例えば基準レベル）を 1 として、音環境の音レベルに合わせて段階的に閾値を設定し、7 段階のスコア（第 2 特徴量）を付けるようにする。基準レベルや閾値は、適宜設定変更可能である。

10

【0165】

図 29 は、音の継続度の計算を説明するための図である。図 29 に示す例では、音の尖度（Kurtosis）を用いて継続度が求められる。まず、音特徴抽出部 302 は、パワースペクトルの最大値（図 29 に示す円で囲む部分）に対して、相関値を、相互情報量を使って計算する。

【0166】

次に、音特徴抽出部 302 は、求めた値を MATLAB の hist を用いてヒストグラムを生成する。音特徴抽出部 302 は、作成したヒストグラムに対して、ヒストグラムの形状を把握するための尖度（Kurtosis）を計算する。

【0167】

最後に、音特徴抽出部 302 は、計算された尖度の値に応じて、以下のとおり 7 つのスコアを付ける。

条件：スコア値

Kurtosis 4.5 : 1

4.5 > Kurtosis 3.5 : 2

3.5 > Kurtosis 2.5 : 3

2.5 > Kurtosis 1.5 : 4

1.5 > Kurtosis 1.0 : 5

1.0 > Kurtosis 0.5 : 6

0.5 > Kurtosis : 7

20

30

【0168】

表現部 304 は、各特徴量に対して、7 段階のスコア値を 0 ~ 255 の値に置換し、色表現する。例えば、表現部 304 は、（第 1 特徴量：第 2 特徴量：第 3 特徴量）が（5 : 2 : 1）であるとすると、色（G : R : B）について（ $255 \times 1 : 255 \times (2/5) : 255 \times (1/5)$ ）で表現可能である。

【0169】

<音模様の実験>

ここで、発明者らが行なった音模様に関する実験について説明する。図 30 は、実験地における音模様を表現する所定領域を示す図である。図 30 に示すように、本実験地は、図 4 に示す実験地と同じ場所であり、所定領域は、マイクロフォンアレイ 20 近傍の領域 AR10 である。音データは、最長で 2 秒間のデータであり、ある程度振幅を持った環境音とする。また、領域 AR10 は、5 m x 5 m の各ブロック領域に分割され、各ブロック領域に、RGB 値が求められる。

40

【0170】

図 31 は、平常時の音模様を示す図である。図 32 は、イベント時の音模様を示す図である。図 31 及び図 32 に示すように、様々な音が発生している（イベント時の）場合には、カラフルな（濃淡の差が大きい）音模様となり、あまり音の変化のない平常時には、青や黄緑が主体の（濃淡の差が小さい）音模様になる。つまり、音模様は、音環境の状況の変化を検知できることが分かる。よって、上述した 3 つの特徴量を用いる音模様は、音環境の状態を検知、及び把握するために、十分に役立つ効果的なモニタリング手法である

50

と言える。

#### 【0171】

上述した例では、3つの特徴量を7段階のスコアで表現し、それぞれの値をRGB値に置き換えて、音模様が描かれる。7段階とする理由は、第1特徴量として音の高低度を算出する際に、高、中、低と3段階に分けてヒストグラムを作成し、そのヒストグラムの大きさを0か1に変換されるため、(高, 中, 低)に対する値が(0, 0, 1)から(1, 1, 1)までの7つに場合分けされるからである。第1特徴量と同様に、他の特徴量も、大きさに応じた7段階のスコアを用いる。

#### 【0172】

ここで、音の大小の値と継続度の尖度値に関して、7段階に分類する閾値設定は、音環境で取得される音データによって変動するので、それぞれの音環境に応じた場合分けで閾値が決められればよい。

10

#### 【0173】

また、音模様を学習データとして用いて、教師データには音模様を計測した日付を用いることにより、音模様を利用した音環境の理解、例えば、どの日に類似しているのか、いつもと違う音模様になっているのかなどが把握できる識別モデルの構築が可能になると考えられる。ここで、識別モデルには、例えば単純ベイズ分類器や判別分析(フィッシャー判別など(参照C.M.Bishop "Pattern Recognition and Machine Learning," Springer, 2006))が利用できる。

#### 【0174】

また、他のセンサ、例えば、振動計や画像などの融合により、上述した音模様や識別モデルは、建物の老朽化(倒壊するまでの過程)のモニタリングに利用することが可能である。

20

#### 【0175】

また、当システムは、音模様を日付ごとに対応付けてデータベースに保存することにより、所定の音模様に対する類似の音模様を検索することが可能になる。このとき、ブロックごとに類似の音模様が検索されてもよい。

#### 【0176】

<動作>

次に、第2実施形態における情報処理システム1の動作について説明する。図33は、第2実施形態における音模様表現処理の一例を示すフローチャートである。図33に示すステップS502で、取得部210は、音データを取得する。

30

#### 【0177】

ステップS504で、推定部306は、所定期間毎に、取得された音データの音源位置を推定する。

#### 【0178】

ステップS506で、音特徴抽出部302は、推定された音源位置に対応する音データの各特徴量を抽出する。第1特徴量を抽出するため、図20に示すステップS102の分類処理が実行される。

#### 【0179】

ステップS508で、表現部304は、抽出された各特徴量に対し、それぞれ各色を示す値に置換し、音を色で表現する。これにより、上述したように、音環境を音模様として色で視覚的に表現し、容易に把握することができるようになる。

40

#### 【0180】

なお、上述した第1実施形態及び第2実施形態における処理フローに含まれる各処理ステップは、処理内容に矛盾が生じない範囲で、任意に順番を変更して又は並列に実行することができるとともに、各処理ステップ間に他のステップを追加してもよい。また、便宜上1ステップとして記載されているステップは、複数ステップに分けて実行することができる一方、便宜上複数ステップに分けて記載されているものは、1ステップとして把握することができる。

50

【0181】

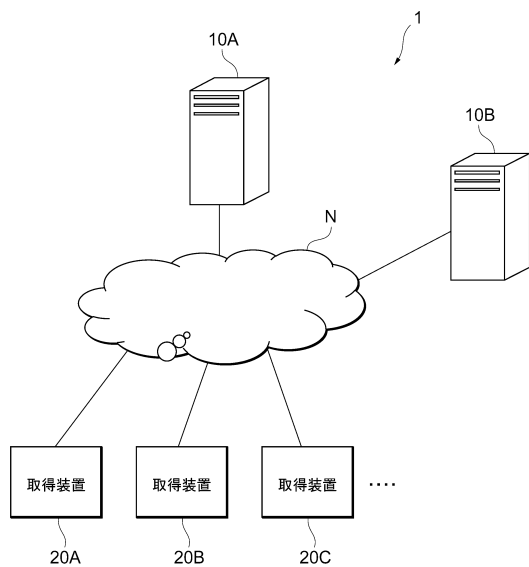
以上、本発明の実施形態について述べたが、本発明は既述の実施形態に限定されるものではなく、本発明の技術的思想に基づいて各種の変形及び変更が可能である。

【符号の説明】

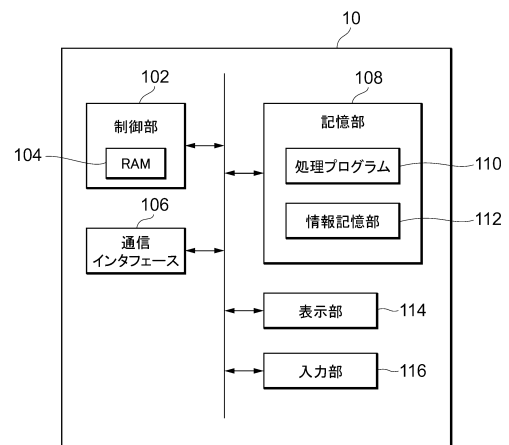
【0182】

- 1 情報処理システム
- 10 情報処理装置
- 20 取得装置
- 102 制御部
- 108 記憶部
- 206 処理制御部
- 210 取得部
- 212 抽出部
- 214 分類部
- 224 検索部
- 226 付与部
- 228 解析部
- 234 可視化部

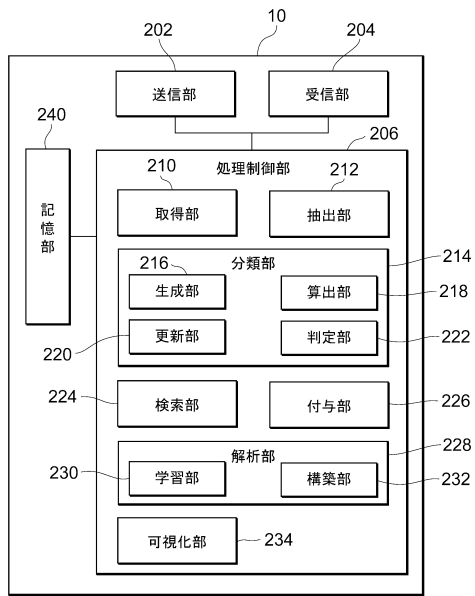
【図1】



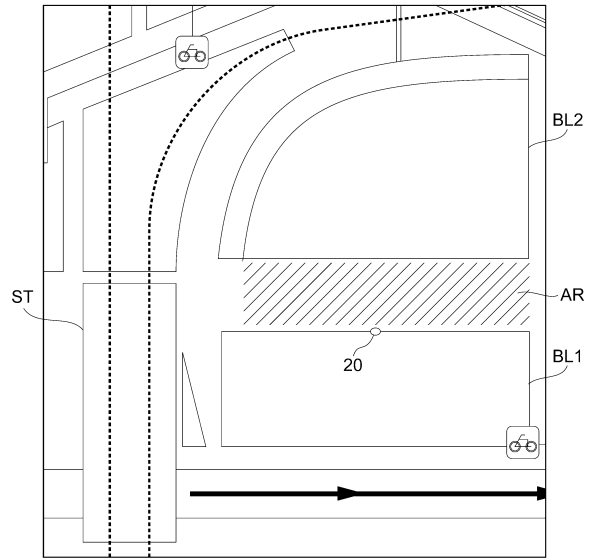
【図2】



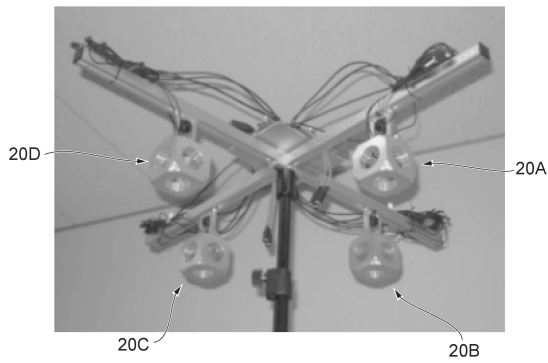
【図3】



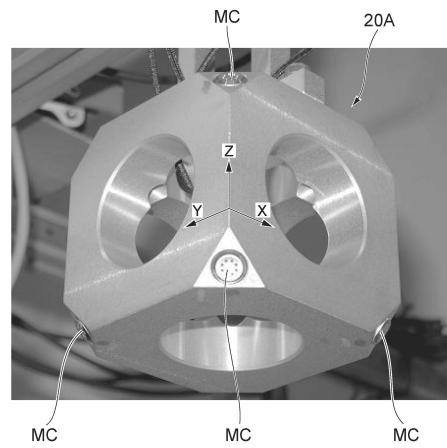
【図4】



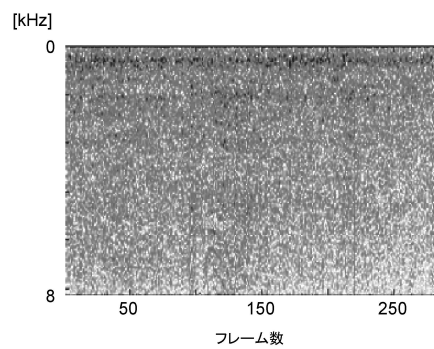
【図5】



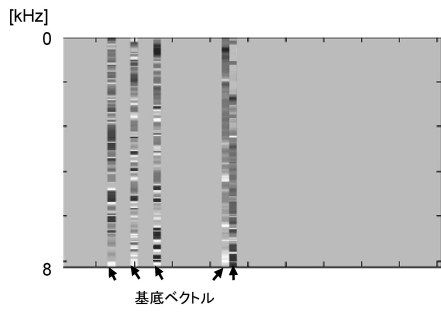
【図6】



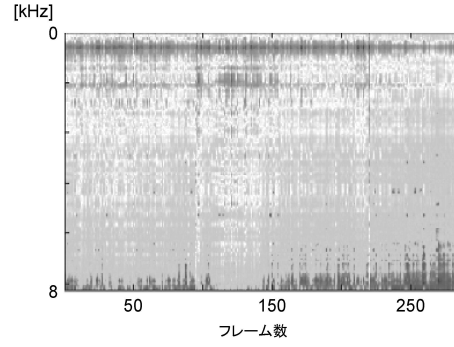
【図7】



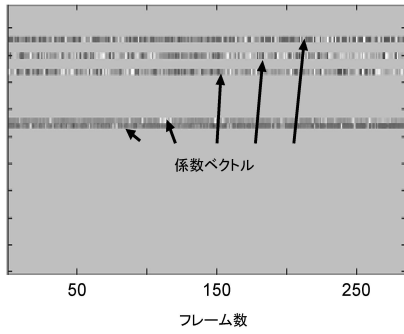
【図8】



【図10】



【図9】



【図11】

	7/7	7/8	7/9	7/10	7/11	圧縮率 (Avg.)
クラスタ数/データ数	112/205	146/293	90/270	128/187	374/596	46.2%

【図12】

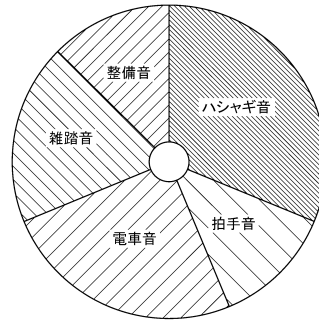
	7/14	7/15	7/16	7/17	7/18	正解率 (Avg.)
誤解答数/データ数	33/289	39/307	12/182	16/221	13/180	91%

【図13】

	7/14	7/15	7/16	7/17	7/18	正解率 (Avg.)
誤解答数/データ数	18/289	25/307	13/182	15/221	11/180	93%

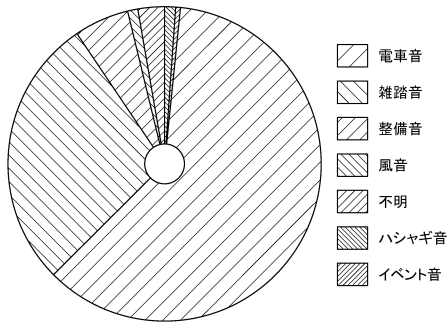
【図15】

音の発生割合(10時から13時)



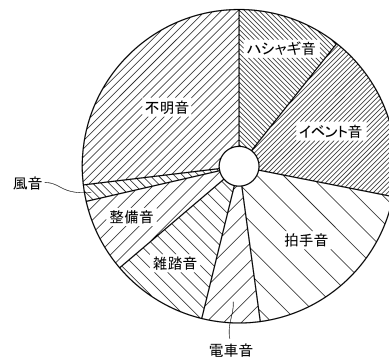
【図14】

音の発生割合(全時間帯)



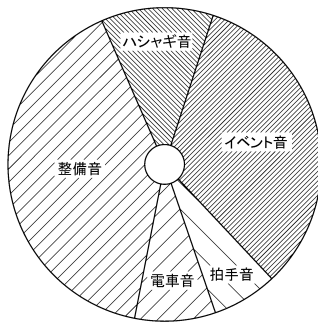
【図16】

音の発生割合(14時から18時)



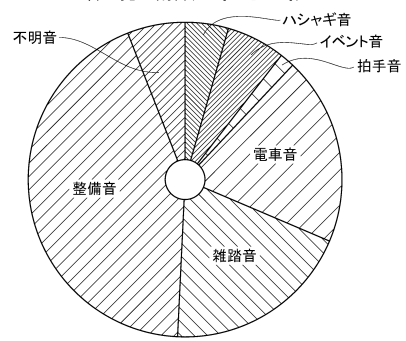
【図17】

音の発生割合(14時から18時)



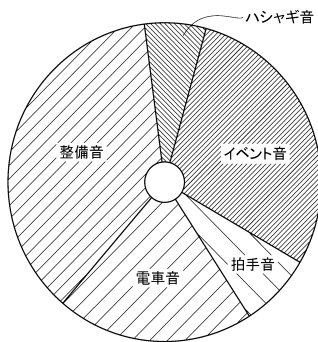
【図19】

音の発生割合(14時から18時)

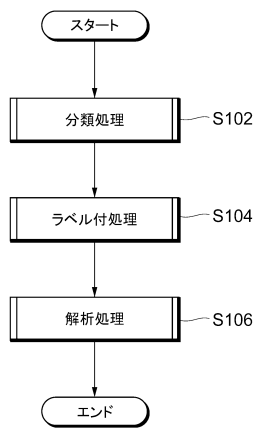


【図18】

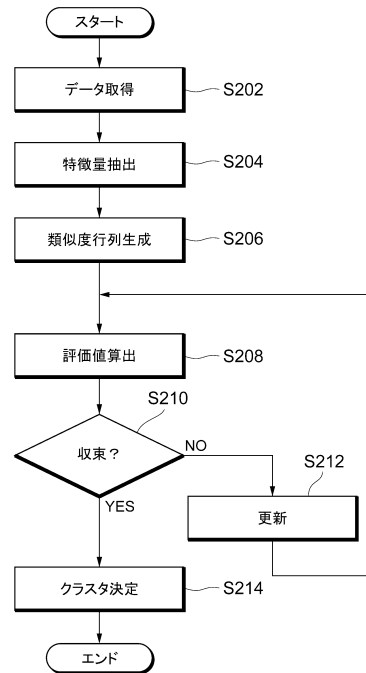
音の発生割合(10時から13時)



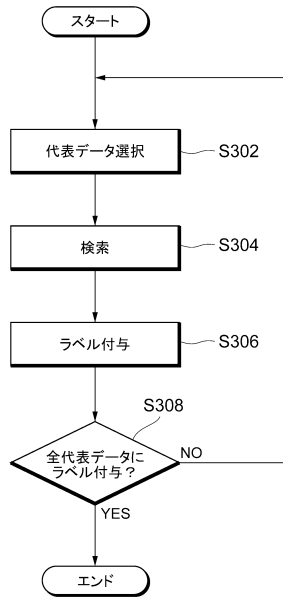
【図20】



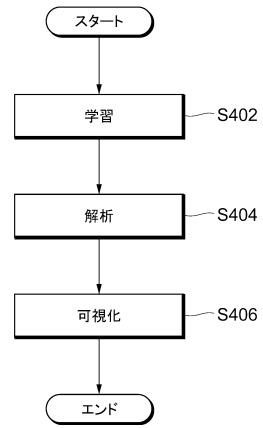
【図21】



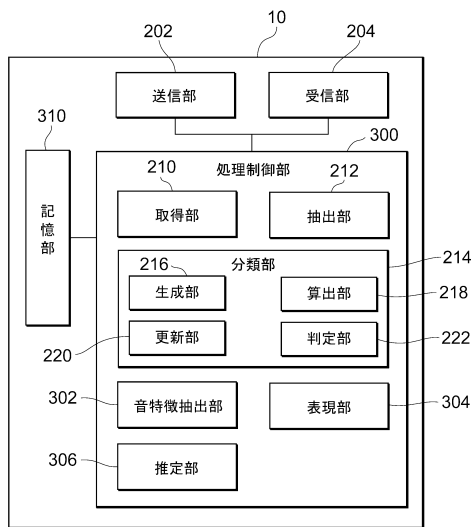
【図 2 2】



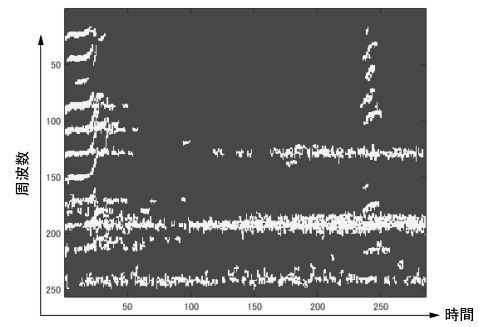
【図 2 3】



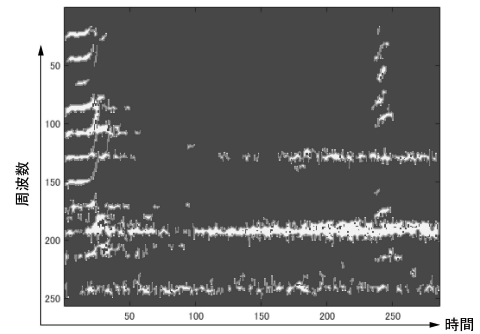
【図 2 4】



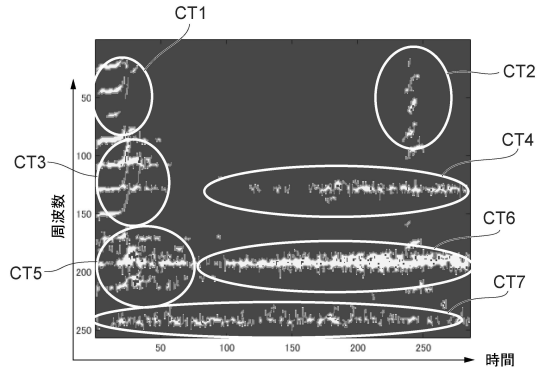
【図 2 5】



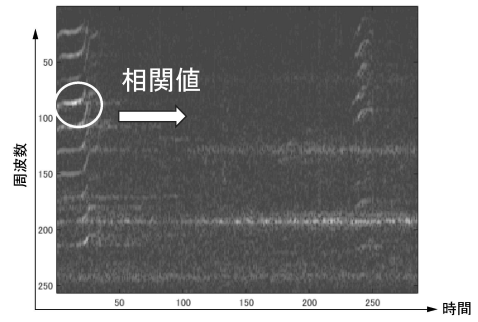
【図 2 6】



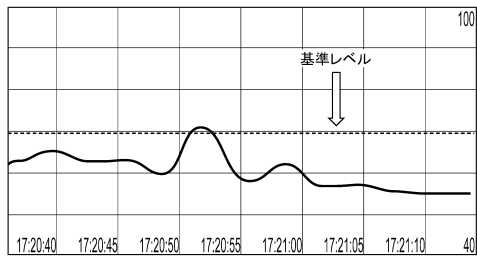
【図 27】



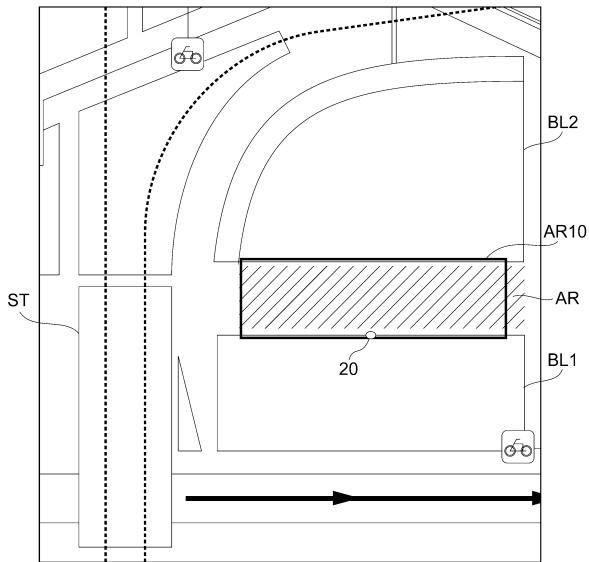
【図 29】



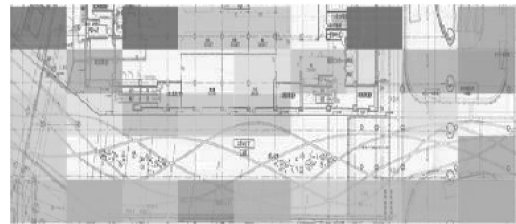
【図 28】



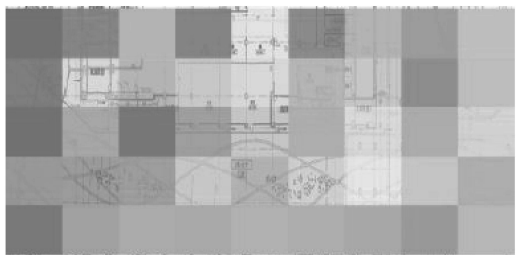
【図 30】



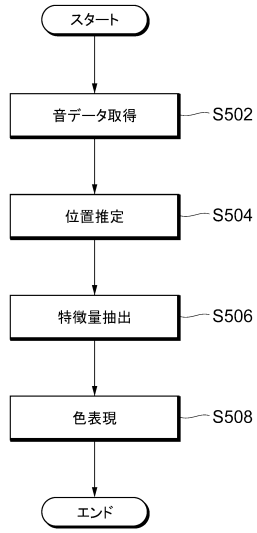
【図 31】



【図 32】



【図 33】



## フロントページの続き

(72)発明者 車谷 浩一

東京都江東区青海2-3-26 国立研究開発法人産業技術総合研究所臨海副都心センター内

(72)発明者 幸島 明男

東京都江東区青海2-3-26 国立研究開発法人産業技術総合研究所臨海副都心センター内

審査官 西村 直史

(56)参考文献 特開2009-182876(JP,A)

米国特許出願公開第2011/0252947(US,A1)

河本 満他, 音環境理解のための音データ解析の効率化手法, 電子情報通信学会2016年総合大会講演論文集, 電子情報通信学会, 2016年 3月 1日, 117頁

田中 希武他, 音環境計測による環境音理解に関する手法, 電子情報通信学会2014年総合大会講演論文集 情報・システム講演論文集2, 日本, 一般社団法人電子情報通信学会, 2014年 3月 4日, 84頁

草間 かおり他, MusCat: 楽曲の印象表現に基づいた一覧表示の一手法, 情報処理学会研究報告, 社団法人情報処理学会, 2009年 8月15日, Vo.2009-MUS-81 No.19, 1-6頁

(58)調査した分野(Int.Cl., DB名)

G06F 16/00 - 16/958