

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7314221号  
(P7314221)

(45)発行日 令和5年7月25日(2023.7.25)

(24)登録日 令和5年7月14日(2023.7.14)

(51)国際特許分類 F I  
G 0 6 F 16/683(2019.01) G 0 6 F 16/683

請求項の数 20 (全22頁)

(21)出願番号	特願2021-148570(P2021-148570)	(73)特許権者	505205812 ネイバー コーポレーション NAVER Corporation 大韓民国 13561 キョンギ-ドソ ンナム-シ ブンダン-グ ブルジョン- ロ 6 グリーンファクトリー
(22)出願日	令和3年9月13日(2021.9.13)	(74)代理人	100107766 弁理士 伊東 忠重
(65)公開番号	特開2022-48130(P2022-48130A)	(74)代理人	100070150 弁理士 伊東 忠彦
(43)公開日	令和4年3月25日(2022.3.25)	(74)代理人	100135079 弁理士 宮崎 修
審査請求日	令和3年9月13日(2021.9.13)	(72)発明者	バク ジョンウン 大韓民国 13561 キョンギ-ドソ ンナム-シ ブンダン-グ ブルジョンロ 最終頁に続く
(31)優先権主張番号	10-2020-0117628		
(32)優先日	令和2年9月14日(2020.9.14)		
(33)優先権主張国・地域又は機関	韓国(KR)		

(54)【発明の名称】 音源を検出するための電子装置およびその作動方法

(57)【特許請求の範囲】

【請求項1】

電子装置の作動方法であって、前記電子装置のプロセッサが、  
マルチメディアコンテンツのフィンガープリントを予め設定された時間間隔によって複数の検索区間に分割する段階、  
前記マルチメディアコンテンツの検索区間のフィンガープリントを、音源のフィンガープリントとマッチングする段階、

前記検索区間のうちの少なくとも1つがマッチングされる検出区間を有する少なくとも1つの音源を検出する段階、

前記マルチメディアコンテンツ内の前記検出区間の時間位置および前記音源内の前記検出区間の時間位置を示す位置情報を決定する段階、および

前記音源と関連する情報および前記位置情報を提供する段階  
を実行する、電子装置の作動方法。

10

【請求項2】

前記電子装置のプロセッサが、  
前記マルチメディアコンテンツの開始点から前記マルチメディアコンテンツ内の前記検出区間の開始点までの時間オフセットと前記音源の開始点から前記音源内の前記検出区間の開始点までの時間オフセットとのオフセット差に基づいて、前記マルチメディアコンテンツの前記フィンガープリントに対して前記音源内の前記検出区間を整理させる段階、および前記マルチメディアコンテンツの検索区間の前記フィンガープリントと前記音源内の前記

20

検出区間のフィンガープリントとを比較して、前記音源の信頼度を検出する段階をさらに実行する、請求項 1 に記載の電子装置の作動方法。

【請求項 3】

前記音源と関連する情報および前記位置情報を提供する段階は、前記音源と関連する情報および前記位置情報とともに、前記音源の信頼度を提供する段階を含む、請求項 2 に記載の電子装置の作動方法。

【請求項 4】

前記信頼度を検出する段階は、前記マルチメディアコンテンツの検索区間の前記フィンガープリントと前記音源内の前記検出区間のフィンガープリントとの比較演算により、前記検出区間に対して比較区間を生成する段階、

10

前記比較区間を複数のビット区間に分割する段階、前記ビット区間のビットエラーレートをそれぞれ計算する段階、前記ビットエラーレートを前記ビット区間の点数にそれぞれ変換する段階、および前記点数の和から前記信頼度を検出する段階を含む、前記比較演算は排他的論理和である、請求項 2 に記載の電子装置の作動方法。

【請求項 5】

前記点数に変換する段階は、前記ビットエラーレートのうちの少なくとも 1 つが閾値を超過すれば、前記少なくとも 1 つのビットエラーレートに対応して 0 を点数として付与する段階、および前記ビットエラーレートのうちの残りが前記閾値以下であれば、前記残りのビットエラーレートに対応して計算される点数を付与する段階を含む、請求項 4 に記載の電子装置の作動方法。

20

【請求項 6】

前記音源と関連する情報および前記位置情報を提供する段階は、前記マルチメディアコンテンツが複数の音源と関連付いていれば、前記音源と関連する情報および前記位置情報をリストで提供する段階を含む、請求項 1 に記載の電子装置の作動方法。

【請求項 7】

前記マルチメディアコンテンツは、オーディオデータを含む、請求項 1 に記載の電子装置の作動方法。

30

【請求項 8】

前記電子装置のプロセッサが、前記音源と関連する情報が選択されれば、前記音源を提供する段階、または前記音源と関連する情報が選択されれば、前記音源と関連付いた他のマルチメディアコンテンツを提供する段階のうち少なくとも 1 つをさらに実行する、請求項 1 に記載の電子装置の作動方法。

【請求項 9】

前記検索区間のうちの少なくとも 1 つは、複数の音源にマッチングされる、請求項 1 に記載の電子装置の作動方法。

40

【請求項 10】

請求項 1 または請求項 9 のうちのいずれか一項に記載の電子装置の作動方法を前記電子装置に実行させる、コンピュータプログラム。

【請求項 11】

請求項 1 または請求項 9 のうちのいずれか一項に記載の方法を前記電子装置に実行させるためのプログラムが記録されている、非一時的なコンピュータ読み取り可能な記録媒体。

【請求項 12】

電子装置であって、メモリ、および

50

前記メモリに連結され、前記メモリに記録された少なくとも1つの命令を実行するように構成されたプロセッサを含み、

前記プロセッサは、

マルチメディアコンテンツのフィンガープリントを予め設定された時間間隔によって複数の検索区間に分割し、

前記マルチメディアコンテンツの検索区間のフィンガープリントを、音源のフィンガープリントとマッチングし、

前記検索区間のうちの少なくとも1つがマッチングされる検出区間を有する少なくとも1つの音源を検出し、

前記マルチメディアコンテンツ内の前記検出区間の時間位置および前記音源内の前記検出区間の時間位置を示す位置情報を決定し、

前記音源と関連する情報および前記位置情報を提供するように構成される、装置。

【請求項13】

前記プロセッサは、

前記マルチメディアコンテンツの開始点から前記マルチメディアコンテンツ内の前記検出区間の開始点までの時間オフセットと前記音源の開始点から前記音源内の前記検出区間の開始点までの時間オフセットとのオフセット差に基づいて、前記マルチメディアコンテンツの前記フィンガープリントに対して前記音源内の前記検出区間を整理させ、

前記マルチメディアコンテンツの検索区間の前記フィンガープリントと前記音源内の前記検出区間のフィンガープリントとを比較して、前記音源の信頼度を検出するように構成される、

請求項12に記載の装置。

【請求項14】

前記プロセッサは、

前記検出区間に対応する前記音源の信頼度を検出し、

前記音源と関連する情報および前記位置情報とともに、前記信頼度を提供するように構成される、

請求項13に記載の装置。

【請求項15】

前記プロセッサは、

前記マルチメディアコンテンツの検索区間の前記フィンガープリントと前記音源内の前記検出区間のフィンガープリントとの比較演算により、前記検出区間に対して比較区間を生成し、

前記比較区間を複数のビット区間に分割し、

前記ビット区間のビットエラーレートをそれぞれ計算し、

前記ビットエラーレートを前記ビット区間の点数にそれぞれ変換し、

前記点数の和から前記信頼度を検出するように構成され、前記比較演算は排他的論理和である、

請求項13に記載の装置。

【請求項16】

前記プロセッサは、

前記ビットエラーレートのうちの少なくとも1つが閾値を超過すれば、前記少なくとも1つのビットエラーレートに対応して0を点数として付与し、

前記ビットエラーレートのうちの残りが前記閾値以下であれば、前記残りのビットエラーレートに対応して計算される点数を付与するように構成される、

請求項15に記載の装置。

【請求項17】

前記プロセッサは、

前記マルチメディアコンテンツが複数の音源と関連付いていれば、前記音源と関連する

10

20

30

40

50

情報および前記位置情報をリストで提供するように構成される、  
請求項 1 2 に記載の装置。

【請求項 1 8】

前記マルチメディアコンテンツは、  
オーディオデータを含む、  
請求項 1 2 に記載の装置。

【請求項 1 9】

前記プロセッサは、  
前記音源と関連する情報が選択されれば、前記音源または前記音源と関連付いた他のマルチメディアコンテンツのうちの少なくとも 1 つを提供するように構成される、  
請求項 1 2 に記載の装置。

10

【請求項 2 0】

前記検索区間のうちの少なくとも 1 つは、  
複数の音源にマッチングされる、  
請求項 1 2 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

多様な実施形態は、マルチメディアコンテンツ ( multimedia content ) に使用された少なくとも 1 つの音源 ( audio source ) を検出するための電子装置およびその作動方法に関する。

20

【背景技術】

【0 0 0 2】

音源検出技術とは、マルチメディアコンテンツに使用された音源を検出する技術である。一般的に、サーバには、複数の音源が登録されており、音源のフィンガープリント ( fingerprint ) がそれぞれ記録されている。このようなサーバは、音源検出技術を利用して、マルチメディアコンテンツのフィンガープリントに基づいて、登録された音源からマルチメディアコンテンツに使用された音源を検出する。これにより、サーバは、音源に関する情報と、音源内でマルチメディアコンテンツに使用された部分の開始位置を提供する。

30

【0 0 0 3】

しかし、このようなサーバは、マルチメディアコンテンツに使用された音源を検出するための動作性能が低いという問題を抱えている。具体的に、サーバが、マルチメディアコンテンツの全体のフィンガープリントと登録された音源のフィンガープリントとを比較しなければならぬため、サーバの演算量が増加し、サーバの動作効率性が低下する。さらに、サーバが、音源内からマルチメディアコンテンツに使用された部分を正確に検出することに困難がある。

【発明の概要】

【発明が解決しようとする課題】

【0 0 0 4】

多様な実施形態は、マルチメディアコンテンツに使用された少なくとも 1 つの音源を効率的に検出することができる、電子装置およびその作動方法を提供する。

40

【0 0 0 5】

多様な実施形態は、マルチメディアコンテンツのフィンガープリントを部分的に利用して、音源内でマルチメディアコンテンツに使用された部分を効率的に検出することができる、電子装置およびその作動方法を提供する。

【0 0 0 6】

多様な実施形態は、音源内でマルチメディアコンテンツに使用された部分に対し、音源内の時間位置だけでなく、マルチメディアコンテンツ内の時間位置を検出することにより、音源内でマルチメディアコンテンツに使用された部分をより正確に特定することができ

50

る、電子装置およびその作動方法を提供する。

【0007】

多様な実施形態は、音源内でマルチメディアコンテンツに使用された部分の音源内の時間位置とマルチメディアコンテンツ内の時間位置に基づいて、音源に対する信頼度を検出することができる、電子装置およびその作動方法を提供する。

【課題を解決するための手段】

【0008】

多様な実施形態に係る電子装置の作動方法は、前記電子装置のプロセッサが、マルチメディアコンテンツのフィンガープリントを予め設定された時間間隔によって複数の検索区間に分割する段階、前記検索区間のうちの少なくとも1つがマッチングされる検出区間を有する少なくとも1つの音源を検出する段階、前記マルチメディアコンテンツ内の前記検出区間の時間位置および前記音源内の前記検出区間の時間位置を示す位置情報を決定する段階、および前記音源と関連する情報および前記位置情報を提供する段階を含んでよい。

10

【0009】

多様な実施形態に係るコンピュータプログラムは、前記作動方法を前記電子装置に実行させるために非一時的なコンピュータ読み取り可能な記録媒体に記録されてよい。

【0010】

多様な実施形態に係る非一時的なコンピュータ読み取り可能な記録媒体は、前記作動方法を前記電子装置に実行させるためのプログラムが記録されている。

【0011】

多様な実施形態に係る電子装置は、メモリ、および前記メモリに連結され、前記メモリに記録された少なくとも1つの命令を実行するように構成されたプロセッサを含み、前記プロセッサは、マルチメディアコンテンツのフィンガープリントを予め設定された時間間隔によって複数の検索区間に分割し、前記検索区間のうちの少なくとも1つがマッチングされる検出区間を有する少なくとも1つの音源を検出し、前記マルチメディアコンテンツ内の前記検出区間の時間位置および前記音源内の前記検出区間の時間位置を示す位置情報を決定し、前記音源と関連する情報および前記位置情報を提供するよう構成されてよい。

20

【発明の効果】

【0012】

多様な実施形態によると、電子装置は、マルチメディアコンテンツに使用された少なくとも1つの音源を効率的に検出することができる。具体的に、電子装置は、マルチメディアコンテンツのフィンガープリントの検索区間のうちの1つから時間範囲を拡張させながら、音源内でマルチメディアコンテンツにマッチングされる検出区間を効率的に検出することができる。また、電子装置は、音源内の検出区間の時間位置だけでなく、マルチメディアコンテンツ内の検出区間の時間位置を検出することにより、音源およびマルチメディアコンテンツ内で検出区間をより正確に特定することができる。さらに、電子装置は、検出区間に対するマルチメディアコンテンツの開始点からの時間オフセットと音源の開始点からの時間オフセットのオフセット差に基づいてマルチメディアコンテンツと音源とを比較することにより、音源に対する信頼度を検出することができる。これにより、電子装置は、利用者のために、音源と関連する情報と位置情報だけでなく、信頼度を提供することができる。

30

40

【図面の簡単な説明】

【0013】

【図1】多様な実施形態における、電子装置を示した図である。

【図2】図1のプロセッサの動作特徴を説明するための例示図である。

【図3a】図1のプロセッサの動作特徴を説明するための例示図である。

【図3b】図1のプロセッサの動作特徴を説明するための例示図である。

【図3c】図1のプロセッサの動作特徴を説明するための例示図である。

【図3d】図1のプロセッサの動作特徴を説明するための例示図である。

【図4】図1のプロセッサを詳しく示した図である。

50

【図5】多様な実施形態における、電子装置の作動方法を示した図である。

【図6】図5の音源の信頼度検出段階を詳しく示した図である。

【図7】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図8】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図9】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図10】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図11】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図12】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図13】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図14】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

10

【図15】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図16】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【図17】多様な実施形態における、電子装置の作動方法を説明するための例示図である。

【発明を実施するための形態】

【0014】

以下、本文書の多様な実施形態について、添付の図面を参照しながら説明する。

【0015】

図1は、多様な実施形態における、電子装置100を示した図である。図2および図3 a~dは、図1のプロセッサ160の動作特徴を説明するため例示図である。図4は、図1のプロセッサ160を詳しく示した図である。

20

【0016】

図1を参照すると、多様な実施形態に係る電子装置100は、連結端子110、通信モジュール120、入力モジュール130、出力モジュール140、メモリ150、またはプロセッサ160のうちの少なくともいずれか1つを含んでよい。一実施形態によっては、電子装置100の構成要素のうちの少なくともいずれか1つが省略されても、少なくとも1つの他の構成要素が追加されてもよい。一実施形態によっては、電子装置100の構成要素のうちの少なくともいずれか2つが、1つの統合された回路で実現されてよい。例えば、電子装置100は、サーバ(server)、スマートフォン(smart phone)、携帯電話、ナビゲーション、PC、ノート型PC、デジタル放送用端末、PDA(personal digital assistants)、PMP(portable multimedia player)、タブレット、ゲームコンソール(game console)、ウェアラブルデバイス(wearable device)、IoT(internet of things)デバイス、家電機器、医療機器、またはロボット(robot)のうちの少なくともいずれか1つを含んでよい。

30

【0017】

連結端子110は、電子装置100で外部装置102と物理的に連結されてよい。例えば、外部装置102は、他の電子装置を含んでよい。このために、連結端子110は、少なくとも1つのコネクタを含んでよい。例えば、コネクタは、HDMIコネクタ、USBコネクタ、SDカードコネクタ、またはオーディオコネクタのうちの少なくともいずれか1つを含んでよい。

40

【0018】

通信モジュール120は、電子装置100で外部装置102、104との通信を実行してよい。通信モジュール120は、電子装置100と外部装置102、104との間に通信チャンネルを樹立し、通信チャンネルを介して外部装置102、104との通信を実行してよい。ここで、外部装置102、104は、衛星、基地局、または他の電子装置のうちの少なくともいずれか1つを含んでよい。通信モジュール120は、有線通信モジュールまたは無線通信モジュールのうちの少なくともいずれか1つを含んでよい。有線通信モジュールは、連結端子102を介して外部装置102と有線で接続し、有線で通信してよい。無線通信モジュールは、近距離通信モジュールまたは遠距離通信モジュールのうちの少なくともいずれか1つを含んでよい。近距離通信モジュールは、外部装置102と近距離

50

通信方式で通信してよい。例えば、近距離通信方式は、ブルートゥース (Bluetooth)、Wi-Fi Direct、または赤外線通信 (IrDA: infrared data association) のうちの少なくともいずれか1つを含んでよい。遠距離通信モジュールは、外部装置104と遠距離通信方式で通信してよい。ここで、遠距離通信モジュールはネットワーク190を介して外部装置104と通信してよい。例えば、ネットワーク190は、セルラネットワーク、インターネット、またはLAN (local area network) やWAN (wide area network) のようなコンピューターネットワークのうちの少なくともいずれか1つを含んでよい。

**【0019】**

入力モジュール130は、電子装置100の少なくとも1つの構成要素に使用される信号を入力してよい。入力モジュール130は、利用者が電子装置100に信号を直接入力するように構成される入力装置、周辺環境を感知して信号を発生するように構成されるセンサ装置、または画像を撮影して画像データを生成するように構成されるカメラモジュールのうちの少なくともいずれか1つを含んでよい。例えば、入力装置は、マイクロフォン (microphone)、マウス (mouse)、またはキーボード (keyboard) のうちの少なくともいずれか1つを含んでよい。一実施形態において、センサ装置は、タッチを感知するように設定されたタッチ回路 (touch circuitry)、またはタッチによって発生する力の強度を測定するように設定されたセンサ回路のうちの少なくともいずれか1つを含んでよい。

**【0020】**

出力モジュール140は、情報を出力してよい。出力モジュール140は、情報を視覚的に表示するように構成される表示モジュール、または情報を聴覚的に再生するように構成されるオーディオモジュールのうちの少なくとも1つを含んでよい。例えば、表示モジュールは、ディスプレイ、ホログラム装置、またはプロジェクタのうちの少なくともいずれか1つを含んでよい。一例として、表示モジュールは、入力モジュール130のタッチ回路またはセンサ回路のうちの少なくともいずれか1つと組み立てられて、タッチスクリーンとして実現されてよい。例えば、オーディオモジュールは、スピーカまたはレシーバのうちの少なくともいずれか1つを含んでよい。

**【0021】**

メモリ150は、電子装置100の少なくとも1つの構成要素によって使用される多様なデータを記録してよい。例えば、メモリ150は、揮発性メモリまたは不揮発性メモリのうちの少なくともいずれか1つを含んでよい。データは、少なくとも1つのプログラム、およびこれと関連する入力データまたは出力データを含んでよい。プログラムは、メモリ150に少なくとも1つの命令を含むソフトウェアとして記録されてよく、例えば、オペレーティングシステム、ミドルウェア、またはアプリケーションのうちの少なくともいずれか1つを含んでよい。

**【0022】**

プロセッサ160は、メモリ150のプログラムを実行して、電子装置100の少なくとも1つの構成要素を制御してよい。これにより、プロセッサ160は、データ処理または演算を実行してよい。このとき、プロセッサ160は、メモリ150に記録された命令を実行してよい。プロセッサ160は、マルチメディアコンテンツ (multimedia content) に使用された少なくとも1つの音源 (audio source) を検出しようとしてよい。ここで、マルチメディアコンテンツは、画像データまたはオーディオデータのうちの少なくとも1つで構成されてよい。一例として、マルチメディアコンテンツは、画像データとオーディオデータで構成され、ミュージックビデオやネットワークを介して共有される動画などを含んでよい。他の例として、マルチメディアコンテンツは、オーディオデータで構成され、ポッドキャスト、放送局などで生成されてよい。また、図2に示すように、マルチメディアコンテンツのオーディオデータには少なくとも1つの音源が使用されてよく、各音源の少なくとも一部が使用されてよい。

**【0023】**

10

20

30

40

50

多様な実施形態によると、プロセッサ160は、マルチメディアコンテンツに対応して、音源の少なくとも一部を検出区間200として検出してよい。このとき、検出区間200に対するオフセット差が定義されてよい。オフセット( $T_m - T_a$ )は、マルチメディアコンテンツの開始点( $T_{m0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_m$ )と音源の開始点( $T_{a0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_a$ )との差を示してよい。ここで、オフセット差( $T_m - T_a$ )としてある値が定義されてよく、一定の範囲内の値が定義されてもよい。一例として、オフセット差( $T_m - T_a$ )は、マルチメディアコンテンツの開始点( $T_{m0}$ )からの時間オフセット( $T_m$ )と音源の開始点( $T_{a0}$ )からの時間オフセット( $T_a$ )との差を中心とする範囲内の値が定義されてよい。オフセット差( $T_m - T_a$ )として一定の範囲内の値が定義される場合、同じ音源に対する多様な再生速度が考慮されてよい。

10

#### 【0024】

第1例として、図3aに示すように、検出区間200は、音源の全体領域であってよく、マルチメディアコンテンツの一部領域に使用されてよい。ここで、音源の開始点( $T_{a0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_a$ )は0であるため、オフセット差( $T_m - T_a$ )は、マルチメディアコンテンツの開始点( $T_{m0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_m$ )が決定されてよい(オフセット差 =  $T_m$ )。第2例として、図3bに示すように、検出区間200は、音源の一部領域であってよく、マルチメディアコンテンツの一部領域に使用されてよい。ここで、マルチメディアコンテンツの開始点( $T_{m0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_m$ )であるため、オフセット差( $T_m - T_a$ )は、音源の開始点( $T_{a0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_a$ )から決定されてよい(オフセット差 =  $-T_a$ )。第3例として、図3cに示すように、検出区間200は、音源の一部領域であってよく、マルチメディアコンテンツの一部領域に使用されてよい。ここで、音源の開始点( $T_{a0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_a$ )は0であるため、オフセット差( $T_m - T_a$ )は、マルチメディアコンテンツの開始点( $T_{m0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_m$ )が決定されてよい(オフセット差 =  $T_m$ )。第4例として、図3dに示すように、検出区間200は、音源の一部領域であってよく、マルチメディアコンテンツの全体領域に使用されてよい。ここで、マルチメディアコンテンツの開始点( $T_{m0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_m$ )であるため、オフセット差( $T_m - T_a$ )は、音源の開始点( $T_{a0}$ )から検出区間200の開始点( $T_{d0}$ )までの時間オフセット( $T_a$ )から決定されてよい(オフセット差 =  $-T_a$ )。

20

30

#### 【0025】

多様な実施形態によると、プロセッサ160は、オフセット差( $T_m - T_a$ )を利用して、音源の信頼度(*confidence*)を検出してよい。信頼度は、検出された音源がマルチメディアコンテンツに使用されたものであるかに対する正確度を示すものであり、信頼度が高いほど正確度が高くてよい。具体的に、プロセッサ160は、オフセット差( $T_m - T_a$ )に基づいて、マルチメディアコンテンツに対して検出区間200を整理させてよい。その後、プロセッサ160は、マルチメディアコンテンツと検出区間200とを比較して、音源の信頼度を検出してよい。一実施形態によると、プロセッサ160は、ビット演算により、音源の信頼度を検出してよい。プロセッサ160は、マルチメディアコンテンツのフィンガープリントと検出区間200のフィンガープリントとの比較演算により、検出区間200のビットエラーレート(*bit error rate: BER*)を計算してよい。プロセッサ160は、ビットエラーレートを点数(*score*)にそれぞれ変換してよい。プロセッサ160は、予め定められたスコア関数(*score function*)を利用して、ビットエラーレートを点数にそれぞれ変換してよい。プロセッサ160は、点数の和から信頼度を検出してよい。プロセッサ160は、予め定められたコンフィデンス関数(*confidence function*)を利用して、点

40

50



数の和から信頼度を検出してよい。

【0026】

多様な実施形態によると、プロセッサ160は、図4に示すように、API(application programming interface)461、プロセスAPI(process-API)463、制御部465、コンテンツ取得部467、フィンガープリント部469、マッチング部471、比較部473、またはクラスタリング部475のうちの少なくとも1つを含んでよい。一実施形態によっては、プロセッサ160の構成要素のうちの少なくともいずれか1つが省略されても、少なくとも1つの他の構成要素が追加されてもよい。一実施形態によっては、プロセッサ160の構成要素のうちの少なくともいずれか2つが、1つの統合された回路で実現されてよい。

10

【0027】

API461は、利用者の要請を検出してよい。プロセスAPI463は、利用者の要請に基づいて、命令語を生成してよい。制御部465は、プロセッサ160の構成要素のうちの少なくとも1つを制御してよい。このとき、制御部465は、プロセッサ160の構成要素のうちの少なくとも2つを仲介する役割を実行してよく、プロセッサ160の構成要素のうちの少なくとも1つのための作業を実行してよい。コンテンツ取得部467は、命令語に基づいて、マルチメディアコンテンツを取得してよい。フィンガープリント部469は、マルチメディアコンテンツのフィンガープリントを取得してよい。このとき、フィンガープリント部469は、マルチメディアコンテンツのオーディオデータからフィンガープリントを直接抽出してよい。マッチング部471は、マルチメディアコンテンツのフィンガープリントに基づいて、少なくとも1つの音源を検出してよい。このとき、メモリ150には、複数の音源が予め登録されており、登録された音源のフィンガープリントがそれぞれ記録されていてよい。マッチング部471は、マルチメディアコンテンツのフィンガープリントと登録された音源のフィンガープリントをマッチングさせることにより、登録された音源のフィンガープリントのうちの少なくとも1つを検出してよい。比較部473は、マルチメディアコンテンツのフィンガープリントと検出された音源のフィンガープリントとを比較して、検出された音源の信頼度を検出してよい。クラスタリング部475は、検出された音源に基づいて、マルチメディアコンテンツに対する比較対象またはマルチメディアコンテンツの比較結果のうちの少なくとも1つを、検出された音源と同一あるいは類似の音源を包括するように拡張させてよい。具体的に、クラスタリング部475は、検出された音源と同一あるいは類似の音源の情報を取得し、マルチメディアコンテンツに対する比較対象を、検出された音源と同一あるいは類似の音源に拡張させてよい。一方、クラスタリング部475は、比較部473の比較結果に基づいて、検出された音源と同一あるいは類似の音源をまとめてよい。

20

30

【0028】

図5は、多様な実施形態における、電子装置100の作動方法を示した図である。図6は、図5の音源の信頼度検出段階(段階550)を詳しく示した図である。図7~17は、多様な実施形態における、電子装置100の作動方法を説明するための例示図である。

【0029】

図5を参照すると、段階510で、電子装置100は、マルチメディアコンテンツのフィンガープリント710を複数の検索区間720に分割してよい。ここで、マルチメディアコンテンツは、画像データまたはオーディオデータのうちの少なくとも1つで構成されてよい。一例として、マルチメディアコンテンツは、画像データとオーディオデータで構成され、ミュージックビデオ、ネットワークを介して共有される動画などを含んでよい。他の例として、マルチメディアコンテンツは、オーディオデータで構成され、ポッドキャストや放送局などで生成されてよい。また、オーディオデータには、少なくとも1つの音源が使用されてよく、各音源の少なくとも一部が含まれてよい。プロセッサ160は、マルチメディアコンテンツのフィンガープリント710を取得してよい。一実施形態によると、プロセッサ160は、マルチメディアコンテンツのオーディオデータからフィンガープリント710を直接抽出してよい。例えば、利用者によってマルチメディアコンテンツ

40

50

が選択されれば、プロセッサ 160 は、マルチメディアコンテンツのオーディオデータからフィンガープリント 710 を抽出してよい。他の実施形態によると、プロセッサ 160 は、外部装置 102、104 からマルチメディアコンテンツのフィンガープリント 710 をクエリとして受信してよい。ここで、フィンガープリントは、オーディオデータに対する時間による周波数分布を示してよい。プロセッサ 160 は、図 7 に示すように、マルチメディアのフィンガープリント 710 を予め設定された時間間隔によって複数の検索区間 720 に分割してよい。一例として、時間間隔は、数秒の範囲内で定められてよい。

#### 【0030】

段階 520 で、電子装置 100 は、検索区間 720 のうちの少なくとも 1 つがマッチングされる検出区間 1010 を有する少なくとも 1 つの音源を検出してよい。このとき、メモリ 150 には、複数の音源が予め登録されており、登録された音源のフィンガープリント 910 がそれぞれ記録されていてよい。プロセッサ 160 は、図 8 に示すように、検索区間 720 のそれぞれを登録された音源のフィンガープリント 910 と比較してよい。これにより、プロセッサ 160 は、検索区間 720 のうちの 1 つに基づいて、登録された音源のフィンガープリント 910 のうちの少なくとも 1 つを検出してよい。このとき、プロセッサ 160 は、図 9 に示すように、検索区間 720 のうちの 1 つから時間範囲を拡張させながら、マルチメディアコンテンツのフィンガープリント 710 と検出された音源のフィンガープリント 910 とを比較してよい。これにより、プロセッサ 160 は、図 10 に示すように、検出された音源のフィンガープリント 910 から、検索区間 720 のうちの少なくとも 1 つがマッチングされる検出区間 1010 を検出できるようになる。

#### 【0031】

段階 530 で、電子装置 100 は、マルチメディアコンテンツおよび検出された音源内で検出区間 1010 の位置情報を決定してよい。位置情報は、マルチメディアコンテンツのフィンガープリント 710 内の検出区間 1010 の時間位置、および検出された音源のフィンガープリント 910 内の検出区間 1010 の時間位置を示してよい。プロセッサ 160 は、検出区間 1010 の位置情報に基づいて、検出区間 1010 に対するオフセット差 ( $T_m - T_a$ ) を検出してよい。オフセット差 ( $T_m - T_a$ ) は、マルチメディアコンテンツのフィンガープリント 710 の開始点 ( $T_{m0}$ ) から検出区間 1010 の開始点 ( $T_{d0}$ ) までの時間オフセット ( $T_m$ ) と、検出された音源のフィンガープリント 910 の開始点 ( $T_{a0}$ ) から検出区間 1010 の開始点 ( $T_{d0}$ ) までの時間オフセット ( $T_a$ ) の差を示してよい。

#### 【0032】

段階 540 で、電子装置 100 は、検出区間 1010 に対するオフセット差 ( $T_m - T_a$ ) を基づいて、マルチメディアコンテンツのフィンガープリント 710 に対して検出区間 1010 を整列させてよい。プロセッサ 160 は、図 10 に示すように、検出区間 1010 と検出区間 1010 にマッチングされた少なくとも 1 つの検索区間 720 とが互いに対応するように、検出区間 1010 を整列させてよい。プロセッサ 160 は、少なくとも 1 つの検索区間 720 の開始点に検出区間 1010 の開始点を整列させてよい。

#### 【0033】

段階 550 で、電子装置 100 は、マルチメディアコンテンツのフィンガープリント 710 と検出区間 1010 とを比較して、検出された音源の信頼度を検出してよい。信頼度は、検出された音源がマルチメディアコンテンツに使用されたものであるかに対する正確度を示すものであり、信頼度が高いほど正確度が高くてよい。プロセッサ 160 は、検出区間 1010 と検出区間 1010 にマッチングされた少なくとも 1 つの検索区間 720 とを比較して、検出された音源の信頼度を検出してよい。一実施形態によると、プロセッサ 160 は、少なくとも 1 つの検索区間 720 に対する検出区間 1010 のビット演算により、検出された音源の信頼度を検出してよい。これについては、図 6 を参照しながらより詳しく説明する。

#### 【0034】

図 6 を参照すると、段階 651 で、電子装置 100 は、マルチメディアコンテンツのフ

インガープリント710と検出区間1010との比較演算により、比較区間1110を生成してよい。一例として、比較演算は、排他的論理和(XOR)を含んでよい。プロセッサ160は、検出区間1010と検出区間1010にマッチングされた少なくとも1つの検索区間720との比較演算により、図11に示すように比較区間1110を生成してよい。

#### 【0035】

段階653で、電子装置100は、比較区間1110を複数のビット区間1210に分割してよい。プロセッサ160は、図12に示すように、予め設定された時間間隔によって比較区間1110を複数のビット区間1210に分割してよい。一例として、時間間隔は、約1秒であってよい。

10

#### 【0036】

段階655で、電子装置100は、ビット区間1210のビットエラーレートをそれぞれ計算してよい。プロセッサ160は、ビット区間1210のそれぞれに対して連続するビットで計算し、ビット区間1210のそれぞれのビットからビットエラーレートを計算してよい。ここで、各ビットエラーレートは、0から1までの値で表現され、ビットエラーレートが低いほど類似性が高くてよい。類似性は、検出区間1010と検出区間1010にマッチングされた少なくとも1つの検索区間720の類似性を示してよい。すなわち、ビットエラーレートが0であるということは、検出区間1010と少なくとも1つの検索区間720が同一であることを意味してよい。一例として、図13に示すようにビット区間1210のビットエラーレートが計算された場合、これは、マルチメディアコンテンツで513~551秒までに検出された音源の検出区間1010が使用されたことを示してよい。他の例として、図14に示すようにビット区間1210のビットエラーレートが計算された場合、これは、マルチメディアコンテンツで複数の音源の検出区間1010が使用され、同じ時間範囲にも複数の音源の検出区間1010が使用されたことを示してよい。このような場合、プロセッサ160は、検出区間1010のビットエラーレートのうちで最も高いビットエラーレートを抽出してよい。

20

#### 【0037】

段階657で、電子装置100は、ビットエラーレートをビット区間1210の点数にそれぞれ変換してよい。プロセッサ160は、予め定められたスコア関数を利用して、ビットエラーレートを点数にそれぞれ変換してよい。ここで、図15に示すように、ビットエラーレートが低いほど高い点数に変換され、ビットエラーレートが高いほど低い点数に変換されてよい。このとき、ビットエラーレートのうちの少なくとも1つが閾値を超過すれば、プロセッサ160は、少なくとも1つのビットエラーレートに対応して0を点数として付与してよい。一方、ビットエラーレートのうちの残りが閾値以下であれば、プロセッサ160は、残りのビットエラーレートに基づいて点数をそれぞれ計算し、残りのビットエラーレートに対応して計算された点数をそれぞれ付与してよい。例えば、スコア関数は、以下の数式(1)のように表現されてよい。

30

#### 【0038】

$$y = \max(0, 10 \cdot ((1 - x))^2 - (10 - \text{閾値})^2) \cdot \dots \cdot (1)$$

40

ここで、 $x$ はビットエラーレートを、 $y$ は点数を示し、閾値は、0超過0.5以下であってよく、一例として、0.35以上0.45以下であってよい。閾値が0に近い値で設定されるほど、音源に対する誤検出の可能性は低下するが、ノイズのある音源に対する検出の可能性は高まる可能性がある。これに反し、閾値が0.5に近い値で設定されるほど、ノイズのある音源に対する検出の可能性が低下するが、音源に対する誤検出の可能性は高まる可能性がある。

#### 【0039】

段階659で、電子装置100は、点数の和から検出された音源に対する信頼度を検出してよい。信頼度は、検出された音源がマルチメディアコンテンツに使用されたものであるかに対する正確度を示すものであり、信頼度が高いほど正確度が高くてよい。プロセッ

50

サ160は、予め定められたコンフィデンス関数を利用して、点数の和から信頼度を検出してよい。ここで、図16に示すように、信頼度は、0から1までの値で表現されてよい。点数の和が一定の範囲内にある場合、点数の和が信頼度に大きく影響を及ぼすため、点数の和が大きいほど信頼度が著しく高くてよい。一方、点数の和が一定の範囲外である場合、点数の和が信頼度に及ぼす影響が減少してよい。例えば、コンフィデンス関数は、以下の数式(2)のように表現されてよい。

【0040】

$$y = \tanh(\text{加重値} \times x) \cdots (2)$$

ここで、 $x$ は点数の和を、 $y$ は信頼度を示し、加重値は、スコア関数の閾値または後述する信頼度に対する基準値のうちの少なくとも1つによって決定されてよく、例えば、0.1以上0.2以下であってよい。

10

【0041】

この後、電子装置100は、図5にリターンして、段階560に進んでよい。

【0042】

再び図5を参照すると、段階560で、電子装置100は、検出された音源と関連する情報、位置情報、および信頼度を提供してよい。音源と関連する情報は、音源の識別子、名称、またはアーティストのうちの少なくとも1つを含んでよい。位置情報は、マルチメディアコンテンツのフィンガープリント710内の検出区間1010の時間位置、および検出された音源のフィンガープリント910内の検出区間1010の時間位置を示してよい。プロセッサ160は、図17に示すように、マルチメディアコンテンツに対応して検出された音源と関連する情報、位置情報、および信頼度を提供してよい。ここで、マルチメディアコンテンツから複数の音源が検出された場合、プロセッサ160は、音源のリストとして、検出された音源と関連する情報、位置情報、および信頼度を提供してよい。一例として、プロセッサ160は、検出された音源の信頼度とは関係なく、検出された音源の関連する情報、位置情報、および信頼度を提供してよい。他の例として、検出された音源の信頼度が基準値以上であれば、プロセッサ160は、検出された音源の関連する情報、位置情報、および信頼度を提供してよい。言い換えれば、検出された音源の信頼度が基準値未満であれば、プロセッサ160は、検出された音源の関連する情報、位置情報、および信頼度を提供しなくてもよい。プロセッサ160は、外部装置102、104からのクエリに対する応答として、検出された音源と関連する情報、位置情報、および信頼度を提供してよい。一実施形態によると、プロセッサ160は、外部装置102、104に、検出された音源と関連する情報、位置情報、および信頼度を送信してよい。他の実施形態によると、プロセッサ160は、出力モジュール140から、検出された音源と関連する情報、位置情報、および信頼度を直接出力してよい。

20

30

【0043】

多様な実施形態によると、利用者は、マルチメディアコンテンツに使用された音源を確認し、これを多様に活用してよい。一例として、マルチメディアコンテンツが放送や公演などの動画である場合、利用者は、マルチメディアコンテンツに使われた音源に基づいて、マルチメディアコンテンツのキューシート(cue sheet)を取得してよい。他の例として、利用者は、マルチメディアコンテンツに使用された音源の著作権保護または著作権精算のために活用してよい。

40

【0044】

多様な実施形態によると、検出された音源と関連する情報、位置情報、および信頼度を提供した後、電子装置100は、検出された音源と関連付いた多様なサービスを提供してよい。一実施形態によると、プロセッサ160は、外部装置102、104に検出された音源を提供してよい。外部装置102、104によって検出された音源と関連する情報が選択されれば、プロセッサ160は、外部装置102、104に検出された音源を提供してよい。他の実施形態によると、プロセッサ160は、検出された音源と関連付いた他のマルチメディアコンテンツを提供してよい。外部装置102、104によって検出された音源と関連する情報が選択されれば、プロセッサ160は、検出された音源と関連する情

50

報に基づいて、他のマルチメディアコンテンツを検索し、外部装置 102、104 に検索されたマルチメディアコンテンツを提供してよい。また他の実施形態によると、プロセッサ 160 は、検出された音源と関連付いた付加情報を提供してよい。外部装置 102、104 によって検出された音源と関連する情報が選択されれば、プロセッサ 160 は、検出された音源と関連する情報に基づいて、例えば、ニュースやソーシャルネットワークサービス (social network service: SNS) などを利用して付加情報を検索して、外部装置 102、104 に検索された付加情報を提供してよい。

#### 【0045】

多様な実施形態によると、電子装置 100 は、マルチメディアコンテンツに使用された少なくとも 1 つの音源を効率的に検出してよい。具体的に、電子装置 100 は、マルチメディアコンテンツのフィンガープリント 710 の検索区間 720 のうちの 1 つから時間範囲を拡張させながら、音源内でマルチメディアコンテンツにマッチングされる検出区間 1010 を効率的に検出してよい。また、電子装置 100 は、音源内の検出区間 1010 の時間位置だけでなく、マルチメディアコンテンツ内の検出区間 1010 の時間位置を検出することにより、音源およびマルチメディアコンテンツ内で検出区間 1010 をより正確に特定してよい。さらに、電子装置 100 は、検出区間 1010 に対するマルチメディアコンテンツの開始点 ( $T_{m0}$ ) からの時間オフセット ( $T_m$ ) と音源の開始点 ( $T_{a0}$ ) からの時間オフセット ( $T_a$ ) のオフセット差 ( $T_m - T_a$ ) に基づいてマルチメディアコンテンツと音源とを比較することにより、音源に対する信頼度を検出してよい。これにより、電子装置 100 は、利用者のために、音源と関連する情報と位置情報だけでなく、信頼度を提供することができる。

#### 【0046】

多様な実施形態に係る電子装置 100 の作動方法は、マルチメディアコンテンツのフィンガープリント 710 を予め設定された時間間隔によって複数の検索区間 720 に分割する段階 (段階 510)、検索区間 720 のうちの少なくとも 1 つがマッチングされる検出区間 1010 を有する少なくとも 1 つの音源を検出する段階 (段階 520)、マルチメディアコンテンツ内の検出区間 1010 の時間位置および音源内の検出区間 1010 の時間位置を示す位置情報を決定する段階 (段階 530)、および音源と関連する情報および位置情報を提供する段階 (段階 560) を含んでよい。

#### 【0047】

多様な実施形態によると、電子装置 100 の作動方法は、マルチメディアコンテンツの開始点 ( $T_{m0}$ ) からの時間オフセット ( $T_m$ ) と音源の開始点 ( $T_{a0}$ ) からの時間オフセット ( $T_a$ ) のオフセット差 ( $T_m - T_a$ ) に基づいて、フィンガープリント 710 に対して検出区間 1010 を整列させる段階 (段階 540)、およびフィンガープリント 710 と検出区間 1010 とを比較して、音源の信頼度を検出する段階 (段階 550) をさらに含んでよい。

#### 【0048】

多様な実施形態によると、音源と関連する情報および位置情報を提供する段階 (段階 560) は、音源と関連する情報および位置情報とともに、信頼度を提供する段階を含んでよい。

#### 【0049】

多様な実施形態によると、信頼度を検出する段階 (段階 550) は、フィンガープリント 710 と検出区間 1010 との比較演算により、比較区間 1110 を生成する段階 (段階 651)、比較区間 1110 を複数のビット区間 1210 に分割する段階 (段階 653)、ビット区間 1210 のビットエラーレートをそれぞれ計算する段階 (段階 655)、ビットエラーレートをビット区間 1210 の点数にそれぞれ変換する段階 (段階 657)、および点数の和から信頼度を検出する段階 (段階 659) を含んでよい。

#### 【0050】

多様な実施形態によると、点数に変換する段階 (段階 657) は、ビットエラーレートのうちの少なくとも 1 つが閾値を超過すれば、少なくとも 1 つのビットエラーレートに対

10

20

30

40

50

応して0を点数として付与する段階、およびビットエラーレートのうちの残りが閾値以下であれば、残りのビットエラーレートに対応して計算される点数を付与する段階を含んでよい。

【0051】

多様な実施形態によると、音源と関連する情報および位置情報を提供する段階（段階560）は、マルチメディアコンテンツが複数の音源と関連付いていれば、音源と関連する情報および位置情報をリストで提供する段階を含んでよい。

【0052】

多様な実施形態によると、マルチメディアコンテンツは、画像データまたはオーディオデータのうちの少なくとも1つで構成されてよい。

10

【0053】

多様な実施形態によると、電子装置100の作動方法は、音源と関連する情報が選択されれば、音源を提供する段階、または音源と関連する情報が選択されれば、音源と関連付いた他のマルチメディアコンテンツを提供する段階のうちの少なくとも1つをさらに含んでよい。

【0054】

多様な実施形態によると、検索区間720のうちの少なくとも1つは、複数の音源にマッチングされてよい。

【0055】

多様な実施形態に係る電子装置100は、メモリ150、およびメモリ150に連結され、メモリ150に記録された少なくとも1つの命令を実行するように構成されたプロセッサ160を含んでよい。

20

【0056】

多様な実施形態によると、プロセッサ160は、マルチメディアコンテンツのフィンガープリント710を予め設定された時間間隔によって複数の検索区間720に分割し、検索区間720のうちの少なくとも1つがマッチングされる検出区間1010を有する少なくとも1つの音源を検出し、マルチメディアコンテンツ内の検出区間1010の時間位置および音源内の検出区間1010の時間位置を示す位置情報を決定し、音源と関連する情報および位置情報を提供するよう構成されてよい。

【0057】

多様な実施形態によると、プロセッサ160は、マルチメディアコンテンツの開始点（ $T_{m0}$ ）からの時間オフセット（ $T_m$ ）と音源の開始点（ $T_{a0}$ ）からの時間オフセット（ $T_a$ ）のオフセット差（ $T_m - T_a$ ）に基づいて、フィンガープリント710に対して検出区間1010を整列させ、フィンガープリント710と検出区間1010とを比較して、音源の信頼度を検出するように構成されてよい。

30

【0058】

多様な実施形態によると、プロセッサ160は、検出区間1010に対応する音源の信頼度を検出し、音源と関連する情報および位置情報とともに、信頼度を提供するように構成されてよい。

【0059】

多様な実施形態によると、プロセッサ160は、フィンガープリント710と検出区間1010との比較演算により、比較区間1110を生成し、比較区間1110を複数のビット区間1210に分割し、ビット区間1210のビットエラーレートをそれぞれ計算し、ビットエラーレートをビット区間1210の点数にそれぞれ変換し、点数の和から信頼度を検出するように構成されてよい。

40

【0060】

多様な実施形態によると、プロセッサ160は、ビットエラーレートのうちの少なくとも1つが閾値を超過すれば、少なくとも1つのビットエラーレートに対応して0を点数として付与し、ビットエラーレートのうちの残りが閾値以下であれば、残りのビットエラーレートに対応して計算される点数を付与するように構成されてよい。

50

## 【 0 0 6 1 】

多様な実施形態によると、プロセッサ 1 6 0 は、マルチメディアコンテンツが複数の音源と関連付いていれば、音源と関連する情報および位置情報をリストで提供するように構成されてよい。

## 【 0 0 6 2 】

多様な実施形態によると、マルチメディアコンテンツは、画像データまたはオーディオデータのうちの少なくとも 1 つで構成されてよい。

## 【 0 0 6 3 】

多様な実施形態によると、プロセッサ 1 6 0 は、音源と関連する情報が選択されれば、音源または音源と関連付けられた他のマルチメディアコンテンツのうちの少なくとも 1 つを提供するように構成されてよい。

10

## 【 0 0 6 4 】

多様な実施形態によると、検索区間 7 2 0 のうちの少なくとも 1 つは、複数の音源にマッチングされてよい。

## 【 0 0 6 5 】

上述した装置は、ハードウェア構成要素、ソフトウェア構成要素、および/またはハードウェア構成要素とソフトウェア構成要素との組み合わせによって実現されてよい。例えば、実施形態で説明された装置および構成要素は、プロセッサ、コントローラ、ALU ( arithmetic logic unit )、デジタル信号プロセッサ、マイクロコンピュータ、FPGA ( field programmable gate array )、PLU ( programmable logic unit )、マイクロプロセッサ、または命令を実行して応答することができる様々な装置のように、1 つ以上の汎用コンピュータまたは特殊目的コンピュータを利用して実現されてよい。処理装置は、オペレーティングシステム ( OS ) および OS 上で実行される 1 つ以上のソフトウェアアプリケーションを実行してよい。また、処理装置は、ソフトウェアの実行に応答し、データにアクセスし、データを記録、操作、処理、および生成してもよい。理解の便宜のために、1 つの処理装置が使用されるとして説明される場合もあるが、当業者は、処理装置が複数個の処理要素および/または複数種類の処理要素を含んでもよいことが理解できるであろう。例えば、処理装置は、複数個のプロセッサまたは 1 つのプロセッサおよび 1 つのコントローラを含んでよい。また、並列プロセッサのような、他の処理構成も可能である。

20

30

## 【 0 0 6 6 】

ソフトウェアは、コンピュータプログラム、コード、命令、またはこれらのうちの 1 つ以上の組み合わせを含んでもよく、思うままに動作するように処理装置を構成したり、独立的または集成的に処理装置に命令したりしてよい。ソフトウェアおよび/またはデータは、処理装置に基づいて解釈されたり、処理装置に命令またはデータを提供したりするために、いかなる種類の機械、コンポーネント、物理装置、コンピュータ記録媒体または装置に具現化されてよい。ソフトウェアは、ネットワークによって接続されたコンピュータシステム上に分散され、分散された状態で記録されても実行されてもよい。ソフトウェアおよびデータは、1 つ以上のコンピュータ読み取り可能な記録媒体に記録されてよい。

## 【 0 0 6 7 】

実施形態に係る方法は、多様なコンピュータ手段によって実行可能なプログラム命令の形態で実現されてコンピュータ読み取り可能な媒体に記録されてよい。ここで、媒体は、コンピュータ実行可能なプログラムを継続して記録するものであっても、実行またはダウンロードのために一時記録するものであってもよい。また、媒体は、単一または複数のハードウェアが結合した形態の多様な記録手段または格納手段であってよく、あるコンピュータシステムに直接接続する媒体に限定されることはなく、ネットワーク上に分散して存在するものであってもよい。媒体の例としては、ハードディスク、フロッピー ( 登録商標 ) ディスク、および磁気テープのような磁気媒体、CD-ROM および DVD のような光媒体、フロプティカルディスク ( floptical disk ) のような光磁気媒体、および ROM、RAM、フラッシュメモリなどを含み、プログラム命令が記録されるよう

40

50

に構成されたものであってよい。また、媒体の他の例として、アプリケーションを配布するアプリケーションストアやその他の多様なソフトウェアを供給または配布するサイト、サーバなどで管理する記録媒体または格納媒体が挙げられる。

【0068】

本文書の多様な実施形態およびこれに使用された用語は、本文書に記載された技術を特定の実施形態だけに対して限定するためのものではなく、該当の実施例の多様な変更、均等物、および/または代替物を含むものと理解されなければならない。図面の説明と関連し、類似する構成要素に対しては類似する参照符号を付与した。単数の表現は、文脈上で明らかに異なるように意味しない限り、複数の表現を含んでよい。本文書において、「AまたはB」、「Aおよび/またはBのうち少なくとも1つ」、「A、B、またはC」、または「A、B、および/またはCのうち少なくとも1つ」などの表現は、ともに羅列される項目のすべての可能な組み合わせを含んでよい。「第1」、「第2」、「1番目」、または「2番目」などの表現は、該当の構成要素を順序または重要度とは関係なく修飾するものであり、ある構成要素を他の構成要素と区分するために使用されるものに過ぎず、該当の構成要素を限定するためのものではない。ある(例:第1)構成要素が他の(例:第2)構成要素に「(機能的にまたは通信的に)連結されて」いるか「接続されて」いると記載されるときには、前記ある構成要素が前記他の構成要素に直接に連結されている場合はもちろん、他の構成要素(例:第3構成要素)を介して連結されている場合も含まれる。

【0069】

本文書で使用される用語「モジュール」は、ハードウェア、ソフトウェア、またはファームウェアで構成されたユニットを含み、例えば、ロジック、論理ブロック、部品、または回路などの用語と互換的に使用されてよい。モジュールは、一体で構成された部品、または1つまたはそれ以上の機能を実行する最小単位またはその一部であってよい。例えば、モジュールは、ASIC(application-specific integrated circuit)で構成されてよい。

【0070】

多様な実施形態によると、記載した構成要素のそれぞれの構成要素(例:モジュールまたはプログラム)は、単数または複数の個体を含んでよい。多様な実施形態によると、上述した該当の構成要素のうち1つ以上の構成要素または段階が省略されてもよいし、1つ以上の他の構成要素または段階が追加されてもよい。大体的にまたは追加的に、複数の構成要素(例:モジュールまたはプログラム)は、1つの構成要素として統合されてよい。このような場合、統合された構成要素は、複数の構成要素それぞれの構成要素の1つ以上の機能を、統合される前に複数の構成要素のうち該当の構成要素によって実行されるときと同一または類似するように実行してよい。多様な実施形態によると、モジュール、プログラム、または他の構成要素によって実行される段階は、順次的に、並列的に、反復的に、または発見的に実行されても、段階のうち1つ以上が他の順序で実行されても、省略されても、または1つ以上の他の段階が追加されてもよい。

【符号の説明】

【0071】

- 460: プロセッサ
- 461: API
- 462: プロセスAPI
- 465: 制御部
- 467: コンテンツ取得部
- 469: フィンガープリント部
- 471: マッチン部
- 473: 比較部
- 475: クラスタリング部

10

20

30

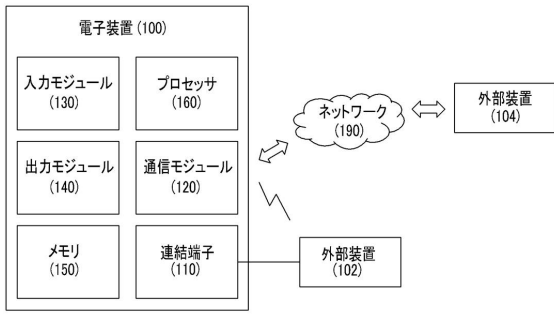
40

50

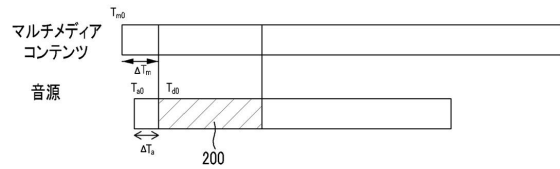


【図面】

【図 1】

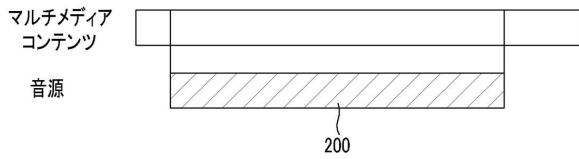


【図 2】

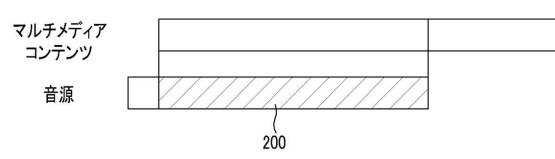


10

【図 3 a】

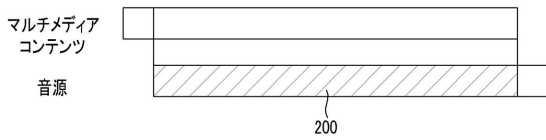


【図 3 b】

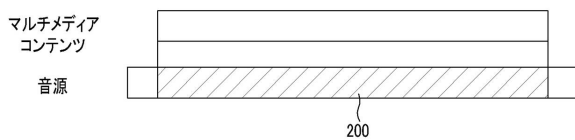


20

【図 3 c】



【図 3 d】

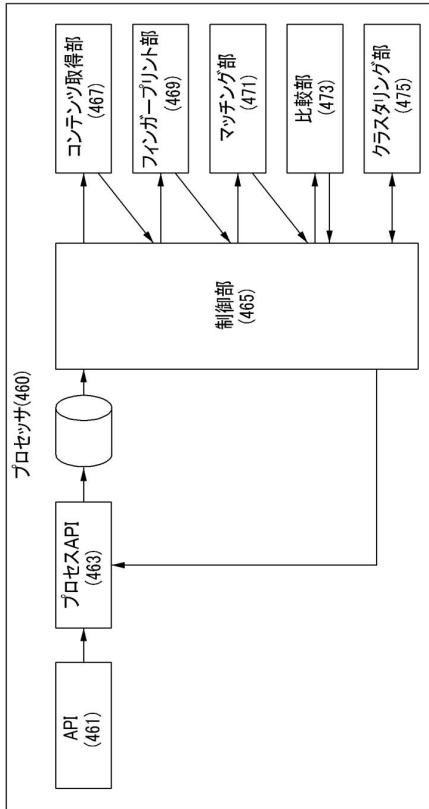


30

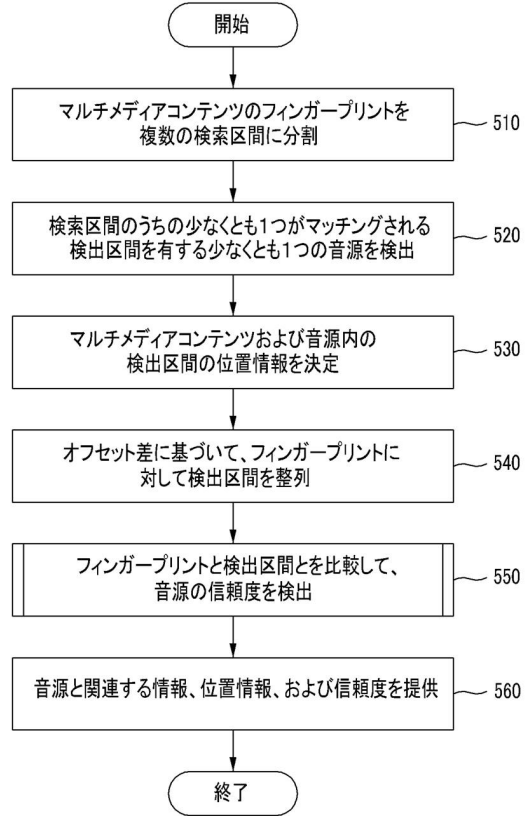
40

50

【図4】



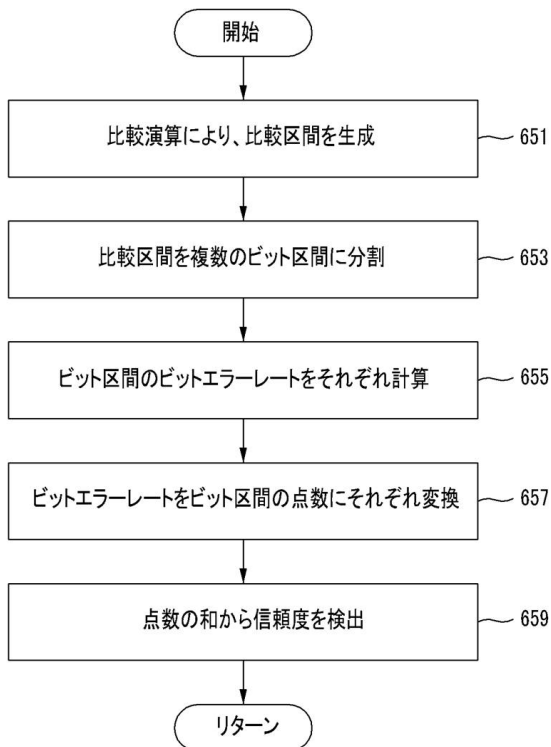
【図5】



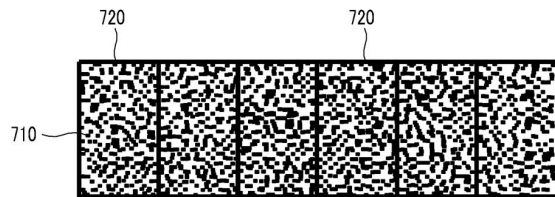
10

20

【図6】



【図7】

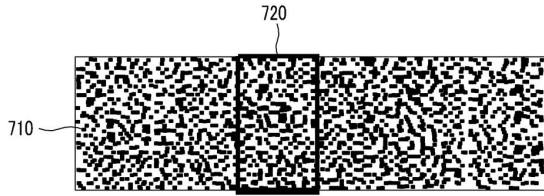


30

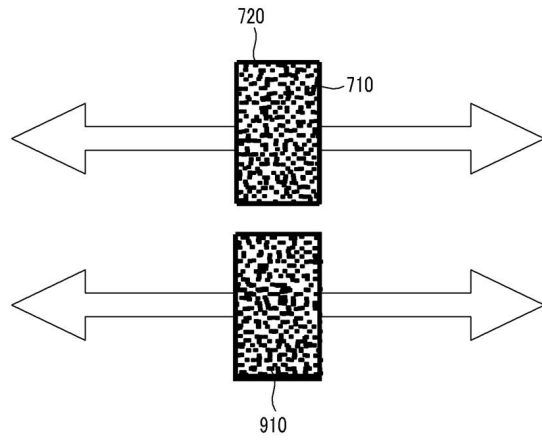
40

50

【 図 8 】



【 図 9 】

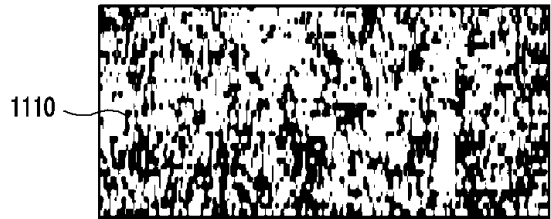


10

【 図 1 0 】



【 図 1 1 】



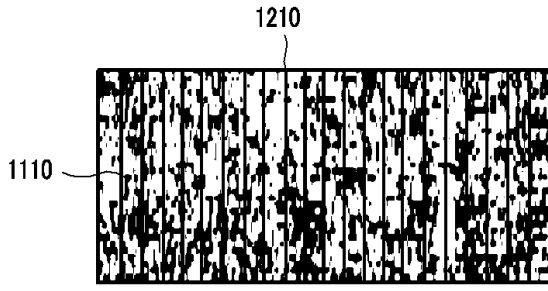
20

30

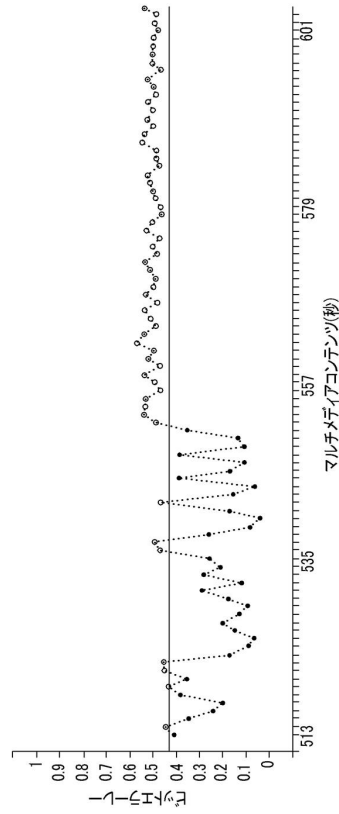
40

50

【図 1 2】



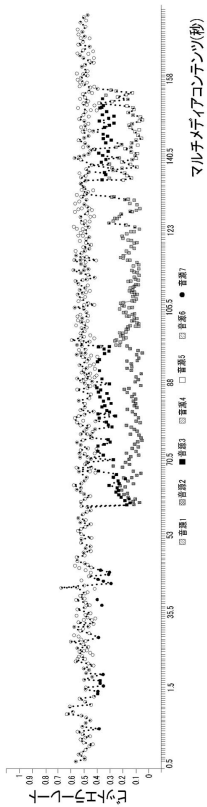
【図 1 3】



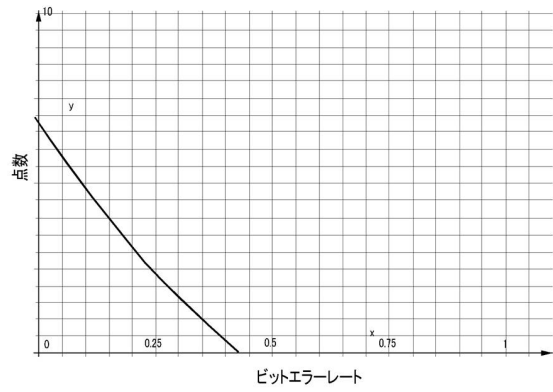
10

20

【図 1 4】



【図 1 5】

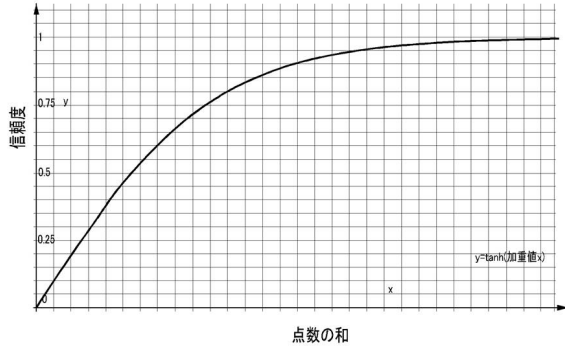


30

40

50

【 図 16 】



【 図 17 】

#	Video Range	Music Range	Music meta	Confidence
1	00:00:27 ~ 00:00:35	00:02:33 ~ 00:02:41	曲'A' - アーティスト1	0.749075
2	00:00:47 ~ 00:00:55	00:02:53 ~ 00:03:01	曲'A' - アーティスト1	0.335756
3	00:06:45 ~ 00:07:20	00:00:09 ~ 00:00:44	曲'B' - アーティスト2	0.999999
4	00:08:33 ~ 00:09:11	00:00:00 ~ 00:00:38	曲'C' - アーティスト1	0.999999
5	00:12:03 ~ 00:12:21	00:02:34 ~ 00:02:52	曲'B' - アーティスト2	0.999999

10

20

30

40

50

## フロントページの続き

- 6 グリーンファクトリー内
- (72)発明者 キム デファン  
大韓民国 13561 キョンギ - ド ソンナム - シ ブンダン - グ プルジョンロ 6 グリーンファクトリー内
- (72)発明者 ソ ドンウ  
大韓民国 13561 キョンギ - ド ソンナム - シ ブンダン - グ プルジョンロ 6 グリーンファクトリー内
- (72)発明者 キム ドンファン  
大韓民国 13561 キョンギ - ド ソンナム - シ ブンダン - グ プルジョンロ 6 グリーンファクトリー内
- (72)発明者 チョン ジス  
大韓民国 13561 キョンギ - ド ソンナム - シ ブンダン - グ プルジョンロ 6 グリーンファクトリー内
- (72)発明者 イ ジャンヒ  
大韓民国 13561 キョンギ - ド ソンナム - シ ブンダン - グ プルジョンロ 6 グリーンファクトリー内
- 審査官 酒井 恭信
- (56)参考文献 特表2018 - 523419 (JP, A)  
特開2020 - 025322 (JP, A)  
特開2019 - 216425 (JP, A)  
特表2020 - 525856 (JP, A)  
特表2015 - 505992 (JP, A)  
特表2017 - 537537 (JP, A)  
米国特許出願公開第2019 / 0213214 (US, A1)
- (58)調査した分野 (Int.Cl., DB名)  
G06F 16 / 00 - 16 / 958  
G10L 25 / 00 - 25 / 93