

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2013-101248

(P2013-101248A)

(43) 公開日 平成25年5月23日(2013.5.23)

(51) Int.Cl.		F I		テーマコード (参考)
G 1 O L 13/08 (2013.01)		G 1 O L	13/08 1 3 2	5 D O 2 O
G 1 O L 19/00 (2013.01)		G 1 O L	19/00 3 1 2 F	5 D O 6 2
G 1 O L 13/033 (2013.01)		G 1 O L	13/02 1 2 2 Z	
G 1 O L 13/00 (2006.01)		G 1 O L	13/08 1 2 4	
H O 4 R 3/00 (2006.01)		G 1 O L	13/00 1 O O Z	
審査請求 未請求 請求項の数 16 O L (全 20 頁) 最終頁に続く				

(21) 出願番号 特願2011-245357 (P2011-245357)
 (22) 出願日 平成23年11月9日 (2011.11.9)

(71) 出願人 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100095957
 弁理士 亀谷 美明
 (74) 代理人 100096389
 弁理士 金本 哲男
 (74) 代理人 100101557
 弁理士 萩原 康司
 (74) 代理人 100128587
 弁理士 松本 一騎
 (72) 発明者 森 麻紀
 東京都港区港南1丁目7番1号 ソニー株式会社社内

最終頁に続く

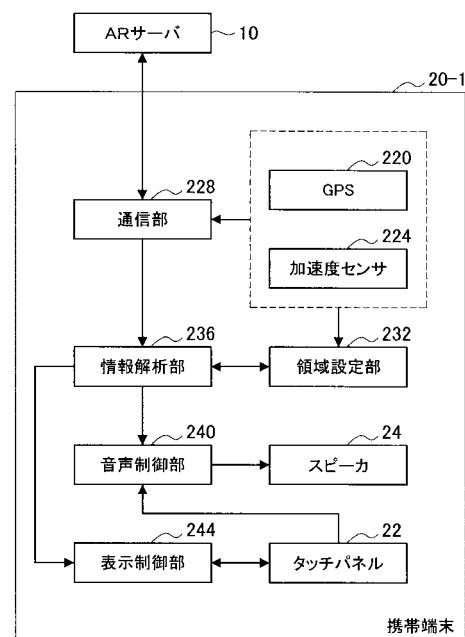
(54) 【発明の名称】 音声制御装置、音声制御方法、およびプログラム

(57) 【要約】

【課題】 音声出力される複数の情報からの迅速な情報抽出を支援するための音声制御装置、音声制御方法、およびプログラムを提供する。

【解決手段】 周囲の位置と対応付けられている情報を取得する情報取得部と、前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、を備える音声制御装置。

【選択図】 図 4



【特許請求の範囲】**【請求項 1】**

周囲の位置と対応付けられている情報を取得する情報取得部と、
前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、
を備える、音声制御装置。

【請求項 2】

前記音声制御部は、前記複数の情報が異なる音特性で出力されるように音声制御を行う、請求項 1 に記載の音声制御装置。

【請求項 3】

前記音声制御部は、前記複数の情報が同時に出力されるように音声制御を行う第 1 の形態と、前記複数の情報の各々が順次に出力されるように制御を行う第 2 の形態とで音声制御を切り替える、請求項 1 に記載の音声制御装置。

【請求項 4】

前記第 2 の形態において第 1 のユーザ操作が行われると、前記第 1 のユーザ操作が行われた際に音声出力された情報と対応付けられている位置を示す表示を制御する表示制御部と、
をさらに備える、請求項 3 に記載の音声制御装置。

【請求項 5】

前記音声制御部は、前記第 2 の形態において、第 2 のユーザ操作に応じて音声出力の対象情報を切り替える、請求項 3 に記載の音声制御装置。

【請求項 6】

前記音声制御部は、前記複数の情報のうちで、前記音声制御装置に近い位置と対応付けられている情報ほど大きな音量で出力されるように音声制御を行う、請求項 2 に記載の音声制御装置。

【請求項 7】

前記音声制御部は、前記複数の情報の各々の音声出力を異なる周波数または音質で制御する、請求項 2 に記載の音声制御装置。

【請求項 8】

前記情報取得部は、前記音声制御装置の特定の部位が向けられている方向を基準とする領域内の位置と対応付けられている情報を取得する、請求項 1 に記載の音声制御装置。

【請求項 9】

前記領域内の位置と対応付けられている情報の数が下限値以上、上限値未満、または所定数となるように前記領域を設定する領域設定部をさらに備える、請求項 8 に記載の音声制御装置。

【請求項 10】

前記音声制御部は、前記複数の情報に含まれるキーワードの出力タイミングが分散するように音声制御を行う、請求項 1 に記載の音声制御装置。

【請求項 11】

前記音声制御部は、ステレオ音声を出力する装着型の音声出力装置からの音声出力を制御し、

前記音声制御装置は、

前記複数の情報の各々について仮想音源位置を設定する音源設定部と、

前記複数の情報の各々について、前記音源設定部により設定された仮想音源位置が音源位置としてユーザに知覚されるように音声処理を行う音声処理部と、
をさらに備える、請求項 1 に記載の音声制御装置。

【請求項 12】

前記音源設定部は、前記複数の情報の各々に対応付けられている位置を前記複数の情報の各々の仮想音源位置に設定する、請求項 11 に記載の音声制御装置。

【請求項 13】

10

20

30

40

50

前記音源設定部は、前記複数の情報の各々の仮想音源位置を異なる位置に設定する、請求項 1 1 に記載の音声制御装置。

【請求項 1 4】

表示画面において、前記複数の情報の各表示が、前記音源設定部により設定された仮想音源位置に応じた位置に配置されるように表示を制御する表示制御部をさらに備える、請求項 1 1 に記載の音声制御装置。

【請求項 1 5】

周囲の位置と対応付けられている情報を取得するステップと、
複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行うステップと、
を含む、音声制御方法。

10

【請求項 1 6】

コンピュータを、
周囲の位置と対応付けられている情報を取得する情報取得部と、
前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、
として機能させるための、プログラム。

【発明の詳細な説明】

【技術分野】

20

【0 0 0 1】

本開示は、音声制御装置、音声制御方法、およびプログラムに関する。

【背景技術】

【0 0 0 2】

近日、撮像装置およびディスプレイを備えた携帯端末の普及に伴い、A R (A u g m e n t e d R e a l i t y : 拡張現実) 技術を用いたアプリケーションが注目を集めている。例えば、A R 技術を用いたアプリケーションとして、携帯機器の位置に基づき、現実空間の撮像画像に写真 / テキスト / 音声などのタグ情報を付加する A R アプリケーションが挙げられる。

【0 0 0 3】

30

ユーザは、このような A R アプリケーションにより、現実空間に付加されたタグ情報を閲覧および操作することが可能となる。しかし、現実空間に付加されるタグ情報の数が増えると、画面の視認性が低下するので、ユーザが所望の情報を見つけることが困難になってしまう。この点について、携帯端末と対象位置との関係に応じてユーザに音声案内を行うための音響制御装置が特許文献 1 に開示されている。

【先行技術文献】

【特許文献】

【0 0 0 4】

【特許文献 1】特開 2 0 1 0 - 4 9 1 5 8 号公報

【発明の概要】

40

【発明が解決しようとする課題】

【0 0 0 5】

しかし、携帯端末の周囲に対応付けられているタグ情報を音声で順次に出力する場合、タグ情報の数が増えると、ユーザが興味のあるタグ情報が出力されるまでの時間が長期化するので、ユーザは所望のタグ情報を抽出し難い。

【0 0 0 6】

そこで、本開示では、音声出力される複数の情報からの迅速な情報抽出を支援することが可能な、新規かつ改良された音声制御装置、音声制御方法、およびプログラムを提案する。

【課題を解決するための手段】

50

【 0 0 0 7 】

本開示によれば、周囲の位置と対応付けられている情報を取得する情報取得部と、前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、を備える音声制御装置が提供される。

【 0 0 0 8 】

また、本開示によれば、周囲の位置と対応付けられている情報を取得するステップと、複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行うステップと、を含む音声制御方法が提供される。

【 0 0 0 9 】

また、本開示によれば、コンピュータを、周囲の位置と対応付けられている情報を取得する情報取得部と、前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、として機能させるためのプログラムが提供される。

10

【 発明の効果 】

【 0 0 1 0 】

以上説明したように本開示によれば、音声出力される複数の情報からの迅速な情報抽出を支援することが可能である。

【 図面の簡単な説明 】

【 0 0 1 1 】

【 図 1 】 本開示の実施形態による音声制御システムの構成を示した説明図である。

20

【 図 2 】 タグ情報が付加された現実空間の撮像画像を示した説明図である。

【 図 3 】 第 1 の実施形態による携帯端末の利用形態の具体例を示した説明図である。

【 図 4 】 本開示の第 1 の実施形態による携帯端末の構成を示した機能ブロック図である。

【 図 5 】 領域設定部により設定される対象領域の具体例を示した説明図である。

【 図 6 】 携帯端末の周囲のタグ情報の密集度が高い場合の対象領域の具体例を示した説明図である。

【 図 7 】 携帯端末の周囲のタグ情報の密集度が低い場合の対象領域の具体例を示した説明図である。

【 図 8 】 タグ情報の再構成を示した説明図である。

【 図 9 】 音声制御部による音声制御の具体例を示した説明図である。

30

【 図 1 0 】 タグ情報のリストの具体例を示した説明図である。

【 図 1 1 】 本開示の第 1 の実施形態による携帯端末の動作を示したフローチャートである。

【 図 1 2 】 本開示の第 2 の実施形態による携帯端末および音声出力装置を示した説明図である。

【 図 1 3 】 第 2 の実施形態による携帯端末の構成を示した機能ブロック図である。

【 図 1 4 】 第 2 の実施形態による表示制御例を示した説明図である。

【 図 1 5 】 携帯端末のハードウェア構成を示した説明図である。

【 発明を実施するための形態 】

【 0 0 1 2 】

40

以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【 0 0 1 3 】

また、本明細書及び図面において、実質的に同一の機能構成を有する複数の構成要素を、同一の符号の後に異なるアルファベットを付して区別する場合もある。ただし、実質的に同一の機能構成を有する複数の構成要素の各々を特に区別する必要がない場合、同一符号のみを付する。

【 0 0 1 4 】

また、以下に示す項目順序に従って本開示を説明する。

50

1. 音声制御システムの基本構成
2. 第1の実施形態
 - 2-1. 第1の実施形態の概要
 - 2-2. 第1の実施形態による携帯端末の構成
 - 2-3. 第1の実施形態による携帯端末の動作
3. 第2の実施形態
4. 携帯端末のハードウェア構成
5. むすび

【0015】

<< 1. 音声制御システムの基本構成 >>

10

本開示による技術は、一例として「2. 第1の実施形態」～「3. 第2の実施形態」において詳細に説明するように、多様な形態で実施され得る。また、各実施形態による音声制御装置（携帯端末20）は、

A. 周囲の位置と対応付けられている情報を取得する情報取得部（通信部228、情報解析部236）と、

B. 前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部（240）と、

を備える。

【0016】

以下では、まず、このような各実施形態において共通する音声制御システムの基本構成について図1を参照して説明する。

20

【0017】

図1は、本開示の実施形態による音声制御システムの構成を示した説明図である。図1に示したように、本開示の実施形態による音声制御システムは、ARサーバ10と、携帯端末20と、を備える。

【0018】

このARサーバ10および携帯端末20は、通信網12を介して接続されており、通信網12を介して各種情報を通信することが可能である。なお、通信網12は、通信網12に接続されている装置から送信される情報の有線、または無線の伝送路である。例えば、通信網12は、インターネット、電話回線網、衛星通信網などの公衆回線網や、Ethernet（登録商標）を含む各種のLAN（Local Area Network）、WAN（Wide Area Network）などを含んでもよい。また、通信網12は、IP-VPN（Internet Protocol-Virtual Private Network）などの専用回線網を含んでもよい。

30

【0019】

ARサーバ10は、位置情報とタグ情報とを対応付けて記憶しており、通信網12を介して位置情報とタグ情報を携帯端末20に送信する。例えば、ARサーバ10は、携帯端末20から携帯端末20の現在位置を示す位置情報を受信すると、携帯端末20の周囲の位置を示す位置情報およびタグ情報を携帯端末20に送信する。さらに、ARサーバ10は、携帯端末20から携帯端末20の特定の部位が向けられている方向を示す方向情報も受信した場合、携帯端末20の現在位置から、方向情報の示す方向側の位置を示す位置情報およびタグ情報を携帯端末20に送信してもよい。

40

【0020】

なお、タグ情報としては、例えば写真や動画のような画像情報、テキスト情報、または音声情報などが想定される。また、タグ情報は、対応付けられている位置情報の示す位置に関する宣伝情報であってもよいし、対応付けられている位置情報の示す位置に関してアップロードされた一般ユーザからの提供情報であってもよい。

【0021】

携帯端末20は、現在位置の周囲に関する位置情報およびタグ情報を取得し、タグ情報をユーザに提供する。例えば、携帯端末20は、現在位置を示す位置情報をARサーバ1

50

0 に送信し、A R サーバ 1 0 から現在位置の周囲に関するタグ情報および位置情報を取得し、タグ情報を現実空間の撮像画像に付加してタッチパネル 2 2 に表示することができる。

【 0 0 2 2 】

また、携帯端末 2 0 は、音声を出力する音声出力部の一例であるスピーカ 2 4、および音声を収音する音声入力部の一例であるマイクロフォン 2 6 を備える。本明細書においては、携帯端末 2 0 においてスピーカ 2 4 が配されている側の端部を上端部 U S (U p p e r S i d e) と称し、マイクロフォン 2 6 が配されている側の端部を下端部 L S (L o w e r S i d e) と称する。

【 0 0 2 3 】

なお、図 1 においては携帯端末 2 0 の一例としてスマートフォンを示しているが、携帯端末 2 0 はスマートフォンに限定されない。例えば、携帯端末 2 0 は、P D A (P e r s o n a l D i g i t a l A s s i s t a n t s)、携帯電話、携帯用音楽再生装置、携帯用映像処理装置、または携帯用ゲーム機器であってもよい。

【 0 0 2 4 】

(本実施形態の背景)

スマートフォンのような携帯端末の普及に伴い、上述したように、現実空間の撮像画像に写真 / テキスト / 音声などのタグ情報を付加するアプリケーションが注目を集めている。このアプリケーションによれば、例えば、図 2 に示したようなタグ情報が付加された現実空間の撮像画像が表示されるので、ユーザは、現実空間からは得られないタグ情報を閲覧および操作することが可能となる。

【 0 0 2 5 】

しかし、現実空間に付加されるタグ情報の数が多くなると、画面の視認性が低下するので、ユーザが所望の情報を見つけることが困難になってしまう。また、タグ情報を音声により出力することも考えられるが、携帯端末の周囲に対応付けられているタグ情報を音声で順次に出力する場合、タグ情報の数が多くなると、ユーザが興味のあるタグ情報が出力されるまでの時間が長期化するので、ユーザは所望のタグ情報を抽出し難い。

【 0 0 2 6 】

そこで、上記事情を一着眼点にして本実施形態を創作するに至った。本実施形態によれば、音声出力される複数のタグ情報からの迅速な情報抽出を支援することが可能である。以下、このような効果を有する本開示の第 1 の実施形態および第 2 の実施形態について詳細に説明する。

【 0 0 2 7 】

< < 2 . 第 1 の実施形態 > >

< 2 - 1 . 第 1 の実施形態の概要 >

ユーザは、様々な雑音が存在する状況においても、自分が興味のあるワードや会話は自然と聞き分けることができることが知られている (カクテルパーティ効果) 。本開示の第 1 の実施形態による携帯端末 2 0 - 1 は、このカクテルパーティ効果を利用することにより、タグ情報を音声出力によりユーザに提供する場合に、ユーザが興味のあるタグ情報の迅速に抽出することを可能とする。以下、図 3 を参照して、第 1 の実施形態による携帯端末 2 0 - 1 の利用形態の具体例を説明する。

【 0 0 2 8 】

図 3 は、第 1 の実施形態による携帯端末 2 0 - 1 の利用形態の具体例を示した説明図である。図 3 に示したように、ユーザが携帯端末 2 0 - 1 のスピーカ 2 4 を耳にかざすと、携帯端末 2 0 - 1 は、周囲の位置と対応付けられているタグ情報を同時に音声出力する。具体的には、携帯端末 2 0 - 1 は、携帯端末 2 0 - 1 の特定の部位として下端部 L S が向けられている方向を基準とする対象領域を設定し、対象領域内の位置 P 1 ~ P 3 と対応付けられているタグ情報 T A ~ T C を同時に出力する。すなわち、携帯端末 2 0 - 1 は、タグ情報 T A 「 K K K のアルバム最高 ! 」、タグ情報 T B 「 F F F C a f e のケーキ美味しい ! 」、タグ情報 T C 「 B B B で女優 A 発見 ! 」を同時に音声出力する。

【0029】

これにより、ユーザは、例えば女優Aに興味がある場合、カクテルパーティ効果に基づいて女優Aに関するタグ情報を迅速に聞きとると共に、女優Aが携帯端末20-1の下端部LSが向けられている方向で発見されたことを知ることができる。なお、携帯端末20-1は、タグ情報が対応付けられている位置と現在位置との距離が遠いほどタグ情報を小さな音量で出力してもよい。この場合、ユーザは、所望のタグ情報が対応付けられている位置と現在位置との距離感も把握することが可能となる。

【0030】

< 2-2. 第1の実施形態による携帯端末の構成 >

以上、本開示の第1の実施形態の概要を説明した。続いて、本開示の第1の実施形態による携帯端末20-1の構成について詳細に説明する。

10

【0031】

図4は、本開示の第1の実施形態による携帯端末20-1の構成を示した機能ブロック図である。図4に示したように、本開示の第1の実施形態による携帯端末20-1は、タッチパネル22と、スピーカ24と、GPS220と、加速度センサ224と、通信部228と、領域設定部232と、情報解析部236と、音声制御部240と、表示制御部244と、を備える。

【0032】

(GPS)

GPS (Global Positioning System) 220は、人工衛星から送信される航法メッセージを受信し、航法メッセージの受信結果に基づいて携帯端末20-1の現在位置を取得する。なお、GPS220は、携帯端末20-1の現在位置を取得するための構成の一例に過ぎず、他の構成により携帯端末20-1の現在位置を取得することも可能である。例えば、携帯端末20-1は、周囲の無線通信装置から送信されるWi-Fi電波の受信強度に基づいて現在位置を推定してもよいし、ユーザ入力に基づいて現在位置を取得してもよい。

20

【0033】

(加速度センサ)

加速度センサ224は、携帯端末20-1の加速度を測定する。加速度センサ224による測定結果に基づき、携帯端末20-1の向き、例えば携帯端末20-1の下端部LSの向いている方向を推定することが可能となる。

30

【0034】

(通信部)

通信部228は、ARサーバ10とのインタフェースであり、ARサーバ10と各種情報を送受信する。例えば、通信部228は、GPS220によって取得された携帯端末20-1の位置情報をARサーバ10に送信し、携帯端末20-1の周囲と対応付けられているタグ情報を位置情報と共にARサーバ10から受信する。また、通信部228は、携帯端末20-1の位置情報に加え、加速度センサ224による測定結果もARサーバ10に送信してもよい。この場合、ARサーバ10から、携帯端末20-1の下端部LSが向いている方向の位置と対応付けられているタグ情報を受信することが可能となる。このように、通信部228、または後述する情報解析部236は、音声出力するためのタグ情報を取得する情報取得部として機能する。

40

【0035】

なお、本実施形態においては携帯端末20-1がARサーバ10からタグ情報を取得する例を説明するが、本実施形態はかかる例に限定されない。例えば、携帯端末20-1は、位置情報とタグ情報が対応付けられているデータベースを有し、当該データベースから携帯端末20-1の周囲と対応付けられているタグ情報を取得してもよい。

【0036】

(領域設定部)

領域設定部232は、GPS220から供給される携帯端末20-1の位置情報、およ

50

び加速度センサ 224 から得られる測定結果に基づいて、タグ情報の音声出力に関する対象領域を設定する。領域設定部 232 により設定された対象領域内の位置情報と対応付けられているタグ情報が音声出力の対象となる。ここで、図 5 を参照し、領域設定部 232 により設定される対象領域について具体的に説明する。

【0037】

図 5 は、領域設定部 232 により設定される対象領域の具体例を示した説明図である。領域設定部 232 は、一例として、携帯端末 20-1 の下端部 LS が向けられている方向を基準とする、半径 r および角度 で表わされる扇状領域を対象領域として設定してもよい。

【0038】

ここで、同時に音声出力されるタグ情報が多すぎると、カクテルパーティ効果が弱まり、ユーザが興味のあるタグ情報を聞き分けることが難しくなると考えられる。一方、同時に音声出力されるタグ情報が少なすぎると、ユーザが興味のあるタグ情報が音声出力される可能性も低くなるので、情報抽出の効率が悪くなる。

【0039】

そこで、領域設定部 232 は、対象領域の広さを周囲のタグ情報の密集度合いに応じて変更してもよい。例えば、領域設定部 232 は、対象領域内の位置情報と対応付けられているタグ情報の数が下限値以上、上限値未満、または所定数となるように対象領域を設定してもよい。以下、この点について図 6 および図 7 を参照してより具体的に説明する。

【0040】

図 6 は、携帯端末 20-1 の周囲のタグ情報の密集度が高い場合の対象領域の具体例を示した説明図である。図 7 は、携帯端末 20-1 の周囲のタグ情報の密集度が低い場合の対象領域の具体例を示した説明図である。領域設定部 232 は、図 6 に示したように、周囲のタグ情報の密集度が高い場合には半径 r や角度 を小さくすることにより対象領域を狭くしてもよい。一方、領域設定部 232 は、図 7 に示したように、周囲のタグ情報の密集度が低い場合には半径 r や角度 を大きくすることにより対象領域を広くしてもよい。

【0041】

かかる構成により、ユーザによるタグ情報の聞き分けの容易性、および情報抽出の効率性を確保することが可能となる。

【0042】

(情報解析部)

情報解析部 236 は、通信部 228 により受信された周囲のタグ情報から、領域設定部 232 により設定された対象領域内の位置情報と対応付けられているタグ情報を抽出する。また、情報解析部 236 は、抽出したタグ情報が対応付けられている位置と、携帯端末 20-1 との距離および角度を解析し、解析結果を音声制御部 240 に供給する。

【0043】

(音声制御部)

音声制御部 240 は、情報解析部 236 により抽出されたタグ情報のスピーカ 24 からの音声出力を制御する。具体的には、音声制御部 240 は、情報解析部 236 により抽出されたタグ情報が同時にスピーカ 24 から出力されるように音声制御を行う(第 1 の制御形態)。かかる構成により、ユーザは、同時に出力される複数のタグ情報からカクテルパーティ効果に基づいて効率的に興味のあるタグ情報を聞き分けることが可能となる。また、音声制御部 240 は、タグ情報の音声出力に際して以下の制御例 A ~ 制御例 E を行ってもよい。

【0044】

- 制御例 A

音声制御部 240 は、タグ情報が対応付けられている位置と携帯端末 20-1 との距離に応じてタグ情報の音量を制御する。例えば、音声制御部 240 は、タグ情報が対応付けられている位置と携帯端末 20-1 との距離が遠いほど当該タグ情報の音量を小さくしてもよい。かかる構成により、ユーザは、タグ情報が対応付けられている位置と現在位置と

10

20

30

40

50

の距離感を把握することが可能となる。

【 0 0 4 5 】

- 制御例 B

音声制御部 240 は、複数のタグ情報の各々が異なる音特性で出力されるように音声制御を行う。例えば、音声制御部 240 は、複数のタグ情報の各々が異なる基本周波数または音質で出力されるように音声制御を行ってもよい。かかる構成により、より確実にカクテルパーティ効果を達成することが可能となる。

【 0 0 4 6 】

- 制御例 C

音声制御部 240 は、タグ情報が音声情報でない場合、タグ情報から音声情報を生成する。例えば、音声制御部 240 は、タグ情報がテキスト情報である場合、テキスト情報に対応する音声情報を生成してもよい。また、音声制御部 240 は、タグ情報が写真タグである場合、写真タグに付加されているコメントに対応する音声情報を生成してもよい。かかる構成により、より多数のタグ情報を音声出力の対象とすることができる。

【 0 0 4 7 】

- 制御例 D

音声制御部 240 は、各タグ情報からキーワードを抽出し、各タグ情報のキーワードの出力タイミングが分散するように音声制御を行う。例えば、タグ情報 T A 「 K K K のアルバム最高！」のキーワードが「 K K K 」であり、タグ情報 T B 「 F F F C a f e のケーキ美味しい！」のキーワードが「 F F F C a f e 」であり、タグ情報 T C 「 B B B で女優 A 発見！」のキーワードが「 女優 A 」であるとする。この場合、これらタグ情報をそのまま出力すると、タグ情報 T A のキーワード「 K K K 」とタグ情報 T B のキーワードが「 F F F C a f e 」が同時に出力される。そこで、音声制御部 240 は、図 8 に示したように、タグ情報 T A のキーワード「 K K K 」とタグ情報 T B のキーワードが「 F F F C a f e 」の出力タイミングがずれるよう、タグ情報 T B の内容を再構成してもよい。かかる構成により、各タグ情報のキーワードの聞き分けをより容易にすることが可能である。

【 0 0 4 8 】

- 制御例 E

音声制御部 240 は、ユーザの嗜好に基づき、ユーザが興味を持っていると考えられるワードの出力タイミングが早くなるようにタグ情報の内容を再構成してもよい。例えば、ユーザが「 女優 A 」に興味を持っている場合、音声制御部 240 は、タグ情報 T C の内容を、「 女優 A 、 B B B で発見 」のように、「 女優 A 」の出力タイミングが早くなるように再構成してもよい。かかる構成により、ユーザは、早期に自分の興味のあるタグ情報を抽出することが可能となる。

【 0 0 4 9 】

また、音声制御部 240 は、タップ操作のような所定のユーザ操作に基づいて、情報解析部 236 により抽出されたタグ情報の各々が順次に出力されるように制御を行う第 2 の形態に音声制御を切り替える。このとき、ユーザは、所望のタグ情報が音声出力されているときに第 1 のユーザ操作を行うことによりタグ情報を選択することが可能である。また、音声制御部 240 は、フリック操作のような第 2 のユーザ操作に応じて音声出力対象のタグ情報を切り替えてもよい。以下、このような一連の動作について図 9 を参照してより具体的に説明する。

【 0 0 5 0 】

図 9 は、音声制御部 240 による音声制御の具体例を示した説明図である。図 9 に示したように、複数のタグ情報が同時に出力されている状態においてユーザによりタップ操作が行われると、音声制御部 240 は、複数のタグ情報を順次に出力する形態に音声制御を切り替える。ここで、音声制御部 240 は、携帯端末 20 - 1 の近くに対応付けられているタグ情報から順次に出力されるように音声制御を行ってもよい。なお、図中の文字の大きさは音量の大きさに対応している。

【 0 0 5 1 】

10

20

30

40

50

その後、所望のタグ情報が音声出力されているときに携帯端末 20 - 1 を振るシェイク操作を行うことにより、所望のタグ情報を選択することができる。例えば、タグ情報 T B の出力時にユーザがシェイク操作を行った場合、表示制御部 244 が、タグ情報 T B が対応付けられている位置をタッチパネル 22 に表示させる。

【0052】

このように、ユーザは、まずカクテルパーティ効果に基づいて興味のあるタグ情報を含む母集団を絞り、その後、興味のあるタグ情報を当該母集団から特定し、興味のあるタグ情報の詳細情報を得ることが可能となる。

【0053】

(表示制御部)

表示制御部 244 は、各種表示画面を生成し、生成した表示画面をタッチパネル 22 に表示させる。特に、本実施形態による表示制御部 244 は、複数のタグ情報を順次に出力される第 2 の形態においてシェイク操作のようなユーザ操作が行われると、シェイク操作が行われた際に音声出力されたタグ情報と対応付けられている位置を示す表示を制御する(図 9)。

【0054】

なお、上記では、タグ情報の母集団からユーザが興味のあるタグ情報を特定するための方法として複数のタグ情報を順次に出力する第 2 の形態を説明したが、ユーザが興味のあるタグ情報を当該母集団から特定するための方法はかかる例に限定されない。例えば、表示制御部 244 は、複数のタグ情報が同時に出力される第 1 の形態において所定のユーザ操作が行われた場合、ユーザ操作が行われた時点での複数のタグ情報を固定し、図 10 に示したように、固定した複数のタグ情報のリストをタッチパネル 22 に表示させてもよい。かかる構成により、ユーザは、タグ情報のリストから所望のタグ情報を特定することが可能となる。

【0055】

< 2 - 3 . 第 1 の実施形態による携帯端末の動作 >

以上、本開示の第 1 の実施形態による携帯端末 20 - 1 の構成を説明した。続いて、図 11 を参照し、本開示の第 1 の実施形態による携帯端末 20 - 1 の動作を説明する。

【0056】

図 11 は、本開示の第 1 の実施形態による携帯端末 20 - 1 の動作を示したフローチャートである。図 11 に示したように、まず、携帯端末 20 - 1 の GPS 220 が携帯端末 20 - 1 の現在位置を測定し、加速度センサ 224 が携帯端末 20 - 1 の下端部 L S の向きを測定する(S304)。そして、通信部 228 が、携帯端末 20 - 1 の位置情報を AR サーバ 10 に送信することにより、AR サーバ 10 から携帯端末 20 - 1 の周囲と対応付けられているタグ情報を取得する(S308)。

【0057】

一方、領域設定部 232 は、GPS 220 から供給される携帯端末 20 - 1 の位置情報、および加速度センサ 224 から得られる測定結果に基づいて、タグ情報の音声出力に関する対象領域を設定する。その後、情報解析部 236 が、携帯端末 20 - 1 の周囲と対応付けられているタグ情報から、対象領域内のタグ情報を抽出する(S316)。そして、音声制御部 240 は、対象領域内のタグ情報がスピーカ 24 から同時に出力されるように音声制御を行う(S320)。

【0058】

その後、タッチパネル 22 に対してユーザによりタップ操作が行われると(S324)、音声制御部 240 は、対象領域内のタグ情報が順次に出力されるように音声制御の形態を切り替える(S328)。そして、ユーザによりシェイク操作が行われると(S332)、表示制御部 244 が、シェイク操作の際に音声出力されたタグ情報と対応付けられている位置を示す表示画面をタッチパネル 22 に表示させる(S336)。

【0059】

以上説明したように、本開示の第 1 の実施形態によれば、複数のタグ情報を同時に出力

10

20

30

40

50

することにより、ユーザは、カクテルパーティ効果に基づいて興味のあるタグ情報を含む母集団を迅速に絞り込むことが可能となる。

【 0 0 6 0 】

< 3 . 第 2 の実施形態 > >

続いて、本開示の第 2 の実施形態を説明する。本開示の第 2 の実施形態は、音声出力をスピーカ 2 4 でなく装着型の音声出力装置から行う点で第 1 の実施形態と異なる。以下、このような本開示の第 2 の実施形態について詳細に説明する。

【 0 0 6 1 】

図 1 2 は、本開示の第 2 の実施形態による携帯端末 2 0 - 2 および音声出力装置 3 0 を示した説明図である。図 1 2 に示したように、第 2 の実施形態による携帯端末 2 0 - 2 は、音声出力装置 3 0 に音声信号を送信し、音声出力装置 3 0 が、携帯端末 2 0 - 2 から受信した音声信号をステレオ音声で出力する。

【 0 0 6 2 】

なお、図 1 2 においては携帯端末 2 0 - 2 と音声出力装置 3 0 が近接無線通信により接続される例を示しているが、携帯端末 2 0 - 2 と音声出力装置 3 0 は有線により接続されてもよい。また、図 1 2 においては装着型の音声出力装置 3 0 の一例としてヘッドホンを示しているが、音声出力装置 3 0 は、ヘッドホンに限定されず、例えばイヤホンであってもよい。

【 0 0 6 3 】

(第 2 の実施形態による携帯端末の構成)

続いて、図 1 3 を参照し、第 2 の実施形態による携帯端末 2 0 - 2 の構成を説明する。図 1 3 は、第 2 の実施形態による携帯端末 2 0 - 2 の構成を示した機能ブロック図である。図 1 3 に示したように、第 2 の実施形態による携帯端末 2 0 - 2 は、タッチパネル 2 2 と、GPS 2 2 0 と、加速度センサ 2 2 4 と、通信部 2 2 8 と、領域設定部 2 3 4 と、情報解析部 2 3 6 と、音声制御部 2 4 0 と、表示制御部 2 4 6 と、音源設定部 2 4 8 と、音声処理部 2 5 2 と、を備える。タッチパネル 2 2、GPS 2 2 0、加速度センサ 2 2 4、通信部 2 2 8、および音声制御部 2 4 0 などの構成について第 1 の実施形態と共通部分が多いので、ここでの詳細な説明を省略する。

【 0 0 6 4 】

領域設定部 2 3 4 は、GPS 2 2 0 から供給される携帯端末 2 0 - 1 の位置情報、および加速度センサ 2 2 4 から得られる測定結果に基づいて、タグ情報の音声出力に関する対象領域を設定する。ここで、第 1 の実施形態と異なり、ユーザはスピーカ 2 4 を耳にかざす必要が無いので、領域設定部 2 3 4 は、携帯端末 2 0 - 2 の下端部 L S でなく、例えば上端部 U S が向けられている方向を基準にして対象領域を設定する。

【 0 0 6 5 】

音源設定部 2 4 8 は、情報解析部 2 3 6 により抽出された複数のタグ情報の各々について仮想的な音源位置（以下、仮想音源位置）を設定する。例えば、音源設定部 2 4 8 は、複数のタグ情報の各々に対応付けられている位置を仮想音源位置に設定してもよい。また、音源設定部 2 4 8 は、タグ情報に対応付けられている位置と異なる位置を複数のタグ情報の各々の仮想音源位置に設定してもよい。このように、複数のタグ情報の各々に異なる仮想音源位置を設定することにより、カクテルパーティ効果が一層高まることが期待される。

【 0 0 6 6 】

音声処理部 2 5 2 は、複数のタグ情報の各々について、音源設定部 2 4 8 により設定された仮想音源位置が音源位置としてユーザに知覚されるように音声処理を行う。この処理はユーザの向きに応じて行われるので、ユーザの向きを検出するためのセンサをユーザが装着してもよいし、携帯端末 2 0 - 2 の上端部 U S の向きをユーザの向きとして扱ってもよい。

【 0 0 6 7 】

近接通信部 2 5 6 は、音声処理部 2 5 2 により処理された音声信号を音声出力装置 3 0

に送信する。そして、音声出力装置 30 は、近接通信部 256 から受信した音声信号をステレオ音声で出力する。

【0068】

表示制御部 246 は、音声出力装置 30 から複数のタグ情報が同時に出力されている間、各タグ情報をタッチパネル 22 に表示させる。ここで、表示制御部 246 は、各タグ情報が、図 14 に示したように、現在位置を基準とする音源設定部 248 により設定された仮想音源位置に表示されるように表示制御を行ってもよい。かかる構成により、ユーザは、聴覚により知覚されるタグ情報の位置を、視覚を通じても知覚することが可能となる。

【0069】

<< 4 . 携帯端末のハードウェア構成 >>

以上、本開示の実施形態について詳細に説明した。上述した本開示の実施形態による携帯端末 20 による情報処理は、ソフトウェアと携帯端末 20 のハードウェアが協働することにより実現される。以下、このような携帯端末 20 のハードウェア構成について図 15 を参照して説明する。

【0070】

図 15 は、携帯端末 20 のハードウェア構成を示した説明図である。図 15 に示したように、携帯端末 20 は、CPU (Central Processing Unit) 201 と、ROM (Read Only Memory) 202 と、RAM (Random Access Memory) 203 と、入力装置 208 と、出力装置 210 と、ストレージ装置 211 と、ドライブ 212 と、撮像装置 213 と、通信装置 215 とを備える。

【0071】

CPU 201 は、演算処理装置および制御装置として機能し、各種プログラムに従って携帯端末 20 内の動作全般を制御する。また、CPU 201 は、マイクロプロセッサであってもよい。ROM 202 は、CPU 201 が使用するプログラムや演算パラメータ等を記憶する。RAM 203 は、CPU 201 の実行において使用するプログラムや、その実行において適宜変化するパラメータ等を一時記憶する。これらは CPU バスなどから構成されるホストバスにより相互に接続されている。

【0072】

入力装置 208 は、マウス、キーボード、タッチパネル、ボタン、マイクロフォン、スイッチおよびレバーなどユーザが情報を入力するための入力手段と、ユーザによる入力に基づいて入力信号を生成し、CPU 201 に出力する入力制御回路などから構成されている。携帯端末 20 のユーザは、該入力装置 208 を操作することにより、携帯端末 20 に対して各種のデータを入力したり処理動作を指示したりすることができる。

【0073】

出力装置 210 は、例えば、液晶ディスプレイ (LCD) 装置、OLED (Organic Light Emitting Diode) 装置およびランプなどの表示装置を含む。さらに、出力装置 210 は、スピーカおよびヘッドホンなどの音声出力装置を含む。例えば、表示装置は、撮像された画像や生成された画像などを表示する。一方、音声出力装置は、音声データ等を音声に変換して出力する。

【0074】

ストレージ装置 211 は、本実施形態にかかる携帯端末 20 の記憶部の一例として構成されたデータ格納用の装置である。ストレージ装置 211 は、記憶媒体、記憶媒体にデータを記録する記録装置、記憶媒体からデータを読み出す読出し装置および記憶媒体に記録されたデータを削除する削除装置などを含んでもよい。このストレージ装置 211 は、CPU 201 が実行するプログラムや各種データを格納する。

【0075】

ドライブ 212 は、記憶媒体用リーダライタであり、携帯端末 20 に内蔵、あるいは外付けされる。ドライブ 212 は、装着されている磁気ディスク、光ディスク、光磁気ディスク、または半導体メモリ等のリムーバブル記憶媒体 23 に記録されている情報を読み出

10

20

30

40

50

して、RAM 203に出力する。また、ドライブ212は、リムーバブル記憶媒体23に情報を書き込むこともできる。

【0076】

撮像装置213は、光を集光する撮影レンズおよびズームレンズなどの撮像光学系、およびCCD (Charge Coupled Device) またはCMOS (Complementary Metal Oxide Semiconductor) などの信号変換素子を備える。撮像光学系は、被写体から発せられる光を集光して信号変換部に被写体像を形成し、信号変換素子は、形成された被写体像を電氣的な画像信号に変換する。

【0077】

通信装置215は、例えば、通信網12に接続するための通信デバイス等で構成された通信インタフェースである。また、通信装置215は、無線LAN (Local Area Network) 対応通信装置であっても、LTE (Long Term Evolution) 対応通信装置であっても、有線による通信を行うワイヤ通信装置であってもよい。

10

【0078】

<<5. むすび>>

以上説明したように、本開示の実施形態によれば、複数のタグ情報を同時に出力することにより、ユーザは、カクテルパーティ効果に基づいて興味のあるタグ情報を含む母集団を迅速に絞り込むことが可能となる。その後、ユーザは、視覚または聴覚を通じて興味のあるタグ情報を当該母集団から特定し、興味のあるタグ情報の詳細情報を得ることが可能となる。

20

【0079】

また、本開示によれば、タグ情報が一般ユーザにより現在位置と対応付けて提供される情報である場合、ユーザは、タグ情報が出力される音量に基づいて、一般ユーザとの距離感を把握することができる。また、本開示のタグ情報を、相手の位置情報と対応付けられたメールやチャットなどに適用すれば、相手の方向や距離を把握することが可能となる。

【0080】

なお、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

30

【0081】

例えば、本明細書の携帯端末20の処理における各ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はない。例えば、携帯端末20の処理における各ステップは、フローチャートとして記載した順序と異なる順序で処理されても、並列的に処理されてもよい。

【0082】

また、携帯端末20に内蔵されるCPU 201、ROM 202およびRAM 203などのハードウェアを、上述した携帯端末20の各構成と同等の機能を発揮させるためのコンピュータプログラムも作成可能である。また、該コンピュータプログラムを記憶させた記憶媒体も提供される。

40

【0083】

また、以下のような構成も本開示の技術的範囲に属する。

(1)

周囲の位置と対応付けられている情報を取得する情報取得部と、

前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、
を備える、音声制御装置。

(2)

50

前記音声制御部は、前記複数の情報が異なる音特性で出力されるように音声制御を行う、前記（１）に記載の音声制御装置。

（３）

前記音声制御部は、前記複数の情報が同時に出力されるように音声制御を行う第１の形態と、前記複数の情報の各々が順次に出力されるように制御を行う第２の形態とで音声制御を切り替える、前記（１）または（２）に記載の音声制御装置。

（４）

前記第２の形態において第１のユーザ操作が行われると、前記第１のユーザ操作が行われた際に音声出力された情報と対応付けられている位置を示す表示を制御する表示制御部と、

10

をさらに備える、前記（３）に記載の音声制御装置。

（５）

前記音声制御部は、前記第２の形態において、第２のユーザ操作に応じて音声出力の対象情報を切り替える、前記（３）または（４）に記載の音声制御装置。

（６）

前記音声制御部は、前記複数の情報のうちで、前記音声制御装置に近い位置と対応付けられている情報ほど大きな音量で出力されるように音声制御を行う、前記（２）に記載の音声制御装置。

（７）

前記音声制御部は、前記複数の情報の各々の音声出力を異なる周波数または音質で制御する、前記（２）または（６）に記載の音声制御装置。

20

（８）

前記情報取得部は、前記音声制御装置の特定の部位が向けられている方向を基準とする領域内の位置と対応付けられている情報を取得する、前記（１）～（７）のいずれか一項に記載の音声制御装置。

（９）

前記領域内の位置と対応付けられている情報の数が下限値以上、上限値未満、または所定数となるように前記領域を設定する領域設定部をさらに備える、前記（８）に記載の音声制御装置。

（１０）

30

前記音声制御部は、前記複数の情報に含まれるキーワードの出力タイミングが分散するように音声制御を行う、前記（１）～（９）のいずれか一項に記載の音声制御装置。

（１１）

前記音声制御部は、ステレオ音声を出力する装着型の音声出力装置からの音声出力を制御し、

前記音声制御装置は、

前記複数の情報の各々について仮想音源位置を設定する音源設定部と、

前記複数の情報の各々について、前記音源設定部により設定された仮想音源位置が音源位置としてユーザに知覚されるように音声処理を行う音声処理部と、

をさらに備える、前記（１）～（１０）のいずれか一項に記載の音声制御装置。

40

（１２）

前記音源設定部は、前記複数の情報の各々に対応付けられている位置を前記複数の情報の各々の仮想音源位置に設定する、前記（１１）に記載の音声制御装置。

（１３）

前記音源設定部は、前記複数の情報の各々の仮想音源位置を異なる位置に設定する、前記（１１）に記載の音声制御装置。

（１４）

表示画面において、前記複数の情報の各表示が、前記音源設定部により設定された仮想音源位置に応じた位置に配置されるように表示を制御する表示制御部をさらに備える、前記（１１）～（１３）のいずれか一項に記載の音声制御装置。

50

(1 5)

周囲の位置と対応付けられている情報を取得するステップと、
複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行うステップと、
を含む、音声制御方法。

(1 6)

コンピュータを、
周囲の位置と対応付けられている情報を取得する情報取得部と、
前記情報取得部によって複数の情報が取得された場合、前記複数の情報が同時に出力されるように音声制御を行う音声制御部と、
として機能させるための、プログラム。

10

【符号の説明】

【 0 0 8 4 】

1 0 A R サーバ

2 0 携帯端末

2 2 タッチパネル

2 4 スピーカ

2 6 マイクロフォン

3 0 音声出力装置

2 2 0 G P S

2 2 4 加速度センサ

2 2 8 通信部

2 3 2、2 3 4 領域設定部

2 3 6 情報解析部

2 4 0 音声制御部

2 4 4、2 4 6 表示制御部

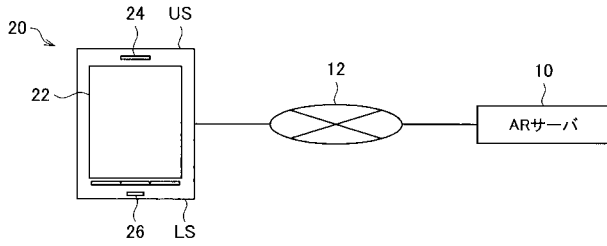
2 4 8 音源設定部

2 5 2 音声処理部

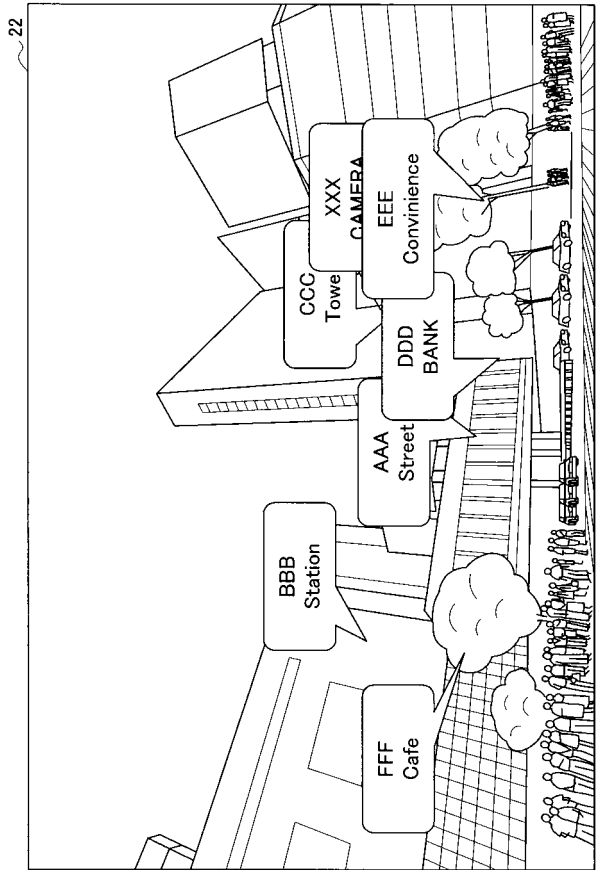
2 5 6 近接通信部

20

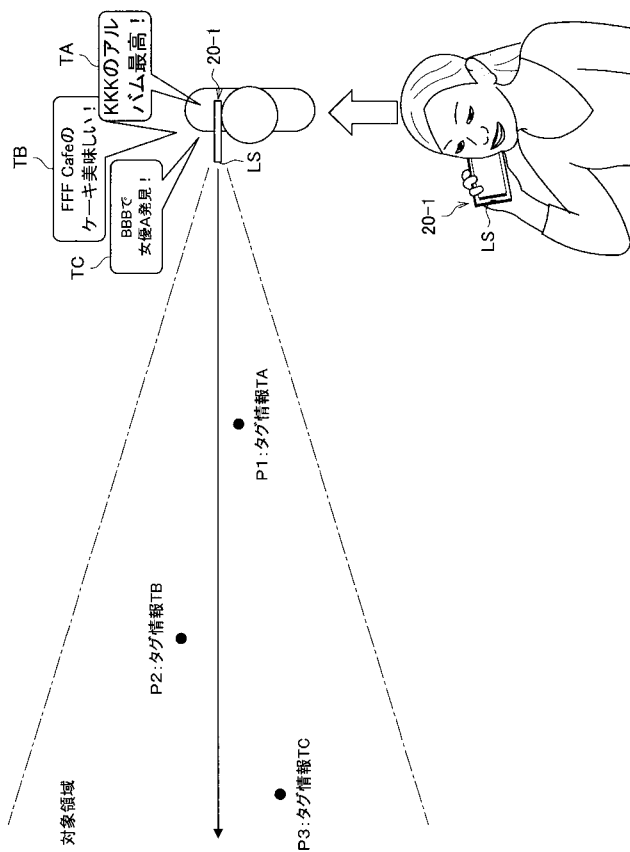
【図 1】



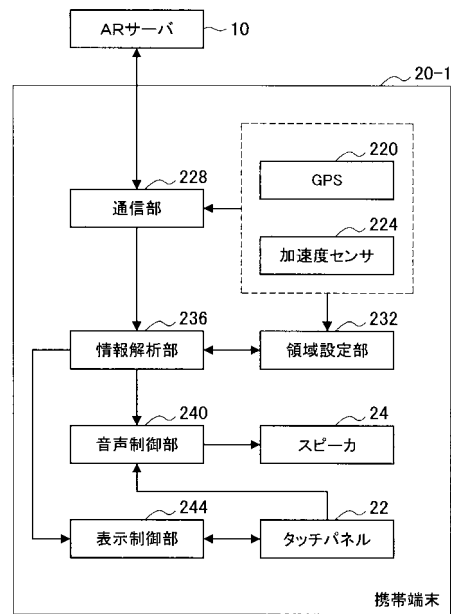
【図 2】



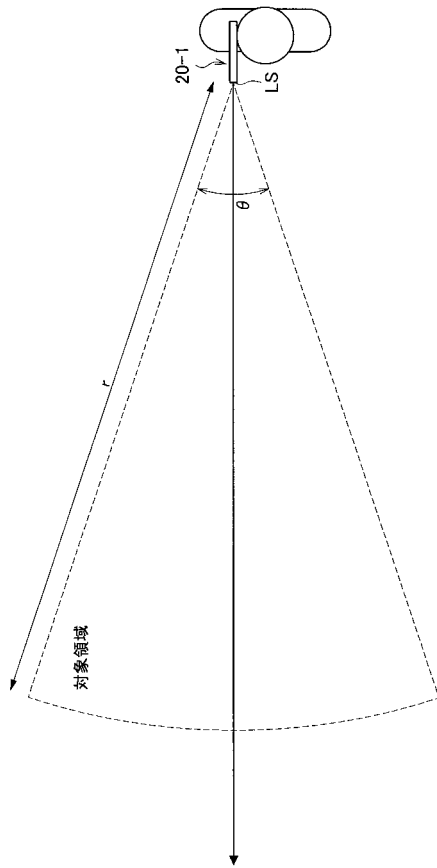
【図 3】



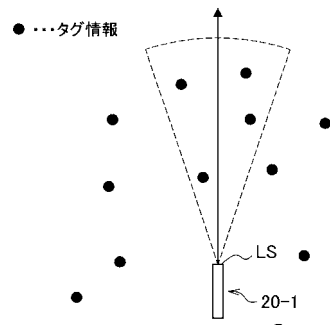
【図 4】



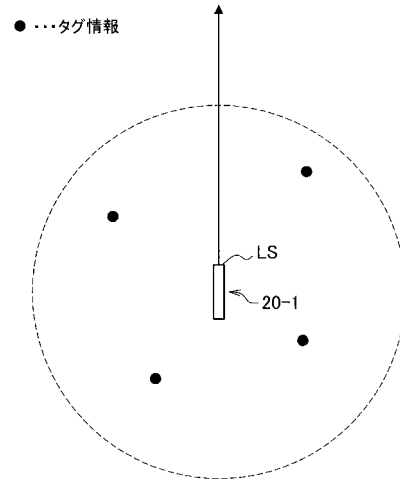
【図 5】



【図 6】



【図 7】

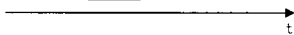


【図 8】

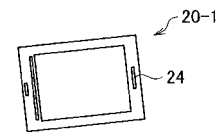
タグ情報TA: KKK のアルバム最高！

タグ情報TA: ケーキ美味しい！ FFF Cafe

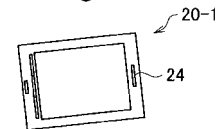
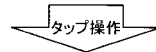
タグ情報TA: BBB で女優 A 発見！



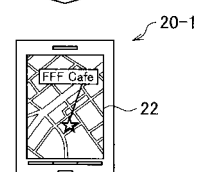
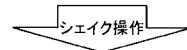
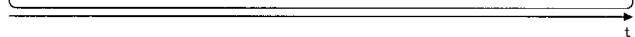
【図 9】



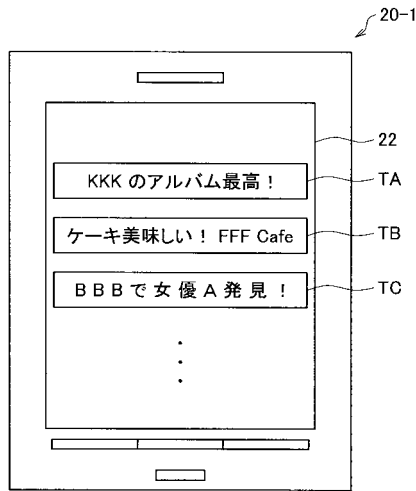
KKK のアルバム最高！
ケーキ美味しい！ FFF Cafe
BBB で女優 A 発見！



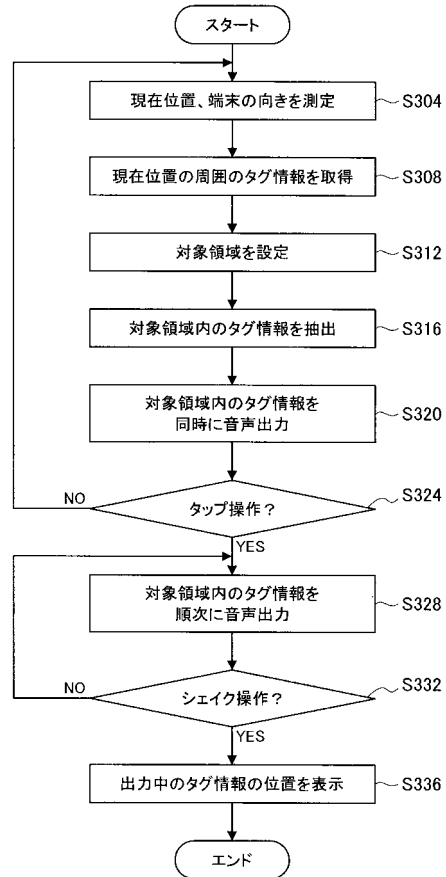
KKK のアルバム最高！ ケーキ美味しい！ FFF Cafe BBB で女優 A 発見！



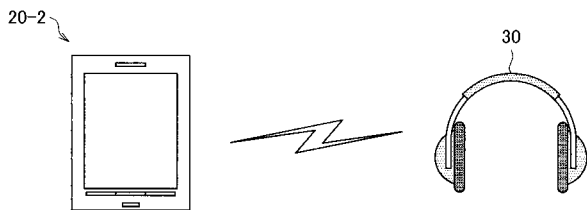
【図 10】



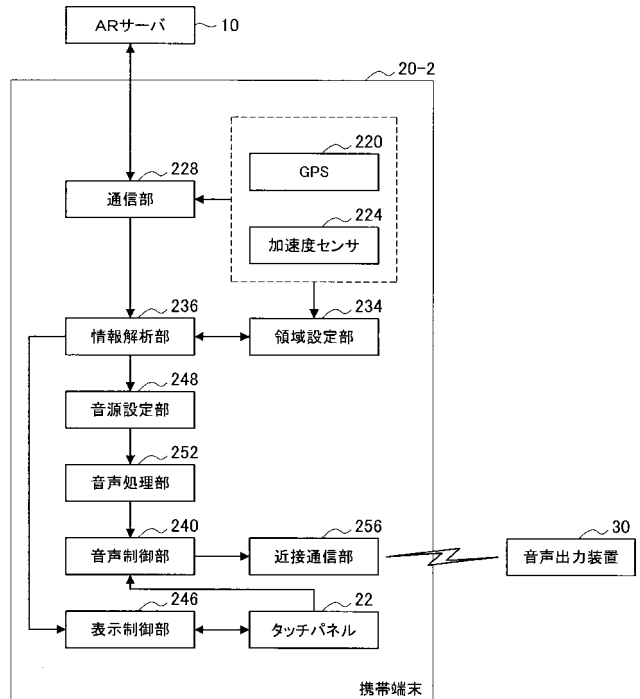
【図 11】



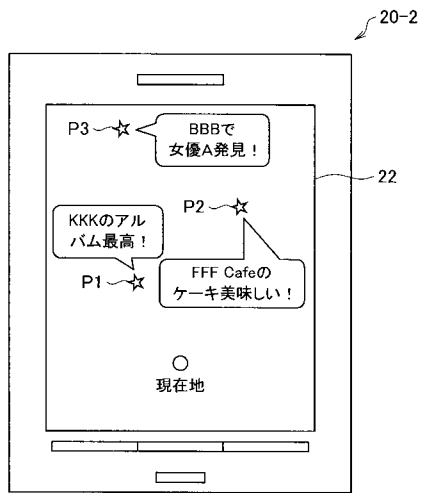
【図 12】



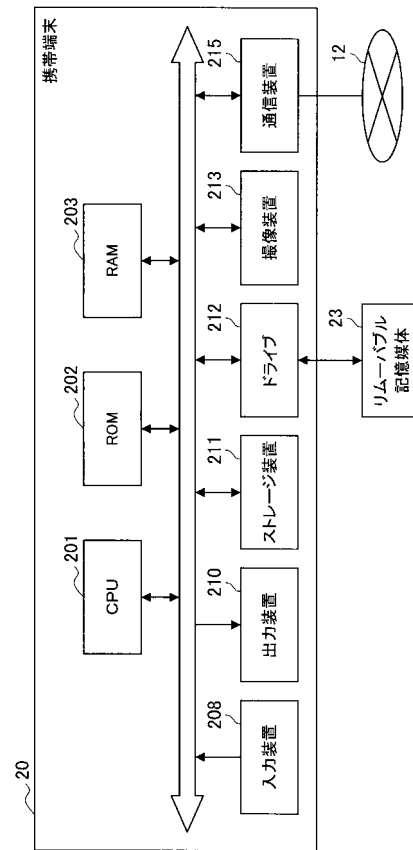
【図 13】



【図 14】



【図 15】



フロントページの続き

(51) Int.Cl.			F I			テーマコード (参考)
H 0 4 S	5/02	(2006.01)	H 0 4 R	3/00	3 1 0	
G 0 6 F	3/16	(2006.01)	H 0 4 S	5/02		D
			G 0 6 F	3/16	3 3 0 C	

(72)発明者 笠原 俊一
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内

(72)発明者 繁田 脩
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内

(72)発明者 鈴木 誠司
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内

(72)発明者 深澤 遼
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内

F ターム(参考) 5D020 AC00
5D062 BB10