(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第5366043号
(P5366043)

(45)発行日　平成25年12月11日(2013.12.11)　(24)登録日　平成25年9月20日(2013.9.20)

(51)Int.Cl.		F I		
G10K 15/04	(2006.01)	G10K	15/04	302F
H04S 1/00	(2006.01)	H04S	1/00	L
G10L 19/00	(2013.01)	G10L	19/00	312Z
H04S 7/00	(2006.01)	H04S	7/00	F

請求項の数 16　(全 30 頁)

(21)出願番号　特願2008-294524 (P2008-294524)
(22)出願日　平成20年11月18日 (2008.11.18)
(65)公開番号　特開2010-122369 (P2010-122369A)
(43)公開日　平成22年6月3日 (2010.6.3)
審査請求日　平成23年10月18日 (2011.10.18)

(73)特許権者　393031586
株式会社国際電気通信基礎技術研究所
京都府相楽郡精華町光台二丁目２番地２
(74)代理人　100090181
弁理士　山田　義人
(72)発明者　米澤　朋子
京都府相楽郡精華町光台二丁目２番地２
株式会社国際電気通信基礎技術研究所内
(72)発明者　山添　大丈
京都府相楽郡精華町光台二丁目２番地２
株式会社国際電気通信基礎技術研究所内

審査官　冨澤　直樹

最終頁に続く

(54)【発明の名称】音声記録再生装置

(57)【特許請求の範囲】
【請求項１】
記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生装置であって、
利用者の発する音声に基づいて音声情報を生成する音声情報生成手段、
記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定手段、
前記音声情報生成手段が生成した音声情報を、前記記録時方向を示す方向情報と対応つけて前記記録装置に記録する記録手段、
再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定手段、および
前記再生時方向に相当する記録時方向を示す方向情報と対応つけて前記記録装置に記録されている音声情報を前記記録装置から再生する第１音声再生手段を備える、音声記録再生装置。
【請求項２】
前記第１音声再生手段は、音声情報を、前記方向情報と利用者から当該音声情報までの距離を示す距離情報とに基づいて音源定位可能に再生する、請求項１記載の音声記録再生装置。
【請求項３】
音声情報の存在を示す象徴化した音声アイコンを、当該音声情報および前記方向情報と対応つけて前記記録装置に記録する音声アイコン記録手段、および
前記再生時方向に相当しない記録時方向を示す方向情報と対応付けて前記記録装置に記

録されている音声情報と対応つけて前記記録装置に記録されている音声アイコンを再生する音声アイコン再生手段をさらに備える、請求項１または２記載の音声記録再生装置。
【請求項４】
前記音声アイコン再生手段は、前記音声アイコンを、前記方向情報と利用者から当該音声情報までの距離を示す距離情報とに基づいて音源定位可能に生成する、請求項３記載の音声記録再生装置。
【請求項５】
前記記録手段は、音声情報を深さの異なる２以上のレイヤに、レイヤを示すレイヤ識別情報とさらに対応つけて記録し、さらに
レイヤの指定を受け付けるレイヤ指定受付手段を備え、
前記第１音声再生手段は、前記再生時方向に相当する記録時方向を示す方向情報と対応つけて前記記録装置に記録された音声情報のうちの前記レイヤ指定受付手段が受け付けたレイヤを示すレイヤ識別情報と対応つけられた音声情報を再生する、請求項１ないし４のいずれかに記載の音声記録再生装置。
【請求項６】
利用者から音声情報までの距離は、レイヤ識別情報が示すレイヤの深さに基づいて決定される、請求項５記載の音声記録再生装置。
【請求項７】
前記再生時方向特定手段は、利用者の顔の向きが示す方向、利用者の視線の向きが示す方向、および利用者の指差す向きが示す方向のいずれかを特定する、請求項１ないし６のいずれかに記載の音声記録再生装置。
【請求項８】
前記記録時方向特定手段は、利用者の顔の向きが示す方向、利用者の視線の向きが示す方向、および利用者の指差す向きが示す方向のいずれかを特定する、請求項１ないし６のいずれかに記載の音声記録再生装置。
【請求項９】
記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生のコンピュータによって実行するプログラムであって、前記プログラムは前記コンピュータを、
利用者の発する音声に基づいて音声情報を生成する音声情報生成手段、
記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定手段、
前記音声情報生成手段が生成した音声情報を、前記記録時方向を示す方向情報と対応つけて前記記録装置に記録する記録手段、
再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定手段、
前記再生時方向に相当する記録時方向を示す方向情報と対応つけて前記記録装置に記録されている音声情報を前記記録装置から再生する音声再生手段として機能させる、音声記録再生プログラム。
【請求項１０】
記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生のコンピュータが実行する方法であって、
利用者の発する音声に基づいて音声情報を生成する音声情報生成ステップ、
記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定ステップ、
前記音声情報生成ステップで生成した音声情報を、前記記録時方向を示す方向情報と対応つけて前記記録装置に記録する記録ステップ、
再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定ステップ、および
前記再生時方向に相当する記録時方向を示す方向情報と対応つけて前記記録装置に記録されている音声情報を前記記録装置から再生する音声再生ステップを含む、音声記録再生方法。
【請求項１１】
記録時の利用者の周辺画像を撮影する撮影手段、および

前記記録時方向を前記撮影手段が撮影した周辺画像上の位置に変換する変換手段をさらに備え、
前記記録手段は、音声情報を前記周辺画像および前記位置情報と対応つけて記録する、請求項１ないし８のいずれかに記載の音声記録再生装置。
【請求項１２】
前記撮影手段が撮影した周辺画像と最後に前記記録装置に記録された音声情報が対応つけられた周辺画像とを比較して、利用者の環境が変化したか否かを判断する第１環境変化判断手段をさらに備え、
前記記録手段は、前記第１環境変化判断手段が環境が変化したと判断した場合に音声情報を、前記撮影手段が撮影した周辺画像と対応つけて記録し、前記第１環境変化手段が環境が変化していないと判断した場合に音声情報を、最後に記録装置に記録された音声情報が対応つけられている周辺画像と対応つけて記録する、請求項１１記載の音声記録再生装置。
【請求項１３】
前記記録装置に最後に音声情報を記録してから所定の時間が経過したか否かを判断する時間経過判断手段、
前記撮影手段が撮影した周辺画像と最後に前記記録装置に記録された音声情報が対応つけられた周辺画像とを比較して利用者の環境が変化したか否かを判断する第２環境変化判断手段をさらに備え、
前記記録手段は、前記時間経過判断手段および前記第２環境変化判断手段の少なくとも一方の判断結果に基づいて、音声情報を、前記撮影手段が撮影した周辺画像と対応つけて記録するか、または音声情報を、最後に記録装置に記録された音声情報が対応つけられている周辺画像と対応つけて記録する、請求項１１記載の音声記録再生装置。
【請求項１４】
前記記録装置に記録されている周辺画像を表示するとともに、当該周辺画像に対応つけて記録されている音声情報を示すアイコンを、音声情報の位置情報に基づいた周辺画像上の位置に表示する表示手段、
前記表示手段に表示されたアイコンの指定を受け付けるアイコン指定手段、
前記アイコン指定手段によって指定が受け付けられたアイコンが示す前記記録装置に記録された音声情報を再生する第２音声再生手段をさらに備える、請求項１１ないし１３のいずれかに記載の音声記録再生装置。
【請求項１５】
前記表示手段は、前記記録手段に記録されている複数の周辺画像を表示し、
前記表示手段に表示された周辺画像の指定を受け付ける画像指定手段をさらに備え、
前記アイコン指定手段は、前記画像指定手段によって指定が受け付けられた周辺画像上に表示されたアイコンの指定を受け付ける、請求項１４記載の音声記録再生装置。
【請求項１６】
前記表示手段に表示された周辺画像上のアイコンの変更位置の指定を受け付ける位置指定手段、
前記位置指定手段が受け付けた変更位置に基づいて当該アイコンの音声情報と対応つけて前記記録装置に記録されている位置情報を変更する変更手段をさらに備える、請求項１４または１５記載の音声記録再生装置。
【発明の詳細な説明】
【技術分野】
【０００１】
この発明は、音声記録再生装置に関し、特に、方向を示す情報と対応つけて音声情報を記録し、記録された音声情報を再生する音声記録再生装置に関する。
【背景技術】
【０００２】
記録手段に記録された音声情報を再生する音声再生装置として、ＩＣレコーダ（ボイス

レコーダ）と呼ばれる装置がある。このようなＩＣレコーダに関する技術としては、特許文献１、特許文献２、および特許文献３などさまざまな工夫がなされたものが存在する。
【特許文献１】特開２００７－３２５３９２号公報［H02J 7/00, G10L 19/00, H01M 10/48］
【特許文献２】特開２００５－２９２６９３号公報［G10L 15/00, G10L 15/22, G10L 15/24, G10L 15/28］
【特許文献３】特開２００５－２６１３３号公報［H01H 15/16, G10L 19/00, H01H 3/52］
【発明の開示】
【発明が解決しようとする課題】
【０００３】
しかし、従来、ＩＣレコーダなどの音声再生装置において、再生の態様について工夫されたものが無かった。つまり、従来のＩＣレコーダでは、装置に設けられた小さな液晶画面に、音声情報の通し番号や記録日時の情報を表示して目的の音声情報に当たりを付けて再生するものであり、どの音声情報を再生する場合にも、体で感じる再生操作の感覚（体感）は同じであった。そのため、音声情報の内容を再生操作の体感と結びつけることによって、音声情報の内容を記憶にとどめやすくしたり、音声情報の内容を思考に取り入れやすくしたりすることはできなかった。
【０００４】
それゆえに、この発明の主たる目的は、新規な音声記録再生装置を提供することである。
【０００５】
また、この発明の他の目的は、音声情報を再生する際の再生操作の体感を音声情報の内容と結びつけて、音声情報の内容を記憶にとどめやすくしたり、音声情報の内容を思考に取り入れやすくしたりできる、音声記録再生装置を提供することである。
【課題を解決するための手段】
【０００６】
本発明は、上記の課題を解決するために、以下の構成を採用した。なお、括弧内の参照符号および補足説明等は、本発明の理解を助けるために後述する実施の形態との対応関係を示したものであって、本発明を何ら限定するものではない。
【０００７】
第１の発明は、記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生装置であって、利用者の発する音声に基づいて音声情報を生成する音声情報生成手段、記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定手段、音声情報生成手段が生成した音声情報を、記録時方向を示す方向情報と対応つけて記録装置に記録する記録手段、再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定手段、および再生時方向に相当する記録時方向を示す方向情報と対応つけて記録装置に記録されている音声情報を記録装置から再生する第１音声再生手段を備える、音声記録再生装置である。
【０００８】
第１の発明では、音声情報生成手段（１２、Ｓ３、Ｓ３５、Ｓ８３、Ｓ１３３）が利用者の発する音声に基づいて音声情報を生成し、記録時方向特定手段（１４、１００、Ｓ５、Ｓ３７、Ｓ８５、Ｓ１３５）が音声情報の記録時に利用者が身体を使って示す方向を特定し、記録手段は、記録装置（１０３）に、音声情報生成手段が生成した音声情報を記録時方向特定手段が特定した記録時方向を示す方向情報と対応つけて記録する。そして、再生時方向特定手段（１４、１００、Ｓ１１、Ｓ６３）が音声情報の再生時に利用者が身体を使って示す方向を特定し、第１音声再生手段（１００、Ｓ１５、Ｓ６７、）が再生時方向特定手段が特定した再生時方向に該当する方向を示す方向情報と対応つけて記録装置に記録された音声情報を記録装置から再生する。
【０００９】

第１の発明によれば、音声情報は利用者が身体を使って示す方向と対応つけて記録されるので、この音声情報を再生する際には、利用者は身体を使って音声情報の記録時に示したのと同じ方向を示すことになり、利用者は再生時に記録時の身体の体感を再体験でき、音声情報の記録時の思考を呼び覚ましやすくしたり、音声情報の内容を記憶にとどめやすくしたり、音声情報の内容を思考に取り入れやすくしたりできる。さらに、利用者が身体を使って示す方向に該当する方向を示す方向情報と対応つけられた音声情報を再生するので、音声情報を再生する際の再生操作の体感を音声情報の内容と結びつけて、音声情報の内容を記憶にとどめやすくしたり、音声情報の内容を思考に取り入れやすくしたりできる。

【００１０】

第２の発明は、第１の発明に従属する発明であって、第１音声再生手段は、音声情報を、方向情報と利用者から当該音声情報までの距離を示す距離情報とに基づいて音源定位可能に再生する音声記録再生装置である。

【００１１】

第２の発明では、第１再生手段は、音声情報を、方向情報と利用者から当該音声情報までの距離を示す距離情報とに基づいて音源定位可能に再生する。

【００１２】

第２の発明によれば、音声情報が音源定位可能に再生されるので、利用者は、音声情報が位置する方向を特定することができる。

第３の発明は、第１または第２の発明に従属する発明であって、音声情報の存在を示す象徴化した音声アイコンを、当該音声情報および方向情報と対応つけて記録装置に記録する音声アイコン記録手段、および再生時方向に相当しない記録時方向を示す方向情報と対応付けて記録装置に記録されている音声情報と対応つけて記録装置に記録されている音声アイコンを再生する音声アイコン再生手段をさらに備える、音声記録再生装置である。

第３の発明によれば、音声アイコンが再生されるので、利用者のそのときの再生時方向に相当しない記録時方向にも音声情報が記録されていることが、容易に分かる。

【００１３】

第４の発明は、第３の発明に従属する発明であって、音声アイコン再生手段は、音声アイコンを、方向情報と利用者から当該音声情報までの距離を示す距離情報とに基づいて音源定位可能に生成する、音声記録再生装置である。

【００１４】

第４の発明では、音声アイコン再生手段は、再生時方向特定手段が特定した再生時方向に該当しない記録時方向を示す方向情報と対応つけて記録装置に記録された音声情報について、当該音声情報の存在を示す音声アイコンを方向情報と距離情報とに基づいて音源定位可能に生成する。

【００１５】

第４の発明によれば、利用者が示す再生時方向以外の記録時方向に位置する音声情報について、当該音声情報の存在を示す音声アイコンを音源定位可能に再生するので、利用者は記録装置に記録されているすべての音声情報の存在を一度に把握することができる。

【００１９】

第５の発明は、第１の発明ないし第４の発明のいずれかに従属する発明であって、記録手段は、音声情報を深さの異なる２以上のレイヤに、レイヤを示すレイヤ識別情報とさらに対応つけて記録し、さらにレイヤの指定を受け付けるレイヤ指定受付手段を備え、第１音声再生手段は、再生時方向に相当する記録時方向を示す方向情報と対応つけて記録装置に記録された音声情報のうちのレイヤ指定受付手段が受け付けたレイヤを示すレイヤ識別情報と対応つけられた音声情報を再生する、音声記録再生装置である。

【００２０】

第５の発明では、記録手段は、音声情報をさらにレイヤを識別するレイヤ識別情報と対応つけて記録する。レイヤ指定受付手段（１００、Ｓ２３、Ｓ５５）はレイヤの指定を受け付け、第１音声再生手段は再生時方向特定手段が特定した再生時方向に該当する記録時

方向を示す方向情報と対応つけて記録装置に記録された音声情報のうちのレイヤ指定受付手段が受け付けたレイヤを示すレイヤ識別情報と対応つけられた音声情報を再生する。つまり、第１音声再生手段は、指定されたレイヤから音声情報を再生する。

【００２１】

第５の発明によれば、音声情報をレイヤに属させてレイヤ毎に区別することができる。

【００２５】

第６の発明は、第５の発明に従属する発明であって、利用者から音声情報までの距離は、レイヤ識別情報が示すレイヤの深さに基づいて決定される、音声記録再生装置である。

【００２６】

第６の発明では、音声情報までの距離は、レイヤ識別情報が示すレイヤの深さ基づいて決定される。

【００２７】

第６の発明によれば、音声情報までの距離をレイヤの深さによって決定することができる。

【００３１】

第７の発明は、第１の発明ないし第６の発明のいずれかに従属する発明であって、再生時方向特定手段は、利用者の顔の向きが示す方向、利用者の視線の向きが示す方向、および利用者の指差す向きが示す方向のいずれかを特定する、音声記録再生装置である。

【００３２】

第７の発明では、再生時方向特定手段は、利用者が顔の向きによって示す方向、視線の向きによって示す方向、および指差す向きによって示す方向のいずれかを再生時方向として特定する。

【００３３】

第７の発明によれば、利用者は、再生する音声情報を、顔の向き、視線の向き、および指差しの向きのいずれかによって指示することができる。

【００３４】

第８の発明は、第１の発明ないし第７の発明のいずれかに従属する発明であって、記録時方向特定手段は、利用者の顔の向きが示す方向、利用者の視線の向きが示す方向、および利用者の指差す向きが示す方向のいずれかを特定する、音声記録再生装置である。

【００３５】

第７の発明では、記録時方向特定手段は、利用者が顔の向きによって示す方向、視線の向きによって示す方向、および指差す向きによって示す方向のいずれかを記録時方向として特定する。

【００３６】

第７の発明によれば、利用者は、記録する音声情報を付加する仮想空間における方向を、顔の向き、視線の向き、および指差しの向きのいずれかによって指示することができる。

第９の発明は、記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生のコンピュータによって実行するプログラムであって、プログラムはコンピュータを、利用者の発する音声に基づいて音声情報を生成する音声情報生成手段、記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定手段、音声情報生成手段が生成した音声情報を、記録時方向を示す方向情報と対応つけて記録装置に記録する記録手段、再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定手段、再生時方向に相当する記録時方向を示す方向情報と対応つけて記録装置に記録されている音声情報を記録装置から再生する音声再生手段として機能させる、音声記録再生プログラムである。

第１０の発明は、記録装置に音声情報を記録し、記録装置から音声情報を再生する、音声記録再生のコンピュータが実行する方法であって、利用者の発する音声に基づいて音声情報を生成する音声情報生成ステップ、記録時に利用者が身体を使って示す記録時方向を特定する記録時方向特定ステップ、音声情報生成ステップで生成した音声情報を、記録時

方向を示す方向情報と対応つけて記録装置に記録する記録ステップ、再生時に利用者が身体を使って示す再生時方向を特定する再生時方向特定ステップ、および再生時方向に相当する記録時方向を示す方向情報と対応つけて記録装置に記録されている音声情報を記録装置から再生する音声再生ステップを含む、音声記録再生方法である。

【００３７】

第１１の発明は、第1の発明ないし第８の発明のいずれかに従属する発明であって、記録時の利用者の周辺画像を撮影する撮影手段、および記録時方向を撮影手段が撮影した周辺画像上の位置に変換する変換手段をさらに備え、記録手段は、音声情報を周辺画像および位置情報と対応つけて記録する、音声記録再生装置である。

【００３８】

第１１の発明では、撮影手段（１７）が音声情報の記録時の利用者の周辺画像を撮影し、変換手段（１００、Ｓ９９、Ｓ１５５）が記録時方向特定手段が特定した方向を撮影手段が撮影した周辺画像上の位置に変換する。そして、記録手段は、音声情報生成手段が生成した音声情報を撮影手段が撮影した周辺画像および変換手段が変換した位置を示す位置情報と対応つけて記録する。

【００３９】

第１１の発明によれば、音声情報を周辺画像上の位置に対応つけて記録することができる。

【００４０】

第１２の発明は、第１１の発明に従属する発明であって、撮影手段が撮影した周辺画像と最後に記録装置に記録された音声情報が対応つけられた周辺画像とを比較して、利用者の環境が変化したか否かを判断する第１環境変化判断手段をさらに備え、記録手段は、第１環境変化判断手段が環境が変化したと判断した場合に音声情報を、撮影手段が撮影した周辺画像と対応つけて記録し、第１環境変化手段が環境が変化していないと判断した場合に音声情報を、最後に記録装置に記録された音声情報が対応つけられている周辺画像と対応つけて記録する、音声記録再生装置。

である。

【００４１】

第１２の発明では、第１環境変化判断手段（１００、Ｓ９１）は撮影手段が撮影した周辺画像と最後に記録装置に記録された音声情報が対応つけられた周辺画像とを比較して利用者の周辺の環境が変化したか否かを判断し、記録手段は、第１環境変化判断手段が環境が変化したと判断した場合に音声情報生成手段が生成した音声情報を、撮影手段が撮影した周辺画像と対応つけて記録し、第１環境変化手段が環境が変化していないと判断した場合に音声情報生成手段が生成した音声情報を、最後に当該記録手段に記録された音声情報が対応つけられている周辺画像と対応つけて記録する。

【００４２】

第１２の発明によれば、利用者の周辺の環境が変化した場合には、変換した環境の周辺画像に音声情報を付加することができる。

【００４３】

第１３の発明は、第１１の発明に従属する発明であって、記録装置に最後に音声情報を記録してから所定の時間が経過したか否かを判断する時間経過判断手段、撮影手段が撮影した周辺画像と最後に記録装置に記録された音声情報が対応つけられた周辺画像とを比較して利用者の環境が変化したか否かを判断する第２環境変化判断手段をさらに備え、記録手段は、時間経過判断手段および第２環境変化判断手段の少なくとも一方の判断結果に基づいて、音声情報を、撮影手段が撮影した周辺画像と対応つけて記録するか、または音声情報を、最後に記録装置に記録された音声情報が対応つけられている周辺画像と対応つけて記録する、音声記録再生装置。

である。

【００４４】

第１３の発明では、時間経過判断手段（１００、Ｓ１４、Ｓ１４５）は記録装置に最後

に音声情報を記録してから所定の時間が経過したか否かを判断し、第２環境変化判断手段（１００、Ｓ１１９）は撮影手段が撮影した周辺画像と最後に記録装置に記録された音声情報が対応つけられた周辺画像とを比較して利用者の周辺の環境が変化したか否かを判断する。そして、記録手段は、時間経過判断手段および第２環境変化判断手段の少なくとも一方の判断結果に基づいて、音声情報生成手段が生成した音声情報を撮影手段が撮影した周辺画像と対応つけて記録するか、音声情報生成手段が生成した音声情報を最後に当該記録装置に記録された音声情報が対応つけられている周辺画像と対応つけて記録する。

【００４５】

第１３の発明によれば、最後に音声情報を記録してからの経過時間、および利用者の周辺の環境が変化の少なくとも一方に基づいて、音声情報を付加する周辺画像を更新するか否かを決定することができる。

【００４６】

第１４の発明は、第１０の発明ないし第１３の発明のいずれかに従属する発明であって、記録装置に記録されている周辺画像を表示するとともに、記録装置に記録されている周辺画像に対応つけて記録されている音声情報を示すアイコンを当該音声情報の位置情報に基づいた周辺画像上の位置に表示する表示手段、表示手段に表示されたアイコンの指定を受け付けるアイコン指定手段、アイコン指定手段によって指定が受け付けられたアイコンが示す記録装置に記録された音声情報を再生する第２音声再生手段をさらに備える、音声記録再生装置である。

【００４７】

第１４の発明では、表示手段（１０６、Ｓ１０１、Ｓ１０３、Ｓ１０７、Ｓ１０９）は記録装置に記録されている周辺画像を表示するとともに、記録装置に記録されている周辺画像に対応つけて記録されている音声情報を示すアイコンを当該音声情報の位置情報に基づいた周辺画像上の位置に表示し、アイコン指定手段（１０５、Ｓ１１１）は表示手段に表示されたアイコンの指定を受け付け、第２再生手段（１００、Ｓ１１３）はアイコン指定手段によって指定が受け付けられたアイコンが示す記録装置に記録された音声情報を再生する。

【００４８】

第１４の発明によれば、周辺画像上に表示されたアイコンを指定することによってそのアイコンに対応する音声情報を再生することができる。

【００４９】

第１５の発明は、第１４の発明に従属する発明であって、表示手段は、記録装置に記録されている複数の周辺画像を表示し、表示手段に表示された周辺画像の指定を受け付ける画像指定手段をさらに備え、アイコン指定手段は、画像指定手段によって指定が受け付けられた周辺画像上に表示されたアイコンの指定を受け付ける、音声再生装置である。

【００５０】

第１５の発明では、表示手段は記録装置に記録されている複数の周辺画像を表示し、画像指定手段（１０５、Ｓ１０５）は表示手段に表示された周辺画像の指定を受け付け、アイコン指定手段は画像指定手段によって指定が受け付けられた周辺画像上に表示されたアイコンの指定を受け付ける。

【００５１】

第１５の発明によれば、周辺画像を指定し、その周辺画像上に表示されたアイコンを指定して、そのアイコンに対応する音声情報を再生することができる。

【００５２】

第１６の発明は、第１４の発明または第１５の発明に従属する発明であって、表示手段に表示された周辺画像上のアイコンの変更位置の指定を受け付ける位置指定手段、位置指定手段が受け付けた変更位置に基づいて当該アイコンの音声情報と対応つけて記録装置に記録されている位置情報を変更する変更手段をさらに備える、音声記録再生装置である。

【００５３】

第１６の発明では、位置指定手段（１０５、Ｓ１１５）は表示手段に表示された周辺画

像上のアイコンの変更位置の指定を受け付け、変更手段（１００、Ｓ１１７）は位置指定手段が受け付けた変更位置に基づいて当該アイコンの音声情報と対応つけて記録装置に記録されている位置情報を変更する。

【００５４】

第１６の発明によれば、周辺画像において音声情報が付加された位置、つまり周辺画像上にアイコンが表示される位置を変更することができる。

【発明の効果】

【００５５】

この発明によれば、音声情報は利用者が身体を使って示す方向と対応つけて記録されるので、この音声情報を再生する際には、利用者は身体を使って音声情報の記録時に示したのと同じ方向を示すことになり、利用者は再生時に記録時の身体の体感を再体験でき、音声情報の記録時の思考を呼び覚ましやすくしたり、音声情報の内容を記憶にとどめやすくしたり、音声情報の内容を思考に取り入れやすくしたりできる。

【００５６】

この発明の上述の目的，その他の目的，特徴および利点は、図面を参照して行う以下の実施例の詳細な説明から一層明らかとなろう。

【発明を実施するための最良の形態】

【００５７】

＜第１の実施形態＞

図１を参照してこの発明の音声記録再生装置１０は、個人利用の携帯情報端末（ＰＤＡ：Personal Digital Assistants）１１と、骨伝導マイク１２と、例えば耳掛け型のステレオヘッドフォン１３と、ステレオヘッドフォン１３に内蔵された３軸地磁気センサ１４とから構成される。骨伝導マイク１２、ステレオヘッドフォン１３、および３軸地磁気センサ１４は携帯情報端末（以下、“ＰＤＡ１１”と記載する。）にそれぞれ接続される。

【００５８】

骨伝導マイク１２はマイクを装着している利用者の音声を内耳道壁に伝達される信号として直接検出するので喧騒の中においても、利用者の音声を拾うことができる。なお、周りの騒音を考慮しなければ、骨伝導マイク１２に替えて通常のマイクを用いてもよい。

【００５９】

３軸地磁気センサ１４としては、バイテック社製の３ＤセンサモジュールＴＤＳ０１Ｖが利用できる。なお、ＴＤＳ０１Ｖには、３軸地磁気センサ１４の他に３軸加速度センサも備えている。

【００６０】

ＰＤＡ１１は、例えば利用者の衣類のポケット１５などに収容される。このＰＤＡ１１は、図２に示すように、ＣＰＵ１００、ＲＯＭ１０１、ＲＡＭ１０２、フラッシュメモリ１０３、時計回路１０４、タッチパネル１０５、ディスプレイ１０６、オーディオ入力Ｉ／Ｆ１０７、オーディオ出力Ｉ／Ｆ１０８、地磁気センサ信号入力Ｉ／Ｆ１０９から構成されている。

【００６１】

ＰＤＡ１１は、ディスプレイ１０６の表示に基づいて、ディスプレイ１０６の上面に配置されたタッチパネル１０５に入力操作を行うことによって操作することができる。オーディオ入力Ｉ／Ｆ１０７は、骨伝導マイク１２からの音声信号をＰＤＡ１１に入力するためのインターフェースである。また、オーディオ出力Ｉ／Ｆ１０８は、ＰＤＡ１１で再生された音声信号をステレオヘッドフォン１３に出力するためのインターフェースである。そして、地磁気センサ信号入力Ｉ／Ｆ１０９は、３軸地磁気センサ１４からの信号をＰＤＡ１１に入力するためのインターフェースである。

【００６２】

このような音声記録再生装置１０によれば、骨伝導マイク１２で拾った利用者の発する音声の情報（音声情報）をＰＤＡ１１のフラッシュメモリ１０３に記録することができ、またフラッシュメモリ１０３に記録された音声情報を再生してステレオヘッドフォン１３

に出力することができる。

【００６３】

そして、この音声記録再生装置１０によれば、利用者の発する音声の情報を仮想空間の任意の位置に付加して記録することができ、また仮想空間に付加された任意の音声の情報を再生することができる。

【００６４】

より具体的には、利用者の音声を記録する際には、３軸地磁気センサ１４でその際の利用者の仮想空間における顔の向きを検出し、この顔の向きを示す方向情報を音声情報と対応つけてフラッシュメモリ１０３に記憶する。そして、フラッシュメモリ１０３に記憶された音声情報を再生する際には、３軸地磁気センサ１４でその際の利用者の仮想空間における顔の向きを検出し、検出された顔の向きから所定の範囲内の方向を示す方向情報と対応つけてフラッシュメモリ１０３に記録されている音声情報を生成する。この際、音声情報は、利用者が仮想空間において音源定位が可能となるように再生される。

【００６５】

また、音声情報の再生の際に、検出された顔の向きから所定範囲外の方向を示す方向情報と対応つけてフラッシュメモリ１０３に記録されている音声情報については、音声情報の代わりに音声情報の存在を示す象徴化された音声（以下、“アイコン音声”と呼ぶ。）が再生される。このアイコン音声も音源定位が可能となるように再生される。

【００６６】

このように、本発明の音声記録再生装置１０によれば、仮想空間における利用者が記録時に顔の向きで指定した位置（方向）に音声情報を付加し、また再生時に顔の向きで指定した位置（方向）に付加されている音声情報を再生することができる。したがって、音声情報を記録したり再生したりする操作の際の顔の向きという体感が音声情報の内容と結びつけられるので、音声情報の内容が記憶にとどめやすくなり、また音声情報の内容が思考に取り入れやすくなる。

【００６７】

また、音声情報とアイコン音声とは、仮想空間において利用者が音源定位することが可能なように再生されるので、ＰＤＡ１１のフラッシュメモリ１０３に記憶されているすべての音声情報のそれぞれの存在を一度に把握することができる。

【００６８】

以下に、図３および図４に示すフロー図などを参照しつつ、音声記録再生装置１０で音声を記録する際および音声を再生する際にＰＤＡ１１のＣＰＵ１００が実行する処理について説明する。なお、音声を記録する際の処理と音声を再生する際の処理とを図３のフロー図と図４のフロー図に別々に示しているが、図３のフロー図の処理と図４のフロー図の処理とは、ＣＰＵ１００によって並行して実行される。つまり、音声の記録と音声の再生は同時に行うことができる。ただし、フロー図には示していないが、音声を記録する際には、利用者の思考の邪魔とならないように、再生される音声の音量は低く抑えられる。なお、図３および図４のフロー図に示す処理は一例であり、各処理ステップの順序はこれを変更しても発明を実施できるものについては順序を変更してもかまわない。

【００６９】

まず、ＰＤＡ１１のＣＰＵ１００は、図３のステップＳ１で、音声の入力があるか否かを判断する。この音声の入力があるか否かの判断は、オーディオ入力Ｉ／Ｆ１０７を介して骨伝導マイク１２から入力される音声信号に基づいて判断される。この判断では、オーディオ入力Ｉ／Ｆ１０７に入力される音声信号から音声の信号（有音）である区間（音声区間）の切り出しを試み、音声区間が切り出せた場合に音声の入力があると判断する。

【００７０】

音声区間の切り出しでは、音声信号のレベルが閾値を上回る時間が０．３秒を超えた位置を発話開始（音声区間開始）とみなし、その後音声信号のレベルが閾値を下回る時間が１．０秒を超えたときに閾値を下回り始めた位置を発話停止（音声区間終了）とみなすことにより音声区間を決定する。ただし、発話が開始されたと判定した後、０．５秒以内に

音声信号のレベルが閾値を下回り、かつ音声信号のレベルが閾値を下回ってからその状態が続く時間が２．０秒以上である場合には、発話が開始されたという判定を覆し音声区間とはみなさない。

【００７１】

このようにして、音声信号から音声区間が切り出されて、音声の入力があると判断すると（ステップＳ１：ＹＥＳ）、次に、ＣＰＵ１００はステップＳ３で、音声信号から音声区間の音声信号を取得して音声情報を生成する。そして生成した音声情報をフラッシュメモリ１０３に保存（記録）する。この音声情報は、例えば、ＭＰ３（MPEG Audio Layer-3）ファイルなどである。

【００７２】

ステップＳ５では、利用者の音声が骨伝道マイク１２に入力された際（音声情報を記録する際）の利用者の仮想空間における顔の向きを示す情報（方向情報）を３軸地磁気センサ１４の出力に基づいて生成し、この方向情報をフラッシュメモリ１０３に先に保存された音声情報と対応つけて保存する。ここで、仮想空間とは、図５に示すように、利用者の頭部を中心とする水平な円弧Ｃであり、円弧Ｃは仮想空間の基準である正面から左右にそれぞれ９０°の広がりを有している。なお、仮想空間の基準となる正面は、音声記録再生装置１０の電源が投入されて、利用者がＰＤＡ１１のタッチパネル１０５に対して所定の操作を行った時に、３軸地磁気センサ１４が検出した利用者の顔が向いている方向である。

【００７３】

そして、利用者の仮想空間における顔の向きを示す情報（方向情報）は、仮想空間の基準（正面）線と利用者の顔の向きが示す直線とが成す角度αである。この利用者の顔の向きを示す角度αは、利用者の顔が基準方向（正面方向）を向いているときの３軸地磁気センサ１４の出力を［I_x，I_y，I_z］’とし、顔の向きの測定時の３軸地磁気センサ１４の出力を［M_x，M_y，M_z］’とすると数１および数２で得ることができる。

【００７４】

【数１】

$$\alpha_t = \tan^{-1}\left(\frac{M_y}{M_x}\right) \quad , \quad \alpha_i = \tan^{-1}\left(\frac{I_y}{I_x}\right) \qquad (1)$$

【００７５】

【数２】

$$\alpha = \begin{cases} \alpha_t - \alpha_i & (-180 \le \alpha_t - \alpha_i \le 180) \\ \alpha_t - \alpha_i + 360 & (\alpha_t - \alpha_i \le -180) \\ \alpha_t - \alpha_i - 360 & (\alpha_t - \alpha_i \ge 180) \end{cases} \qquad (2)$$

【００７６】

実験によると、利用者が一方向に顔を向けて音声を入力しているつもりであっても、音声の入力の開始から入力の終了までの間に、角度αに５°から７°のブレがあることが確認された。そのため、音声の入力時の利用者の顔の向きを示す方向情報は、音声の入力の開始から音声の入力の終了までの間（音声区間）の顔の向き（角度α）の平均値とする。なお、３軸地磁気センサ１４からは１００ミリ秒毎に測定値が出力される。

【００７７】

ステップＳ７では、アイコン音声情報を生成して、これを先に保存した音声情報および方向情報と対応つけてフラッシュメモリ１０３に保存する。このアイコン音声情報は、例えば、ＭＰ３（MPEG Audio Layer-3）ファイルなどである。なお、このアイコン音声情報は、あらかじめフラッシュメモリ１０３に記録されているアイコン音声情報の複製である。また、アイコン音声は象徴化された音、例えば、単純な鐘の音の楽音である。ただし、鐘の音の楽音は、利用者の左右の耳の音量差によって音源定位が容易となるように、１００００Ｈｚ以上を含むことが望ましい。

【００７８】

次に、ステップＳ９では、アイコン音声情報に基づいて再生されるアイコン音声の変調パラメータを決定し、これを先に保存した音声情報などと対応つけてフラッシュメモリ１０３に保存する。ここで、変調パラメータとは、利用者が再生されるアイコン音声を識別しやすいように、仮想空間におけるアイコン音声の位置に応じてアイコン音声の周波数をアイコン音声毎に異ならせるためのものである。変調パラメータは、音声の記録時の利用者の顔の向きの角度α（平均値）の絶対値に基づいて、アイコン音声の周波数を数３によって、１．０倍から２．０倍の間で変換して決定される。この周波数変換により最大で１オクターブの差があるアイコン音声が生成される。

【００７９】

【数３】

$$speed = \left(\sqrt[12]{2}\right)^{|\alpha|/90*12} \qquad (3)$$

【００８０】

このようにして、利用者が発した音声をＰＤＡ１１に記録することによって、フラッシュメモリ１０３には、１つの音声情報について、図６（ａ）に示すように、音声情報、方向情報、アイコン音声情報、変調パラメータが記録される。

【００８１】

一方、音声情報を再生する際には、ＰＤＡ１１のＣＰＵ１００は、図４のステップＳ１１で、３軸地磁気センサ１４の出力に基づいて、仮想空間において利用者の顔が向いている方向を特定する。つまり、利用者の顔が仮想空間の基準からどれだけの角度（α）の方向を向いているかを特定する。

【００８２】

次に、ステップＳ１３では、フラッシュメモリ１０３に記憶されているすべての音声情報について、仮想空間における音声情報が位置する方向が、ステップＳ１１で特定した利用者の顔の向きの方向に対してどれだけの相対的方向（角度δ）（図７参照）を有しているかを算出する。この相対的方向（角度δ）は、フラッシュメモリ１０３に音声情報と対応つけて記録されている方向情報とステップＳ１１で特定された利用者の顔の向きの角度の情報に基づいて算出される。

【００８３】

そして、ステップＳ１５では、フラッシュメモリ１０３に記録されている音声情報のうち、ステップＳ１３で算出された相対的方向が、例えば、±１０°の範囲以内である音声情報を再生してオーディオ出力Ｉ／Ｆ１０８を介してステレオヘッドフォン１３に出力する。このとき、音声情報は、仮想空間における音声情報の位置を利用者が特定することが可能なように、つまり音源定位が可能なように再生される。

【００８４】

具体的には、利用者の左右の耳のそれぞれに対する音声の到達の時間差と音声の音量差とに基づいて音源定位が可能なように音声情報を再生する。

【００８５】

図７に示すように、利用者の左右の耳の間の距離をｈ［ｃｍ］、左右の耳から仮想空間である円弧Ｃまでの距離をｃ［ｃｍ］とすると、音源（音声情報）から左耳までの距離ｄＬ［ｃｍ］および右耳までの距離ｄＲ［ｃｍ］はそれぞれ、数４および数５で決定される。そして、音声の左右の耳への到達にかかる時間は、音速３４［ミリ秒／ｃｍ］を用いて、それぞれ、ｄＬ／３４、ｄＲ／３４で導出され、これらに基づいて時間差が算出される。

【００８６】

なお、図７（ａ）は利用者の顔が仮想空間の正面の方向を向いている場合を示しており、図７（ｂ）は利用者の顔が仮想空間の正面から角度αの方向を向いている場合を示している。また、これらの図において、角度γは仮想空間の正面の方向から音声情報が位置する方向までの角度を示しており、角度δは利用者の顔が向いている方向から音声情報が位置する方向までの角度を示している。

【００８７】

利用者の左右の耳の間の距離ｈ［ｃｍ］および利用者から円弧Ｃまでの距離ｃ［ｃｍ］はＰＤＡ１１のＲＯＭ１０１に記録されている。これは他の実施形態でも同じである。

【００８８】

【数４】

$$dL=\sqrt{\left\{\left(c+\frac{h}{2}\right)\cos\delta''+\frac{h}{2}\right\}^{2}+\left\{\left(c+\frac{h}{2}\right)\sin\delta''\right\}^{2}} \qquad (4)$$

（ただし、$\delta''=\delta+90$）

【００８９】

【数５】

$$dR=\sqrt{\left\{\left(c+\frac{h}{2}\right)\cos\delta''-\frac{h}{2}\right\}^{2}+\left\{\left(c+\frac{h}{2}\right)\sin\delta''\right\}^{2}} \qquad (5)$$

（ただし、$\delta''=\delta+90$）

【００９０】

一方、音量差は、一般的な音源であれば、常に同じ音量であっても定位が変わると左右の耳に届く音量［ｄｂ］の総合は常に同じであるとは限らないが、本実施形態では回り込み音声などの計算は行わず、単純に距離による減衰率“ａｔｎ”のみに基づいて算出する。減衰率の初期値（ATN_{init}）を“０．８c^{2}”とすると、音源と耳との距離“ｄ”とに基づく減衰率“ａｔｎ”は、数６で算出される。ただし、$d_{min}=c$である。例えば、ｃ＝１５、ｈ＝１５とすれば、$ATN_{init}=180$である。

【００９１】

【数６】

$$atn=\frac{ATN_{init}}{d^{2}} \qquad (6)$$

【００９２】
次に、ステップＳ１７では、フラッシュメモリ１０３に記録されている音声情報のうち、ステップＳ１３で算出された相対的方向が、例えば、±１０°の範囲外である音声情報について、アイコン音声情報を変調パラメータに基づいて変調して再生し、オーディオ出力Ｉ／Ｆ１０８を介してステレオヘッドフォン１３に出力する。アイコン音声情報も上述したように、利用者の左右の耳のそれぞれに対する音声の到達の時間差と音声の音量差とに基づいて音源定位が可能なように再生する。なお、ステップＳ１５における音声情報の再生とステップＳ１７におけるアイコン音声情報の再生とは同時に並行して実施される。また、音声情報とアイコン音声情報とは、５００ミリ秒の無音区間を挟みながら繰り返し再生される。

【００９３】
このようにして、第１の実施形態の音声記録再生装置１０では、図８に示すように、仮想空間の円弧Ｃにおいて、利用者の顔の向いている方向から±１０°の範囲以内に位置する音声情報（図中で黒塗りの音声アイコン２０１で示される音声情報）はそのまま音声情報が音源定位可能に再生され、利用者の顔の向いている方向から±１０°の範囲外に位置する音声情報（図中で白抜きの音声アイコン２０２で示される音声情報）は音声情報に替えてアイコン音声情報が音源定位可能に再生される。

＜第２の実施形態＞

第２の実施形態の音声記録再生装置１０では、音声情報を付加する仮想空間を３次元空間とする。つまり、仮想空間における利用者の顔の向きの情報（方向情報）として、第１実施形態と同じ水平方向の角度に加え、垂直方向の角度を取り入れる。また、レイヤという概念を取り入れ、仮想空間における音声情報の利用者（の頭の中心）からの距離を考慮する。なお、この実施形態では、レイヤとは、仮想空間において利用者の頭部の中心を同心として一定の距離をおいて位置する半球面状の層である。

【００９４】
具体的には、新たに音声情報を記録する際（仮想空間に付加する際）には、当該音声情報が属するレイヤを最も浅い（距離が近い）レイヤとし、新たな音声情報を記録する際にすでに記録されている音声情報については、新たな音声情報を記録する際に最後の音声情報を記録した時刻から所定の時間が経過していれば、当該音声情報が属するレイヤをより深い（距離が遠い）レイヤに更新する。つまり、すでに記録されている音声情報のレイヤをより利用者から遠いレイヤとする。

【００９５】
したがって、第２の実施形態の音声記録再生装置１０では、図９に示すように、仮想空間において利用者によって指定されたレイヤに属し（図９の例ではレイヤ２が指定されている。）、利用者の顔が向いている方向から水平方向に±１０°の範囲以内で、かつ垂直方向に±１０°の範囲以内に位置する音声情報（図中で黒塗りの音声アイコン２０１で示される音声情報）はそのまま音声情報を音源定位可能に再生し、利用者によって指定されたレイヤ以外に属する音声情報、および利用者によって指定されたレイヤに属するが利用者の顔が向いている方向から水平方向に±１０°の範囲外で、かつ垂直方向に±１０°の範囲外に位置する音声情報（図中で白抜きの音声アイコン２０２で示される音声情報）は音声情報に替えてアイコン音声情報を音源定位可能に再生する。

【００９６】
なお、利用者の左右の耳の間の距離をｈ［ｃｍ］とすると、隣接するレイヤ間の距離はｈ／２［ｃｍ］である。また、利用者の耳から最近のレイヤまでの距離はｃ［ｃｍ］である。

【００９７】
図９において、音声アイコン２０２は、利用者の顔の向きより下方に位置する音声情報についてはより小さく、顔の向きより上方に位置する音声情報についてはより大きくなるように示している。

【００９８】

なお、音声記録再生装置１０を構成するＰＤＡ１１、骨伝導マイク１２、ステレオヘッドフォン１３、および３軸地磁気センサ１４の構成は、第１の実施形態のものと同じである。

【００９９】

以下に、図１０および図１１に示すフロー図などを参照しつつ、音声記録再生装置１０で音声を記録する際および音声を再生する際にＰＤＡ１１のＣＰＵ１００が実行する処理について説明する。なお、音声を記録する際の処理と音声を再生する際の処理とを図１０のフロー図と図１１のフロー図に別々に示しているが、図１０のフロー図の処理と図１１のフロー図の処理とは、ＣＰＵ１００によって並行して実行される。つまり、音声の記録と音声の再生は同時に行うことができる。ただし、フロー図には示していないが、音声を記録する際には、利用者の思考の邪魔とならないように、再生される音声の音量は低く抑えられる。なお、図１０および図１１のフロー図に示す処理は一例であり、各処理ステップの順序はこれを変更しても発明を実施できるものについては順序を変更してもかまわない。

【０１００】

なお、図１０のフロー図の処理を実行して、利用者が発した音声をＰＤＡ１１に記録すると、フラッシュメモリ１０３には、１つの音声情報について、図６（ｂ）に示すように、レイヤ識別情報、音声情報、方向情報、アイコン音声情報、変調パラメータ、時刻情報が記録される。

【０１０１】

まず、ＰＤＡ１１のＣＰＵ１００は、図１０のステップＳ２１で、音声の入力があるか否かを判断する。この音声の入力があるか否かの判断は、オーディオ入力Ｉ／Ｆ１０７を介して骨伝導マイク１２から入力される音声信号に基づいて判断される。この判断では、オーディオ入力Ｉ／Ｆ１０７に入力される音声信号から音声の信号（有音）である区間（音声区間）の切り出しを試み、音声区間が切り出せた場合に音声の入力があると判断する。なお、音声区間の切り出しの方法は、第１の実施形態と同じである。

【０１０２】

音声信号から音声区間が切り出されて、音声の入力があると判断すると（ステップＳ２１：ＹＥＳ）、次に、ＣＰＵ１００はステップＳ２３で、音声区間の音声信号に基づいて音声認識を行って、利用者によって入力された音声がレイヤを指定する音声であるか否かを判断する。利用者はレイヤを指定する際にはあらかじめ決められたフレーズである、例えば、“レイヤ２を指定”などと発話するように求められており、ＣＰＵ１００は、音声区間の音声信号の音声がこのフレーズに合致するか否かを音声認識によって判断し、入力された音声がレイヤを指定するものであるか否かを判断する。そして、レイヤを指定する音声であると判断すると（ステップＳ２３：ＹＥＳ）、ステップＳ２５の音声の再生に処理を引き継ぐ。なお、ステップＳ２５の音声の再生の処理は、図１１のステップＳ５９以降の処理である。

【０１０３】

一方、レイヤを指定する音声でないと判断すると（ステップＳ２３：ＮＯ）、続いて、ステップＳ２７において、時計回路１０４から現在（音声情報の記録時（より正確には音声区間の開始時））の時刻の情報を取得し、これをフラッシュメモリ１０３に記録する。

【０１０４】

そして、ステップＳ２９では、最後に（前回）フラッシュメモリ１０３に記録された音声情報に対応つけられて記録されている時刻の情報を参照し、ステップＳ２７で時計回路１０４から取得した現在の時刻が、最後に音声情報を記録した時刻から、例えば、１時間が経過しているか否かを判断する。なお、フラッシュメモリ１０３にすでに音声情報と対応つけて記録された時刻の情報が存在しない場合は、つまりいまだに音声情報が記録されていない場合は、１時間が経過していないと判断する。

【０１０５】

１時間が経過していると判断すると（ステップＳ２９：ＹＥＳ）、ステップＳ３１で、

フラッシュメモリ１０３にすでに記録されている音声情報のレイヤを示すレイヤ識別情報を更新する。つまり、レイヤ識別情報が示すレイヤ数（レイヤ１、レイヤ２、レイヤ３など）を１つだけインクリメントする。なお、レイヤ数が大きいレイヤほど利用者からレイヤの距離は遠くなる。一方、１時間が経過していないと判断すると（ステップＳ２９：ＮＯ）、ステップＳ３１をスキップする。

【０１０６】

ステップＳ３３では、現在記録しようとしている音声情報のレイヤ識別情報、つまり初期値であるレイヤ１を示す情報を、ステップＳ２７で保存した時刻の情報に対応つけてフラッシュメモリ１０３に保存する。

【０１０７】

次に、ＣＰＵ１００はステップＳ３５で、音声信号から音声区間の音声信号を取得して音声情報を生成する。そして生成した音声情報を先に保存した時刻の情報およびレイヤ識別情報と対応つけてフラッシュメモリ１０３に保存（記録）する。この音声情報は、例えば、ＭＰ３（MPEG Audio Layer-3）ファイルなどである。

【０１０８】

ステップＳ３７では、利用者の音声が骨伝道マイク１２に入力された際（音声情報を記録する際）の利用者の仮想空間における顔の向きを示す情報（方向情報）を３軸地磁気センサ１４の出力に基づいて生成し、フラッシュメモリ１０３に先に保存された音声情報などと対応つけて保存する。ここで、仮想空間とは、図９に示したように、利用者の頭部を中心とする半球面の層の連なりである。なお、仮想空間の基準となる正面は、音声記録再生装置１０の電源が投入されて、利用者がＰＤＡ１１のタッチパネル１０５に対して所定の操作を行った時に、３軸地磁気センサ１４が検出した利用者の顔が向いている方向である。

＜第３の実施形態＞

第３の実施形態の音声記録再生装置１０ａでは、図１３に示すように、第１の実施形態および第２の実施形態の音声記録再生装置１０の構成に加え、魚眼レンズを有するカメラ１６を備えている。このカメラ１６は、音声の記録時の利用者の正面方向の周辺画像を撮影するものであり、利用者の胴体の前方に位置するように図示しない紐によって首から下げられる。一方、ＰＤＡ１１ａは、図１４に示すように、第１の実施形態および第２の実施形態のＰＤＡ１１の構成に加え、カメラ１６からの画像信号を入力するための画像信号入力Ｉ／Ｆ１１０を備えている。なお、音声記録再生装置１０ａおよびＰＤＡ１１ａについて、音声記録再生装置１０およびＰＤＡ１１と同じ構成については同じ符号を用いている。

【０１０９】

そして、利用者の仮想空間における顔の向きを示す情報（方向情報）は、仮想空間の基準（正面）線と利用者の顔の向きが示す直線とが成す、水平方向の角度α（図５参照）および垂直方向の角度β（図１２参照）である。なお、先述したように、音声の入力の開始から音声の入力の終了までの間に利用者の顔の向きにブレが生じるため、そのため、音声の入力時の利用者の顔の向きを示す方向情報は、音声の入力の開始から音声の入力の終了までの間（音声区間）の水平方向の顔の向き（角度α）の平均値および垂直方向の顔の向き（角度β）の平均値とする。

【０１１０】

ステップＳ３９では、アイコン音声情報を生成して、これを先に保存した音声情報、方向情報などと対応つけてフラッシュメモリ１０３に保存する。なお、このアイコン音声情報は、先述したように、あらかじめフラッシュメモリ１０３に記録されている鐘の音の楽音の音声情報の複製である。

【０１１１】

次に、ステップＳ４１では、アイコン音声情報に基づいて再生されるアイコン音声の変調パラメータを決定し、これを先に保存した音声情報などと対応つけてフラッシュメモリ１０３に保存する。ここで、変調パラメータとは、先述したように、利用者が再生される

アイコン音声を識別しやすいように、仮想空間におけるアイコン音声の位置に応じてアイコン音声の周波数を変化させるためのものである。

【０１１２】

一方、音声情報を再生する際には、ＰＤＡ１１のＣＰＵ１００は、図１１のステップ５１で、音声の入力があるか否かを判断する。この判断の方法は、図１０のステップ２１における判断と同じである。そして、音声の入力がないと判断すると（ステップＳ５１：ＮＯ）、ステップＳ５３で、再生する音声情報の属するレイヤ数を初期値に設定する。つまり、レイヤ１が指定されたとみなす。

【０１１３】

一方、音声の入力があると判断すると（ステップＳ５１：ＹＥＳ）、続いて、ステップＳ５５で、入力された音声はレイヤを指定する音声であるか否かを音声認識により判断する。この判断の方法は、図１０のステップＳ２３における判断と同じである。そして、レイヤを指定する音声でない、つまり記録すべき音声であると判断すると（ステップＳ５５：ＮＯ）、ステップＳ５７の音声の記録に処理を引き継ぐ。なお、音声の記録の処理は、図１０のフロー図のステップＳ２７以降の処理である。

【０１１４】

一方、レイヤを指定する音声であると判断すると（ステップＳ５５：ＹＥＳ）、ステップＳ５９で、入力された音声の音声認識の結果に基づいてレイヤ数を特定し、ステップＳ６１で特定されたレイヤ数を再生するべき音声情報が属するレイヤ数として設定する。つまり、音声認識されたレイヤ数が指定されたとみなす。

【０１１５】

次に、ステップＳ６３で、３軸地磁気センサ１４の出力に基づいて、仮想空間において利用者の顔が向いている方向を特定する。つまり、利用者の顔が仮想空間の基準から水平方向にどれだけの角度（α）の方向を向いており、垂直方向にどれだけの角度（β）の方向を向いているかを特定する。

【０１１６】

ステップＳ６５では、フラッシュメモリ１０３に記憶されているすべての音声情報について、仮想空間における音声情報が位置する方向の、ステップＳ６３で特定した利用者の顔の向きの方向に対する相対的方向を算出する。この相対的方向は、フラッシュメモリ１０３に音声情報と対応つけて記録されている方向情報（水平方向角度および垂直方向角度）とステップＳ６３で特定された利用者の顔の向きの角度の情報に基づいて算出される。

【０１１７】

そして、ステップＳ６７では、フラッシュメモリ１０３に記録されている音声情報のうち、ステップＳ５３またはステップＳ６１で設定（指定）されたレイヤに属し、ステップＳ６５で算出された相対的方向が水平方向および垂直方向ともに±１０°の範囲以内である音声情報を再生してオーディオ出力Ｉ／Ｆ１０８を介してステレオヘッドフォン１３に出力する。このとき、音声情報は、仮想空間における音声情報の位置を利用者が特定することが可能なように、つまり音源定位が可能なように再生される。具体的には、利用者の左右の耳のそれぞれに対する音声の到達の時間差と音声の音量差とに基づいて音源定位が可能なように音声情報を再生する。

【０１１８】

音声の到達の時間差と音声の音量差との算出方法は、第１実施形態で説明したものと同じ方法で実現できるが、音声情報を音源定位が可能なように再生する処理は、大内誠，岩谷幸雄，鈴木陽一，棟方哲弥による“三次元音響ＶＲエデュテイメントシステムによる視覚障害者の空間認識能力訓練効果”，第３回科学技術フォーラムＦＩＴ２００４情報技術レターズ，ｐｐ．２８３－２８４，２００４や、岩永信之，坂本憲成，小林亙，尾上孝雄，白川功による“組込みシステム向けヘッドフォンステレオ頭外音場拡大手法とその実装”，第１７回ディジタル信号処理シンポジウム，Ｂ２－４，２００２で提案されている手法に基づく、音波の音量・伝達時間・周波数特性の差分や変化などにより実現される。

【０１１９】

次に、ステップＳ６９では、フラッシュメモリ１０３に記録されている音声情報のうち、指定されたレイヤ以外に属する音声情報、および指定されたレイヤに属すがステップＳ６５で算出された相対的方向が水平方向および垂直方向ともに±１０°の範囲外である音声情報についてアイコン音声情報を変調パラメータに基づいて変調して再生してオーディオ出力Ｉ／Ｆ１０８を介してステレオヘッドフォン１３に出力する。このアイコン音声情報も上述したように、上述の手法に基づいて音源定位が可能なように再生される。なお、ステップＳ６７における音声情報の再生とステップＳ６９におけるアイコン音声情報の再生とは同時に並行して実施される。また、音声情報とアイコン音声情報とは、５００ミリ秒の無音区間を挟みながら繰り返し再生される。

＜第３の実施形態＞

第３の実施形態の音声再生装置１０ａでは、図１３に示すように、第１の実施形態および第２の実施形態の音声再生装置１０の構成に加え、魚眼レンズを有するカメラ１６を備えている。このカメラ１６は、音声の記録時の利用者の正面方向の周辺画像を撮影するものであり、利用者の胴体の前方に位置するように図示しない紐によって首から下げられる。一方、ＰＤＡ１１ａは、図１４に示すように、第１の実施形態および第２の実施形態のＰＤＡ１１の構成に加え、カメラ１６からの画像信号を入力するための画像信号入力Ｉ／Ｆ１１０を備えている。なお、音声再生装置１０ａおよびＰＤＡ１１ａについて、音声再生装置１０およびＰＤＡ１１と同じ構成については同じ符号を用いている。

【０１２０】

第３の実施形態の音声記録再生装置１０ａでは、音声情報をカメラ１６で撮影した利用者の周辺の画像（周辺画像）に付加して記録する。より具体的には、音声情報を記録する際の利用者の顔の向きを検出し、この顔の向きを周辺画像における位置に変換し、音声情報をこの周辺画像における位置の情報（位置情報）や周辺画像の情報（周辺画像情報）とともに記録する。なお、新規に音声情報を記録する際には、最後に（前回）記録した音声情報が付加された周辺画像と新規に音声情報を記録する際にカメラ１６で撮影した周辺画像を比較し、利用者の環境が変化したと判断する場合には新たにカメラ１６で撮影した周辺画像に新規の音声情報を付加し、環境が変化したと判断しない場合には最後に（前回）記録した音声情報が付加された周辺画像に新規の音声情報を付加する。つまり、第３の実施形態では、カメラ１６で撮影した周辺画像がレイヤの役割を果たす。

【０１２１】

この周辺画像の情報（周辺画像情報）は、図６（ｃ）に示すように、当該周辺画像情報を識別する画像識別情報と対応つけてフラッシュメモリ１０３に記憶され、また音声情報も画像識別情報と対応つけて記憶される。このことにより、音声情報が付加される周辺画像が特定される。

【０１２２】

以下に、図１５に示すフロー図を参照して、音声記録再生装置１０ａで音声を記録する際にＰＤＡ１１aのＣＰＵ１００が実行する処理について説明する。なお、図１５に示す処理を実行して利用者が発した音声をＰＤＡ１１ａに記録すると、フラッシュメモリ１０３には、１つの音声情報について、図６（ｃ）に示すように、画像識別情報、音声情報、位置情報、時刻情報が記録される。また、これらとは別個にフラッシュメモリ１０３には、画像識別情報、周辺画像情報が記録される。さらに、図１５のフロー図に示す処理は一例であり、各処理ステップの順序はこれを変更しても発明を実施できるものについては順序を変更してもかまわない。

【０１２３】

まず、ＰＤＡ１１ａのＣＰＵ１００は、図１５のステップＳ８１で、音声の入力があるか否かを判断する。この音声の入力があるか否かの判断は、オーディオ入力Ｉ／Ｆ１０７を介して骨伝導マイク１２から入力される音声信号に基づいて判断される。この判断では、オーディオ入力Ｉ／Ｆ１０７に入力される音声信号から音声の信号（有音）である区間（音声区間）の切り出しを試み、音声区間が切り出せた場合に音声の入力があると判断する。なお、音声区間の切り出しの方法は、第１の実施形態などと同じである。

【０１２４】
音声信号から音声区間が切り出されて、音声の入力があると判断すると（ステップＳ８１：ＹＥＳ）、次に、ＣＰＵ１００はステップＳ８３で、音声信号から音声区間の音声信号を取得して音声情報を生成する。そして生成した音声情報をフラッシュメモリ１０３に保存（記録）する。この音声情報は、例えば、ＭＰ３（MPEG Audio Layer-3）ファイルなどである。
【０１２５】
次に、ステップＳ８５で、３軸地磁気センサ１４の出力に基づいて、仮想空間において利用者の顔が向いている方向を特定し、方向情報を生成する。つまり、利用者の顔が仮想空間の基準から水平方向にどれだけの角度（α）の方向を向いており、垂直方向にどれだけの角度（β）の方向を向いているかを特定する。ここで、仮想空間とは魚眼レンズを有するカメラ１６によって撮影される１８０°の画角を持った周辺画像で示される空間である。なお、この方向情報（水平方向の角度αおよび垂直方向の角度β）は、ＲＡＭ１０２に記録される。
【０１２６】
ステップＳ８７では、画像信号入力Ｉ／Ｆ１１０を介してカメラ１６から画像信号を取得して、利用者の体が向いている方向の周辺画像情報を生成しこれを一旦ＲＡＭ１０２に記録する。この周辺画像情報は、例えば、ＪＰＥＧファイルなどである。
【０１２７】
次に、ステップＳ８９では、時計回路１０４からから現在（音声情報の記録時（より正確には音声区間の開始時））の時刻の情報を取得し、これをステップＳ８３で保存した音声情報に対応つけてフラッシュメモリ１０３に記録する。
【０１２８】
そして、ステップＳ９１では、ステップＳ８７で生成してＲＡＭ１０２に一旦記録した周辺画像情報と、フラッシュメモリ１０３に最後に（前回）記録された周辺画像情報とを比較し、利用者の周辺の環境に変化があるか否かを判断する。最後（前回）の周辺画像情報は、画像識別情報が対応つけられた音声情報のうち、対応つけられた時刻情報が示す時刻が最新である音声情報を特定すれば、この音声情報に対応つけられた画像識別情報に基づいて決定できる。（ステップＳ８３で保存した音声情報は画像識別情報がいまだ対応つけられていないため除外される。）
なお、周辺画像情報をフラッシュメモリ１０３に記録する際に時計回路１０４から取得した現在の時刻の情報を周辺画像情報とともに記録し、この時刻の情報に基づいて最後（前回）の周辺画像情報を特定するようにしてもよい。
【０１２９】
なお、周辺画像情報に基づく環境の変化の有無の判断は、例えば、２つの周辺画像情報における輝度やカラー情報の単位時間当たりの変化を集計し、変化量が所定の閾値以上である場合に、環境が変化していると判断することによって実現することができる。また、特開平６－２５１１４７号公報（映像特徴処理方法）に開示されている動画像におけるシーン変化検出方法も利用することができる。
【０１３０】
そして、利用者の周辺の環境に変化があると判断すると（ステップＳ９１：ＹＥＳ）、ステップＳ９３で、ＲＡＭ１０２に一旦記録されている（新規の）周辺画像情報をフラッシュメモリ１０３に記録する。このとき、周辺画像情報には当該周辺画像情報を識別するための画像識別情報が割り当てられ、周辺画像情報はこの画像識別情報と対応つけてフラッシュメモリ１０３に記録される。
【０１３１】
次に、ステップＳ９５では、ステップＳ９３でフラッシュメモリ１０３に記録した周辺画像情報の画像識別情報をステップＳ８３で保存した音声情報と対応つけてフラッシュメモリ１０３に記録する。このことによって、新規に記録した周辺画像情報と音声情報とが対応つけられる。

【０１３２】
一方、利用者の周辺の環境に変化がないと判断すると（ステップＳ９１：ＮＯ）、最後に（前回）フラッシュメモリ１０３に記録された（既存の）周辺画像情報（ステップＳ９１の判断に利用された一方の周辺画像情報）の画像識別情報をステップＳ８３で保存した音声情報と対応つけてフラッシュメモリ１０３に記録することによって、最後に（前記）記録した周辺画像情報と音声情報とを対応つける。
【０１３３】
そして、ステップＳ９９では、ステップＳ８５で生成されてＲＡＭ１０２に記録された方向情報である水平方向の角度αおよび垂直方向の角度βを、ステップＳ９５またはステップＳ９７で音声情報が対応つけられた周辺画像（情報）における位置を示す座標（ｘ，ｙ）に変換し、これらの位置情報（座標（ｘ，ｙ））を先に保存された音声情報と対応つけてフラッシュメモリ１０３に記録する。なお、顔の向きを示す角度αおよび角度βから画像上の座標（ｘ，ｙ）への変換は、Hirotake Yamazoe, Akira Utsumi, Kenichi Hosaka, Masahiko Yachida, “A body-mounted camera system for head-pose estimation and user-view image synthesis,” Available online at www.sciencedirect.com ScienceDirect Image and Vision Computing 25 (2007) 1848-1855に開示されている手法を用いて実現することができる。
【０１３４】
次に、図１６に示すフロー図を参照して、音声記録再生装置１０ａで音声情報を視覚的に閲覧して音声の再生などをする際にＰＤＡ１１aのＣＰＵ１００が実行する処理について説明する。なお、図１６のフロー図に示す処理は一例であり、各処理ステップの順序はこれを変更しても発明を実施できるものについては順序を変更してもかまわない。
【０１３５】
利用者がＰＤＡ１１ａのタッチパネル１０５に対して、音声情報を閲覧するための操作を行うと、ＣＰＵ１００は、ステップＳ１０１で、フラッシュメモリ１０３に記録されているすべての周辺画像（情報）（レイヤ）をディスプレイ１０６に３次元で表示する。このとき、図１７に示すように、フラッシュメモリ１０３に記録された時刻が古いものを下（奥あるいは遠く）に表示し、時刻が新しい物をこの上に重ねて手前（近く）に表示する。また、各周辺画像の右上には、その周辺画像を選択するための付箋が表示される。
【０１３６】
なお、周辺画像情報がフラッシュメモリ１０３に記録された時刻の新旧は、周辺画像情報に対応つけられている音声情報の時刻情報に基づいて特定できるが、先述したように、周辺画像情報自体に記録時の時刻を示す情報を対応つけて記録するようにしてもよい。
【０１３７】
そして、ステップＳ１０３では、ディスプレイ１０６に表示した周辺画像の上に当該周辺画像の周辺画像情報の画像識別情報と対応つけられた音声情報を示すアイコン２００を表示する。このとき、アイコン２００は対応する音声情報と対応つけてフラッシュメモリ１０３に記録された位置情報（ｘ座標、ｙ座標）によって特定される周辺画像上の位置に表示される。
【０１３８】
次に、利用者がタッチパネル１０５に対して周辺画像（レイヤ）を選択する操作、つまりディスプレイ１０６に表示されている付箋のいずれかを選択する操作（付箋をクリック（タップ）する操作）を行うと、ＣＰＵ１００は、ステップＳ１０５で、レイヤが選択されたと判断する。
【０１３９】
そして、ステップＳ１０７で、選択された周辺画像（レイヤ）のみをディスプレイ１０６に３次元で表示し、ステップＳ１０９で選択された周辺画像に付加された音声情報のアイコン２００、つまり選択された周辺画像の画像識別情報と対応つけて記録された音声情報のアイコン２００をディスプレイ１０６に表示された周辺画像の上に表示する。この際もアイコン２００は対応する音声情報と対応つけてフラッシュメモリ１０３に記録された

位置情報に基づいた位置に表示される。

【０１４０】

このように、ディスプレイ１０６に周辺画像とその上にアイコン２００が表示された状態で、利用者がタッチパネル１０５に対してアイコン２００をクリック（タップ）する操作を行うと、ＣＰＵ１００は、ステップＳ１１１でアイコン２００がクリックされたと判断し（ステップＳ１１１：ＹＥＳ）、ステップＳ１１３で、クリックされたアイコン２００に対応する音声情報を再生する。

【０１４１】

また、利用者がタッチパネル１０５に対してアイコン２００をドラッグしてその位置を変更する操作を行うと、ＣＰＵ１００は、ステップＳ１１５でアイコン２００のドラッグ操作が行われたと判断し（ステップＳ１１５：ＹＥＳ）、ステップＳ１１７で、利用者のドラッグ操作に応じてアイコン２００の周辺画像上の位置を変更する。また、このとき、変更されたアイコン２００の位置に応じて、フラッシュメモリ１０３に記録されている位置情報（ｘ座標、ｙ座標）を変更する。

【０１４２】

さらに、利用者がタッチパネル１０５に対して周辺画像をドラッグしてその向きを変える操作を行うと、ＣＰＵ１００は、ステップＳ１１９で周辺画像のドラッグ操作が行われたと判断し（ステップＳ１１９：ＹＥＳ）、利用者のドラッグ操作に応じてディスプレイ１０６に表示される周辺画像の向きを変更する。

【０１４３】

なお、利用者がタッチパネル１０５に対して音声情報の閲覧を終了する操作を行うと、ＣＰＵ１００は、ステップＳ１２３で、閲覧の終了操作が行われたと判断して処理を終了する。

＜第４の実施形態＞

第３の実施形態の音声記録再生装置１０ａでは、利用者の周辺の環境が変化した場合に、音声情報を付加する周辺画像を更新したが、第４の実施形態の音声記録再生装置１０ａでは、所定の時間が経過した場合にも音声情報を付加する周辺画像を更新する。なお、音声記録再生装置１０ａの構成は、第３の実施形態のものと同じである。

【０１４４】

以下に、図１８に示すフロー図を参照して、音声記録再生装置１０ａで音声を記録する際にＰＤＡ１１aのＣＰＵ１００が実行する処理について説明する。なお、図１８のフロー図に示す処理を実行して利用者が発した音声をＰＤＡ１１ａに記録すると、フラッシュメモリ１０３には、１つの音声情報について、図６（ｄ）に示すように、画像識別情報、音声情報、位置情報、時刻情報が記録される。また、これらとは別個にフラッシュメモリ１０３には、画像識別情報、周辺画像情報が記録される。さらに、図１８のフロー図に示す処理は一例であり、各処理ステップの順序はこれを変更しても発明を実施できるものについては順序を変更してもかまわない。

【０１４５】

まず、ＰＤＡ１１ａのＣＰＵ１００は、図１８のステップＳ１３１で、音声の入力があるか否かを判断する。この音声の入力があるか否かの判断は、オーディオ入力Ｉ／Ｆ１０７を介して骨伝導マイク１２から入力される音声信号に基づいて判断される。この判断では、オーディオ入力Ｉ／Ｆ１０７に入力される音声信号から音声の信号（有音）である区間（音声区間）の切り出しを試み、音声区間が切り出せた場合に音声の入力があると判断する。なお、音声区間の切り出しの方法は、第１の実施形態などと同じである。

【０１４６】

音声信号から音声区間が切り出されて、音声の入力があると判断すると（ステップＳ１３１：ＹＥＳ）、次に、ＣＰＵ１００はステップＳ１３３で、音声信号から音声区間の音声信号を取得して音声情報を生成する。そして生成した音声情報をフラッシュメモリ１０３に保存（記録）する。この音声情報は、例えば、ＭＰ３（MPEG Audio Layer-3）ファイルなどである。

【０１４７】
次に、ステップＳ１３５で、３軸地磁気センサ１４の出力に基づいて、仮想空間において利用者の顔が向いている方向を特定し、方向情報を生成する。つまり、利用者の顔が仮想空間の基準から水平方向にどれだけの角度（α）の方向を向いており、垂直方向にどれだけの角度（β）の方向を向いているかを特定する。ここで、仮想空間とは魚眼レンズを有するカメラ１６によって撮影される１８０°の画角を持った周辺画像で示される空間である。なお、この方向情報（水平方向の角度αおよび垂直方向の角度β）は、ＲＡＭ１０２に記録される。
【０１４８】
ステップＳ１３７では、画像信号入力Ｉ／Ｆ１１０を介してカメラ１６から画像信号を取得して、利用者の体が向いている方向の周辺画像情報を生成しこれを一旦ＲＡＭ１０２に記録する。この周辺画像情報は、例えば、ＪＰＥＧファイルなどである。
【０１４９】
次に、ステップＳ１３９では、時計回路１０４からから現在（音声情報の記録時（より正確には音声区間の開始時））の時刻の情報を取得し、これをステップＳ１３３で保存した音声情報に対応つけてフラッシュメモリ１０３に記録する。
【０１５０】
そして、ステップＳ１４１では、ステップＳ１３９でフラッシュメモリ１０３に記録した時刻情報と、フラッシュメモリ１０３に画像識別情報と対応つけられた音声情報に対応つけて最後に記録された時刻情報（つまり、前回音声情報を記録したときの時刻を示す時刻情報）とを比較し、現在の時刻（ステップＳ１３９でフラッシュメモリ１０３に記録した時刻情報が示す時刻（正確には、音声区間の始まりの時刻））が前回音声情報をフラッシュメモリ１０３に記録してから、例えば、３０分以上経過しているか否かを判断する。
【０１５１】
３０分以上経過していないと判断すると（ステップＳ１４１：ＮＯ）、最後に（前回）フラッシュメモリ１０３に記録された周辺画像情報の画像識別情報をステップＳ１３３で保存した音声情報と対応つけてフラッシュメモリ１０３に記録する。このことによって、最後に（前回）記録した（既存の）周辺画像情報と音声情報とが対応つけられる。
【０１５２】
一方、３０分以上経過していると判断すると（ステップＳ１４１：ＹＥＳ）、さらにステップＳ１４５で、前回音声情報をフラッシュメモリ１０３に記録してから、例えば、６０分以上経過しているか否かを判断する。
【０１５３】
そして、６０分以上経過していると判断すると（ステップＳ１４５：ＹＥＳ）、ステップＳ１４７で、ＲＡＭ１０２に一旦記録されている周辺画像情報をフラッシュメモリ１０３に記録する。このとき、周辺画像情報には当該周辺画像情報を識別するための画像識別情報が割り当てられ、周辺画像情報はこの画像識別情報と対応つけてフラッシュメモリ１０３に記録される。
【０１５４】
次に、ステップＳ１４９では、ステップＳ１４７でフラッシュメモリ１０３に記録した周辺画像情報の画像識別情報をステップＳ１３３で保存した音声情報と対応つけてフラッシュメモリ１０３に記録することによって、新規に記録した周辺画像情報と音声情報とを対応つける。
【０１５５】
一方、６０分以上経過していないと判断すると（ステップＳ１４５：ＮＯ）、ステップＳ１５１で、ステップＳ１３７で生成してＲＡＭ１０２に一旦記録した周辺画像情報と、フラッシュメモリ１０３に最後に（前回）記録された周辺画像情報とを比較し、利用者の周辺の環境に変化があるか否かを判断する。
【０１５６】
そして、利用者の周辺の環境に変化があると判断すると（ステップＳ１５１：ＹＥＳ）

、ステップＳ１４７およびステップＳ１４９で先述の処理を実行する。
【０１５７】
一方、利用者の周辺の環境に変化がないと判断すると（ステップＳ１５１：ＮＯ）、最後に（前記）フラッシュメモリ１０３に記録された周辺画像情報（ステップＳ１５１の判断に利用された一方の周辺画像情報）の画像識別情報をステップＳ１３３で保存した音声情報と対応つけてフラッシュメモリ１０３に記録することによって、最後に（前回）記録した（既存の）周辺画像情報と音声情報とを対応つける。
【０１５８】
そして、ステップＳ１５５では、ステップＳ１３５で生成されてＲＡＭ１０２に記録された方向情報である水平方向の角度αおよび垂直方向の角度βを、ステップＳ１４３、ステップＳ１４９、或いはステップＳ１５３で音声情報が対応つけられた周辺画像（情報）における座標（ｘ，ｙ）に変換し、これらの位置情報（座標（ｘ，ｙ））を先に保存された音声情報と対応つけてフラッシュメモリ１０３に記録する。
【０１５９】
なお、音声記録再生装置１０ａで音声情報を視覚的に閲覧して音声の再生などをする際にＰＤＡ１１ａのＣＰＵ１００が実行する処理は、第３の実施形態において図１６のフロー図を用いて説明した処理と同じであるので、重複する説明を省略する。
＜その他の実施形態＞
その他の実施形態として、第２の実施形態におけるレイヤ識別情報（レイヤ数）によって規定され時間とともに仮想空間における利用者からの距離が遠くなるレイヤ（“時間レイヤ”と呼ぶ）と、第３および第４の実施形態における周辺画像によって規定されるレイヤ（“画像レイヤ”と呼ぶ）とを混在させる形態が考えられる。この場合、周辺画像情報に撮影時の時刻の情報（時刻情報）やレイヤ識別情報（レイヤ数）を対応つけて記憶するとよい。このようにすれば、時間レイヤと画像レイヤとからなるレイヤの集合に、時刻情報が示す時刻に応じて通し番号となったレイヤ数を割り当てて一連のレイヤとして取り扱うことができる。
【０１６０】
また、この場合は、周辺画像情報に対応つける音声情報には、音声の記録時の利用者の顔の向きの情報（方向情報）を対応つけて記録し、再生時にこの方向情報にもとづいて音源定位可能に再生するようにするとよい。さらに、第３および第４の実施形態のように音声情報を視覚的に閲覧する場合には、時間レイヤに対応させて、あらかじめ用意した擬似的な周辺画像をディスプレイ１０６に表示しその上に音声情報に対応するアイコン２００を表示するようにするとよい。
【０１６１】
なお、上述した実施の形態は、本発明の技術思想の範囲内で種々に変更してかまわない。例えば、上述の実施形態では、利用者の顔の向きを３軸地磁気センサのみで検出しているが、３軸地磁気センサと３軸加速度センサとの２つで検出するようにしてもよい。また、利用者の顔を撮影した画像を処理することによって顔の向きを検出するようにしてもよい。
【０１６２】
さらに、利用者の顔の向きによって仮想空間における方向を指定したが、これに替えて、利用者の視線の向きや指差しの向きによって方向を指定するようにしてもよい。なお、利用者の視線の向きの特定には、H. Yamazoe, A. Utsumi, T. Yonezawa, S. Abe. “Remote Gaze Direction Estimation with a Single Camera Based on Facial-Feature Tracking,” ETRA2008, pp.245-250, 2008で提案されている手法を用いることができる。
【０１６３】
また、利用者による音声の入力には、骨伝導マイク１２を用いることとしたが、これに替えて、ＮＡＭ（非可聴つぶやき）マイクを用いるようにしてもよい。ＮＡＭマイクとは、声帯振動を伴わない無声呼気音が発話器官の運動による音響的フィルタ特性により調音されて人体頭部の主に軟部組織を伝導したものを収集することができるマイクロフォンで

ある。このＮＡＭマイクを利用すれば、電車内などの人目が気になる場所においても、気兼ねなく音声記録再生装置１０、１０ａに音声を入力することができる。

【０１６４】

なお、人目が気になる場所においても気兼ねなく音声を入力するためには、喉元に装着して声帯の振動を音声にして伝える方式であるスロートマイクを採用してもよい。

【０１６５】

また、上述の実施形態では、レイヤを円弧または半球面とし、時間の経過とともに、時間の経過とともに音声情報は利用者から距離が遠いレイヤに属することとしたが、レイヤは必ずしも利用者を中心とする円弧や半球面でなくてもよく、また、レイヤごとに利用者からの距離が同心円を中心として遠くなる構成でなくてもよい。つまり、レイヤは平面であり、利用者の周りに等距離で位置する構成としてもよく、さらには、レイヤは空間的広がりを有さず、単に属する音声情報をまとめる概念的な区別単位であってもよい。

【図面の簡単な説明】

【０１６６】

【図１】本発明の音声記録再生装置の利用状況の一例を示す図解図である。

【図２】本発明の音声記録再生装置を構成する携帯情報端末の構成の一例を示すブロック図である。

【図３】携帯情報端末のＣＰＵが音声の記録時に実行する処理の一例を示すフロー図である。

【図４】携帯情報端末のＣＰＵが音声の再生時に実行する処理の一例を示すフロー図である。

【図５】仮想空間における利用者の顔が向いている方向を説明する図解図である。

【図６】携帯情報端末のフラッシュメモリに記録されるデータの構成の一例を示す図解図である。

【図７】音声情報の音源定位について説明する図解図である。

【図８】音声情報の再生について説明する図解図である。

【図９】音声情報の再生について説明する図解図である。

【図１０】携帯情報端末のＣＰＵが音声の記録時に実行する処理の一例を示すフロー図である。

【図１１】携帯情報端末のＣＰＵが音声の再生時に実行する処理の一例を示すフロー図である。

【図１２】仮想空間における利用者の顔が向いている方向を説明する図解図である。

【図１３】本発明の音声記録再生装置の利用状況の一例を示す図解図である。

【図１４】本発明の音声記録再生装置を構成する携帯情報端末の構成の一例を示すブロック図である。

【図１５】携帯情報端末のＣＰＵが音声の記録時に実行する処理の一例を示すフロー図である。

【図１６】携帯情報端末のＣＰＵが音声情報の閲覧・再生時に実行する処理の一例を示すフロー図である。

【図１７】携帯情報端末のディスプレイにおける表示の一例を示す図解図である。

【図１８】携帯情報端末のＣＰＵが音声の記録時に実行する処理の一例を示すフロー図である。

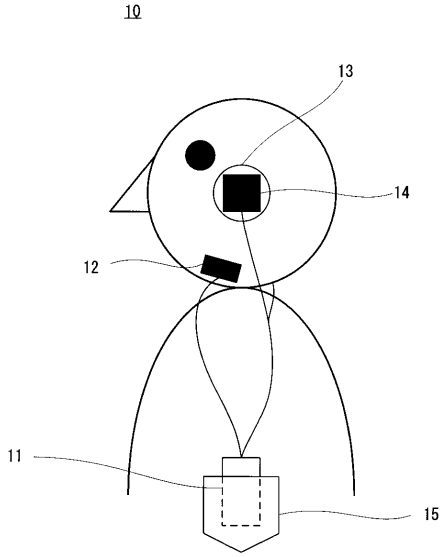
【符号の説明】

【０１６７】

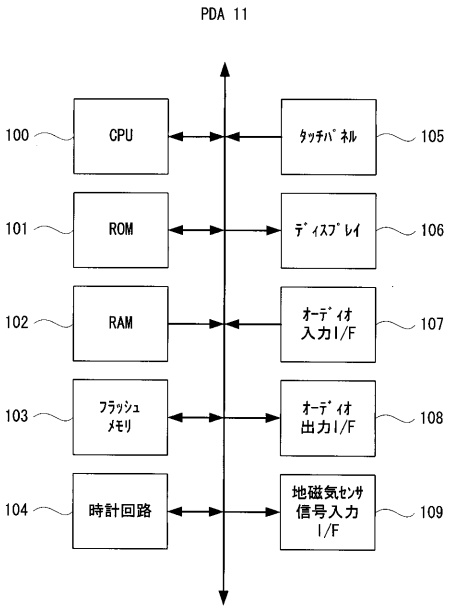
１０　…音声記録再生装置
１１　…ＰＤＡ
１２　…骨伝導マイク
１３　…ステレオヘッドフォン
１４　…３軸地磁気センサ
１６　…カメラ

１００　…ＣＰＵ
１０３　…フラッシュメモリ
１０４　…時計回路

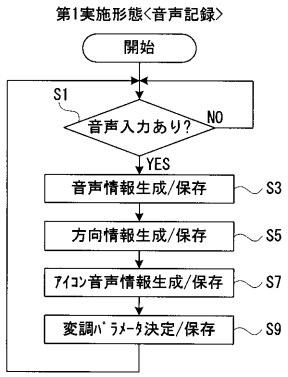
【図１】



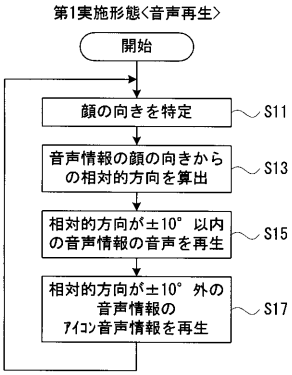
【図２】



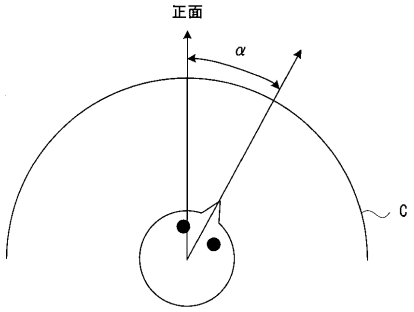
【図３】


第1実施形態<音声記録>
開始
S1
音声入力あり?
NO
YES
音声情報生成/保存
S3
方向情報生成/保存
S5
ｱｲｺﾝ音声情報生成/保存
S7
変調ﾊﾟﾗﾒｰﾀ決定/保存
S9

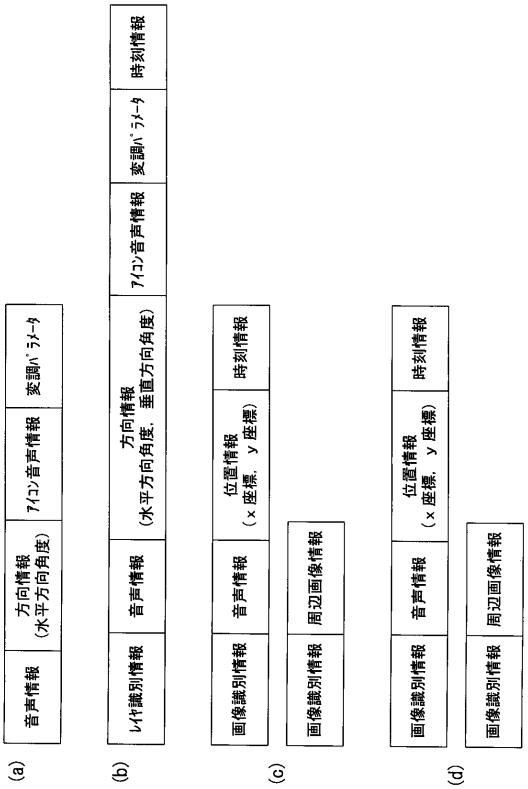
【図４】


第1実施形態<音声再生>
開始
顔の向きを特定
S11
音声情報の顔の向きからの相対的方向を算出
S13
相対的方向が±10°以内の音声情報の音声を再生
S15
相対的方向が±10°外の音声情報のｱｲｺﾝ音声情報を再生
S17

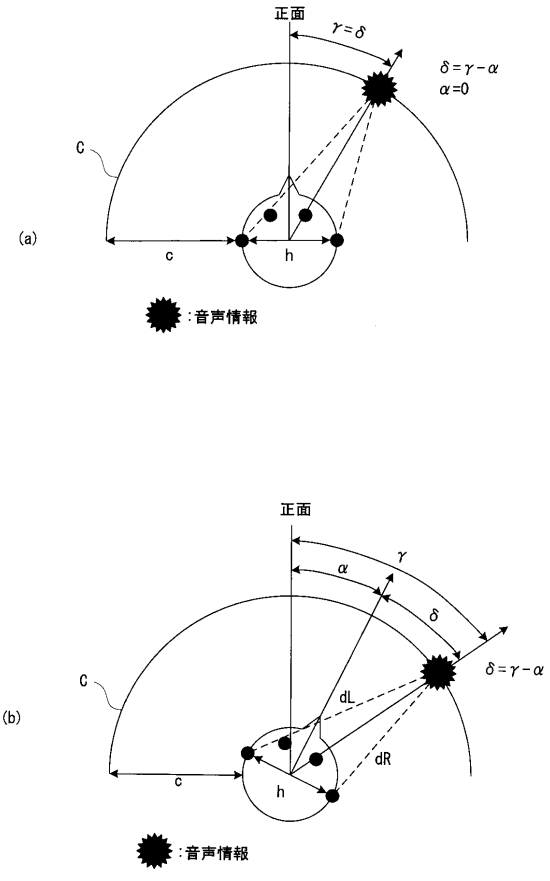
【図５】


正面
α
C

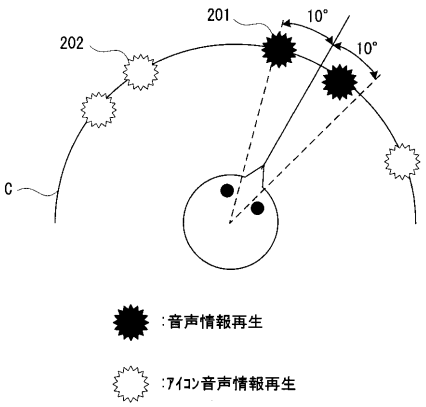
【図６】


(a)
音声情報
方向情報
(水平方向角度)
ｱｲｺﾝ音声情報
変調ﾊﾟﾗﾒｰﾀ
(b)
ﾚｲﾔ識別情報
音声情報
方向情報
(水平方向角度，垂直方向角度)
ｱｲｺﾝ音声情報
変調ﾊﾟﾗﾒｰﾀ
時刻情報
(c)
画像識別情報
音声情報
位置情報
(x座標，y座標)
時刻情報
周辺画像情報
(d)
画像識別情報
音声情報
位置情報
(x座標，y座標)
時刻情報
周辺画像情報

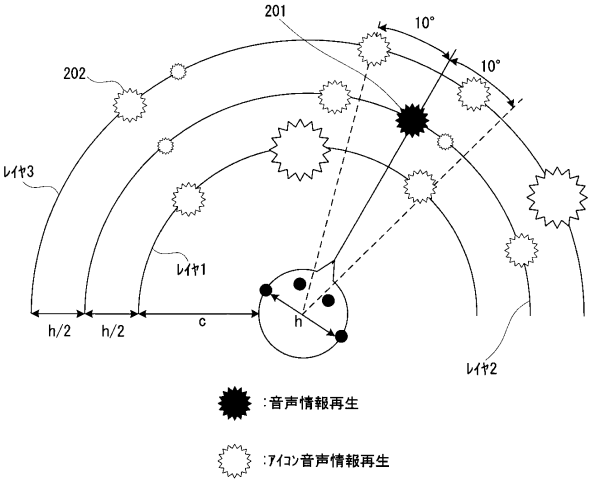
【図７】


正面
γ=δ
δ=γ-α
α=0
C
(a)
c
h
:音声情報
正面
γ
α
δ
δ=γ-α
C
(b)
dL
dR
c
h
:音声情報

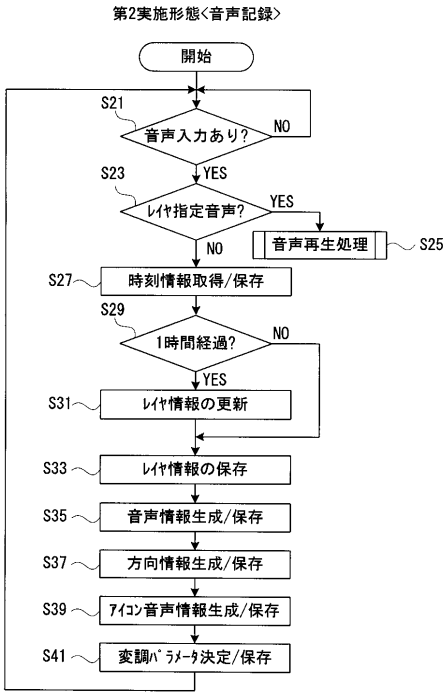
【図８】


201
10°
10°
202
C
:音声情報再生
:ｱｲｺﾝ音声情報再生

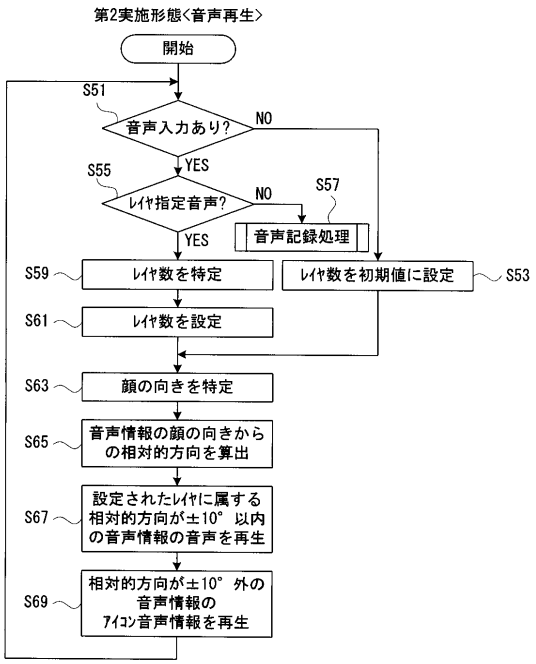
【図９】


201
10°
10°
202
ﾚｲﾔ3
ﾚｲﾔ1
h/2
h/2
c
h
ﾚｲﾔ2
:音声情報再生
:ｱｲｺﾝ音声情報再生

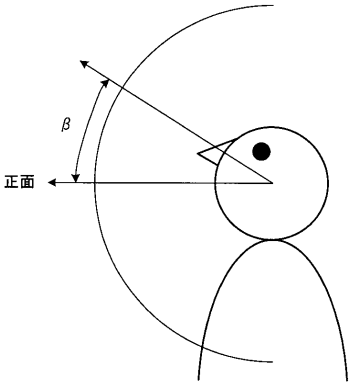
【図１０】


第2実施形態<音声記録>
開始
S21
音声入力あり?
NO
YES
S23
ﾚｲﾔ指定音声?
YES
NO
音声再生処理
S25
S27
時刻情報取得/保存
S29
1時間経過?
NO
YES
S31
ﾚｲﾔ情報の更新
S33
ﾚｲﾔ情報の保存
S35
音声情報生成/保存
S37
方向情報生成/保存
S39
ｱｲｺﾝ音声情報生成/保存
S41
変調ﾊﾟﾗﾒｰﾀ決定/保存

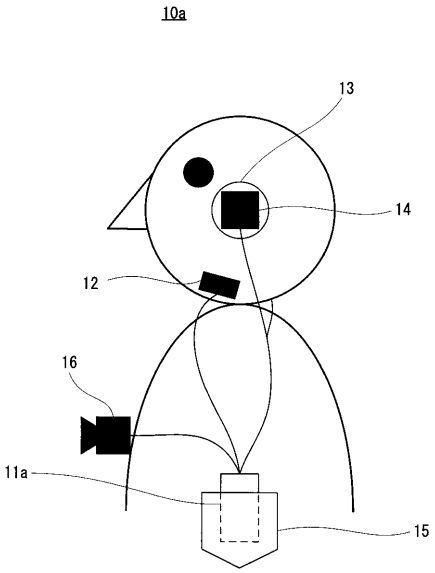
【図１１】


第2実施形態<音声再生>
開始
S51
音声入力あり?
NO
YES
S55
ﾚｲﾔ指定音声?
NO
S57
YES
音声記録処理
S59
ﾚｲﾔ数を特定
ﾚｲﾔ数を初期値に設定
S53
S61
ﾚｲﾔ数を設定
S63
顔の向きを特定
S65
音声情報の顔の向きから
の相対的方向を算出
S67
設定されたﾚｲﾔに属する
相対的方向が±10°以内
の音声情報の音声を再生
S69
相対的方向が±10°外の
音声情報の
ｱｲｺﾝ音声情報を再生

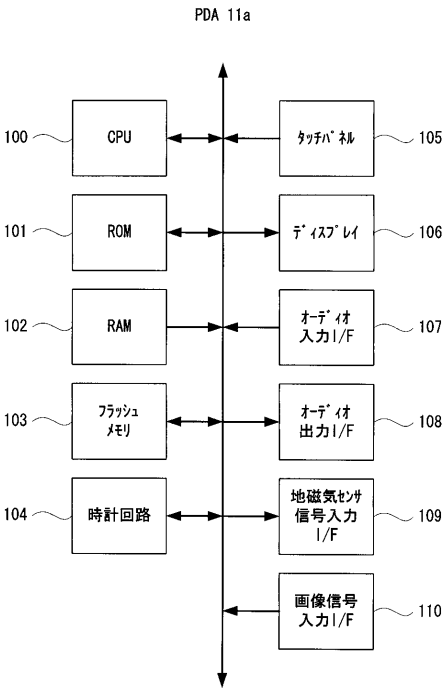
【図１２】



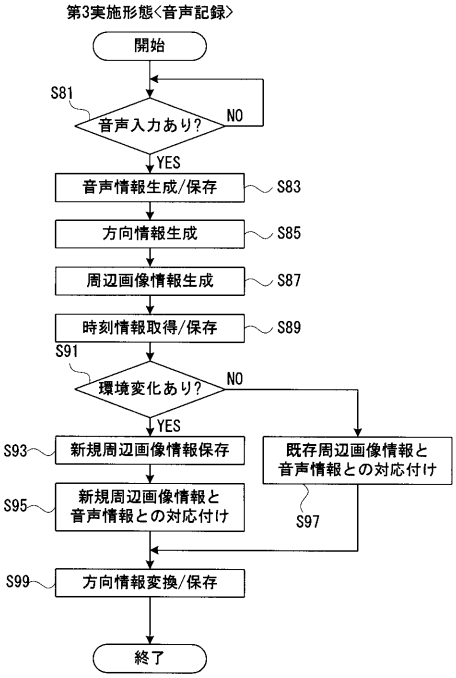
【図１３】



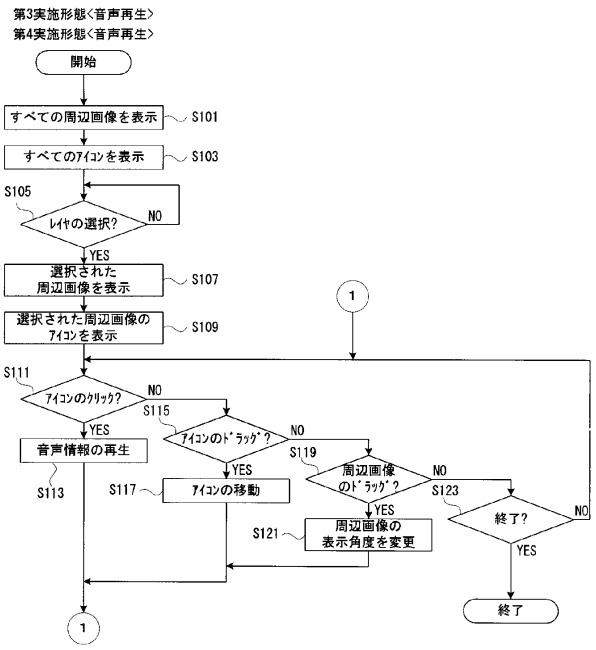
【図１４】



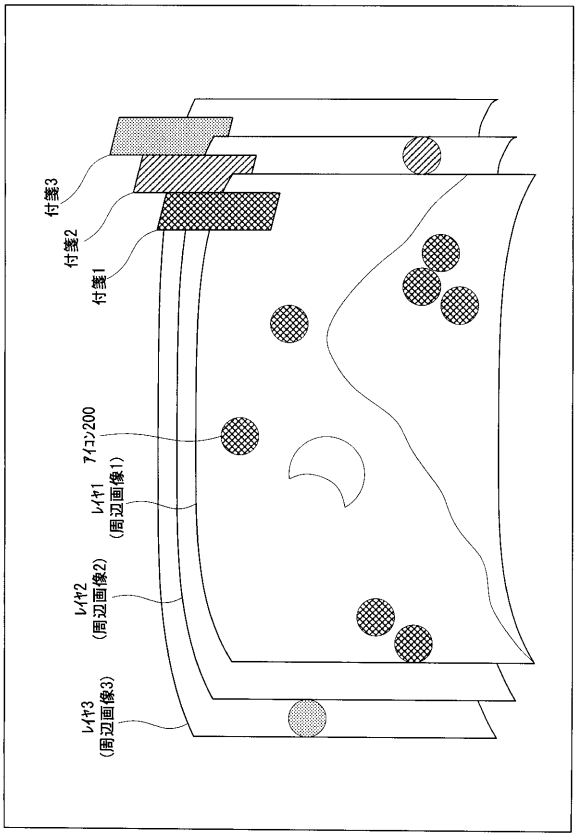
【図１５】



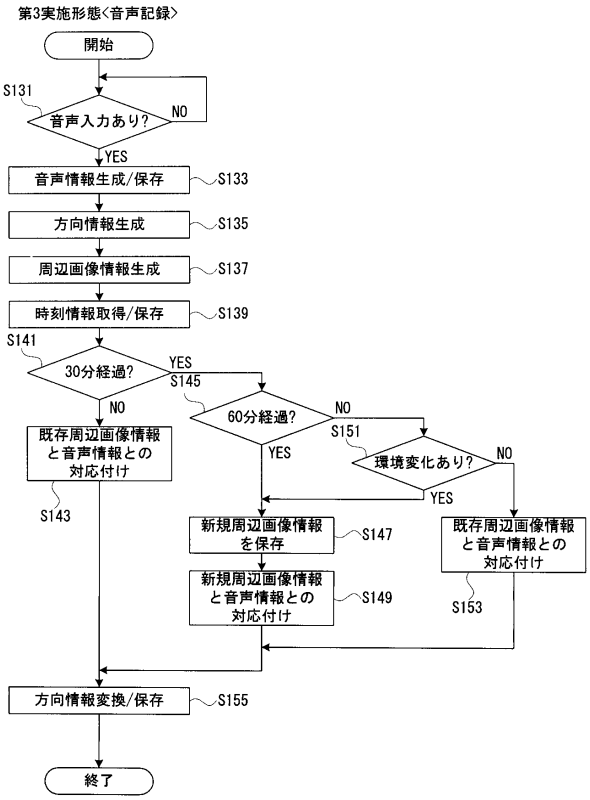
【図１６】


第3実施形態<音声再生>
第4実施形態<音声再生>
開始
すべての周辺画像を表示
S101
すべてのｱｲｺﾝを表示
S103
S105
ﾚｲﾔの選択?
NO
YES
選択された
周辺画像を表示
S107
選択された周辺画像の
ｱｲｺﾝを表示
S109
1
S111
ｱｲｺﾝのｸﾘｯｸ?
NO
YES
音声情報の再生
S113
S115
ｱｲｺﾝのﾄﾞﾗｯｸﾞ?
NO
YES
S117
ｱｲｺﾝの移動
S119
周辺画像
のﾄﾞﾗｯｸﾞ?
NO
YES
S121
周辺画像の
表示角度を変更
S123
終了?
NO
YES
終了
1

【図１７】


付箋3
付箋2
付箋1
ｱｲｺﾝ200
ﾚｲﾔ1
(周辺画像1)
ﾚｲﾔ2
(周辺画像2)
ﾚｲﾔ3
(周辺画像3)

【図１８】


第3実施形態<音声記録>
開始
S131
音声入力あり?
NO
YES
音声情報生成/保存
S133
方向情報生成
S135
周辺画像情報生成
S137
時刻情報取得/保存
S139
S141
30分経過?
YES
NO
既存周辺画像情報
と音声情報との
対応付け
S143
S145
60分経過?
NO
YES
S151
環境変化あり?
NO
YES
新規周辺画像情報
を保存
S147
新規周辺画像情報
と音声情報との
対応付け
S149
既存周辺画像情報
と音声情報との
対応付け
S153
方向情報変換/保存
S155
終了

フロントページの続き

(56)参考文献 特開２００５－３４１０９２（ＪＰ，Ａ）
特開２００６－１４６４１５（ＪＰ，Ａ）
特開２００４－１４７２８３（ＪＰ，Ａ）
特開２００８－０６５６７５（ＪＰ，Ａ）
特表２００９－５１７９３６（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｋ １５／０４
Ｇ１０Ｌ １９／００
Ｈ０４Ｓ １／００
Ｈ０４Ｓ ７／００
Ｇ１１Ｂ ２０／００－２０／０４
Ｇ１１Ｂ ２０／１０－２０／１６