

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5383056号  
(P5383056)

(45) 発行日 平成26年1月8日(2014.1.8)

(24) 登録日 平成25年10月11日(2013.10.11)

(51) Int. Cl.	F I	
<b>G 1 O L 19/00 (2013.01)</b>	G 1 O L 19/00	3 1 2 Z
<b>G 1 O L 15/00 (2013.01)</b>	G 1 O L 15/00	2 0 0 B
<b>G 1 O L 15/28 (2013.01)</b>	G 1 O L 15/28	4 0 0
<b>G 1 O L 15/22 (2006.01)</b>	G 1 O L 15/22	4 6 0 Z
<b>G 1 O L 21/0216 (2013.01)</b>	G 1 O L 21/02	1 0 1 Z
請求項の数 8 (全 13 頁) 最終頁に続く		

(21) 出願番号	特願2008-25678 (P2008-25678)	(73) 特許権者	000005326 本田技研工業株式会社 東京都港区南青山二丁目1番1号
(22) 出願日	平成20年2月5日(2008.2.5)	(74) 代理人	110000246 特許業務法人O F H特許事務所
(65) 公開番号	特開2008-197650 (P2008-197650A)	(72) 発明者	中臺 一博 埼玉県和光市本町8-1、株式会社ホンダ ・リサーチ・インスティテュート・ジャパ ン内
(43) 公開日	平成20年8月28日(2008.8.28)	(72) 発明者	長谷川 雄二 埼玉県和光市本町8-1、株式会社ホンダ ・リサーチ・インスティテュート・ジャパ ン内
審査請求日	平成22年11月26日(2010.11.26)		
(31) 優先権主張番号	60/901, 081		
(32) 優先日	平成19年2月14日(2007.2.14)		
(33) 優先権主張国	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 音データ記録再生装置および音データ記録再生方法

(57) 【特許請求の範囲】

【請求項1】

音データを取得する音データ取得部と、  
音源が存在する方向を特定する音源定位部と、  
音源ごとの音データを分離する音源分離部と、  
音源ごとの時系列の音データを格納する時系列データ格納部と、  
所定の時刻における所定の音源の方向を示す、音に関するストリームデータを格納するストリームデータ格納部と、  
前記時系列データ格納部および前記ストリームデータ格納部に接続され、データの処理を行うデータ処理部と、  
時刻と当該時刻における音源についての前記ストリームデータへのリンクとを対応付けるリストを格納するストリームリスト格納部と、  
音データを再生する音データ再生部と、  
ストリームデータを表示する表示部と、  
を備え、  
前記表示部は、前記ストリームデータに含まれた所定の時刻における所定の音源の方向を示す第1の表示を時間軸上に表示し、  
前記音データ再生部は、前記表示部によって時間軸上に表示された前記第1の表示の一つが選択されたことに応じて、当該選択された第1の表示に対応する所定の時刻における所定の音源の音データを再生する、

音データ記録再生装置であって、  
前記時間軸上の一の時刻が選択されたことに応じて、  
前記データ処理部は、当該選択された時刻におけるストリームデータを前記ストリー  
ムデータ格納部から取り出し、  
前記表示部は、当該取り出された前記対応するストリームデータに基づいて、前記選  
択された時刻に存在した音源およびその方向を示す第2の表示を表示し、  
前記第2の表示の中央部にはマイクロホンが表示され、前記選択された時刻に存在し  
た音源の方向が、前記表示されたマイクロホンの位置に対する方向として示され、  
前記第2の表示により表示された前記音源の一つが選択されることにより、前記第1  
の表示の一つが選択される、  
 音データ記録再生装置。

10

## 【請求項2】

ストリームデータが、時系列の音データへのリンクに関するデータを含み、  
前記データ処理部は、前記第2の表示に表示された前記音源の一つが選択されたこと  
に応じて、前記リンクにより当該選択された音源のストリームデータを参照し、当該選択さ  
れた音源の方向が所定の角度範囲にあるときは、当該選択された音源の音データを前記時  
系列データ格納部から取り出し、  
前記音データ再生部は、当該取り出された前記所定の時刻における前記所定の音源の音  
データを再生する、

請求項1に記載の音データ記録再生装置。

20

## 【請求項3】

音データから音声を認識し音声のテキスト情報を生成する音声認識部をさらに備え、前記表示部が前記音声のテキスト情報をさらに表示する請求項1または2に記載の音データ記録再生装置。

## 【請求項4】

画像データを取得する画像データ取得部をさらに備え、前記時系列データ格納部が、時系列の画像データをさらに格納し、前記ストリームデータ格納部に格納されるデータが、所定の時間において、所定の対象の方向を示す画像に関するストリームデータをさらに含む請求項1から3のいずれか一項に記載の音データ記録再生装置。

## 【請求項5】

音データ記録再生装置によって音データを再生する方法であって、  
 音データを取得し、  
 音源が存在する方向を特定し、  
 音源ごとの音データを分離し、  
 音源ごとの時系列の音データを、時系列データ格納部に格納し、  
 所定の時間において、所定の音源の方向を示す、音に関するストリームデータを作成し、  
 ストリームデータ格納部に格納し、  
 時刻と当該時刻における音源についての前記ストリームデータへのリンクとを対応付けるリストをストリームリスト格納部に格納し、  
 前記ストリームデータに含まれた所定の時刻における所定の音源の方向を示す第1の表示を時間軸上に表示し、

30

前記時間軸上に表示された前記第1の表示の一つが選択されたことに応じて、当該選択された第1の表示に対応する所定の時刻における所定の音源の音データを再生する、

音データ記録再生方法であって、

前記時間軸上の一の時刻が選択されたことに応じて、

当該選択された時刻におけるストリームデータを前記ストリームデータ格納部から取り出し、

当該取り出された前記対応するストリームデータに基づいて、前記選択された時刻に存在した音源およびその方向を示す第2の表示を表示し、

前記第2の表示の中央部にはマイクロホンが表示され、前記選択された時刻に存在し

50

た音源の方向が、前記表示されたマイクロホンの位置に対する方向として示され、  
前記第 2 の表示により表示された前記音源の一つが選択されることにより、前記第 1  
の表示の一つが選択される、  
音データ記録再生方法。

【請求項 6】

ストリームデータが、時系列の音データへのリンクに関するデータを含み、  
前記第 2 の表示に表示された前記音源の一つが選択されたことに応じて、前記リンクに  
より当該選択された音源のストリームデータを参照し、当該選択された音源の方向が所定  
の角度範囲にあるときは、当該選択された音源の音データを前記時系列データ格納部から  
取り出し、

10

当該取り出された前記所定の時刻における前記所定の音源の音データを再生する、  
請求項 5 に記載の音データ記録再生方法。

【請求項 7】

音データから音声を認識し音声のテキスト情報を生成し、前記音声のテキスト情報を表示する請求項 5 または 6 に記載の音データ記録再生方法。

【請求項 8】

画像データを取得し、時系列の画像データを前記時系列データ格納部に格納し、前記ストリームデータ格納部に格納されるデータが、所定の時間において、所定の対象の方向を示す画像に関するストリームデータをさらに含む請求項 5 から 7 のいずれか一項に記載の音データ記録再生方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記録された音データの中から所望の音データを容易に再生することのできる音データ記録再生装置および音データ記録再生方法に関する。

【背景技術】

【0002】

記録された音データの中から所望の音データを再生することは、記録された画像データの中から所望の画像データを再生することに比較して困難である。

【0003】

その第 1 の理由は、音データを時間に沿って概観するのが困難だからである。画像データの場合、データを早送りすること、あるいは、特定の時間間隔でデータをピックアップし同時に表示することで、所望の画像データを比較的容易に特定し、再生することができる。画像を早送りしても画像の色調などが変化することはなく、画像のコマ数を減らしても大きな認識誤りは発生しない。これに対して、音データの場合、データを早送りすると認識するのが困難となる。また、ピックアップした短時間の音データは、再生しても認識することができない。

30

【0004】

第 2 の理由は、音源の弁別が困難であるからである。画像データの場合、画面上で二つの物体が重なっていても前面の物体は正しく認識することができ、二つの物体を弁別することができる。これに対して、音データの場合、複数話者の発言が重なっている場合などに、発言内容を理解することは困難である。

40

【0005】

これまでに、たとえば、音源の位置を推定し、画像上に推定された音源の位置を表示するシステムが開発されている（たとえば、特許文献 1）。

【0006】

しかし、音源を弁別するとともに、音源ごとの音データを時間に沿って概観することができるように構成された音データ記録再生装置および音データ記録再生方法は開発されていない。

【特許文献 1】特開平 2 0 0 3 - 1 1 1 1 8 3 号公報

50

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0007】

したがって、記録された音データの中から所望の音データを容易に再生することができるように、音源ごとの音データを時間に沿って概観することができるように構成された音データ記録再生装置および音データ記録再生方法に対するニーズがある。

## 【課題を解決するための手段】

## 【0008】

本発明による音データ記録再生装置は、音データを取得する音データ取得部と、音源が存在する方向を特定する音源定位部と、音源ごとの音データを分離する音源分離部と、を備える。本装置は、音源ごとの時系列の音データを格納する時系列データ格納部と、所定の時間において、所定の音源の方向を示す、音に関するストリームデータを格納するストリームデータ格納部と、前記時系列データ格納部および前記ストリームデータ格納部に接続され、データの処理を行うデータ処理部と、をさらに備える。本装置は、音データを再生する音データ再生部と、ストリームデータを表示する表示部と、をさらに備え、前記表示部によって表示されたストリームデータが選択されると、前記音データ再生部が、選択されたストリームデータに関する音データを再生するように構成される。

10

## 【0009】

本発明による音データ記録再生装置によって音データを再生する方法は、音データを取得し、音源が存在する方向を特定し、音源ごとの音データを分離する。さらに、音源ごとの時系列の音データを、時系列データ格納部に格納し、所定の時間において、所定の音源の方向を示す、音に関するストリームデータを作成し、ストリームデータ格納部に格納する。さらに、ストリームデータを表示し、表示されたストリームデータが選択されると、選択されたストリームデータに関する音データを再生する。

20

## 【0010】

本発明によれば、所定の時間において、所定の音源の方向を示す、音に関するストリームデータが表示されるので、音源ごとの音データを時間に沿って概観ことができ、記録された音データの中から所望の音データを容易に再生することができる。

## 【0011】

本発明の実施形態によれば、音データから音声を認識し音声のテキスト情報を生成し、前記音声のテキスト情報を表示する。

30

## 【0012】

本実施形態によれば、たとえば、耳の不自由な人も、記録された音データの中から所望の音データ（音声）を容易に再生し利用することができる。

## 【0013】

本発明の実施形態によれば、画像データを取得し、時系列の画像データを前記時系列データ格納部に格納し、前記ストリームデータ格納部に格納されるデータが、所定の時間において、所定の対象の方向を示す画像に関するストリームデータをさらに含む。

## 【0014】

本実施形態によれば、音源（話者）の音データ（話の内容）と画像データ（話者の表情）を関連付けて再生することもできる。

40

## 【発明を実施するための最良の形態】

## 【0015】

図1は、本発明の一実施形態による音データ記録再生装置の構成を示す図である。

## 【0016】

音データ記録再生装置において、音データ取得部101が音データを取得する。取得された音データは、音源定位部103に送られ、音源定位部103は、音源が存在する方向を特定する。また、取得された音データは、音源分離部105に送られ、音源分離部105は、音源ごとの音を分離する。また、画像データ取得部107が画像データを取得する。

50

## 【 0 0 1 7 】

音データ記録再生装置は、時系列データ格納部 1 0 9、ストリームデータ格納部 1 1 1 およびストリームリスト格納部 1 1 3 の 3 種類のメモリを備える。これらのメモリのデータ構造については後で説明する。

## 【 0 0 1 8 】

音データ記録再生装置は、時系列データ格納部 1 0 9、ストリームデータ格納部 1 1 1 およびストリームリスト格納部 1 1 3 に接続されたデータ処理部 1 1 5 ならびに時系列データ格納部 1 0 9 およびデータ処理部 1 1 5 に接続された音声認識部 1 2 1 を備える。データ処理部 1 1 5 は、表示・入力部 1 1 7 および音データ再生部 1 1 9 にさらに接続されている。データ処理部 1 1 5 は、上記 3 種類のメモリに格納されたデータを使用して処理 10  
を行い、表示・入力部 1 1 7 が音データを概観的に表示し、表示・入力部 1 1 7 からのオペレータ入力にしたがって、音データ再生部 1 1 9 が所望の音データを再生するように構成されている。音声認識部 1 2 1 は、データ処理部 1 1 5 からの指示により時系列データ格納部 1 0 9 に格納された音データから音声を認識し、音声のテキスト情報を生成し、時系列データ格納部 1 0 9 に格納する。表示・入力部 1 1 7 は、表示部および入力部の別個の筐体であってもよい。

## 【 0 0 1 9 】

音データ取得部 1 0 1 は、一例として、8 個のマイクと、マイクが採取した音データを処理する音響信号処理用プロセッサとを含む。音響信号処理用プロセッサは、たとえば、 16 チャンルの音データを所定の周波数で採取することができる。8 個のマイクは、7 個 20  
の通常のマイクと 1 個のサラウンド用マイクからなる。7 個の通常のマイクは、球状の形をした台に取り付けられており、その台は、それぞれのマイクが設置されている方向からの音を採取しやすいように設計されている。

## 【 0 0 2 0 】

本発明による音データ記録再生装置は、たとえば、会議室など所定の位置で音データを取得し再生する場合に使用してもよい。その場合に、音データ取得部 1 0 1 は、該所定の位置に配置される。他の実施形態において、本発明による音データ記録再生装置は、移動体における音データを取得し再生する場合に使用してもよい。その場合に、音データ取得部 1 0 1 は、該移動体に取り付けられる。移動体には、一例として、車両、監視ロボット 30  
および人間などが含まれる。たとえば、GPS (Global Positioning System、全地球測位システム) と組み合わせて、移動体の位置において音源の方向を認識することができる。

## 【 0 0 2 1 】

音源定位部 1 0 3 は、音源が存在する方向を特定するために、たとえば、Steered beam former による定位を行い、カルマンフィルタにより精度を上げる (Masamitsu Murase, Shun'ichi Yamamoto, Jean-Marc Valin, Kazuhiro Nakadai, Kentaro Yamada, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno: Multiple Moving Speaker Tracking by Microphone Array on Mobile Robot, Proceedings of the Ninth European Conference on Speech Communication and Technology (Interspeech-2005), 249-252, Lisboa, Sep. 2005 p.10)。Steered beam former による定位は、以下のステップにより、マイクペア 40  
間での相互相関に基づいた定位を行う。

## 【 0 0 2 2 】

1) マイクの座標を定義した座標系の原点を中心とする球を考え、その表面に 2 5 6 2 個の点を等間隔に配置する。

2) 各点に対して、全てのマイクのペアについて相互相関の和を求める。この和が最大となる点の方向が音源の方向であると推定する。

3) 推定された相互相関の値を全て 0 とする。

4) 2) および 3) を繰り返し、全ての音源方向を推定する。これにより、推定した音源方向を時間軸上で走査し、方向の近いものを 1 つの音源としてラベル付けする。

## 【 0 0 2 3 】

10

20

30

40

50

カルマンフィルタは、移動音源のような動的に変化するシステムを観測する場合に、過去の観測値から現在及び未来の内部状態を予測するためのものである。たとえば、過去の話者の状態から現在の話者の状態を推定する。これにより、単一の話者が移動する場合に、単一話者へ確実に同一のラベルを付与することができる。

【 0 0 2 4 】

音源分離部 1 0 5 は、音源ごとの音データを分離するために、たとえば、Geometric Source Separation(GSS)による音源分離を行いPost-Filterによって雑音抑圧処理を行う ( J.-M. Valin, J. Rouat, F. Michaud: Enhanced Robot Audition Based on Microphone Array Source Separation with Post-Filter, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems ( IROS), pp.2123-2128, 2004. )。

10

【 0 0 2 5 】

周波数 $k$ において、音響信号 $(k)$ から観測信号 $x(k)$ への変換は線形であると仮定して、音源信号から観測信号への伝達関数を式 ( 1 ) によって表す。

【 数 1 】

$$x(k) = A(k) s(k) + n(k) \quad (1)$$

ここで $A(k)$ は変換を表す行列であり、 $n(k)$ はノイズである。これより推定した音源信号 $y(k)$ は式 ( 2 ) で表される。

【 数 2 】

$$y(k) = W(k) x(k) \quad (2)$$

20

$x(k)$ は観測信号であるため、音源信号を求めることは変換行列 $W(k)$ を求める問題に帰着する。異なる音源間信号の独立性を仮定することにより式 ( 3 ) が、音源とマイクの幾何学的制限により式 ( 4 ) が得られる。

【 数 3 】

$$R_{yy}(k) - \text{diag}[R_{yy}(k)] = 0 \quad (3)$$

$$W(k) A(k) = I \quad (4)$$

これらの式 ( 3 ) および ( 4 ) は制限が強いため、 $W$ の近似解を確率的勾配法により求める。

30

【 0 0 2 6 】

Post-Filterによる雑音抑圧は、GSSによる分離音から雑音を取り除くものである。この手法では雑音を定常性雑音と非定常性雑音に分けて推定し除去する。定常性雑音はMCRA (Minima Controlled Recursive Average)により計算する。非定常性雑音は、GSSの過程で別のチャンネルから漏洩したものと仮定し、適応的に干渉成分の推定を行う。

【 0 0 2 7 】

時系列データ格納部 1 0 9、ストリームデータ格納部 1 1 1 およびストリームリスト格納部 1 1 3 のデータ構造について以下に説明する。

【 0 0 2 8 】

図 2 は、時系列データ格納部 1 0 9、ストリームデータ格納部 1 1 1 およびストリームリスト格納部 1 1 3 の音データに関するデータ構造を示す図である。

40

【 0 0 2 9 】

時系列データ格納部 1 0 9 には、それぞれの時刻における音の波形データが格納される。音の波形データは、分離されていない波形データと音源ごとに分離された波形データを含む。音源ごとに分離されたデータには、音源の方向のデータが付与される。それぞれの時刻におけるデータは、時間方向にリンクされており、たとえば、時刻  $t = 2$  におけるデータから、時刻  $t = 1$  または時刻  $t = 3$  のデータを参照することができる。

【 0 0 3 0 】

ストリームデータ格納部 1 1 1 には、音に関するストリームデータが格納される。音に

50

関するストリームは、一定時間連続する所定の音源の音である。音に関するストリームデータは、一定時間連続する所定の音源の音について、時刻ごとの該音源の方向（位置）を示すデータである。具体的にストリームデータは、開始時刻、終了時刻、音源の識別（たとえば、人物名）、該所定の時間内のそれぞれの時刻における音源の方向に関するデータを含む。音源の方向は、たとえば、水平面および鉛直面内の角度で表現される。ストリームデータは、音の波形データへのリンクをさらに含む。たとえば、所定の時刻におけるストリームデータから該時刻に対応する音の波形データを参照することができる。

【0031】

図3は、時系列データ格納部109、ストリームデータ格納部111およびストリームリスト格納部113の画像データに関するデータ構造を示す図である。

10

【0032】

時系列データ格納部109には、それぞれの時刻における画像データ（動画情報）が格納される。画像データには、対象（たとえば人物）の方向のデータが付与される。人物の識別は、一般的な顔認識の技術（たとえば、特開2002-216129号公報）を使用して行う。それぞれの時刻におけるデータは、時間方向にリンクされており、たとえば、時刻 $t = 2$ におけるデータから、時刻 $t = 1$ または時刻 $t = 3$ のデータを参照することができる。

【0033】

ストリームデータ格納部111には、画像に関するストリームデータが格納される。画像に関するストリームは、一定時間連続する所定の対象の画像である。画像に関するストリームデータは、一定時間連続する所定の対象の画像について、時刻ごとの該対象の方向（位置）を示すデータである。具体的にストリームデータは、開始時刻、終了時刻、対象の識別（たとえば、人物名）、それぞれの時刻における対象の方向に関するデータを含む。対象の方向は、たとえば、水平面および鉛直面内の角度で表現される。ストリームデータは、画像データ（動画情報）へのリンクをさらに含む。たとえば、所定の時刻におけるストリームデータから該時刻に対応する画像データを参照することができる。

20

【0034】

ストリームデータ格納部111において、音に関するストリームデータと画像に関するストリームデータが存在する。音に関するストリームデータにおける音源の方向と画像に関するストリームデータの対象の方向とが、所定の時間以上一致する場合、上記2個のストリームデータを統合し、統合ストリームデータとしてもよい。統合ストリームデータを使用すれば、たとえば、音源（話者）の音データ（話の内容）と画像データ（話者の表情）を関連付けて再生することもできる。

30

【0035】

ストリームリスト格納部113におけるデータは、ストリームデータのリストに関するデータである。ストリームデータのリストは、概観表示用のものであり、間引いた時刻における総ストリーム数、各ストリームの音源の方向、各ストリームのタイプ（音に関するストリーム、画像に関するストリーム、統合ストリーム）、各ストリームデータへのリンクを含む。たとえば、所定の時刻におけるストリームデータのリストから該時刻に対応するストリームデータを参照することができる。図2および図3に示すように、ストリームリスト格納部113におけるデータに基づいて、音源または対象の識別（人物名）と時間との関数、または音源または対象の方向（角度）と時間との関数として、それぞれのストリームを、表示・入力部117に表示することができる。図2および図3において、3個のストリームデータが、実線、点線および一点鎖線によって示されている。

40

【0036】

図4は、本実施形態による音データ記録再生装置のインタフェース機能を概念的に示した図である。本実施形態による音データ記録再生装置のインタフェースは、音データをわかりやすく再生することができるように、概観表示（Over first）、ズーム（Zoom）、フィルタ（Filter）および要求による詳細表示（Details on demand）のための機能を備える。

50

## 【 0 0 3 7 】

図 4 ( a ) は、概観表示の内容を示す図である。縦軸は音源の識別 ( 人物名 ) を表し、横軸は時間を表す。この表示は、ストリームリスト格納部 1 1 3 におけるデータに基づいて行うことができる。図 4 ( a ) に示した概観表示によって、ユーザは、音データを時間に沿って概観することができる。

## 【 0 0 3 8 】

図 4 ( b ) は、ズームされた内容を示す図である。図 4 ( b ) は、具体的に、ズームされた時刻における音源およびその方向を示す。これらのデータは、ストリームデータ格納部 1 1 1 におけるストリームデータに含まれる。ユーザは、図 4 ( a ) に示した外観表示の横軸における特定の時刻を指定することにより、該時刻のデータを容易に取り出す ( ズームする ) ことができる。ストリームデータのリストは、各ストリームデータへのリンクを含むので、上記のデータの取り出しを容易に行うことができる。

10

## 【 0 0 3 9 】

図 4 ( c ) は、フィルタ処理の内容を示す図である。図 4 ( c ) は、具体的に、音源の範囲を指定することで、BGM ( バックグラウンドミュージック ) の音源を除去することを示す。

## 【 0 0 4 0 】

図 4 ( d ) は、詳細表示の内容を示す図である。ユーザは、たとえば、図 4 ( b ) に示された所定の時刻における音源 ( 発話者 ) から所定の音源 ( 発話者 ) を指定して、その音データ ( 発言内容 ) のみを再生することができる。

20

## 【 0 0 4 1 】

図 5 は、本発明の一実施形態による音データ記録再生方法を示す流れ図である。

## 【 0 0 4 2 】

図 5 におけるステップ S 0 1 0 において、音データ取得部 1 0 1 が音データを取得し、画像データ取得部 1 0 7 が画像データを取得する。

## 【 0 0 4 3 】

図 5 におけるステップ S 0 2 0 において、音源定位部 1 0 3 が、音源が存在する方向を特定する。

## 【 0 0 4 4 】

図 5 におけるステップ S 0 3 0 において、音源分離部 1 0 5 が、音源ごとの音を分離する。

30

## 【 0 0 4 5 】

図 5 におけるステップ S 0 4 0 において、音データおよび画像データが、時系列データ格納部 1 0 9 に格納される。時系列データ格納部は、音声認識部 1 2 1 によって生成された音声のテキスト情報をさらに格納してもよい。

## 【 0 0 4 6 】

図 5 におけるステップ S 0 5 0 において、データ処理部 1 1 5 が、時系列データ格納部 1 0 9 に格納された音データおよび画像データに基づいて、ストリームデータを作成し、ストリームデータ格納部 1 1 1 に格納する。

## 【 0 0 4 7 】

図 5 におけるステップ S 0 6 0 において、データ処理部 1 1 5 が、ストリームデータ格納部 1 1 1 に格納されたストリームデータに基づいて、ストリームデータのリストを作成し、ストリームリスト格納部 1 1 3 に格納する。

40

## 【 0 0 4 8 】

図 5 におけるステップ S 0 7 0 において、表示・入力部 1 1 7 が、ストリームデータを時間軸とともに表示する ( 図 4 ( a ) ) 。

## 【 0 0 4 9 】

図 5 におけるステップ S 0 8 0 において、表示・入力部 1 1 7 が、表示された時間軸上においてユーザによって選択された時刻における、音源およびその方向を表示する ( 図 4 ( b ) ) 。

50



## 【0050】

図5におけるステップS090において、音データ再生部119が、ユーザによって選択された音源の音データを再生する(図4(d))。

## 【0051】

ステップS070、S080およびステップS090に代えて、表示・入力部117が、ストリームデータにしたがって、たとえば、図4(b)に示すような画面によって、時間ごとに変化する音源およびその方向を、リアルタイムに、または、リアルタイムと同じか増加もしくは減少させた速度で表示するように構成してもよい。

## 【0052】

図6は、表示・入力部117の表示画面の一例を示す図である。表示画面は、たとえば、操作パネル部、音源方向表示部およびストリーム表示部からなる。操作パネル部から、ユーザによる音データ記録再生装置に対する指示が入力される。表示部と入力部は、図6に示すように1筐体の1画面から構成してもよい。あるいは、1筐体の2画面または画面を使用するかまたは使用しない2筐体によって構成してもよい。

10

## 【0053】

表示画面は、音声認識部121によって生成された音声のテキスト情報を表示するテキスト情報表示部(不図示)をさらに備えてもよい。表示・入力部117は、音データ再生部119が、ユーザによって選択された音源の音データ(音声)を再生する(図4(d))際に、音声のテキスト情報をテキスト情報表示部に表示する。テキスト情報表示部において、認識尤度の高い音声のテキスト情報を強調して濃い色で表示し、認識尤度の低い音声のテキスト情報を薄い色で表示するようにしてもよい。音声のテキスト情報を表示することにより、たとえば、耳の不自由な人も、記録された音データの中から所望の音データ(音声)を容易に再生し利用することができる。

20

## 【0054】

図7は、表示画面のストリーム表示部の詳細を示す図である。ストリーム表示部は、ストリームリスト格納部113またはストリームデータ格納部111のデータに基づいて、ストリームデータを表示する。横軸は、時間であり、縦軸は水平面内における角度を示す。ストリーム表示部の表示内容は、図4(a)の表示と対応する。図7には、実線で示したストリームAおよび点線で示したストリームBのストリームデータが表示されている。ユーザは、時間軸(横軸)上の時刻を指定することにより、再生時刻を指定することができる。時系列データ格納部109のデータに基づいて、ストリーム表示部に、収録された音の波形データをあわせて表示してもよい。

30

## 【0055】

図8は、表示画面の音源方向表示部の詳細を示す図である。音源方向表示部は、ストリームデータ格納部111のデータに基づいて、指定された再生時刻における音源およびその方向を表示する。表示画面は、たとえば、水平面を示し、中心に配置されたマイクの位置に対する音源の方向を示す。図8において、Aと付された線がストリームAの音源の方向を示し、Bと付された線がストリームBの音源の方向を示す。この表示は、図4(b)の表示と対応する。音源表示部の画面において、たとえばクリックなどの操作により角度範囲を限定してその範囲の音源の音データを再生するようにしてもよい。この操作は、図4(c)の操作と対応する。さらに、図8において、Aと付された線またはBと付された線をクリックすることによって、音源を指定しその音源の音データのみを再生するようにしてもよい。この操作は、図4(d)の操作と対応する。

40

## 【0056】

本発明の実施形態の特徴を以下に説明する。

## 【0057】

本発明の実施形態によれば、ストリームデータが、時系列の音データへのリンクに関するデータを含む。

## 【0058】

本実施形態によれば、時系列の音データへのリンクにより、選択された所定の時刻にお

50

ける音データを容易に再生することができる。

【0059】

本発明の実施形態によれば、ストリームデータのリストがストリームリスト格納部に格納される。

【0060】

本実施形態によれば、音源ごとの音データを時間に沿って概観できるように表示するのが容易になる。

【0061】

本発明の実施形態によれば、ストリームデータとともに時間軸を表示し、表示された時間軸上の点を選択されると、対応するストリームデータを前記ストリームデータ格納部から取り出し、該対応するストリームデータに基づいて、選択された所定の時刻における音源およびその方向を表示する。

10

【0062】

本実施形態によれば、ユーザは、ストリームデータによって、記録された音データを時間に沿って概観することができ、ストリームデータとともに表示された時間軸上の点を選択することによって、選択された所定の時刻における音源およびその方向のデータを容易に取り出すことができる。

【0063】

本発明の実施形態によれば、前記表示部に表示された所定の時刻における音源から所定の音源が選択されると、該所定の音源の音データを前記時系列データ格納部から取り出し、該対応する音データを再生する。

20

【0064】

本実施形態によれば、ユーザは、表示された所定の時刻における音源から所定の音源を選択することによって、該所定の音源の音データを容易に取り出すことができる。

【図面の簡単な説明】

【0065】

【図1】本発明の一実施形態による音データ記録再生装置の構成を示す図である。

【図2】時系列データ格納部、ストリームデータ格納部およびストリームリスト格納部の音データに関するデータ構造を示す図である。

【図3】時系列データ格納部、ストリームデータ格納部およびストリームリスト格納部の画像データに関するデータ構造を示す図である。

30

【図4】本実施形態による音データ記録再生装置のインタフェース機能を概念的に示した図である。

【図5】本発明の一実施形態による音データ記録再生方法を示す流れ図である。

【図6】表示・入力部の表示画面の一例を示す図である。

【図7】表示画面のストリーム表示部の詳細を示す図である。

【図8】表示画面の音源方向表示部の詳細を示す図である。

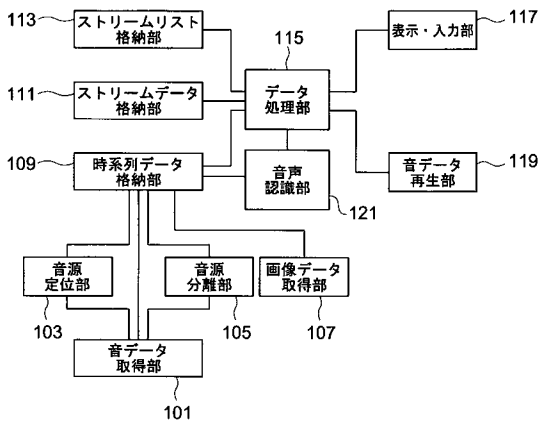
【符号の説明】

【0066】

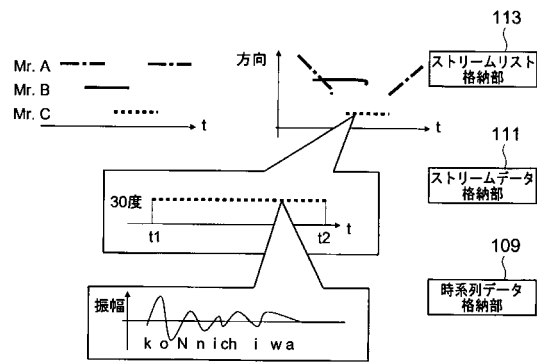
101...音データ取得部、103...音源定位部、105...音源分離部、109...時系列データ格納部、111...ストリームデータ格納部、113...ストリームリスト格納部、115...データ処理部

40

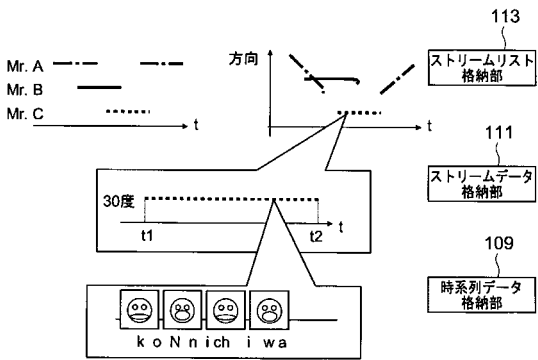
【図1】



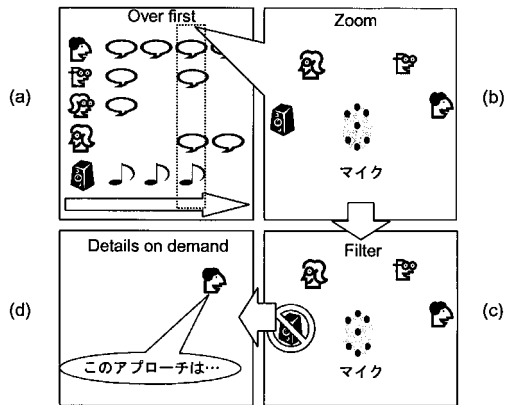
【図2】



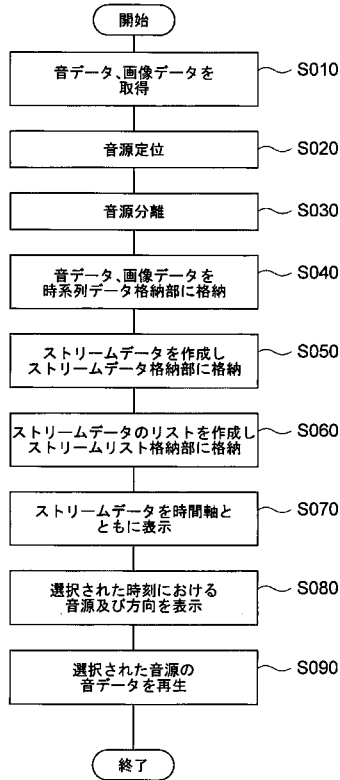
【図3】



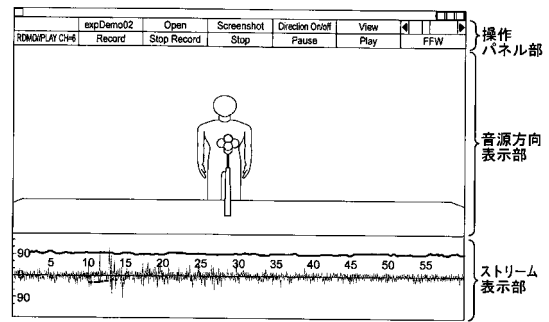
【図4】



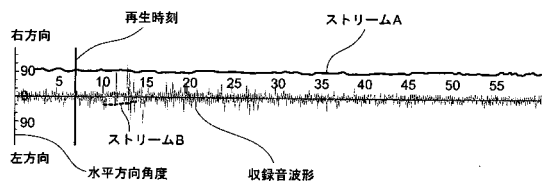
【 図 5 】



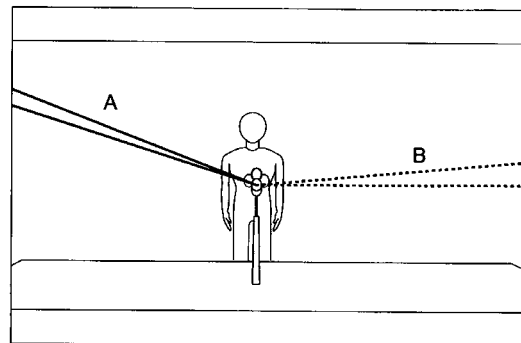
【 図 6 】



【 図 7 】



【 図 8 】



## フロントページの続き

(51)Int.Cl.		F I		
<b>G 1 0 L</b>	<b>21/0208</b>	<b>(2013.01)</b>	G 1 0 L	21/02 1 0 2 B
<b>G 1 0 L</b>	<b>21/028</b>	<b>(2013.01)</b>	G 1 0 L	21/02 1 0 2 A
<b>G 1 0 L</b>	<b>21/0272</b>	<b>(2013.01)</b>	G 1 0 L	21/02 2 0 1 B
<b>H 0 4 R</b>	<b>3/00</b>	<b>(2006.01)</b>	G 1 0 L	21/02 2 0 2 A
			G 1 0 L	21/02 2 0 2 B
			H 0 4 R	3/00 3 2 0

(72)発明者 辻野 広司  
埼玉県和光市本町 8 - 1、株式会社ホンダ・リサーチ・インスティテュート・ジャパン内

(72)発明者 奥乃 博  
埼玉県和光市本町 8 - 1、株式会社ホンダ・リサーチ・インスティテュート・ジャパン内

審査官 毛利 太郎

(56)参考文献 国際公開第 2 0 0 6 / 1 3 7 7 3 2 ( W O , A 1 )

特開 2 0 0 4 - 1 8 5 5 1 4 ( J P , A )  
 特開 2 0 0 3 - 2 7 0 0 3 4 ( J P , A )  
 特開平 1 1 - 1 4 3 4 6 0 ( J P , A )  
 特開 2 0 0 7 - 3 2 9 7 0 2 ( J P , A )  
 特開 2 0 0 1 - 0 5 6 7 0 0 ( J P , A )  
 特開 2 0 0 1 - 1 3 4 6 1 3 ( J P , A )  
 特開 2 0 0 5 - 2 9 5 0 1 5 ( J P , A )  
 特開 2 0 0 6 - 1 2 3 1 6 1 ( J P , A )  
 特開平 0 8 - 2 9 7 6 2 4 ( J P , A )  
 特開 2 0 0 5 - 0 4 3 6 2 8 ( J P , A )

(58)調査した分野(Int.Cl. , D B 名)

G 1 0 L 1 9 / 0 0 - 1 9 / 2 6  
 G 1 0 L 1 5 / 0 0 - 1 5 / 3 4  
 G 1 0 L 2 1 / 0 0 - 2 1 / 1 8  
 H 0 4 R 3 / 0 0