

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第3890692号  
(P3890692)

(45) 発行日 平成19年3月7日(2007.3.7)

(24) 登録日 平成18年12月15日(2006.12.15)

(51) Int.C1.

F 1

<b>G 1 O K</b>	<b>15/04</b>	<b>(2006.01)</b>	G 1 O K	15/04	3 0 2 D
<b>G 1 O L</b>	<b>13/02</b>	<b>(2006.01)</b>	G 1 O L	13/02	1 2 2 B
<b>G 1 O L</b>	<b>15/00</b>	<b>(2006.01)</b>	G 1 O L	15/00	2 0 0 C
			G 1 O L	15/00	2 0 0 G

請求項の数 14 (全 23 頁)

(21) 出願番号

特願平9-234127

(22) 出願日

平成9年8月29日(1997.8.29)

(65) 公開番号

特開平11-73192

(43) 公開日

平成11年3月16日(1999.3.16)

審査請求日

平成16年1月21日(2004.1.21)

(73) 特許権者 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(74) 代理人 100086841

弁理士 脇 篤夫

(74) 代理人 100102635

弁理士 浅見 保男

(72) 発明者 濱谷 健二

東京都品川区北品川6丁目7番35号 ソニー株式会社内

審査官 荘原 雄一

最終頁に続く

(54) 【発明の名称】情報処理装置及び情報配信システム

## (57) 【特許請求の範囲】

## 【請求項 1】

第1のオーディオ情報を、上記第1のオーディオ情報よりボーカル部を抽出したボーカル情報と、上記第1のオーディオ情報よりボーカル部を取り除いた伴奏情報とに分離する楽曲情報分離手段と、

上記ボーカル情報について第1の言語における音声認識を行って第1の言語文字情報を生成する音声認識手段と、

上記第1の言語文字情報をについて第2の言語への翻訳処理を行って第2の言語文字情報を生成する翻訳手段と、

上記第2の言語文字情報を利用して上記第2の言語により発音される翻訳ボーカル情報を生成し、この翻訳ボーカル情報と上記伴奏情報を合成することにより、第2のオーディオ情報を生成する情報合成手段と

を備えていることを特徴とする情報処理装置。

## 【請求項 2】

上記楽曲情報分離手段、上記音声認識手段、上記翻訳手段、及び上記情報合成手段により生成される情報のうち、少なくとも何れか1種類の情報を格納することができる情報記憶手段が設けられる

ことを特徴とする請求項1に記載の情報処理装置。

## 【請求項 3】

上記伴奏情報、上記第1の言語文字情報、上記第2の言語文字情報、及び上記第2のオ

10

20

ーディオ情報のうちから少なくとも 1 つの情報を選択可能な選択操作手段と、  
上記選択操作手段により選択された情報を上記情報記憶手段から読み出して出力する情報出力手段と、  
を備えていること特徴とする請求項 2 に記載の情報処理装置。

**【請求項 4】**

表示手段が備えられ、  
上記選択操作手段は、上記伴奏情報、上記第 1 の言語文字情報、上記第 2 の言語文字情報、及び上記第 2 のオーディオ情報のうちから所望の情報を指定するための操作に利用する操作画像を上記表示手段に対して行うように構成されている  
ことを特徴とする請求項 3 に記載の情報処理装置。 10

**【請求項 5】**

上記表示手段は、上記伴奏情報、上記第 1 の言語文字情報、上記第 2 の言語文字情報、及び上記第 2 のオーディオ情報の各情報を獲得するための処理が完了するごとに、これら各情報に対応する項目を上記操作画像として順次表示していくように構成されている  
ことを特徴とする請求項 4 に記載の情報処理装置。

**【請求項 6】**

第 1 のオーディオ情報を選択して出力可能に構成された情報送信装置と、  
上記情報送信装置と通信可能とされることにより、上記情報送信装置から出力された上記第 1 のオーディオ情報を受信する受信動作とが可能とされると共に、情報出力動作として、少なくとも上記第 1 のオーディオ情報に基づいて獲得した情報を外部に対して送信出力可能とされる情報伝送装置と、 20

情報記憶手段が備えられると共に、上記情報伝送装置と通信可能とされることで、情報記憶動作として、少なくとも上記情報伝送装置から送信出力された情報を上記情報記憶手段に対して記憶可能とされる端末装置と

を備えて当該情報配信システムが構成され、  
この情報配信システムにおいて備えられる情報処理系として、  
上記情報送信装置から出力された第 1 のオーディオ情報について、ボーカル情報と伴奏情報とに分離する楽曲情報分離手段と、

上記ボーカル情報について音声認識を行って第 1 の言語文字情報を生成する音声認識手段と、 30

上記第 1 の言語文字情報について翻訳処理を行って第 2 の言語文字情報を生成する翻訳手段と、

上記第 2 の言語文字情報を利用して翻訳言語により発音される翻訳ボーカル情報を生成し、この翻訳ボーカル情報と上記伴奏情報を合成することにより、第 2 のオーディオ情報を生成する情報合成手段と

が備えられることを特徴とする情報配信システム。

**【請求項 7】**

上記端末装置は、  
上記楽曲情報分離手段、上記音声認識手段、上記翻訳手段、及び上記情報合成手段により生成される情報のうち、少なくとも何れか 1 種類の情報を、上記情報記憶手段に格納するように構成されていることを特徴とする請求項 6 に記載の情報配信システム。 40

**【請求項 8】**

上記端末装置は、  
上記伴奏情報、上記第 1 の言語文字情報、上記第 2 の言語文字情報、及び上記第 2 のオーディオ情報のうちから少なくとも 1 つの情報を選択可能な選択操作手段と、  
上記選択操作手段により選択された情報を上記情報記憶手段から読み出して出力する情報出力手段と、  
を備えていること特徴とする請求項 7 に記載の情報配信システム。

**【請求項 9】**

上記情報伝送装置及び上記端末装置の少なくとも何れか一方に表示手段が備えられ、 50

上記選択操作手段は、上記伴奏情報、上記第1の言語文字情報、上記第2の言語文字情報、及び上記第2のオーディオ情報のうちから所望の情報を指定するための操作に利用する操作画像を上記表示手段に対して行うように構成されている

ことを特徴とする請求項8に記載の情報配信システム。

#### 【請求項10】

上記表示手段は、

上記伴奏情報、上記第1の言語文字情報、上記第2の言語文字情報、及び上記第2のオーディオ情報の各情報を獲得するための処理が完了するごとに、これら各情報に対応する項目を上記操作画像として順次表示していくように構成されている

ことを特徴とする請求項9に記載の情報配信システム。

10

#### 【請求項11】

第1のオーディオ情報を、上記第1のオーディオ情報よりボーカル部を抽出したボーカル情報と、上記第1のオーディオ情報よりボーカル部を取り除いた伴奏情報とに分離する楽曲情報分離ステップと、

上記ボーカル情報について第1の言語における音声認識を行って第1の言語文字情報を生成する音声認識ステップと、

上記第1の言語文字情報について第2の言語への翻訳処理を行って第2の言語文字情報を生成する翻訳ステップと、

上記第2の言語文字情報をを利用して、上記第2の言語により発音される翻訳ボーカル情報を生成する翻訳ボーカル生成ステップと、

20

上記翻訳ボーカル情報と上記伴奏情報を合成することにより、第2のオーディオ情報を生成する情報合成ステップと

を有する情報処理方法。

#### 【請求項12】

上記楽曲情報分離ステップ、上記音声認識ステップ、上記翻訳ステップ、及び上記情報合成ステップにより生成される情報のうち、少なくとも何れか1種類の情報を格納する情報記憶ステップ

を有する請求項11に記載の情報処理方法。

#### 【請求項13】

上記伴奏情報、上記第1の言語文字情報、上記第2の言語文字情報、及び上記第2のオーディオ情報のうちから少なくとも1つの情報を選択する選択操作ステップと、

上記選択操作ステップにより選択された情報を出力する情報出力ステップと

を有する請求項12に記載の情報処理方法。

30

#### 【請求項14】

上記伴奏情報、上記第1の言語文字情報、上記第2の言語文字情報、及び上記第2のオーディオ情報を獲得する処理が完了するごとに、これら各情報に対応する項目を順次表示していく表示ステップ

を有する請求項13に記載の情報処理方法。

#### 【発明の詳細な説明】

##### 【0001】

40

##### 【発明の属する技術分野】

本発明は、例えば情報が蓄積される情報格納装置から情報伝送装置に情報を配信し、更に情報伝送装置にて受信した情報を出力することで、端末装置においてその情報をコピーすることができるようとした情報配信システム、及びこのような情報配信システムに備えられて、所要の情報処理を行う情報処理装置に関するものである。

##### 【0002】

##### 【従来の技術】

先に本出願人により、例えばサーバに大量の楽曲データ（オーディオデータ）や映像データ等の情報をデータベースとして格納しておくと共に、この大量の情報のうちから必要とされる情報を多数の中間サーバ装置に配信することにより、この中間サーバ装置から、ユ

50

ーザが個人で所有する携帯端末装置に対して指定の情報をコピー（ダウンロード）できるようにした情報配信システムが提案されている。

#### 【0003】

##### 【発明が解決しようとする課題】

例えば上記のような情報配信システムにおいて、楽曲データを携帯端末装置にダウンロードする場合のサービスの形態について考えてみた場合、一般的には、楽曲単位もしくはアルバム単位の複数楽曲のオーディオ信号をデジタル情報化し、このデジタル情報化された楽曲をサーバ装置から中間サーバ装置を介して携帯端末装置に伝送することになる。

このようにデジタル情報化された情報を送信するのであれば、単にデジタル情報化された楽曲情報だけでなく、例えば情報配信システム内において、例えばある楽曲のデジタルデータを素材として扱って所要の情報処理を施すことにより、1つの楽曲情報から付随して生成される二次的な各種派生情報を、携帯端末装置のユーザに対して提供することは可能である。このような派生情報をユーザに提供できるようにすれば、情報配信システムとしての利用価値はより高められることになる。10

#### 【0004】

##### 【課題を解決するための手段】

本発明は上記したような課題を考慮して、第1のオーディオ情報を、上記第1のオーディオ情報よりボーカル部を抽出したボーカル情報と、上記第1のオーディオ情報よりボーカル部を取り除いた伴奏情報とに分離する楽曲情報分離手段と、上記ボーカル情報について第1の言語における音声認識を行って第1の言語文字情報を生成する音声認識手段と、上記第1の言語文字情報について第2の言語への翻訳処理を行って第2の言語文字情報を生成する翻訳手段と、上記第2の言語文字情報をを利用して上記第2の言語により発音される翻訳ボーカル情報を生成し、この翻訳ボーカル情報と上記伴奏情報を合成することにより、第2のオーディオ情報を生成する情報合成手段とを備えて情報処理装置を構成することとした。20

#### 【0005】

また、第1のオーディオ情報を選択して出力可能に構成された情報送信装置と、上記情報送信装置と通信可能とされることにより、上記情報送信装置から出力された上記第1のオーディオ情報を受信する受信動作とが可能とされると共に、情報出力動作として、少なくとも上記第1のオーディオ情報に基づいて獲得した情報を外部に対して送信出力可能とされる情報伝送装置と、情報記憶手段が備えられると共に、上記情報伝送装置と通信可能とされることで、情報記憶動作として、少なくとも上記情報伝送装置から送信出力された情報を上記情報記憶手段に対して記憶可能とされる端末装置とを備えて当該情報配信システムを構成することとした。30

そして、この情報配信システムにおいて備えられる情報処理系として、上記情報送信装置から出力された第1のオーディオ情報について、ボーカル情報と伴奏情報とに分離する楽曲情報分離手段と、上記ボーカル情報について音声認識を行って第1の言語文字情報を生成する音声認識手段と、上記第1の言語文字情報について翻訳処理を行って第2の言語文字情報を生成する翻訳手段と、上記第2の言語文字情報を利用して翻訳言語により発音される翻訳ボーカル情報を生成し、この翻訳ボーカル情報と上記伴奏情報を合成することにより、第2のオーディオ情報を生成する情報合成手段とを備えることとした。40

#### 【0006】

上記した構成によれば、情報配信システムにおいて例えばボーカル入りの楽曲情報について情報処理を施して得られる派生情報をとして、カラオケの楽曲情報、ボーカルの歌詞情報（音声認識処理により得られる一次言語文字情報）、他の言語に翻訳されたボーカルの歌詞情報（元の歌詞情報に対して行った翻訳処理により得られる二次言語文字情報）、及び音声合成処理により生成した翻訳言語で歌うボーカルによる楽曲情報（合成楽曲情報）の各々が生成され、これら各情報を携帯端末装置にダウンロードすることが可能となる。

#### 【0007】

##### 【発明の実施の形態】

以下、本発明の実施の形態について図1～図10を参照して説明する。

なお、以降の説明は次の順序により行うこととする。

<1. 情報配信システムの構成例>

(1-a. 情報配信システムの概要)

(1-b. 情報配信システムを構成する各装置の構成)

(1-c. ポーカル分離部の構成例)

(1-d. 音声認識翻訳部の構成例)

(1-e. 音声合成部の構成例)

(1-f. 基本的なダウンロード動作及びダウンロード情報の利用例)

<2. 派生情報のダウンロード>

【0008】

<1. 情報配信システムの構成例>

(1-a. 情報配信システムの概要)

図1は、本発明の実施の形態としての情報配信システムの構成を概略的に示している。

この図において、サーバ装置1は、後述するようにして配信用データ（例えば、オーディオ情報、テキスト情報、画像情報、映像情報等）をはじめとする所要の情報が格納される大容量の記録媒体を備えており、少なくとも通信網4を介して多数の中間伝送装置2と相互通信可能に構成されている。例えば、サーバ装置1は上記通信網4を介して中間伝送装置2から送信されてくる要求情報を受信し、この要求情報が指定する情報を記録媒体に格納されている情報から検索する。

【0009】

なお、上記のような要求情報は、例えば後述する携帯端末装置3のユーザが、携帯端末装置3又は中間伝送装置2に対して所望の情報をリクエストするための操作を行うことによって発生させることができるものとされている。そして、検索して得られた情報を通信網4を介して中間伝送装置2に対して送信する。

【0010】

また、本実施の形態では、後述するようにしてサーバ装置1から中間伝送装置2を介してアップロードした情報を携帯端末装置3によりコピー（ダウンロード）したり、中間伝送装置2を利用して携帯端末装置3に対して充電を行うのにあたり、ユーザに対して課金が行われるのであるが、この課金処理に従ってユーザから料金を徴収するために課金通信網5が設けられる。この課金通信網5は、例えば各ユーザが当該情報配信システムの利用料金を支払うために契約した金融機関などと接続される。

【0011】

中間伝送装置2は、例えば図のような形態により携帯端末装置3が装着可能とされ、主として、サーバ装置1より送信されてきた情報を通信制御端子201にて受信し、この受信情報を携帯端末装置3に対して出力する機能を有する。また、本実施の形態の中間伝送装置2には、携帯端末装置3に対して充電を行うための充電回路が備えられる。

【0012】

本実施の形態の携帯端末装置3は、中間伝送装置2に対して装着（接続）されることで、中間伝送装置2との相互通信、及び中間伝送装置2からの電力供給が可能なようになっている。そして、携帯端末装置3は、上記のようにして中間伝送装置2から出力された情報を携帯端末装置内に内蔵された所定種類の記録媒体に対して格納するようにされる。また、必要があれば携帯端末装置3に内蔵の充電池に対して中間伝送装置2から充電を行うことも可能とされる。

【0013】

このように、本実施の形態の情報配信システムは、サーバ装置1に格納されている大量の情報の中から、携帯端末装置3のユーザがリクエストした情報を携帯端末装置3の記録媒体にコピーすることができるといいわゆるデータ・オン・デマンドを実現するシステムとされる。

【0014】

10

20

30

40

50

なお、上記通信網 4 としては特に限定されるものではなく、例えば I S D N (Integrated services digital network) 、 C A T V (Cable Television, Community Antenna Television) 、通信衛星、電話回線、ワイヤレス通信等を利用することが考えられる。

また、通信網 4 としてはオン・デマンドを行うために双方向通信が必要であるが、例えば既存の通信衛星等を採用した場合には一方向のみの通信となるため、このような場合には、他方向には他の通信網 4 を用いるという 2 種類以上の通信網を併用してもかまわない。また、サーバ装置 1 から中間伝送装置 2 へ通信網 4 を介して直接情報を送信するためにはサーバ装置 1 から全ての中間伝送装置 2 への回線の接続等のインフラに費用がかかるばかりでなく、要求情報がサーバ装置 1 に一極集中し、それに応じて各々の中間伝送装置にデータを送信するためサーバ装置 1 に負荷がかかる可能性がある。そこでサーバ装置 1 と中間伝送装置 2 の間にデータを一時記憶する代理サーバ 6 を設けるようにして回線長の節約を図ると共に、代理サーバ 6 に予め所定のデータをダウンロードしておき、代理サーバ 6 と中間伝送装置 2 とのデータ交信のみで要求情報に応じた情報をダウンロードできるようしてもよい。10

#### 【 0 0 1 5 】

次に、図 2 の斜視図を参照して中間伝送装置 2 、及びこの中間伝送装置 2 に対して接続される携帯端末装置 3 についてより詳細に説明する。なお、この図において図 1 と同一部分には同一符号を付している。

#### 【 0 0 1 6 】

中間伝送装置 2 は、例えば各駅にある売店、コンビニエンスストア、公衆電話、各家庭等に配され、この場合には、本体の前面部において、その動作に応じて適宜所要の内容について表示を行う表示部 202 と、例えば所望の情報の選択その他の所要の操作を行うためのキー操作部 203 等が設けられている。20

また、本体上面部に設けられた通信制御端子 201 は、図 1 でも説明したように、サーバ装置 1 と通信網 4 を介してサーバ装置と相互通信を行うための制御端子として設けられる。。

#### 【 0 0 1 7 】

中間伝送装置 2 には、携帯端末装置 3 を装着するための端末装着部 204 が設けられている。例えばこの端末装着部 204 においては、情報入出力端子 205 と電源供給端子 206 が設けられている。端末装着部 204 に対して携帯端末装置 3 が装着された状態では、情報入出力端子 205 は携帯端末装置 3 の情報入出力端子 306 と接続され、電源供給端子 206 は携帯端末装置 3 の電源入力端子 307 と接続されるようになっている。30

#### 【 0 0 1 8 】

携帯端末装置 3 においては、例えば本体の前面部に表示部 301 、及びキー操作部 302 が設けられている。表示部 301 は、例えばユーザがキー操作部 302 に対して行った操作や動作に応じた所要の表示が行われる。また、この場合のキー操作部 302 としては、リクエストする情報を選択するためのセレクトキー 303 と、選択したリクエスト情報を確定するための決定キー 304 、及び動作キー 305 等が設けられる。本実施の形態の携帯端末装置 3 は、内部の記録媒体に格納された情報について再生を行うことが可能とされているが、上記動作キー 305 はこのような情報について再生操作を行うために設けられる。40

#### 【 0 0 1 9 】

また、携帯端末装置 3 の底面部には、情報入出力端子 306 及び電源入力端子 307 が備えられている。前述のように携帯端末装置 3 が中間伝送装置 2 に対して装着されることで、情報入出力端子 306 及び電源入力端子 307 は、それぞれ中間伝送装置 2 の情報入出力端子 205 及び電源供給端子 206 と接続される。これにより、携帯端末装置 3 と中間伝送装置 2 との情報の入出力が可能とされると共に、中間伝送装置 2 内の電源回路を利用した携帯端末装置 3 への電源供給（及び充電）が可能とされる。

また、携帯端末装置 3 の上面部にはオーディオ出力端子 309 及びマイク端子 310 が設けられると共に、側面部には外部のディスプレイ装置、キーボード、モデム、又はターミ50

ナルアダプタ等を接続可能なコネクタ308が設けられているが、これについては後述する。

#### 【0020】

なお、中間伝送装置2に設けられている表示部202及びキー操作部203は省略して中間伝送装置2が担当する機能を削減し、代わって、携帯端末装置3の表示部301及びキー操作部302により同様の操作が行えるようにしてもかまわない。

また、図2(及び図1)においては携帯端末装置3の本体部が中間伝送装置2に対して脱着可能な構成を探っているが、少なくとも中間伝送装置2側との情報入出力、電源入力が可能であればよいため、携帯端末1の底面、側面、或いは先端部等の所要の位置から小型装着部を有する電源供給線及び情報入出力線が伸長され、小型装着部を中間伝送装置に装着されるものであってもよい。10

また、一つの中間伝送装置2に対して複数のユーザが各々の携帯端末装置3を有してアクセスを行う可能性が考えられるので、一つの中間伝送装置2に複数の携帯端末装置3が装着あるいは接続可能なように構成することも考えられる。

#### 【0021】

##### (1 - b . 情報配信システムを構成する各装置の構成)

次に、図3のブロック図を参照して、本実施の形態の情報配信システムを形成する各装置(サーバ装置1、中間伝送装置2、及び携帯端末装置3)の内部構成について説明する。なお、図1及び図2と同一部分には同一符号を付している。

#### 【0022】

先ず、サーバ装置1から説明する。

図3に示すサーバ装置1は、制御部101、記憶部102、検索部103、照合処理部104、課金処理部105、インターフェイス部106を備えて構成されており、これら各機能回路部はバスラインB1を介してデータの送受信が可能なように接続されている。

制御部101は、例えばマイクロコンピュータ等を備えて構成され、通信網4からインターフェイス部106を介して供給された各種情報に応答して、サーバ装置1における各機能回路部に対する制御を実行する。

#### 【0023】

インターフェイス部106は、通信網4(この図では代理サーバ6の図示は省略している)を介して、中間伝送装置2と相互通信を行うために設けられる。なお、送信時の伝送プロトコルについては独自のプロトコルであってもよいし、又はインターネットで汎用となっているTCP/IP(Transmission control protocol/internet protocol)等でパケット化されてデータ送信されるものであってもよい。30

#### 【0024】

検索部103は、制御部101の制御によって、記憶部102に格納されているデータから所要のデータを検索する処理を実行するために設けられる。例えば、この検索処理は、例えば中間伝送装置2から送信され、通信網4からインターフェイス部106を介して制御部101に入力された要求情報に基づいて行われる。

#### 【0025】

記憶部102は、例えば大容量の記録媒体と、この記録媒体を駆動するためのドライバ装置等を備えて構成され、前述した配信用データの他、携帯端末装置3ごとに設定した端末ID、及び課金設定情報などのユーザ関連データをはじめとする所要の情報がデータベース化されて格納されている。

ここで、記憶部102に備えられる記録媒体としては、現在の放送用機器に用いられる磁気テープ等も考えられるが、本システムの特徴の一つであるオン・デマンド機能を実現するためには、ランダムアクセス可能なハードディスク、ICメモリ、光ディスク、光磁気ディスク等を採用することが好ましい。

#### 【0026】

また、記憶部102に格納されるデータは、大量な複数のデータを記録する必要があるためデジタル圧縮されていることが望ましい。圧縮方法としてはATRAC(Adaptive Tran40

10

20

30

40

50

sform Acoustic Coding)、ATRAC2、TwInVQ(Transform domain Weighted Interleave Vector Quantization)等(商標)様々な手法が考えられるが、例えば中間伝送装置側で伸張可能な圧縮手法であるならば特に限定されるものではない。

#### 【0027】

照合処理部103は、例えば要求情報等と共に送信されてきた携帯端末装置の端末IDと、本実施の形態の情報配信システムを現在利用可能な携帯端末装置の端末ID(例えば記憶部102にユーザ関連データとして格納されている)とについて照合を行い、その照合結果を制御部101に出力する。例えば制御部101ではその照合結果に基づいて、要求情報送信先の中間伝送装置2に対して接続されている携帯端末装置3に対して、当該情報配信システム利用の許可・不許可を設定するようにされる。10

#### 【0028】

また、課金処理部105は、制御部101の制御によって、携帯端末装置3を所有するユーザによる情報配信システムの利用内容に応じた金額を課金するための処理を行う。例えば、通信網4を介して中間伝送装置2からサーバ装置1に対して、情報コピーや充電のための要求情報が供給されると、制御部101では、これに応答して必要な情報の送信供給や充電許可のためのデータを送信出力するが、制御部101では、これらの情報に基づいて実際の利用状況を把握した上で、所定規則に従ってその利用内容に見合った課金金額が課金処理部105にて設定されるように制御を行う。

#### 【0029】

次に、中間伝送装置2について説明する。20

図3に示す中間伝送装置2においてはキー操作部202、表示部203、制御部207、記憶部208、インターフェイス部209、電源供給部(充電回路含む)210、装着判別部211、及びボーカル分離部212が、それぞれバスラインB2により接続されて構成されている。

#### 【0030】

制御部207は、マイクロコンピュータ等を備えて構成され、必要に応じて中間伝送装置2内部の各機能回路部の動作を制御する。

この場合、インターフェイス部209は、通信制御端子201と情報入出力端子205間に設けられており、通信網4を介したサーバ装置1との相互通信、及び携帯端末装置3との相互通信が可能とされる。つまり、このインターフェイス部209を介在するようにしてサーバ装置1と携帯端末装置3が通信可能な環境が得られることになる。30

記憶部208は、例えばメモリなどにより構成され、サーバ装置1又は携帯端末装置3から送信された所要の情報を一時保持する。この記憶部208に対する書き込み及び読み出し制御は、制御部207により実行される。

#### 【0031】

ボーカル分離部212は、例えばサーバ装置1からアップロードされた配信情報のうち、所要のボーカル入りの楽曲情報について、ボーカルパートの情報(ボーカル情報)と、ボーカルパート以外の伴奏のパートの情報(カラオケ情報)とに分離して出力可能に構成される。なお、ボーカル分離部212の内部構成例については後述するため、ここでの詳しい説明は省略する。40

#### 【0032】

電源供給部210は、例えばスイッチングコンバータ等を備えて構成され、図示しない商用交流電源を入力して所定電圧の直流電源を生成して、中間伝送装置2の各機能回路部に対して動作電源として供給する。また、この電源供給部210には、携帯端末装置3の充電池に対して充電を行うための充電回路が備えられ、電源供給端子206から携帯端末装置3の電源入力端子307を介して充電電力を供給可能に構成されている。

#### 【0033】

装着判別部211は、当該中間伝送装置2の端末装着部204に対する携帯端末装置3の装着/非装着の状態を判別する部位とされる。この装着判別部211は、例えばフォトイントラップタやメカスイッチなどの機構を備えて構成されてもよいし、例えば、電源供給端50

子 2 0 6 や情報入出力端子 2 0 5 などに含められて、中間伝送装置 2 に携帯端末装置 3 が適正に装着されることにより得られる所定端子の導通状態を検出するようにしてよい。

#### 【 0 0 3 4 】

キー操作部 2 0 2 は、例えば図 2 に示したように各種キーが設けられて構成されており、このキー操作部 2 0 2 に対して行われた操作情報はバスライン B 2 を介して制御部 2 0 7 に対して供給される。制御部 2 0 7 では供給された操作情報に応じて適宜所要の制御処理を実行する。

表示部 2 0 3 は、先に図 1 あるいは図 2 に示したようにして本体に表出するようにして設けられ、例えば液晶ディスプレイや C R T (Cathode-Ray Tube) などの表示デバイス及びその表示駆動回路等を備えて構成される。この表示部 2 0 3 の表示動作は制御部 2 0 7 により制御される。10

#### 【 0 0 3 5 】

続いて、携帯端末装置 3 について説明する。

図 3 に示す携帯端末装置 3 は、先に図 2 にて説明したようにして中間伝送装置 2 に対して装着されることにより、中間伝送装置 2 と、情報入出力端子 2 0 5 - 3 0 6 を介してデータの通信が可能なように接続されると共に、電源供給端子 2 0 6 - 電源入力端子 3 0 7 を介して、中間伝送装置 2 の電源供給部 2 1 0 から電力が供給される。

#### 【 0 0 3 6 】

この図に示す携帯端末装置 3 では、制御部 3 1 1 、 R O M 3 1 2 、 R A M 3 1 3 、信号処理回路 3 1 4 、 I / O ポート 3 1 7 , 3 1 9 、音声認識部 3 2 1 、音声合成部 3 2 2 、キー操作部 3 0 1 及びキー操作部 3 0 2 が備えられ、これら各機能回路部がバスライン B 3 により接続されている。20

この場合も、制御部 3 1 1 はマイクロコンピュータ等を備えて構成され、携帯端末装置 3 内の各機能回路部の動作についての制御を実行する。

また、 R O M 3 1 2 には、例えば制御部 3 1 1 が所要の制御処理を実行するのに必要なプログラムデータや、各種データベース等の情報が格納されているものとされる。 R O M 3 1 3 には、中間伝送装置 2 と通信すべき所要のデータや、制御部 3 1 2 の処理により発生したデータが一時保持される。

#### 【 0 0 3 7 】

I / O ポート 3 1 7 は、情報入出力端子 3 0 6 を介して中間伝送装置 2 と相互通信を行うために設けられる。当該携帯端末装置 3 から送信する要求情報や、ダウンロードされるデータは、この I / O ポート 3 1 7 を介して入出力される。30

#### 【 0 0 3 8 】

この携帯端末装置 3 に設けられる記憶部 3 2 0 は、所定の記録媒体について記録再生を行うためのドライバ等を備えて構成されるものであり、サーバ装置 1 から中間伝送装置 2 を介してダウンロードした情報を格納するために設けられる。なお、この記憶部 3 2 0 に採用される記録媒体も特に限定されるものではないが、この場合にもランダムアクセス性を考慮すれば、ハードディスク、光ディスク、 I C メモリ等のランダムアクセスが可能な記録媒体を採用することが好ましい。

#### 【 0 0 3 9 】

音声認識翻訳部 3 2 1 では、中間伝送装置 2 のボーカル分離部 2 1 2 において生成され、携帯端末装置 3 に伝送されたボーカル情報とカラオケ情報のうち、ボーカル情報を入力し、先ず、ボーカル情報について音声認識処理を行って、元のボーカルにより歌われている歌詞の文字情報（第 1 言語歌詞情報）を生成する。ここで、例えばボーカルが英語により歌っているのであれば、英語についての音声認識が行われて、第 1 言語歌詞情報としては英語の歌詞による文字情報が得られることになる。

続いて、音声認識翻訳部 3 2 1 では、上記のようにして生成した第 1 言語歌詞情報を利用して翻訳処理を行って、他の所定言語に翻訳された第 2 言語歌詞情報を生成する。例えば第 2 言語として日本語が設定されていれば、第 2 言語歌詞情報は日本語の歌詞による文字情報とされる。40

## 【0040】

音声合成部322では、先ず、上記第2言語歌詞情報に基づいて、翻訳処理後の第2言語の歌詞により歌われる新規ボーカル情報（オーディオデータ）を生成する。この際、元のボーカル情報を利用することで、オリジナルのボーカルの声質は損なわずに、第2言語に翻訳した歌詞により歌われる新規ボーカル情報を生成することができる。続いて、上記新規ボーカル情報と、このボーカル情報に対応するカラオケデータを合成することによって合成楽曲情報を生成する。

この合成楽曲情報は、同じ歌手がオリジナルの楽曲とは異なる言語で歌っている楽曲情報となる。

## 【0041】

10

このように本実施の形態の情報配信システムでは、オリジナルの楽曲データから、少なくとも、カラオケ情報（オーディオデータ）、オリジナルの言語と翻訳言語による2種類の言語による歌詞情報（文字情報データ）、及び第2言語により歌われる合成楽曲情報（オーディオデータ）を派生情報として獲得することができる。そして、これらの情報はユーザが利用するコンテンツとして管理された状態で、携帯端末装置3の記憶部320に対して、他の通常のダウンロードデータと共に格納することが可能とされている。

なお、上記音声認識翻訳部321及び音声合成部322の内部構成例については後述する。

## 【0042】

20

本実施の形態では、記憶部320に格納されたデータのうち、オーディオデータについては当該携帯端末装置3により再生出力することが可能とされている。このため、携帯端末装置3には信号処理回路314が設けられる。

信号処理回路314は、例えば記憶部320から読み出されたオーディオデータをバスラインB3を介して入力して所要の信号処理を行う。ここで、記憶部320に格納されているオーディオデータが所定形式に従って圧縮処理をはじめとする所定のエンコードが施されているのであれば、信号処理回路314では入力された圧縮オーディオデータについて伸張処理及び所定のデコード処理を施して、D/Aコンバータ315に出力する。D/Aコンバータ315でアナログオーディオ信号に変換されたオーディオデータは、オーディオ出力端子309に供給される。なお、この図ではオーディオ出力端子309にヘッドフォン8が接続された状態が示されている。

30

## 【0043】

また、携帯端末装置3にはマイク端子310が設けられている。例えば、マイク端子310にマイクロフォン12を接続して音声を吹き込んだとすると、この音声信号がA/Dコンバータ316を介してデジタルオーディオ信号に変換されて信号処理回路314に入力される。

この場合、信号処理回路314では入力されたデジタルオーディオ信号について、例えば圧縮処理及び記憶部320へのデータ書き込みに適合する所要のエンコード処理を施すように動作する。ここでエンコード処理が施されたデータは、例えば制御部311の制御によって記憶部320に対して格納することが可能とされている。あるいは、そのまま信号処理回路314の音声出力系からD/Aコンバータ315を介してオーディオ出力端子309に出力することも可能である。

40

## 【0044】

I/Oポート318は、コネクタ308を利用して外部と接続される機器や装置との入出力を可能とするために設けられる。コネクタ308には、例えばディスプレイ装置、キーボード、モ뎀、又はターミナルアダプタ等が接続可能とされるが、これについては、本実施の形態の携帯端末装置3の利用形態例として後述する。

## 【0045】

また、携帯端末装置3に備えられるバッテリ回路部319は、少なくとも充電池を備えると共に、この充電池の電力をを利用して携帯端末装置3内の各機能回路部の動作電源を供給するようにされた電源回路を備えて構成される。また、携帯端末装置3が中間伝送装置2

50

に装着された状態では、電源供給端子 206 - 電源入力端子 307 を介して、電源供給部 210 からバッテリ回路部 319 に対して、携帯端末装置 3 の回路のための動作電源及び充電電力が供給されるようになっている。

#### 【0046】

この図に示す携帯端末装置 3 の表示部 301 及びキー操作部 302 は、例えば図 2 に示したようにして本体に設けられているものであり、この携帯端末装置 3 においても、上記表示部 301 に対する表示制御は制御部 207 により実行される。また、制御部 207 は、上記キー操作部 302 から出力される操作情報に基づいて適宜所要の制御処理を実行することになる。

#### 【0047】

##### (1-c. ボーカル分離部の構成例)

図 3 の中間伝送装置 2 に備えられるボーカル分離部 212 は、例えば図 4 のブロック図のようにして構成される。

図 4において、ボーカルキャンセル部 212 は例えばデジタルフィルタ等を備えて構成され、入力されたボーカル入りの楽曲情報 D1 (オーディオデータ) からボーカルパートの成分をキャンセル (消去) して、伴奏パートだけのオーディオデータであるカラオケ情報 D2 を生成して出力する。ボーカルキャンセル部 212 の詳しい内部構成の説明は省略するが、例えばよく知られている、ステレオ音声のセンターに定位する音声を、(L チャンネルデータ) - (R チャンネルデータ) によりキャンセルする技術が用いられればよい。この際、バンドパスフィルタなどを用いてボーカル音声の帯域のみがキャンセルされて、伴奏楽器の音などは極力キャンセルされないようにすることが可能である。

#### 【0048】

ボーカルキャンセル部 212a で生成されたカラオケ情報 D2 は、ボーカル抽出部 212b 及びデータ出力部 212c に分岐して出力される。ボーカル抽出部 212b では、上記カラオケ情報 D2 及び楽曲情報 D1 を入力して、原理的に [ 楽曲情報 D1 - カラオケ情報 D2 = ボーカル情報 D3 ] の演算処理を行うことで、楽曲情報 D1 からボーカルパートのみのオーディオデータであるボーカル情報 D3 を抜き出してデータ出力部 212c に対して出力する。

#### 【0049】

データ出力部 212c では、入力されたカラオケ情報 D2 及びボーカル情報 D3 について、例えば所定規則に従って時系列的に配列して送信用データ (D2 + D3) として出力する。この送信用データ (D2 + D3) は中間伝送装置 2 から携帯端末装置 3 に対して送信出力される。

#### 【0050】

##### (1-d. 音声認識翻訳部の構成例)

図 5 は、携帯端末装置 3 に備えられる音声認識翻訳部 321 の一構成例を示すブロック図である。

音響分析部 321a は、中間伝送装置 2 から送信用データ (D2 + D3) として送信されてきたカラオケ情報 D2 とボーカル情報 D3 のうち、ボーカル情報 D3 を入力して音響分析を行い、例えば所定の帯域ごとの音声パワーや、線形予測計数 (LPC)、ケプストラム係数などの音声の特徴パラメータ抽出をする。つまり、フィルタバンク等により音声信号を所定の帯域ごとにフィルタリングし、このフィルタリング結果を整流平滑化することで、所定の帯域ごとの音声のパワーを求めるようにしている。あるいは、入力音声データ (ボーカル情報 D3) について線形予測分析処理を行うことで線形予測係数を求め、更にその線形予測係数からケプストラム係数を求めるようにされる。

上記のようにして音響分析部で求められた特徴パラメータは、直接、あるいは必要に応じてベクトル量子化されて認識処理部 321b に出力される。

#### 【0051】

認識処理部 321b は、音響分析部 13 からの特徴パラメータ (あるいは、特徴パラメータをベクトル量子化して得られるシンボル) に基づき、例えばダイナミックプログラミン

10

20

30

40

50

グ( D P )マッチング法や、隠れマルコフモデル( H M M )などの音声認識アルゴリズムにしたがい、後述する大規模の単語辞書 3 2 1 c を参照して音声認識を行い、例えばボーカル情報 D 3 としての音声に含まれる単語ごとに、音声認識結果として出力する。

#### 【 0 0 5 2 】

単語辞書 3 2 1 c には、音声認識の対象とする単語(オリジナルのボーカルの言語)の標準パターン(あるいはモデルなど)が記憶されている。認識処理部 3 2 1 b では、この単語辞書 3 2 1 c に記憶されている単語を対象として、音声認識を行う。

#### 【 0 0 5 3 】

第 1 言語文格納部 3 2 1 e は、オリジナルのボーカルの言語による文章を数多く記憶している。

10

第 2 言語文格納部 3 2 1 f は、第 1 言語文格納部 3 2 1 e に記憶されている文章を、目的とする言語に翻訳した文章を記憶している。従って、第 1 言語文格納部 3 2 1 e に記憶されている言語の文章と、第 2 言語文格納部 3 2 1 f に記憶されている他言語の文章とは、1 対 1 に対応している。

なお、例えば、第 1 言語文格納部 3 2 1 e には、日本語の文章とともに、その文章に対応する英語の文章が記憶されている第 2 言語文格納部 3 2 1 f のアドレスが記憶されている。これにより、第 1 言語文格納部 3 2 1 e に記憶されている日本語の文章に対応する英語の文章は、第 2 言語文格納部 3 2 1 f から即座に検索することができるようになされている。

#### 【 0 0 5 4 】

20

音声認識の結果により得られた 1 以上の単語列は、翻訳処理部 3 2 1 d に出力される。翻訳処理部 3 2 1 d は、認識処理部 3 2 1 b から音声認識結果としての 1 以上の単語を入力すると、その単語の組み合わせと最も類似する文章を、第 1 言語文格納部 3 2 1 e に記憶されている言語による文章(第 1 言語文)の中から検索する。

#### 【 0 0 5 5 】

上記検索処理は例えば次のようにして行われる。翻訳処理部 3 2 1 d は、音声認識の結果得られた単語(以下、認識単語ともいう)すべてを含む第 1 言語文を、第 1 言語文格納部 3 2 1 e から検索する。そのような文章が存在する場合、翻訳処理部 3 2 1 d は、その第 1 言語文を認識単語の組み合わせに最も類似するものとして、第 1 言語文格納部 3 2 1 e から読み出す。また、第 1 言語文格納部 3 2 1 e に記憶されている第 1 言語文の中に、認識単語をすべて含むものが存在しない場合、翻訳処理部 3 2 1 d は、そのうちのいずれか 1 単語を除いた単語をすべて含む第 1 言語文を検索する。そのような第 1 言語文が存在する場合、翻訳処理部 3 2 1 d は、その第 1 言語文を、認識単語の組み合わせにもっとも類似するものとして、第 1 言語文格納部 3 2 1 e から読み出す。また、そのような第 1 言語文が存在しない場合、翻訳処理部 3 2 1 d は、認識単語のうちいずれか 2 単語を除いた単語をすべて含む第 1 言語文を検索する。以下、同様にして認識単語の組み合わせに最も類似する第 1 言語文が検索される。

30

#### 【 0 0 5 6 】

上記のようにして、認識単語の組み合わせに最も類似する第 1 言語文を検索すると、翻訳処理部 3 2 1 d では、この第 1 言語文の文字情報を連結することによって第 1 言語歌詞情報として出力する。この第 1 言語歌詞情報は、派生情報の 1 コンテンツとして記憶部 3 2 0 に格納される。

40

また、翻訳処理部 3 2 1 d は、上記検索により得られた第 1 言語文を利用して、この第 1 言語文に対応する第 2 言語を第 2 言語文格納部 3 2 1 f から検索して対応付けを行う。そして、例えば認識言語単位でこの対応付け処理により得られた第 2 言語文を所定規則に従って連結していくことで、第 1 言語から第 2 言語に翻訳された歌詞の文字情報が得られる。翻訳処理部 3 2 1 d では、これを第 2 言語歌詞情報として出力する。この第 2 言語歌詞情報は、第 1 言語歌詞情報と同様に派生情報の 1 コンテンツとして記憶部 3 2 0 に格納されるとともに、次に説明する音声合成処理部 3 2 2 に入力される。

#### 【 0 0 5 7 】

50

## (1 - e. 音声合成部の構成例)

続いて、図6のブロック図は、携帯端末装置3に備えられる音声合成部322の構成例を示している。

音声分析部322aにおいては、入力されるボーカル情報D3について所要の解析処理(波形分析処理等)を実行することで、ボーカルの声質を特徴づける所定のパラメータ(声質情報)を発生させると共に、時間軸に沿ったボーカルのピッチ情報(即ちボーカルパートのメロディー情報)を生成し、これらの情報をボーカル生成処理部322bに出力する。

音声発生部322dでは、入力された第2言語歌詞情報に基づいて、第2言語による音声合成処理を行い、この合成処理により得られた音声信号データ(第2言語による歌詞を発音した音声信号)をボーカル生成処理部322bに出力する。10

## 【0058】

ボーカル生成処理部322bでは、例えば音声分析部322aから入力された声質情報に基づいて波形変形処理等を行うことによって、先ず、音声発生部322dから入力した音声信号データの声質を、ボーカル情報D3のボーカルと同等の声質となるように処理を行う。つまり、ボーカル情報D3のボーカルの声質を有しながら第2言語により歌詞を発音する音声信号データ(第2言語発音データ)を生成する。

続いて、ボーカル生成処理部322bは、上記第2言語発音データに対して、音声分析部322aから入力したピッチ情報に基づいて、音階(メロディー)を与えていく処理を行う。この処理に際しては、例えば、メロディーの区切りと歌詞との区切りの一一致が図られるように、音声発生部322dから出力される音声信号データと、ピッチ情報とに対して、これより以前のある処理段階においてタイムコードを付加するようにすることが考えられる。つまり、このタイムコードに従って、第2言語発音データを適宜区切っていきながら、ピッチ情報に基づく音階を与えていくことになる。20

このようにして生成された音声信号データは、オリジナルの楽曲の歌手と同一の声質及び同一のメロディーでもって、翻訳後の第2言語の歌詞により歌われているボーカル情報となる。このボーカル情報が、新規ボーカル情報D4として合成部322cに入力される。

## 【0059】

合成部322cでは、入力されたカラオケ情報と上記新規ボーカル情報D4を合成することによって合成楽曲情報D5を生成して出力する。合成楽曲情報D5は、聴感上では、オリジナルの楽曲に対して翻訳後の第2言語の歌詞により歌われている点が異なり、伴奏のパートやボーカルパートの歌手の声質はオリジナルの楽曲と同様とされる。30

## 【0060】

## (1 - f. 基本的なダウンロード動作及びダウンロード情報の利用例)

先ず、上記のようにして構成される本実施の形態の情報配信システムにおける携帯端末装置3に対するデータのダウンロードの基本的な動作について、再度図1～図3を参照して説明する。

## 【0061】

本実施の形態の場合、ユーザが所有する携帯端末装置3に対して所望の情報(例えば楽曲のオーディオデータであれば楽曲単位のデータをいうことになる)をダウンロードするのにあたり、このダウンロードすべき情報をユーザが選択する事が必要とされるが、ダウンロード情報について選択設定を行う方法としては、次のような方法が考えられる。40

## 【0062】

第1は、携帯端末装置3に備えられたキー操作部302の所定のキー(図1、図2参照)をユーザが操作して行う方法である。この場合には、例えば携帯端末装置3内の記憶部320に対して、当該情報配信システムによりダウンロード可能な情報がデータベース化されたメニュー情報が格納されているものとされる。このようなメニュー情報は、例えば以前に当該情報配信システムを利用して何らかの情報をダウンロードしたときに共に得られるようにされればよい。

携帯端末装置3のユーザは、例えば上記メニュー情報に基づいて得られる情報選択用のメ50

ニュー画面を表示部 301 に対して表示させ、この表示内容を見ながらセレクトキー 303 を操作して所望の情報を選択し、決定キー 304 により選択した情報を確定するようになる。

なお、上記セレクトキー及び決定キーとしてジョグダイヤルを用い、ジョグの回転を選択操作とし、ジョグの押圧により決定を行うという操作形態を採れば、情報選択時の操作体系をより簡単にすることができます。

そして、上記のような選択設定操作が携帯端末装置 3 を中間伝送装置 2 に対して装着している状態で行われているのであれば、選択設定操作に応じた要求情報が中間伝送装置 2 (インターフェイス部 209) から通信網 4 を介してサーバ装置 1 に供給されることになる。  
10

#### 【0063】

また、上記のような選択設定操作により得られた設定情報が、携帯端末装置 3 内の RAM 313 (図 3 参照) に対して保持されるように構成すれば、携帯端末装置 3 を中間伝送装置 2 に装着しない状態 (即ち、身近に中間伝送装置 2 が無いような環境) のもとでも、ユーザは、予め任意の機会で情報を選択する操作を行って、この操作により発生した要求情報を携帯端末装置 3 に保持させておくことが可能になる。

この場合には、例えばユーザが携帯端末装置 3 を中間伝送装置 2 に装着したときに、RAM 313 に保持されているダウンロード情報に関する設定情報が、要求情報として中間伝送装置 2 (インターフェイス部 209) から通信網 4 を介してサーバ装置 1 に伝送されることになる。  
20

#### 【0064】

また、これまでの説明は、携帯端末装置 3 に備えられるキー操作部 302 により情報の選択設定操作を行うものであったが、中間伝送装置 2 に対してキー操作部 202 が備えられているのであれば、例えば携帯端末装置 3 が中間伝送装置 2 に装着された状態で、中間伝送装置 2 のキー操作部 202 により同様の操作が可能なように構成してもかまわない。

#### 【0065】

上記した何れの方法により選択設定操作を行ったとしても、携帯端末装置 3 を中間伝送装置 2 に対して装着することにより、選択設定操作に応じた要求情報が携帯端末装置 3 にて発生され、この要求情報が中間伝送装置 2 を介してサーバ装置 1 に対してアップロードされることになる。なお、このアップロード動作は、中間伝送装置 2 の装着判別部 211 における検出情報を開始トリガとするようにしてもよい。また、上記要求情報をサーバ装置に対して送信するときには、これとともに携帯端末装置 3 が保持している端末 ID の情報をも送信するようになる。  
30

#### 【0066】

そして、このようなデータ送信が終了したことが確認されると、サーバ装置 1 では、先ず、照合処理部 104 において要求情報と共に送信された端末 ID について照合を行う。ここで、照合結果として端末 ID が当該情報配信システムを利用可能であることが判別されれば、記憶部 102 に格納されている情報のうちから、送信された要求情報に対応する情報を検索する処理を実行する。

この検索処理は、制御部 101 が検索部 103 を制御することにより、例えば、要求情報に含まれる識別コードと、記憶部 102 に格納されている情報ごとに与えられた識別コードとを照合していくことにより実行されればよい。このようにして、要求情報に対応する情報が検索されることにより、サーバ装置 1 において配信すべき情報の決定が行われることになる。  
40

#### 【0067】

なお、上述の端末 ID の照合処理時において、端末 ID が未登録であったり、残金が足りない等の理由で、送信された端末 ID が情報配信システムを現在利用不可であるとの判断結果が得られたときには、この内容を示すエラー情報を中間伝送装置 2 に送信するようにしてもよい。これにより、中間伝送装置 2 、あるいは携帯端末装置 3 に備えられる表示部 (203、301) においてその警告を表示したり、あるいはスピーカなどの音声出力手  
50

段を設けて、警告音を出力させるような構成をとることが可能になる。

**【0068】**

サーバ装置1では、上述のように要求情報に応じて記憶部102から検索した情報を中間伝送装置2に対して送信する。中間伝送装置2に装着された携帯端末装置3は、中間伝送装置2にて受信した情報を、情報入出力端子205-306を介して取り込んで内部の記憶部320にコピー（ダウンロード）する。

**【0069】**

また、本実施の形態では、携帯端末装置3に情報のダウンロードが行われている間に、中間伝送装置2から携帯端末装置3の充電池に対して自動的に充電が行われるものとされる。  
10

また、例えば携帯端末装置3のユーザの要望として、情報のダウンロードは必要ないが、中間伝送装置2を充電だけのために利用したいというようなことも当然考えられるので、所定の操作を行うことで、中間伝送装置2に対して充電のみを行うことができるようにもされている。

**【0070】**

例えば、上述のようにして、携帯端末装置3に対して情報のダウンロードが終了すると、中間伝送装置2の表示部202あるいは携帯端末装置3の表示部202等に対して、情報のダウンロードの終了が完了したことを告げるメッセージ等が表示される。

そして、携帯端末装置3のユーザがこの表示を確認して、携帯端末装置1を中間伝送装置2から外した後は、携帯端末装置3はダウンロードにより記憶部306に格納したデータを再生するための再生装置として機能する。つまり、ユーザは、携帯端末装置3さえ所持していれば、特に場所や時間を問わず携帯端末装置3に格納した情報を再生して表示したり、あるいは音声として出力させることができる。この際、ユーザは携帯端末装置3に備えられている動作キー305により、その再生動作を任意に操作することが可能とされている。この動作キー305としては、例えば早送り、再生、巻戻し、停止、一時停止キーなどが備えられているものとされる。  
20

**【0071】**

例えば、オーディオデータを再生して視聴したい場合には、図7に示すように携帯端末装置3のオーディオ出力端子308にヘッドフォン8或いはアクティブスピーカSP等を接続することにより、オーディオデータの再生音声を視聴することが可能となる。  
30

**【0072】**

また、例えば図8に示すように、マイクロフォン端子309に対してマイクロフォン12を接続することにより、このマイクロフォン12から入力した音声をA/Dコンバータ316信号処理回路314を介することによりデータ化して、記憶部320に対して格納する、つまりマイク音声を録音することが可能とされる。この場合には、前述した動作キー305として録音キー等が設けられることになる。

さらには、例えばオーディオデータとしてカラオケを再生出力しているのであれば、マイクロフォン端子309に接続したマイクロフォン12により、カラオケに合わせてユーザが歌うことなどもできる。

**【0073】**

また、本実施の形態の携帯端末装置3は、図8に示すように本体に備えられたコネクタ308に対してモニタ装置9、モデム10（又はターミナルアダプタ）を接続可能なコネクタ308、キーボード11を接続可能とされている。  
40

例えば、携帯端末装置3自体によっても、表示部301によりダウンロードした画像データ等を表示出力することは可能であるが、コネクタ308に対してモニタ装置9を接続して、携帯端末装置3から画像データを出力すれば、より大きな画面によって画像を見ることが可能である。また、キーボード22を接続して文字入力等を可能とすることにより、要求する情報の選択を容易にするだけでなく、より複雑なコマンド入力が可能となる。

また、モデム（ターミナルアダプタ）10を接続すれば、中間伝送装置2を利用することなく、サーバ装置1と直接データの送受を可能とすることができます。また、ROM312  
50

に保持させるプログラム等によっては、通信網 4 を介して他のコンピュータ或いは携帯端末装置 3 と通信可能に構成することが可能であり、これにより、ユーザ同士のデータ交換なども容易に行うことができる。また、これらの代わりに無線接続コントローラを用いれば、例えば中間伝送装置 2 と携帯端末装置 3 とを無線接続することも容易に可能となる。

#### 【0074】

##### <2. 派生情報のダウンロード>

これまで説明してきた、本実施の形態の情報配信システムの構成、携帯端末装置に対する情報のダウンロードの基本動作、及び利用形態例を前提として、本実施の形態の特徴となる、派生情報のダウンロードについて、図 9 及び図 10 を参照して説明する。図 9 は、派生情報をダウンロードする際の中間伝送装置 2 及び携帯端末装置 3 の動作の経緯を時間軸に従って示しており、図 10 は、派生情報のダウンロードの経過に従って、例えば携帯端末装置 3 の表示部 301 に表示される表示内容を示している。

10

#### 【0075】

また、ここでいう「派生情報」とは、これまでの説明からわかるように、ボーカル入りのオリジナル楽曲情報から得られる、カラオケ情報、第 1 言語歌詞情報、第 2 言語歌詞情報、及び同じ歌手が第 2 言語により歌う合成楽曲情報とされる。

なお、派生情報のダウンロードに伴う情報配信システムを構成する各装置（サーバ装置 1、中間伝送装置 2、及び携帯端末装置 3）の動作の詳細であるが、ダウンロード時の基本的な動作は図 3 により説明し、派生情報生成のための動作は、図 4、図 5 及び図 6 により既に説明したことから、以降において、システムの動作についての詳しい説明は若干の補足を除いて省略し、主として、時間経過に従った動作の状態遷移について説明を行っていくこととする。

20

#### 【0076】

図 9 には、派生情報のダウンロードに際しての中間伝送装置 2 及び携帯端末装置 3 の動作例が示されている。ここで、図の 内の英数字は、中間伝送装置 2 及び携帯端末装置 3 の時間経過に従った動作順を示しており、以降の説明はこの動作順に従って行うこととする。

#### 【0077】

動作 1： ここでは、先に利用形態として説明した操作方法として、携帯端末装置 3 のキー操作部 302 を操作することにより、ユーザが所望する「楽曲情報の派生情報」を要求するための選択設定操作が行われるものとされる。なお、利用形態として前述したようく、中間伝送装置 2 に設けられたキー操作部 203 により同様の選択設定操作が行われるようになされてもかまわない。

30

#### 【0078】

動作 2： 携帯端末装置 3 は、上記動作 1 として得られた操作情報に従った要求情報、つまり、指定の楽曲情報の派生情報を要求することを示す要求情報を送信出力する。

#### 【0079】

動作 3： 携帯端末装置 3 から要求情報が送信出力され場合、これまでの説明からわかるように、この要求情報を中間伝送装置 2 にて受信し、さらに中間伝送装置 2 から通信網 4 を介してサーバ装置 1 に対して送信する。

40

図 9 には示していないが、サーバ装置 1 では、受信入力した要求情報に対応する楽曲情報を記憶部 102 から検索し、検索した楽曲情報を記憶部 102 から読み出して中間伝送装置 2 に対して送信する。なお、要求情報が派生情報とされる場合であっても、サーバ装置 1 から配信される楽曲情報はオリジナルの楽曲情報であり、この段階では派生情報は発生していない。図 9 では、ここまで段階を動作 3 とする。

#### 【0080】

動作 4： 中間伝送装置 2 では、サーバ装置 1 から送信されてきた楽曲情報を受信して、例えば一旦、記憶部 208 に格納して保持する。即ち、楽曲情報のダウンロードを行う。

動作 5： 中間伝送装置 2 では、上記動作 4 として記憶部 208 に格納した楽曲情報を読み出してボーカル分離部 212 に入力する。ボーカル分離部 212 では、図 4 にて説明し

50

たようにして、上記楽曲情報 D 1 についてカラオケ情報 D 2 とボーカル情報 D 3 に分離する。

動作 6： 上記ボーカル分離部 2 1 2 では、例えば、図 4 により説明したように、最終段のデータ出力部 2 1 2 cにおいて、カラオケ情報 D 2 とボーカル情報 D 3 を送信情報（D 2 + D 3）として出力するようにされる。そして、動作 6 として、中間伝送装置 2 は送信情報（D 2 + D 3）を、携帯端末装置 3 に対して送信する処理を行う。

#### 【 0 0 8 1 】

このように本実施の形態において、中間伝送装置 2 により派生情報を得るための動作としては、ボーカル分離部 2 1 2 での信号処理によってカラオケ情報 D 2 とボーカル情報 D 3 を生成する処理のみを行うようにされる。つまり、以降において生成される各種派生情報は、受信入力したカラオケ情報 D 2 とボーカル情報 D 3（送信情報（D 2 + D 3））に基づいて、全て携帯端末装置 3 側において生成するようにされる。10

即ち、本実施の形態では、ユーザにとってのコンテンツとなる各種派生情報を得るためにあたり、中間伝送装置 2 と携帯端末装置 3 間でその役割が分担されるように構成されるものである。これにより、例えば各種派生情報を得るために中間伝送装置 2 あるいは携帯端末装置 3 の何れかにおいてのみ、その役割を与えるように構成した場合と比較して、中間伝送装置 2 と携帯端末装置 3 間の処理負担を軽減することが可能となる。

#### 【 0 0 8 2 】

動作 7： 携帯端末装置 3 は、上記動作 6 により中間伝送装置 2 から送信された送信情報（D 2 + D 3）を受信入力することになる。20

動作 8： そして、携帯端末装置 3 においては、受信入力した送信情報（D 2 + D 3）から、カラオケ情報 D 2 とボーカル情報 D 3 をそれぞれ独立に得て、先ず、カラオケ情報 D 2 については、記憶部 3 2 0 に対して格納する。

これにより、携帯端末装置 3 にとっては、派生情報のコンテンツとして最初にカラオケ情報 D 2 を獲得したことになるため、携帯端末装置 3 では、続いて図 10 ( a ) に示すように表示部 3 0 1 に対してカラオケボタン B 1 を表示させる。このようなボタン表示は、携帯端末装置 3 において新しい派生情報が得られるごとに逐次表示されるものであり、派生情報のダウンロードの経過をユーザに示すものである。

また、各ボタン表示はユーザが所望のコンテンツを選択して再生するための操作用のインターフェイス画像として利用される。これは、後述する図 10 ( b ) ~ 図 10 ( d ) に追加表示される各ボタン表示についても同様である。30

また、ボーカル情報 D 3 は、音声認識翻訳部 3 2 1 に入力される。

#### 【 0 0 8 3 】

動作 9： 音声認識翻訳部 3 2 1 は、先ず、入力されたボーカル情報 D 3 について図 5 にて説明したようにして音声認識を行うことで、派生情報として第 1 言語歌詞情報（文字情報）を生成する。ここでは、第 1 言語、つまり楽曲情報のボーカル言語として例えば英語が規定されているものとする。従って、ここで生成される第 1 言語歌詞情報としては、英語歌詞情報となる。

音声認識翻訳部 3 2 1 で生成された英語歌詞情報は、記憶部 3 2 0 に対して格納される。これにより、携帯端末装置 3 では 2 番目の派生情報を獲得したことになるため、図 10 ( b ) に示すように、表示部 3 0 1 に対してカラオケボタン B 1 に追加して英語歌詞がコンテンツ化されたことを示す英語歌詞ボタン B 2 の表示を行うようされる。40

#### 【 0 0 8 4 】

動作 10： 音声認識翻訳部 3 2 1 では、動作 9 により生成した第 1 言語歌詞情報（英語歌詞情報）について翻訳を行って第 2 言語歌詞情報を生成する。ここでは、第 2 言語として日本語が設定されているものとする。このため、実際に作成される第 2 言語歌詞情報としては、英語による歌詞を日本語に翻訳した歌詞情報（日本語歌詞情報）となる。

そして、携帯端末装置 3 ではこの日本語歌詞情報を 3 番目に獲得すべき派生情報を記憶部 3 2 0 に格納する。そして、図 10 ( c ) に示すように表示部 3 0 1 に対して日本語歌詞がコンテンツ化されたことを示す日本語歌詞ボタン B 3 を表示させる。50

## 【0085】

動作11： 続いて携帯端末装置3では、音声合成部322による信号処理により、合成楽曲情報D5を生成する。この合成楽曲情報D5は、たとえば図6にて説明したように、カラオケ情報D2、ボーカル情報D3、及び上記動作10により生成された第2言語歌詞情報（この場合は日本語歌詞情報）を利用して生成される。ここでは、第1言語が英語、第2言語が日本語とされていることから、合成楽曲情報D5としては、英詩により歌われるオリジナルの楽曲を、同一の歌手が日本語の歌詞に訳して歌っている楽曲の情報となる。

そして、この合成楽曲情報D5を最後に獲得すべき派生情報として記憶部320に格納し、表示部301に対して図10（d）に示すように合成楽曲がコンテンツ化されたことを示す合成楽曲ボタンB4を表示させる。10

この段階では、派生情報として獲得可能とされる4種類の全てのコンテンツが表示部301にボタン表示されて、派生情報のダウンロードが全て完了したことを示すことになる（なお、別途、ダウンロードの完了を示すメッセージ等が表示されてもよい）。また、実際に、これら全ての派生情報が携帯端末装置3の記憶部320に対して格納済みの状態にある。

そして、上記のようにして携帯端末装置3にダウンロードした派生情報は、例えば、先に図7及び図8により説明したようにして外部に出力して利用することができる。

## 【0086】

なお、実際の使用形態に際しては、細部は適宜変更されてかまわない。例えば、図9による説明では、楽曲情報のダウンロードから派生情報の獲得までが時間的にほぼ連続する一連の動作として扱われていたが、例えば、携帯端末装置3の記憶部320に対して少なくとも送信情報（カラオケ情報D2+ボーカル情報D3）を格納しておき、携帯端末装置3を中間伝送装置2から外した後の任意の機会に、所定の操作によって携帯端末装置3においてカラオケ情報D2以外の残る3つの派生情報のコンテンツを作成して獲得するように構成することも考えられる。20

## 【0087】

また、図9による説明では、オリジナルの英語歌詞を日本語に翻訳して最終的に合成楽曲情報を得るものとして説明したが、特にオリジナル言語（第1言語）及び翻訳言語（第2言語）としての言語は限定されるものではない。さらには、複数言語のオリジナル言語に対応可能と共に、翻訳言語をユーザの指定操作などによって複数言語から選択指定するように構成することも可能とされる。この場合には、音声認識翻訳部321において、対応する言語種類に応じて、単語辞書321cや、第1言語格納部321e及び第2言語格納部321fに格納される言語種類数が増設されることになる。30

## 【0088】

また、図9による派生情報のダウンロード動作としては、オリジナルの楽曲情報は携帯端末装置3にて得られるコンテンツとしては除外されていたが、中間伝送装置2から携帯端末装置3にカラオケ情報D2とボーカル情報D3による送信情報（D2+D3）を送信する際に、共にオリジナルの楽曲情報D1を送信し、携帯端末装置3の記憶部320に対して格納するように構成することも考えられる。40

## 【0089】

更に、図9による説明では、楽曲に関する派生情報を要求すると自動的に4種類の全ての派生情報が獲得されるものとして説明したが、例えばユーザの選択設定操作に従って、4種類の派生情報のコンテンツのうちから一部のコンテンツのみを得るようにすることも可能である。

さらには、例えば4種類の全ての派生情報のうち、所定の一部の派生情報のみを提供可能な簡易な構成による情報配信システムを構築することも可能であり、例えば、派生情報としてカラオケ情報のみを提供するのであれば、ボーカル分離部212におけるボーカルキヤンセル部212cに相当する機能回路部が、情報配信システムを構成する装置の何れか1つに設けられるように構成すればよいことになる。50

**【0090】**

また、本実施の形態では、派生情報を生成するための機能回路部として、ボーカル分離部212のみを中間伝送装置2に設け、残る音声認識翻訳部321及び音声合成部322は携帯端末装置3に設けるようにしているが、これに限定されるものではなく、これら各機能回路部を当該情報配信システムを構成する各装置（サーバ装置1、中間伝送装置2、携帯端末装置3）に対してどのように振り分けて設けるのかについては、実際の適用条件等により変更されてかまわない。

**【0091】****【発明の効果】**

以上説明したように本発明は、情報配信システムにおいて、サーバ装置から配信したオリジナルの楽曲情報をを利用して、その楽曲のカラオケ情報、オリジナルの言語によるボーカルの歌詞情報、他の言語に翻訳されたボーカルの歌詞情報、及び翻訳言語の歌詞によりオリジナルと同一のボーカルにより歌われる合成楽曲情報の各々が生成され、これら各情報を携帯端末装置においてダウンロード情報として獲得することが可能となる。これにより、オリジナルの楽曲情報だけでなく、これをを利用して生成した派生情報を携帯端末装置のコンテンツとすることができますため、情報配信システムとしての利用価値がより高まることになる。

この際、派生情報を生成するための各種機能回路部を、情報配信システムを構成する各装置に適宜振り分けるようにして設けることで、ある1つの装置における動作負担が重くなるのを避けることができる。

**【0092】**

更に、派生情報を獲得するためのダウンロードを行っている際に、順次獲得されていく派生情報の種類に対応する表示を行うことで、たとえばユーザは派生情報のダウンロードの動作の経過を把握することができるとともに、この表示を、各派生情報を呼び出して再生するための操作用インターフェイスとして機能させることで、携帯端末装置のユーザの使い勝手が更に向上されることになる。

**【図面の簡単な説明】**

**【図1】**本発明の実施の形態としての情報配信システムの構成例を概念的に示す説明図である。

**【図2】**中間伝送装置及び携帯端末装置の外観例を示す斜視図である。

**【図3】**本実施の形態の情報配信システムを形成する各装置の内部構成を示すブロック図である。

**【図4】**ボーカル分離部の内部構成例を示すブロック図である。

**【図5】**音声認識翻訳部の内部構成例を示すブロック図である。

**【図6】**音声合成部の内部構成例を示すブロック図である。

**【図7】**携帯端末装置の利用形態例を示す斜視図である。

**【図8】**携帯端末装置の利用形態例を示す斜視図である。

**【図9】**派生情報のダウンロード動作の経緯を示す説明図である。

**【図10】**派生情報のダウンロードに伴う携帯端末装置の表示部の表示形態例を示す説明図である。

**【符号の説明】**

1 サーバ装置、2 中間伝送装置、3 携帯端末装置、4 通信網、5 課金通信網、  
6 代理サーバ、8 ヘッドフォン、9 モニタ装置、10 モデム、11 キーボード  
、12 マイクロフォン、101 制御部、102 記憶部、103 検索部、104  
照合処理部、105 課金処理部、106 インターフェイス部、B1 バスライン、2  
01 通信制御端子、202 キー操作部、203 表示部、204 端末装着部、20  
5 情報入出力端子、206 電源供給端子、207 制御部、208 記憶部、209  
インターフェイス部、210 電源供給部、211 装着判別部、212 ボーカル分  
離部、B2 バスライン、301 表示部、302 キー操作部、303 セレクトキー  
、304 決定キー、305 動作キー、306 情報入出力端子、307 電源入力端

10

20

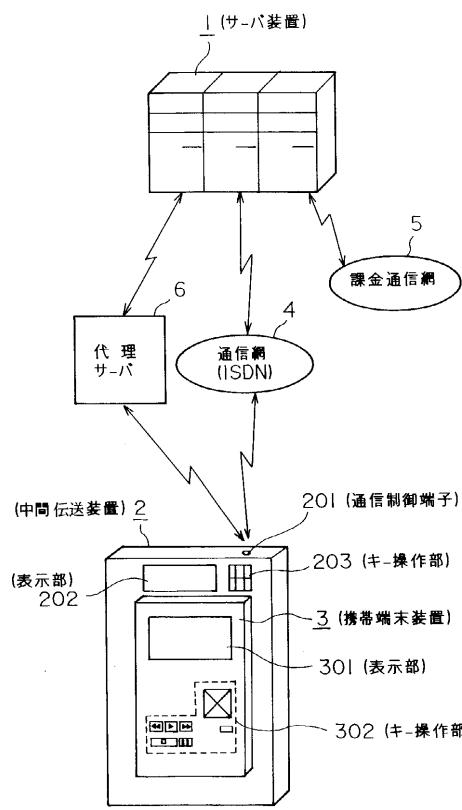
30

40

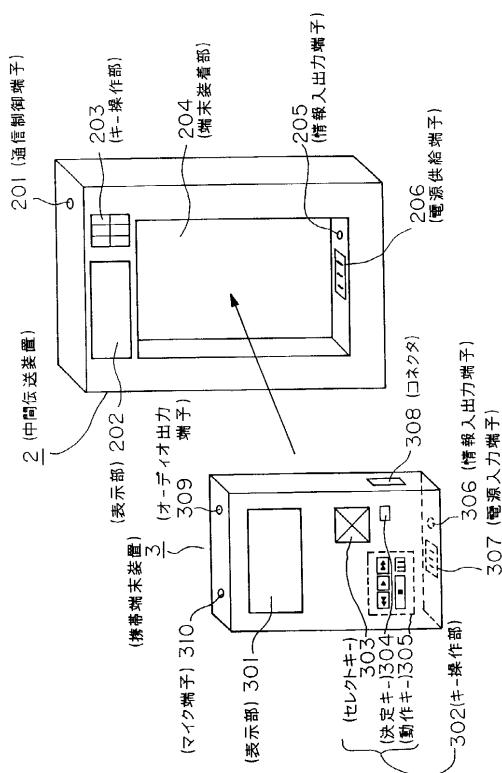
50

子、308 コネクタ、309 オーディオ出力端子、310 マイクロフォン端子、311 制御部、312 ROM、313 RAM、314 信号処理回路、315 D/Aコンバータ、316 A/Dコンバータ、317, 318 I/Oポート、319 バッテリ回路部、320 記憶部、321 音声認識翻訳部、322 音声合成部、B3 バスライン

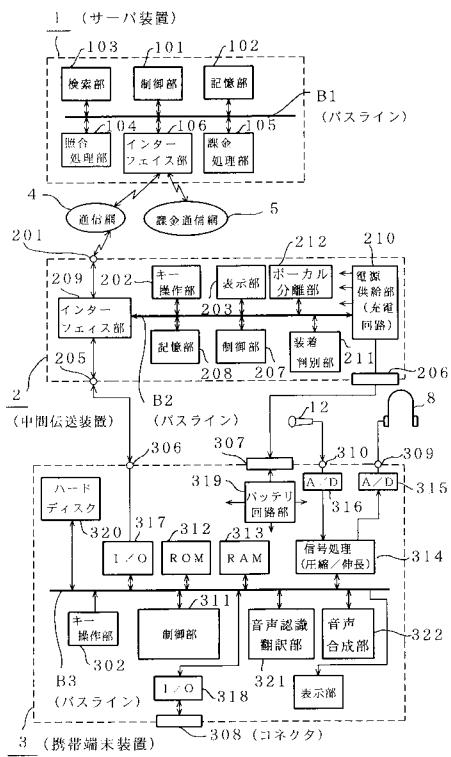
【図1】



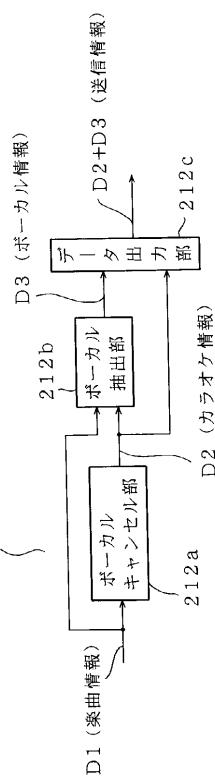
【図2】



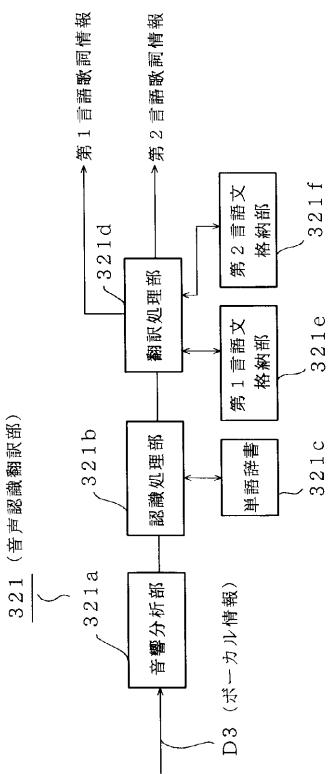
【図3】



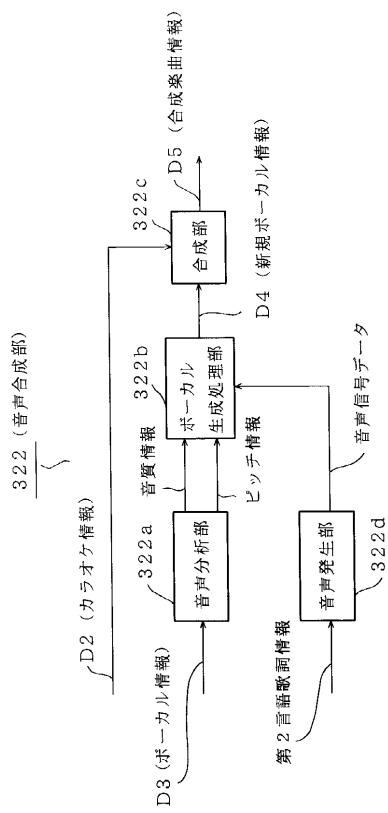
【 図 4 】



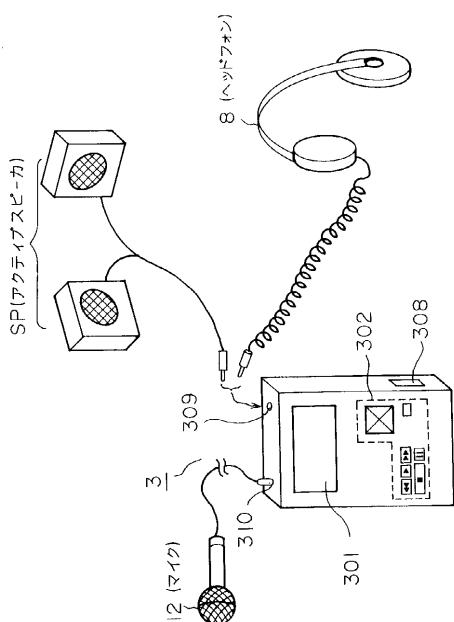
【図5】



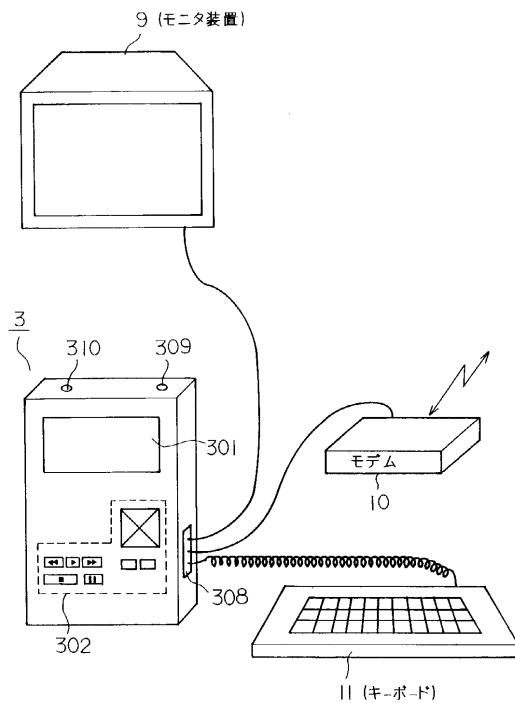
【 义 6 】



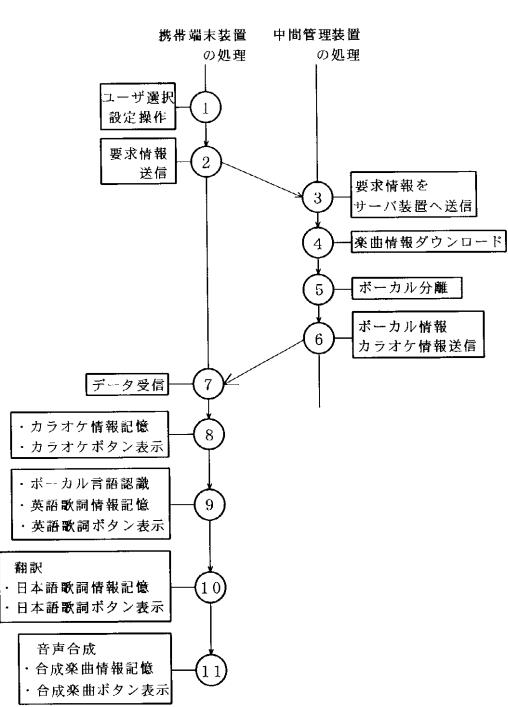
【図7】



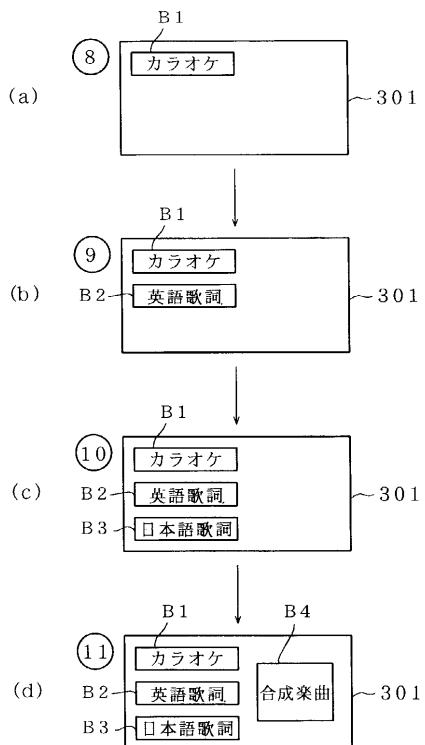
【図8】



【図9】



【図10】



---

フロントページの続き

(56)参考文献 特開平09-121325(JP,A)

特開平08-002015(JP,A)

山本,他,日本語テキスト音声合成を利用した歌唱合成システム,情報処理学会第49回(平成6年後期)全国大会講演論文集,1994年,5G-4,p.2-39~40

(58)調査した分野(Int.Cl., DB名)

G10K 15/04

G10L 13/02

G10L 15/00