

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6512607号  
(P6512607)

(45) 発行日 令和1年5月15日(2019.5.15)

(24) 登録日 平成31年4月19日(2019.4.19)

(51) Int. Cl.		F I			
<b>G 1 O L</b>	<b>19/00</b>	<b>(2013.01)</b>	<b>G 1 O L</b>	19/00	<b>4 0 0 Z</b>
<b>G 1 O K</b>	<b>15/08</b>	<b>(2006.01)</b>	<b>G 1 O L</b>	19/00	<b>3 3 0 B</b>
			<b>G 1 O K</b>	15/08	

請求項の数 8 (全 26 頁)

(21) 出願番号	特願2016-26744 (P2016-26744)	(73) 特許権者	000004226
(22) 出願日	平成28年2月16日(2016.2.16)		日本電信電話株式会社
(65) 公開番号	特開2017-146391 (P2017-146391A)		東京都千代田区大手町一丁目5番1号
(43) 公開日	平成29年8月24日(2017.8.24)	(73) 特許権者	504145342
審査請求日	平成29年11月21日(2017.11.21)		国立大学法人九州大学
			福岡県福岡市西区元岡744
		(74) 代理人	100121706
			弁理士 中尾 直樹
		(74) 代理人	100128705
			弁理士 中村 幸雄
		(74) 代理人	100147773
			弁理士 義村 宗洋
		(72) 発明者	鎌本 優
			東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 環境音合成装置、その方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

環境音分析装置から出力された伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する環境音合成装置であって、

前記環境音分析装置から出力された前記環境音量パラメタを取得するデータ受信部と、1フレーム分の環境音のテンプレート(以下、テンプレートという)と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部と、

前記環境音量パラメタで特定される音量に応じたテンプレートを前記テンプレート記憶部から選択し、前記環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、前記選択したテンプレートに残響を加え、当該残響を加えたテンプレートを合成することで環境音を生成する残響付加音源合成部とを含む、

環境音合成装置。

【請求項2】

請求項1の環境音合成装置であって、

前記残響付加音源合成部は、伝送元に存在する音源の個数に応じた空間の大きさに基づく残響を加える、

環境音合成装置。

【請求項3】

環境音分析装置から出力された伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する環境音合成装置であって、

前記環境音分析装置から出力された前記環境音量パラメタを取得するデータ受信部と、  
1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部と、

前記環境音量パラメタで特定される音量に応じたテンプレートを前記テンプレート記憶部から選択し、前記選択したテンプレートを合成して環境音を生成する残響付加音源合成部とを含む、

環境音合成装置。

【請求項4】

環境音分析装置から出力された伝送元の空間の大きさに基づく音響信号の残響に関する環境残響パラメタを取得して環境音を生成する環境音合成装置であって、

前記環境音分析装置から出力された前記環境残響パラメタを取得するデータ受信部と、  
1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の残響に対応する情報とを対応付けて記憶するテンプレート記憶部と、

前記環境残響パラメタで特定される残響に応じたテンプレートを前記テンプレート記憶部から選択し、前記選択したテンプレートを合成して環境音を生成する残響付加音源合成部とを含む、

環境音合成装置。

【請求項5】

伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する環境音合成方法であって、

データ受信部が、前記環境音量パラメタを取得するデータ受信ステップと、

残響付加音源合成部が、1フレーム分の環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部から前記環境音量パラメタで特定される音量に応じたテンプレートを選択し、前記環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、前記選択したテンプレートに残響を加え、当該残響を加えたテンプレートを合成することで環境音を生成する残響付加音源合成ステップとを含む、

環境音合成方法。

【請求項6】

伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する環境音合成方法であって、

データ受信部が、前記環境音量パラメタを取得するデータ受信ステップと、

残響付加音源合成部が、1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部から、前記環境音量パラメタで特定される音量に応じたテンプレートを選択し、前記選択したテンプレートを合成して環境音を生成する音源合成ステップとを含む、

環境音合成方法。

【請求項7】

伝送元の空間の大きさに基づく音響信号の残響に関する環境残響パラメタを取得して環境音を生成する環境音合成方法であって、

データ受信部が、前記環境残響パラメタを取得するデータ受信ステップと、

残響付加音源合成部が、1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の残響に対応する情報とを対応付けて記憶するテンプレート記憶部から、前記環境残響パラメタで特定される残響に応じたテンプレートを選択し、前記選択したテンプレートを合成して環境音を生成する音源合成ステップとを含む、

環境音合成方法。

## 【請求項 8】

請求項 1 から請求項 4 の何れかの環境音合成装置としてコンピュータを機能させるためのプログラム。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、伝送元で収録された環境音を、伝送先で再生する環境音合成装置、その方法及びプログラムに関する。

## 【背景技術】

## 【0002】

実測データをもとに算出した個人差や、速度・大きさの揺らぎの程度を利用して、単独のユーザと同期するように複数の拍手音を合成し出力する技術が提案されている（非特許文献 1）。また、ある地点の音を別の場所に伝送し再生する技術として、音響符号化技術が知られている。例えば、非特許文献 2 では、聴覚マスキングを巧みに利用し、また楽器の特性を利用して低域の成分を高域にコピーして使うという楽音の特性に合わせたモデルにより、低ビットレートで品質の高い音響符号化技術が提案されている。

## 【0003】

非特許文献 1 は、ユーザと同調する複数の人がその場にいるような環境を仮想的に実現することを目的としたものであり、ユーザの拍手のピッチに合わせて仮想的な拍手音を合成する技術であり、実在する遠隔地の場の状況（拍手音や手拍子）を、別の場所に伝送し再現することはできなかつた。また、声援・掛け声などの拍手音以外の環境音を伝送し再現することは対象としていない。また、拍手音や声援・掛け声などの環境音は純粋な音声や楽器音とは異なり白色雑音に近いため、非特許文献 2 のような従来の音響符号化技術ではうまく表現できず、音質が劣化していた。

## 【0004】

伝送元において収録された拍手や手拍子音、声援・掛け声などの環境音を効率よく伝送し、伝送先で伝送元の場の雰囲気を実現することができる環境音合成装置として、特許文献 1 が知られている。

## 【0005】

特許文献 1 の環境音合成装置では、テンプレート記憶部に 1 フレーム分（一定時間分）の環境音のテンプレートと当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶しておき、音源合成部が受信した環境音量パラメータと同じ音量大きさのテンプレートをテンプレート記憶部から選択し、選択したテンプレートを合成して環境音を生成する。

## 【先行技術文献】

## 【特許文献】

## 【0006】

【特許文献 1】特開 2014 - 63145 号公報

## 【非特許文献】

## 【0007】

【非特許文献 1】西村竜一、宮里勉、「仮想的集団による拍手音の合成」、電子情報通信学会技術研究報告、電子情報通信学会、1999年3月、MVE, マルチメディア・仮想環境基礎、98(684), p.17 24,

【非特許文献 2】Stefan Meltzer and Gerald Moser, "MPEG 4 HE AAC v2 audio coding for today's digital media world," EBU technical review, Jan., 2006.

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0008】

しかしながら、特許文献 1 では、伝送元における残響を考慮していない。そのため、生成された環境音を再生すると、一点から環境音が発せられるように聞こえてしまい、実際

10

20

30

40

50

には一点ではなく所定の空間から発せられる環境音を適切に再現することが難しい。特に、伝送元の空間が広い場合にその傾向が強くなる。

【0009】

そこで本発明では、伝送元において収録された環境音を効率よく伝送し、伝送先で伝送元の場の雰囲気を残響を考慮して再現することができる環境音合成装置、その方法及びプログラムを提供することを目的とする。

【課題を解決するための手段】

【0010】

上記の課題を解決するために、本発明の一態様によれば、環境音合成装置は、環境音分析装置から伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する。環境音合成装置は、環境音分析装置から環境音量パラメタを受信するデータ受信部と、1フレーム分の環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部と、環境音量パラメタで特定される音量に応じたテンプレートをテンプレート記憶部から選択し、選択したテンプレートに環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、残響を加えたテンプレートを合成することで環境音を生成する残響付加音源合成部とを含む。

10

【0011】

上記の課題を解決するために、本発明の他の態様によれば、環境音合成装置は、環境音分析装置から伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する。環境音合成装置は、環境音分析装置から環境音量パラメタを受信するデータ受信部と、1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部と、環境音量パラメタで特定される音量に応じたテンプレートをテンプレート記憶部から選択し、選択したテンプレートを合成して環境音を生成する残響付加音源合成部とを含む。

20

【0012】

上記の課題を解決するために、本発明の他の態様によれば、環境音合成装置は、環境音分析装置から伝送元の空間の大きさに基づく音響信号の残響に関する環境残響パラメタを取得して環境音を生成する。環境音合成装置は、環境音分析装置から環境残響パラメタを受信するデータ受信部と、1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の残響に対応する情報とを対応付けて記憶するテンプレート記憶部と、環境残響パラメタで特定される残響に応じたテンプレートをテンプレート記憶部から選択し、選択したテンプレートを合成して環境音を生成する残響付加音源合成部とを含む。

30

【0013】

上記の課題を解決するために、本発明の他の態様によれば、環境音合成方法は、伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する。環境音合成方法は、データ受信部が、環境音量パラメタを受信するデータ受信ステップと、残響付加音源合成部が、1フレーム分の環境音のテンプレート（以下、テンプレートという）と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部から環境音量パラメタで特定される音量に応じたテンプレートを選択し、選択したテンプレートに環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、残響を加えたテンプレートを合成することで環境音を生成する残響付加音源合成ステップとを含む。

40

【0014】

上記の課題を解決するために、本発明の他の態様によれば、環境音合成方法は、伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する。環境音合成方法は、データ受信部が、環境音量パラメタを受信するデータ受信ステップと、残響付加音源合成部が、1フレーム分の残響を加えた環境音のテンプレート（以下、テンプレートと

50

いう)と当該テンプレートの環境音の音量に対応する情報とを対応付けて記憶するテンプレート記憶部から、環境音量パラメタで特定される音量に応じたテンプレートを選択し、選択したテンプレートを合成して環境音を生成する音源合成ステップとを含む。

【0015】

上記の課題を解決するために、本発明の他の態様によれば、環境音合成方法は、伝送元の空間の大きさに基づく音響信号の残響に関する環境残響パラメタを取得して環境音を生成する。環境音合成方法は、データ受信部が、環境残響パラメタを受信するデータ受信ステップと、残響付加音源合成部が、1フレーム分の残響を加えた環境音のテンプレート(以下、テンプレートという)と当該テンプレートの環境音の残響に対応する情報とを対応付けて記憶するテンプレート記憶部から、環境残響パラメタで特定される残響に応じたテンプレートを選択し、選択したテンプレートを合成して環境音を生成する音源合成ステップとを含む。

【発明の効果】

【0016】

本発明によれば、伝送元において収音された環境音を効率よく伝送し、伝送先で伝送元の場の雰囲気を残響を考慮して再現することができるという効果を奏する。

【図面の簡単な説明】

【0017】

【図1】本発明の環境音伝送システムの構成例を示すブロック図。

【図2】実施例1の環境音分析装置の構成を示すブロック図。

【図3】実施例1の環境音分析装置の動作を示すフローチャート。

【図4】実施例2の環境音分析装置の構成を示すブロック図。

【図5】実施例2の環境音分析装置の動作を示すフローチャート。

【図6】実施例2のパラメタ変換部のパラメタ生成手順を例示する図。

【図7】実施例2の変形例1の環境音分析装置の構成を示すブロック図。

【図8】実施例2の変形例1の環境音分析装置の動作を示すフローチャート。

【図9】実施例3、実施例5の環境音合成装置の構成を示すブロック図。

【図10】実施例3の環境音合成装置の動作を示すフローチャート。

【図11】実施例3の残響付加音源合成部の残響付加方法を説明するための図。

【図12】実施例4の環境音合成装置の構成を示すブロック図。

【図13】実施例4の環境音合成装置の動作を示すフローチャート。

【図14】実施例4の音源合成部の環境音素片テンプレート合成手順を例示する図。

【図15】実施例5の環境音合成装置の動作を示すフローチャート。

【図16】実施例6の環境音分析装置の構成を示すブロック図。

【図17】実施例6の環境音分析装置の動作を示すフローチャート。

【図18】実施例7、8の環境音合成装置の構成を示すブロック図。

【図19】実施例7の環境音合成装置の動作を示すフローチャート。

【図20】実施例8の環境音分析装置の動作を示すフローチャート。

【発明を実施するための形態】

【0018】

以下、本発明の実施の形態について、詳細に説明する。なお、同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

【0019】

拍手や手拍子音、声援・掛け声などの環境音の総音量は、観客の人数が多いほど大きくなる。本発明では、環境音そのものを伝送するのではなく、環境音の音量を表す情報だけを伝送する。そして、伝送先では予め記憶された環境音のテンプレートを、音量を表す情報に応じて変換することにより、伝送元の環境音(に類似した音)を再生する。

【0020】

また、拍手や手拍子音の一拍(一度両手を合わせて打つこと)は、音響パワーの個人差が小さい。また、一拍と一拍の時間間隔(以下、拍手間隔ともいう)の個人差も小さく、

200ms ~ 300ms 程度である。したがって、ある人の拍手音（一拍分）を環境音素片テンプレートとして用意しておき、それを個人差に応じたゆらぎ（200ms ~ 300ms）を持たせた間隔で繰り返し再生することにより、別の人の拍手音に類似した音を構成することができる。

#### 【0021】

##### < 環境音伝送システム >

以下、図1を参照して本発明の環境音伝送システムについて説明する。図1は本発明の環境音伝送システムの構成例を示すブロック図である。図1に示すように、本発明の環境音伝送システムは、伝送元の環境音分析装置と、伝送先の環境音合成装置から構成される。図1Aに示すように、後述する実施例1, 2, 2の環境音分析装置は、入力された音響信号（環境音）の音量に対応する情報（環境音量パラメタ $P_j$ 、以下単にパラメタともいう）を抽出し、出力する。後述する実施例4, 5の環境音合成装置は、予め記憶された環境音のテンプレートを用いて、入力された環境音量パラメタ $P_j$ を用いてテンプレートを選択し、選択したテンプレートを用いて環境音を合成し、出力する。また、図1Bに示すように、後述する実施例6の環境音分析装置は、入力された音響信号（環境音）の音量に対応する情報（環境音量パラメタ $P_j$ 、以下単にパラメタ $P_j$ ともいう）と、音響信号（環境音）の残響に対応する情報（環境残響パラメタ $RP_j$ 、以下単にパラメタ $RP_j$ ともいう）とを抽出し、出力する。後述する実施例7, 8の環境音合成装置は、予め記憶された環境音のテンプレートを用いて、入力された環境残響パラメタ $RP_j$ または環境音量パラメタ $P_j$ を用いて、テンプレートを選択し、選択したテンプレートを用いて環境音を合成し、出力する。以下、実施例1において環境音分析装置1、実施例2において環境音分析装置2、実施例2の変形例1において環境音分析装置2、実施例3において環境音合成装置3、実施例4において環境音合成装置4、実施例5において環境音合成装置5、実施例6において環境音分析装置6、実施例7において環境音合成装置7、実施例8において環境音合成装置8をそれぞれ説明する。また、環境音分析装置1, 2, 2と環境音合成装置3, 4, 5との組み合わせを環境音伝送システム1000、環境音分析装置6と環境音合成装置7, 8との組み合わせを環境音伝送システム2000と呼ぶ。

#### 【実施例1】

#### 【0022】

以下、図2、図3を参照して本発明の実施例1の環境音分析装置について説明する。図2は本実施例の環境音分析装置1の構成を示すブロック図である。図3は本実施例の環境音分析装置1の動作を示すフローチャートである。図2に示すように、本実施例の環境音分析装置1は、收音部11と、音量計算部12と、パラメタ変換部13と、データ送信部14とを備える。

#### 【0023】

##### < 收音部11 >

收音部11は伝送元の音を收音する（S11）。ここでは、收音部11には伝送元の拍手音が入力されるものとする。

#### 【0024】

##### < 音量計算部12 >

音量計算部12は、拍手音の音響信号を取得する。音量計算部12が取得する拍手音の音響信号は、所定のサンプリング周波数でサンプリングされた信号列とする。ここで、 $X_j$ を第 $j$ フレームの音響信号とし、 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ （ $N$ はフレームあたりのサンプル数）とする。例えば8kHzサンプリングのときに1フレーム20msとすると、 $N = 160$ である。なお、遅延が短い方が良ければフレームの長さを短くし、遅延が長くなっても良ければ、フレームの長さを長くすれば良い。音量計算部12は、フレーム毎に、入力された拍手音の音響信号の音量に対応する値（以下、「拍手音量に対応する値」ともいう）を求めて出力する。具体的には、音量計算部12は、フレーム毎に、入力された拍手音の音響信号 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ の平均エネルギー

10

20

30

40

50

【 0 0 2 5 】

【 数 1 】

$$E_j = \frac{x_j(1)^2 + x_j(2)^2 + \dots + x_j(N)^2}{N}$$

【 0 0 2 6 】

を計算する ( S 1 2 ) 。

【 0 0 2 7 】

< パラメタ変換部 1 3 >

パラメタ変換部 1 3 は、音量計算部 1 2 から出力された拍手音量に対応する値を取得する。パラメタ変換部 1 3 は、取得した拍手音量に対応する値を量子化し、環境音量パラメタを出力する。具体的には、パラメタ変換部 1 3 は、平均エネルギー  $E_j$  の取りうる範囲 (例えば  $x_j(i)$  ( $i = 1, 2, \dots, N$ )) が符号付き 16 bit の場合は最小値が 0 で最大値が  $2^30$  となる) をあらかじめ定められた場合の数 (例えば 16 bit) に量子化し、そのインデックスを環境音量パラメタ  $P_j$  として出力する ( S 1 3 ) 。

【 0 0 2 8 】

< データ送信部 1 4 >

データ送信部 1 4 は、パラメタ変換部 1 3 が出力した環境音量パラメタ  $P_j$  を伝送先の環境音合成装置 3 (または 4、5) に送信する ( S 1 4 ) 。

【 0 0 2 9 】

このように、本実施例の環境音分析装置 1 によれば、伝送元において収音された拍手音を効率よく低遅延に伝送することができる。

【 0 0 3 0 】

[実施例 1 の動作例 2]

上述の実施例 1 では、伝送元の環境音の例として拍手音を対象とし、伝送元の拍手音を分析する環境音分析装置 1 の動作例を説明したが、これに限らず拍手音以外の環境音を対象としても良い。例えば、声援や掛け声などを環境音としても良いし、伝送元で収音される音の中から伝送元会場のメインコンテンツの音を除いた音響信号 (雑音を含む) を環境音としても良い。

【 0 0 3 1 】

実施例 1 の動作例 2 における環境音分析装置 1 は、環境音分析装置 1 の収音部 1 1、音量計算部 1 2、パラメタ変換部 1 3、データ送信部 1 4 の各部で取り扱われる拍手音および拍手音量が、環境音及び環境音の音量に置き換わる点を除いては、上述の動作例と同じである。

【 0 0 3 2 】

拍手音や声援・掛け声、雑音などは、いずれも伝送元の会場の雰囲気を決定的な重要な要素である一方で、いろいろな音響信号が混合された白色雑音に近い信号である。前述したようにこれらの音を環境音と呼ぶ。伝送元で環境音が発せられたタイミング及び音量が保たれていれば、信号そのものは伝送元の環境音と全く同じ信号でなくとも、場の雰囲気を再現することができる。そこで、環境音分析装置 1 において、伝送元の環境音の音量に関するパラメタを抽出することで、伝送元において収音された環境音を効率よく低遅延に伝送することができる。

【 実施例 2 】

【 0 0 3 3 】

以下、図 4、図 5、図 6 を参照して本発明の実施例 2 の環境音分析装置について説明する。図 4 は本実施例の環境音分析装置 2 の構成を示すブロック図である。図 5 は本実施例の環境音分析装置 2 の動作を示すフローチャートである。図 6 は本実施例のパラメタ変換部 2 3 のパラメタ生成手順を例示する図である。図 4 に示すように、本実施例の環境音分析装置 2 は、収音部 1 1 と、音量計算部 1 2 と、パラメタ変換部 2 3 と、データ送信部 1

4 とを備える。收音部 1 1、音量計算部 1 2、データ送信部 1 4 は実施例 1 の環境音分析装置 1 における同一番号の各構成部と同じであるから説明を適宜略する。

【 0 0 3 4 】

< 音量計算部 1 2 >

音量計算部 1 2 は、4 8 k H z サンプリングでサンプリングされた信号列であり、1 フレーム 6 サンプル ( N = 6 ) で構成される信号列  $X_j = ( x_j(1), x_j(2), \dots, x_j(6) )$  を取得する。音量計算部 1 2 は、フレーム毎に、入力された拍手音響信号  $X_j = ( x_j(1), x_j(2), \dots, x_j(6) )$  から、平均エネルギー

【 0 0 3 5 】

【数 2】

$$E_j = \frac{x_j(1)^2 + x_j(2)^2 + \dots + x_j(6)^2}{6}$$

10

【 0 0 3 6 】

を計算する ( S 1 2 ) 。

【 0 0 3 7 】

< パラメタ変換部 2 3 >

パラメタ変換部 2 3 は、求めた平均エネルギー  $E_j$  を以下の式により変形した列  $F_j$  を求める。

【 0 0 3 8 】

【数 3】

$$F_j = \begin{cases} -\sqrt{E_j} - 1 & \left( \sqrt{E_j} \text{が奇数の場合} \right) \\ \sqrt{E_j} & \left( \sqrt{E_j} \text{が偶数の場合} \right) \end{cases} \dots (1)$$

20

【 0 0 3 9 】

つまり、図 6 に示すように、ガウス関数や床関数により整数値化された  $F_j$  の取りうる値 ( 0 ~ 3 2 7 6 8 ) のうち、奇数の値に負の符号を与え、さらに 1 を減じる。これにより、 $F_j$  はすべて偶数の値を取ることになる。次に、全部偶数になった  $F_j$  の各々を 2 で割る ( 右に 1 ビットシフトでも構わない ) 。この値を  $G \cdot 7 1 1$  準拠の範囲に収めるために、 $\mu - 1 a w$  を使うのであればさらに 2 で割り ( 右に 1 ビットシフトでもよい ) 値  $G_j$  を求める。そして、 $G$  に対して、ITU - T  $G \cdot 7 1 1$  の符号化処理を行い、 $G_j$  を  $G \cdot 7 1 1$  の符号 ( 番号 ) に変換する。4 8 k H z サンプリング 6 サンプル分を一塊 ( 1 フレーム ) にすると 8 k H z の 1 サンプル分に相当するので、上記の  $G_j$  毎に  $G \cdot 7 1 1$  のシンボル 1 つを割り当てることができる。割り当てられたシンボル列をパラメタ  $P_j$  として出力する ( S 2 3 ) 。パラメタ  $P_j$  は通常の音声と同様に固定電話回線を用いて伝送すると遅延が短くすむ。式 ( 1 ) のかわりに式 ( 2 ) のように対数を用いてもよい。

30

【 0 0 4 0 】

【数 4】

$$F_j = \begin{cases} -\log E_j - 1 & \left( \log E_j \text{が奇数の場合} \right) \\ \log E_j & \left( \log E_j \text{が偶数の場合} \right) \end{cases} \dots (2)$$

【 0 0 4 1 】

また、平方根演算や対数演算は多項式近似 ( テイラー展開など ) で演算量を削減してもよい。

【 0 0 4 2 】

[ 実施例 2 の変形例 1 ]

以下、図 7、図 8 を参照して実施例 2 のパラメタ変換部 2 3 に変更を加えた変形例 1 の環境音分析装置について説明する。図 7 は本変形例の環境音分析装置 2 の構成を示すブロック図である。図 8 は本変形例の環境音分析装置 2 の動作を示すフローチャートである。図 7 に示すように、本変形例の環境音分析装置 2 は、收音部 1 1 と、音量計算部 1

50



2と、パラメタ変換部23と、データ送信部14とを備える。收音部11、音量計算部12、データ送信部14は実施例2の環境音分析装置2における同一番号の各構成部と同じであるから説明を適宜略する。

【0043】

<パラメタ変換部23>

パラメタ変換部23は、図6のようなマッピング演算の代わりに、 $F_j$ の取りうる0~32768の値を直接8bitのシンボルにマッピングするマッピングテーブル23Aを予め備えており、マッピングテーブル23Aを参照してパラメタ $P_j$ を求める(S23)。または、パラメタ変換部23は、 $F_j$ の取りうる0~32768の値をあらかじめビットシフト等により場合の数を減らしてから、マッピングテーブル23Aを用いてパラメタ $P_j$ を求めてもよい。この場合はマッピングテーブル23Aの大きさを削減できる。 $F_j$ はデシベル単位に変換したものをを用いてもよい。

【0044】

実施例2及び変形例1の環境音分析装置は以下の効果を有する。收音された拍手音の音響信号は正の値となるため、 $E_j$ の平方根の値の取りうる範囲は正の整数値、例えば $x_j(n)$  ( $n=1, 2, \dots, N$ )が符号付き16bitの場合は最小値が0で最大値が32768となる。このまま、パラメタ変換部でITU-T G.711の符号化を行うと、符号化効率が悪くなるという問題がある。上記式(1)の変形を行うと、例えば $x_j(n)$  ( $n=1, 2, \dots, N$ )が符号付き16bitの場合は、 $F_j$ の取りうる範囲は-16384から16384になる。そこで、パラメタ変換部において $E_j$ の取りうる範囲が負の整数値から正の整数値の範囲となるように変換した値 $F_j$ を用いることにより、符号化効率を向上させることができ、パラメタ $P_j$ の情報量を削減することができる。つまり、伝送遅延をより少なくすることが可能となる。

【0045】

[実施例2の動作例2]

上述の実施例2および実施例2の変形例1では、伝送元の環境音の例として拍手音を対象とし、伝送元の拍手音を分析する環境音分析装置2(2)の動作例を説明したが、これに限らず拍手音以外の環境音を対象としても良い。例えば、声援や掛け声などを環境音としても良いし、伝送元で收音される音の中から伝送元会場のメインコンテンツの音を除いた音響信号(雑音を含む)を環境音としても良い。

【0046】

実施例2の動作例2においては、環境音分析装置2(2)の收音部11、音量計算部12、パラメタ変換部23または23、データ送信部14の各部分で取り扱われる拍手音および拍手音量が、環境音及び環境音の音量に置き換わる点を除いては、上述の動作例と同じである。

【実施例3】

【0047】

以下、図9、図10を参照して本発明の実施例3の環境音合成装置について説明する。図9は本実施例の環境音合成装置3の構成を示すブロック図である。図10は本実施例の環境音合成装置3の動作を示すフローチャートである。図9に示すように、本実施例の環境音合成装置3は、データ受信部31と、残響付加音源合成部32と、テンプレート記憶部33と、再生部34とを備える。環境音合成装置3は環境音分析装置1(2、2)から伝送元の音響信号の音量に関する環境音量パラメタを取得して環境音を生成する装置である。以下、実施例1、2で詳述した動作例に従い、環境音の例として拍手音を用いて説明を進める。

【0048】

<データ受信部31>

データ受信部31は、環境音分析装置から環境音量パラメタ $P_j$ を受信する(S31)

【0049】

10

20

30

40

50

## &lt; テンプレート記憶部 33 &gt;

テンプレート記憶部 33 には、拍手音の各音量バリエーションに対して複数の拍手音（1 フレーム分）のテンプレートが記憶されている。つまり、テンプレート記憶部 33 には、 $i$  をフレームのインデックスとした場合に、1 フレーム分の拍手音を含む環境音のテンプレート  $T_i$  と当該テンプレートの環境音の音量に対応する情報  $E_i$  とが対応付けて記憶されているものとする。なお、テンプレートの環境音の音量に対応する値は、各テンプレート  $T_i$  を入力として、上記実施例 1 または 2 の音量計算部 12 及びパラメタ変換部 13（23）と同じ方法により求めることができる。なお、実施例 1 または 2 のどの方法を用いるかは、環境音分析装置と環境音合成装置との間で統一しておくものとする。

## 【 0050 】

## &lt; 残響付加音源合成部 32 &gt;

残響付加音源合成部 32 は、入力された環境音量パラメタ  $P_j$  で特定される音量に応じたテンプレートのうちいずれか 1 つをテンプレート記憶部 33 からランダムに選択する。つまり、 $P_j = E_i$  を満たす  $E_i$  に対応づけられているテンプレート  $T_i$  のうち、いずれか 1 つをランダムに選択する。残響付加音源合成部 32 は、選択したテンプレートに環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、テンプレートに残響を加え、残響を加えたテンプレートを、必要に応じて前のフレームと補間をして、1 フレーム分の音響信号を合成して環境音（この動作例では拍手音）を生成する（S32）。ここでは、環境音量パラメタで特定される音量が大きいほど、伝送元の空間の広さと仮定する。例えば、環境音量パラメタ  $P_j$  の値が所定の閾値よりも小さい場合は、狭い空間であることが想定されるので、図 11A のように、テンプレートに短い残響  $H_s$  を積み込む。また、環境音量パラメタ  $P_j$  の値が閾値以上の場合は、広い空間であることが想定されるので、図 11B のように、テンプレートに短い残響  $H_s$  とともに長い残響  $H_l$  を積み込む。所定の閾値は、例えば実験やシミュレーション等により適切な値を調べ、設定すればよい。例えば、20ms のフレームあたり環境音量パラメタに 8bit のバリエーションがあったとすると、400bit/sec で拍手音を伝送できる。なお、音量が大きいほど、残響が長くなるという特性が前述の残響特性に相当する。

## 【 0051 】

## &lt; 再生部 34 &gt;

再生部 34 は、残響付加音源合成部 32 が合成した拍手音を再生する（S34）。

## 【 0052 】

このように、本実施例の環境音合成装置 3 によれば、テンプレート記憶部 33 に拍手音の各音量バリエーションに対して複数のテンプレートを保持しておき、残響付加音源合成部 32 が音量の条件を充たす複数のテンプレートから 1 つのテンプレートをランダムに選択するため、合成された拍手音が定常的なパターンとして聞こえないようにすることができる。さらに、選択したテンプレートに環境音量パラメタで特定される音量に応じた残響を加えるため、伝送先で伝送元の場の雰囲気を残響を考慮して再現することができる。

## 【 0053 】

## [実施例 3 の動作例 2]

実施例 3 では、伝送元の環境音の例として拍手音を対象とし、伝送元の拍手音の音量に関するパラメタを取得して、伝送先で拍手音を生成する環境音合成装置 3 の動作例を説明したが、これに限らず拍手音以外の環境音を対象としても良い。例えば、声援や掛け声や、伝送元で収音される音の中から伝送元会場のメインコンテンツの音を除いた音響信号（雑音を含む）を環境音とし、伝送元の環境音量パラメタが入力され、伝送先で環境音を合成してもよい。

## 【 0054 】

実施例 3 の動作例 2 では、実施例 3 の環境音合成装置 3 のデータ受信部 31 と、残響付加音源合成部 32 と、テンプレート記憶部 33 と、再生部 34 において、拍手音が環境音に置き換わる点を除いては、上述の動作例と同じである。なお、以降において説明する環境音分析装置、環境音合成装置においても同様に拍手音以外の環境音を対象としても良い

10

20

30

40

50

## 【実施例 4】

## 【0055】

以下、図 1 2、図 1 3、図 1 4 を参照して本発明の実施例 4 の環境音合成装置について説明する。図 1 2 は本実施例の環境音合成装置 4 の構成を示すブロック図である。図 1 3 は本実施例の環境音合成装置 4 の動作を示すフローチャートである。図 1 4 は本実施例の残響付加音源合成部 4 2 の環境音素片テンプレート合成手順を例示する図である。図 1 2 に示すように、本実施例の環境音合成装置 4 は、データ受信部 3 1 と、残響付加音源合成部 4 2 と、テンプレート記憶部 4 3 と、再生部 3 4 と、人数推定部 4 5 と、テンプレート音量記憶部 4 6 とを備える。データ受信部 3 1、再生部 3 4 は実施例 3 の環境音合成装置 3 における同一番号の各構成部と同じであるから説明を省略する。

## 【0056】

## &lt;テンプレート記憶部 4 3 &gt;

テンプレート記憶部 4 3 には、一人の人間による一拍分の拍手音（300ms 程度）のテンプレートの複数のバリエーションが記憶されている。本実施例では環境音の例として拍手音を扱うため、拍手音のテンプレートを環境音素片テンプレートのバリエーションのひとつとする。従って、以下では拍手音のテンプレートを環境音素片テンプレートともいう。例えば、異なる人の一拍分の拍手音をそれぞれ異なる環境音素片テンプレートとして記憶しておく。以下、単にテンプレートという場合には、所定フレーム長の複数人による拍手音（環境音）全体を収録したテンプレートを指すものとし、環境音素片テンプレートという場合には、一人の人間による一拍分の拍手音（環境音）のテンプレートを指すものとする。

## 【0057】

## &lt;テンプレート音量記憶部 4 6 &gt;

テンプレート音量記憶部 4 6 には、テンプレート記憶部 4 3 に記憶されている環境音素片テンプレートの音量に対応する情報（具体的には、実施例 1 または 2 の音量計算部 1 2 により計算される、平均エネルギー）が記憶されている。なお、1 人分の拍手音の音量の差は小さいので、テンプレート記憶部 4 3 に記憶されている環境音素片テンプレートのいずれか一つについて計算された平均エネルギーを環境音素片テンプレートの音量に対応する情報として記憶しておいてもよい。また、テンプレート記憶部 4 3 に記憶されている全環境音素片テンプレートの平均エネルギーの平均値を、環境音素片テンプレートの音量に対応する情報としてテンプレート音量記憶部 4 6 に記憶しておいてもよい。あるいは、予め定めた定数を音量に対応する情報としてテンプレート音量記憶部 4 6 に記憶しておいてもよい。

## 【0058】

なお、テンプレート音量記憶部 4 6 に予め環境音素片テンプレートの音量に対応する情報を記憶せず、その都度テンプレート記憶部 4 3 からランダムに選択した環境音素片テンプレートについて計算した平均エネルギーを環境音素片テンプレートの音量に対応する情報として用いてもよい。

## 【0059】

## &lt;人数推定部 4 5 &gt;

人数推定部 4 5 は、環境音量パラメタ  $P_j$  に応じて音量のゲイン調整を行うための構成である。人数推定部 4 5 は、伝送元から出力された環境音量パラメタ  $P_j$  を取得し、当該環境音量パラメタ  $P_j$  から音量に対応する情報  $E_j$  を求める。具体的には、実施例 1 または 2 のパラメタ変換部 1 3（2 3）と逆の処理を行うことにより、音量に対応する情報  $E_j$  を得る。人数推定部 4 5 は、音量に対応する情報  $E_j$  を環境音素片テンプレートの音量に対応する情報で除算した値の整数値（小数点以下を四捨五入、または切り捨てた値）を拍手の人数  $M$  として出力する（S 4 5）。

## 【0060】

## &lt;残響付加音源合成部 4 2 &gt;

残響付加音源合成部 4 2 は、テンプレート記憶部 4 3 から環境音素片テンプレートをランダムに選択して、環境音量パラメタで特定される音量とその音量に応じた残響特性とを用いて、テンプレートに残響を加え、残響を加えた環境音素片テンプレートを合成することで環境音を生成する (S 4 2)。ここでは (環境音量パラメタで特定される音量に応じた) 人数に応じて空間の広さが変化すると仮定する。例えば、10人 (M=10) 程度の拍手であれば10人程度の人が入れる空間の大きさに応じた残響を加え、100人 (M=100) 程度の拍手であれば100人程度の人が入れる空間の大きさに応じた残響を加える。なお、空間の大きさに応じて残響の長さが変化するという特性が前述の残響特性に相当する。例えば、コンサート会場等における、観客間の間隔は50~70cm程度なので、その値から空間の大きさを推定する。例えば、円形、正方形、直線、格子状または、それらの組合せからなる形状に、50~70cmの間隔で人間を配置したとして、空間の大きさを推定する。人数が少ない場合、例えば、環境音量パラメタ  $P_j$  の値が小さく、Mの値が閾値よりも小さい場合は、狭い空間であることが想定されるので、図 1 1 A のように、環境音素片テンプレートに短い残響  $H_s$  を畳み込む。人数が多い場合、例えば、環境音量パラメタ  $P_j$  の値が大きく、Mの値が閾値以上の場合は、広い空間であることが想定されるので、図 1 1 B のように、環境音素片テンプレートに短い残響  $H_s$  とともに長い残響  $H_l$  を畳み込む。

【 0 0 6 1 】

例えば、拍手音の間隔を特許文献 1 と同様とする。例えば、 $M = 1$  の場合、図 1 4 A のように、約 3 0 0 m s ごとにランダムに選択された環境音素片テンプレート  $T_i$  を用いて合成した波形に残響を付加して拍手音として出力する。前述のように合成の時間間隔は約 3 0 0 m s でよいが、より好ましくは 3 0 0 m s を中心として時間間隔に揺らぎを持たせてもよい。時間間隔に揺らぎを持たせることによってさらに自然な拍手音を合成することができる。たとえば 3 0 0 m s を中心としてガウス分布にしたがう乱数により、 $\pm$  数 1 0 m s の揺らぎを持たせればよい。例えば残響付加音源合成部 4 2 は

【 0 0 6 2 】

【 数 5 】

$$Z_i = (i \cdot \tau + \sigma_i) \cdot T_i \quad (i = 0, 1, 2, \dots) \quad \dots(3)$$

$$Y_i = Z_i * H$$

【 0 0 6 3 】

によりテンプレートを変換した拍手音  $Y_i$  ( $i = 0, 1, 2, \dots$ ) を出力する (S 4 2)。なお、式中、 $H$  は残響を示し、前述の通り、Mの値に応じて短い残響  $H_s$  または長い残響  $H_l$  を用いる。M = 1 の場合には、狭い空間であることが想定されるので、短い残響  $H_s$  を用いる。別の表現方法で書くと、時系列テンプレート信号  $T_i = (t_i[1] \ t_i[2] \ \dots \ t_i[P])$  と拍手タイミングを表すインパルス ( $\delta(i \cdot \tau + \sigma_i)$ ) とを用いて、合成音  $Z_i$  を求め、残響  $H$  を畳み込み、 $Y_i$  を求め、出力とする。

【 0 0 6 4 】

【 数 6 】

$$Z_i = \delta(i \cdot \tau + \sigma_i) \cdot T_i \quad (i = 0, 1, 2, \dots) \quad \dots(3)$$

$$Y_i = Z_i * H$$

【 0 0 6 5 】

ここで  $*$  は畳み込み演算を表す。ここで、 $\tau = 300$  m s であり、 $\sigma_i$  は  $-10$  m s から  $10$  m s の範囲で生成した乱数である。また、 $\delta$  は関数ではなく時間方向に揺れている伝達関数(残響)  $H$  を畳み込み、 $Y_i$  を求めてもよい。

【 0 0 6 6 】

【 数 7 】

$$Y_i = H(i \cdot \tau + \sigma_i) \cdot T_i \quad (i = 0, 1, 2, \dots) \quad \dots(3')$$

【 0 0 6 7 】

10

20

30

50

環境音量パラメタによりM人分の拍手を合成する場合は、図14Bのように、時間間隔を約300/M(ms)ごとにランダムに選択された環境音素片テンプレートを用いて合成された波形に残響を付加して拍手音として出力する。人数Mの逆数を使って、時間間隔を約300/M(ms)と設定することで、拍手の人数Mが増えるに従って時間間隔が小さくなるように設定することができる。この場合もガウス分布やラプラス分布に従う乱数によって、揺らぎを持たせることができる。例えば残響付加音源合成部42は、

【0068】

【数8】

$$Z_i = \sum_{m=1}^M \left( \frac{1}{M} m \cdot \tau + \sigma_m \right) \cdot T_i \quad (i=0,1,2,\dots) \quad \dots(4)$$

$$Y_i = Z_i * H$$

【0069】

によりテンプレートを変換し、残響を付加した環境音 $Y_i$ ( $i=0,1,2,\dots$ )を出力する(S42)。

【0070】

このように、本実施例の環境音合成装置4によれば、実施例3のように音量ごとにテンプレートを用意しておく必要がなく、テンプレート記憶部43に記憶しておく環境音素片テンプレートの数も少なくてもよい。環境音合成装置4のメモリ量を削減することができる。さらに、人数に応じて空間の広さを推定することができ、より適切な残響を生成し、より適切に伝送元の場の雰囲気を再現することができると考えられる。なお、本実施例のポイントは、環境音素片テンプレートを用いて合成された波形に残響を付加して拍手音とすることなので、拍手音の間隔については他の方法を用いて設定してもよい。

20

【0071】

[実施例4の動作例2]

実施例4は、伝送元の伝送元の環境音の例として拍手音を対象とし、伝送元の拍手音の音量に関するパラメタを取得して、伝送先で拍手音を生成する環境音合成装置4を説明したが、これに限らず拍手音以外の環境音を対象としても良い。上述では、一人の人間による一拍分の拍手音(300ms程度)のテンプレートを環境音素片テンプレートの例として示したが、これに限らず、たとえば、一人の人間による一拍分の声援、掛け声のテンプレートを環境音素片テンプレートとしてもよい。

30

【0072】

実施例4の動作例2では、実施例4の環境音合成装置4のデータ受信部31と、残響付加音源合成部42と、テンプレート記憶部43と、再生部34と、人数推定部45と、テンプレート音量記憶部46において取り扱われるデータが拍手音から環境音に置き換わる点を除いては、上述の動作例と同じである。

【0073】

なお、残響付加音源合成部42において、式(3)の代わりに、時系列テンプレート信号 $T_i = (t_i[1] \quad t_i[2] \quad \dots \quad t_i[P])$ と環境音タイミングを表すインパルス $(m \cdot \tau + \sigma_m)$ とを用いて、合成音 $Z_i$ を求め、残響 $H$ を畳み込み、 $Y_i$ を求め、 $Y_i$ を出力としても良い。

40

【0074】

【数9】

$$Z_i = \delta(m \cdot \tau + \sigma_i) \cdot T_i \quad (i=0,1,2,\dots) \quad \dots(3'')$$

$$Y_i = Z_i * H$$

【0075】

ここで\*は畳み込み演算を表す。

【0076】

また、テンプレート記憶部43に記憶しておく環境音素片テンプレートの波形のエネルギー

50

ギーをあらかじめ正規化してあってもよい。その場合は、人数推定部45のパラメタに  
 じで、音量(ゲイン)を調整すればよい。この場合もメモリ量を少なくしながらバリエー  
 ションを増やすことができる。

【実施例5】

【0077】

以下、実施例3と異なる部分を中心に説明する。

【0078】

以下、図9、図15を参照して本発明の実施例5の環境音合成装置について説明する。  
 図9は本実施例の環境音合成装置5の構成を示すブロック図である。図15は本実施例の  
 環境音合成装置5の動作を示すフローチャートである。図9に示すように、本実施例の環  
 境音合成装置5は、データ受信部31と、残響付加音源合成部52と、テンプレート記憶  
 部53と、再生部34とを備える。残響付加音源合成部52及びテンプレート記憶部53  
 以外の各構成部は実施例3の環境音合成装置3における同一番号の各構成部と同じである  
 から説明を省略する。

【0079】

<テンプレート記憶部53>

テンプレート記憶部53には、拍手音の各音量バリエーションに対して残響を加えた複  
 数の拍手音(以下「残響付加済の拍手音」ともいう、これを1フレーム分)のテンプレ  
 ートが記憶されている。つまり、テンプレート記憶部53には、 $i$ をフレームのインデッ  
 スとした場合に、1フレーム分の残響付加済の拍手音を含む環境音のテンプレート $T_i$ と  
 当該テンプレートの環境音の音量に対応する情報 $E_i$ とが対応付けて記憶されているも  
 のとする。ここでは、環境音量パラメタで特定される音量が大きいほど、伝送元の空間の  
 広いと仮定する。そのため、音量が大きいほど、伝送元の空間の広く、残響は長くなる。  
 つまり、本実施例のテンプレート記憶部53に記憶されるテンプレートには、既に、残響  
 特性(音量が大きいほど、残響は長くなるという特性)に応じた残響が加えられていると言  
 える。

【0080】

なお、残響付加済の拍手音は、所望の残響を観測できる場所で録音したものでもよいし  
 、残響がない(または少ない)状態で録音した信号に残響を畳み込んだものでもよい。

【0081】

<残響付加音源合成部52>

残響付加音源合成部52は、入力された環境音量パラメタ $P_j$ で特定される音量に応じ  
 た(残響付加済の拍手音の)テンプレートをテンプレート記憶部53から選択し、選択し  
 たテンプレートを合成して環境音を生成し( $S_{52}$ )、出力する。

【0082】

例えば、1つ以上の閾値を設け、テンプレート記憶部53では、閾値と環境音の音量に  
 対応する情報 $E_i$ との大小関係により、テンプレートを複数のグループに分けておく。  
 残響付加音源合成部52は、環境音量パラメタ $P_j$ と閾値との大小関係により、何れのグ  
 ループに含まれるテンプレートを選択するか決定する。

【0083】

(グループ例1)

例えば、二つの閾値 $Th1$ と $Th2$ ( $Th1 < Th2$ )を設け、 $E_i < Th1$ となるテンプレートを短い  
 残響が畳み込まれたテンプレートのグループ(以下DB11ともいう)に、 $E_i > Th2$ となる  
 テンプレートを長い残響が畳み込まれたテンプレートのグループ(以下DB13ともいう)に  
 、 $Th1 < E_i < Th2$ となるテンプレートを中くらいの残響が畳み込まれたテンプレートの  
 グループ(以下DB12ともいう)に分類する。

【0084】

(選択例1-1)

残響付加音源合成部52は、 $P_j < Th1$ のときにDB11からテンプレートを選択し、 $Th1 < P_j < Th2$ のときにDB11及びDB12からテンプレートを選択し、 $Th2 < P_j$ のときにDB11

10

20

30

40

50

、DB1 2及びDB1 3からテンプレートを選択する。

【0085】

(選択例1 - 2)

閾値で完全に分けずに、音量に応じて各DBから選ばれるテンプレートに確率の重みを付けて選択しても良い。

【0086】

例えば、残響付加音源合成部5 2は、 $P_j < Th1$ のときに70パーセントの確率でDB1 1からテンプレートを選択し、20パーセントの確率でDB1 2からテンプレートを選択し、10パーセントの確率でDB1 3からテンプレートを選択する。また、 $Th1 < P_{j_i} < Th2$ のときに80パーセントの確率でDB1 1及びDB1 2からテンプレートを選択し、20パーセントの確率でDB1 3からテンプレートを選択する。また、 $Th2 < P_j$ のときに10パーセントの確率でDB1 1からテンプレートを選択し、20パーセントの確率でDB1 2からテンプレートを選択し、70パーセントの確率でDB1 3からテンプレートを選択する。

【0087】

(グループ例2)

例えば、二つの閾値 $Th1$ と $Th2$ ( $Th1 < Th2$ )を設け、 $E_i < Th1$ となるテンプレートを短い残響が畳み込まれたテンプレートのグループ(以下DB2 1ともいう)に、 $E_i < Th2$ となるテンプレートを短い残響と中くらいの残響とが畳み込まれたテンプレートのグループ(以下DB2 2ともいう)に、 $E_i > Th2$ となるテンプレートを短い残響と中くらいの残響と長い残響とが畳み込まれたテンプレートのグループ(以下DB2 3ともいう)に分類する。

【0088】

(選択例2 - 1)

残響付加音源合成部5 2は、 $P_j < Th1$ のときにDB2 1からテンプレートを選択し、 $Th1 < P_{j_i} < Th2$ のときにDB2 2からテンプレートを選択し、 $Th2 < P_j$ のときにDB2 3からテンプレートを選択する。

【0089】

(選択例2 - 2)

閾値で完全に分けずに、音量に応じて各DBから選ばれるテンプレートに確率の重みを付けて選択しても良い。例えば、選択例1 - 2と同様の方法により選択する。

【0090】

このような構成により、残響付加音源合成部において畳み込み処理に伴う演算量、時間を省くことができる。なお、本実施例と実施例4とを組合せてもよい。

【実施例6】

【0091】

以下、図1 6、図1 7を参照して本発明の実施例6の環境音分析装置について説明する。図1 6は本実施例の環境音分析装置6の構成を示すブロック図である。図1 7は本実施例の環境音分析装置6の動作を示すフローチャートである。図1 6に示すように、本実施例の環境音分析装置6は、收音部1 1と、音量計算部1 2と、パラメタ変換部6 3と、データ送信部6 4と、空間計算部6 5を備える。パラメタ変換部6 3、データ送信部6 4、及び空間計算部6 5以外の各構成部は実施例1の環境音分析装置1における同一番号の各構成部と同じであるから説明を省略する。

【0092】

<空間計算部6 5>

空間計算部6 5は、拍手音の音響信号を取得する。空間計算部6 5は、フレーム毎に入力された拍手音の音響信号 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ の残響に対応する値を求め( $S_{65}$ )、出力する。具体的には、空間計算部6 5は、フレーム毎に入力された拍手音の音響信号 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ の直接音と残響音との平均エネルギー比 $RE_j$ を計算し、残響に対応する値として出力する。例えば、参考文献1の残響制御技術を用いて音響信号から直接音と残響音との平均エネルギー比 $RE_j$ を計算することができる。

30

40

50

(参考文献1) 木下慶介、中谷智広、三好正人、実環境音声処理 音声認識に適した残響除去收音、N T T技術ジャーナル、2007、Vol.19、No.6

また、直接音と残響音との平均エネルギー比 $R E_j$ に代えて、直接音の平均エネルギーと残響音の平均エネルギーとの組合せ(等価な値)を残響に対応する値として出力してもよい。残響に対応する値は、伝送元の空間の残響の特徴を示すような値であれば、上述の値以外の値でもよい。また、図示しないカメラ等を使って(または人手により)、横並びか正方形か円形かなどの配置に関する情報を残響に対応する値の一部として付加しても良い。

【0093】

<パラメタ変換部63及びデータ送信部64>

パラメタ変換部63は、音量計算部12から出力された拍手音量に対応する値及び残響に対応する値を取得する。パラメタ変換部13は、取得した拍手音量に対応する値及び残響に対応する値を量子化し、環境音量パラメタ及び環境残響パラメタを出力する。環境音量パラメタについては、実施例1で説明した通りである。例えば、パラメタ変換部13は、直接音と残響音との平均エネルギー比 $R E_j$ の取りうる範囲をあらかじめ定められた場合の数(例えば16bit)に量子化し、そのインデックスを環境残響パラメタ $R P_j$ として出力する(S63)。

【0094】

データ送信部64は、パラメタ変換部63が出力した環境音量パラメタ $P_j$ 及び環境残響パラメタ $R P_j$ を伝送先の環境音合成装置7に送信する(S64)。環境音合成装置7については実施例7に記載する。例えば、環境音量パラメタ $P_j$ を示すデータなのか、環境残響パラメタ $R P_j$ を示すデータなのかを示すフラグを付けて符号化し、送信すればよい。

【0095】

なお、環境音量パラメタ $P_j$ 及び環境残響パラメタ $R P_j$ そのものではなく、たとえば環境音量パラメタ $P_j$ に基づく値と環境残響パラメタ $R P_j$ に基づく値の比(例:たとえば、部屋の容積から概算収容人数を求める所定のテーブルを用いて、環境音量パラメタ $P_j$ から推定される人数を、環境残響パラメタ $R P_j$ に基づく部屋の容積で特定される概算収容人数で割った値)や、その符号化値を送信することとしても良い。このとき、環境音量パラメタ $P_j$ と環境残響パラメタ $R P_j$ との比を示すデータであることを示すフラグをつけて符号化することとしても良い。もしくは、環境音量パラメタ $P_j$ に基づく値を符号化したものと、環境音量パラメタ $P_j$ に基づく値と環境残響パラメタ $R P_j$ に基づく値の比を符号化したもの、を伝送しても良い。なお、「部屋の容積」は、環境残響パラメタ $R P_j$ に基づき推定することができる。例えば、環境残響パラメタ $R P_j$ が直接音と残響音との平均エネルギー比の場合、環境残響パラメタ $R P_j$ が大きいときには残響が多く部屋の容積が大きいと推定し、環境残響パラメタ $R P_j$ が小さいときには残響が少なく部屋の容積が小さいと推定する。また、環境残響パラメタは、音響信号(環境音)の残響に対応する情報であり、例えば、部屋の容積や部屋の概算収容人数自体を環境残響パラメタとして用いてもよい。

【0096】

このように、本実施例の環境音分析装置6によれば、伝送元において收音された拍手音を残響を考慮した上で効率よく低遅延に伝送することができる。

【実施例7】

【0097】

以下、図18、図19を参照して本発明の実施例7の環境音合成装置について説明する。図18は本実施例の環境音合成装置7の構成を示すブロック図である。図19は本実施例の環境音合成装置7の動作を示すフローチャートである。図18に示すように、本実施例の環境音合成装置7は、データ受信部71と、残響付加音源合成部72と、テンプレート記憶部73と、再生部34とを備える。再生部34は実施例3の環境音合成装置3における同一番号の再生部34と同じであるから説明を省略する。

【0098】



## &lt; データ受信部 7 1 &gt;

データ受信部 7 1 は、環境音分析装置から環境音量パラメタ  $P_j$  及び環境残響パラメタ  $RP_j$  を受信する ( S 7 1 )。

【 0 0 9 9 】

## &lt; テンプレート記憶部 7 3 &gt;

テンプレート記憶部 7 3 には、拍手音の各残響バリエーションに対して複数の拍手音 ( 1 フレーム分、残響バリエーションに対するものなので、当然、残響付加済の拍手音である ) のテンプレートが記憶されている。つまり、テンプレート記憶部 7 3 には、 $i$  をフレームのインデックスとした場合に、1 フレーム分の残響付加済の拍手音を含む環境音のテンプレート  $T_i$  と当該テンプレートの環境音の残響に対応する情報  $R_i$  とが対応付けて記憶されているものとする。なお、当該テンプレートの環境音の残響に対応する情報  $R_i$  は、例えば、空間計算部 6 5 において、残響に対応する値を計算する際に用いた方法と同じ方法を用いて、テンプレートから計算すればよい。

10

【 0 1 0 0 】

なお、残響付加済の拍手音は、所望の残響を観測できる場所で録音したものでよいし、残響がない ( または少ない ) 状態で録音した信号に残響を畳み込んだものでよい。

【 0 1 0 1 】

## &lt; 残響付加音源合成部 7 2 &gt;

残響付加音源合成部 7 2 は、入力された環境残響パラメタ  $RP_j$  で特定される残響に応じたテンプレートのうちいずれか 1 つをテンプレート記憶部 7 3 からランダムに選択する。つまり、 $RP_j = R_i$  を満たす  $R_i$  に対応づけられているテンプレート  $T_i$  のうち、いずれか 1 つをランダムに選択する。残響付加音源合成部 7 2 は、選択したテンプレートを、必要に応じて前のフレームと補間をして、1 フレーム分の音響信号を合成して環境音 ( この動作例では拍手音 ) を生成する ( S 7 2 )。例えば、20 ms のフレームあたり環境残響パラメタに 8 bit のバリエーションがあったとすると、400 bit / sec で拍手音を伝送できる。

20

【 0 1 0 2 】

例えば、1 つ以上の閾値を設け、テンプレート記憶部 7 3 では、閾値と環境音の残響に対応する情報  $R_i$  との大小関係により、テンプレートを複数のグループに分けておく。残響付加音源合成部 5 2 は、環境残響パラメタ  $RP_j$  と閾値との大小関係により、何れのグループに含まれるテンプレートを選択するか決定する。

30

【 0 1 0 3 】

(グループ例)

例えば、二つの閾値  $Th1$  と  $Th2$  ( $Th1 < Th2$ ) を設け、 $R_i < Th1$  となるテンプレートを短い残響が畳み込まれたテンプレートのグループ ( 以下 DB3 1 ともいう ) に、 $R_i > Th2$  となるテンプレートを長い残響が畳み込まれたテンプレートのグループ ( 以下 DB3 3 ともいう ) に、 $Th1 < R_i < Th2$  となるテンプレートを中くらいの残響が畳み込まれたテンプレートのグループ ( 以下 DB3 2 ともいう ) に分類する。

【 0 1 0 4 】

(選択例 3 - 1)

残響付加音源合成部 7 2 は、 $RP_j < Th1$  のときに DB3 1 からテンプレートを選択し、 $Th1 < RP_j < Th2$  のときに DB3 2 からテンプレートを選択し、 $Th2 < RP_j$  のときに DB3 3 からテンプレートを選択する。

40

【 0 1 0 5 】

(選択例 3 - 2)

閾値で完全に分けずに、環境残響パラメタ  $RP_j$  に応じて各 DB から選ばれるテンプレートに確率の重みを付けて選択しても良い。なお、本実施例と実施例 2 及びその変形例とを組合せてもよい。

【 0 1 0 6 】

例えば、残響付加音源合成部 7 2 は、 $RP_j < Th1$  のときに 70 パーセントの確率で DB3 1

50

からテンプレートを選択し、20パーセントの確率でDB3 2からテンプレートを選択し、10パーセントの確率でDB3 3からテンプレートを選択する。また、 $Th1 < R P_{ji} < Th2$ のときに70パーセントの確率でDB3 2からテンプレートを選択し、それぞれ15パーセントの確率でDB3 1、DB3 3からテンプレートを選択する。また、 $Th2 < R P_j$ のときに10パーセントの確率でDB3 1からテンプレートを選択し、20パーセントの確率でDB3 2からテンプレートを選択し、70パーセントの確率でDB3 3からテンプレートを選択する。この例では、 $R P_j$ により各DBに対して所定の重みをつけているが、各DBを同じ重みで利用することとしても良い。部屋の中に人がまばらに分散している場合、つまり、 $R P_j$ から想定される部屋の収容人数に対して、 $P_j$ から想定される人数が小さな値を取る場合、様々な長さの残響が均等に混じったように聞こえるため、各DBから等分の重みでテンプレートを選択することとしても良い。このとき、 $R P_j$ や $P_j$ とそのものではなく、部屋の中で人がどの程度密集しているかに関する値である $R P_j$ と $P_j$ との比（例： $P_j$ を $R P_j$ で割った値、あるいは、その逆数）に着目し、所定の閾値を用いて、たとえば $P_j$ を $R P_j$ で割った値がこの閾値よりも小さな場合には各DBそれぞれから同じ確率でテンプレートを選択することとしても良い（選択するテンプレートの数は、 $P_j$ に基づく個数とする）。なお、受信したデータが、環境音量パラメタ $P_j$ と環境残響パラメタ $R P_j$ との比を示すデータであることを示すフラグがつけられたデータであった場合、受信したデータを $R P_j$ と $P_j$ との比の代りに用いても良い。もしくは、環境音量パラメタ $P_j$ に基づく値を符号化したものと、環境音量パラメタ $P_j$ に基づく値と環境残響パラメタ $R P_j$ に基づく値の比を符号化したもの、を受信して代わりに用いても良い。

#### 【0107】

実施例3、4、5では、音量から残響を推定していたが、環境音の残響に対応する情報（例えば、直接音と間接音とのエネルギー比）を用いることで、より適切に残響の特徴をとらえることができ、伝送元の場の雰囲気をもより適切に再現することができる。例えば、観客などの音源が劇場や映画館などの空間の後ろ側の座席に偏って座っている場合には、音量が小さくても（人数が少なくても）、長い残響が畳み込まれたテンプレートを用いたほうが伝送元の場の雰囲気をより適切に再現することができる。しかし、実施例3、4、5の構成では、短い残響が畳み込まれたテンプレートを選択する可能性が高い。一方、本実施例では長い残響が畳み込まれたテンプレートを選択する可能性が高く、伝送元の場の雰囲気をより適切に再現することができる。また、本実施例の構成であれば、実施例5と同様に畳み込み処理に伴う演算量、時間を省くことができる。なお、本実施例と実施例4とを組合せてもよい。

#### 【0108】

なお、本実施例では、残響付加音源合成部72において、音響信号を合成して環境音を生成する際に、環境音量パラメタ $P_j$ を利用していない。そのため、環境音量パラメタ $P_j$ を伝送しない構成としてもよい。その場合、環境音分析装置6では環境音量パラメタ $P_j$ を求める必要がないため、音量計算部12を備えなくともよい。パラメタ変換部63では、残響に対応する値のみを量子化すればよい。また、本実施例の構成を維持しておき、何からの原因により、環境残響パラメタ $R P_j$ を伝送されずに、環境音量パラメタ $P_j$ のみが伝送されてきた場合に、実施例3、4、5の構成で環境音を生成してもよい。

#### 【0109】

また、テンプレート記憶部73には、拍手音の各残響バリエーションと各音量のバリエーションの組合せに対して複数の拍手音のテンプレートが記憶されている構成としてもよい。この場合、残響付加音源合成部72は、入力された環境残響パラメタ $R P_j$ 及び環境音量パラメタ $P_j$ で特定される残響及び音量の組合せに応じたテンプレートのうちいずれか1つをテンプレート記憶部73からランダムに選択する。つまり、 $R P_j = R_i$ かつ $P_j = E_j$ を満たす $R_i$ 及び $E_j$ に対応づけられているテンプレート $T_i$ のうち、いずれか1つをランダムに選択する。

#### 【実施例8】

#### 【0110】

10

20

30

40

50

以下、図18、図20を参照して本発明の実施例8の環境音合成装置について説明する。図18は本実施例の環境音合成装置8の構成を示すブロック図である。図20は本実施例の環境音合成装置8の動作を示すフローチャートである。図18に示すように、本実施例の環境音合成装置8は、データ受信部71と、残響付加音源合成部82と、テンプレート記憶部33と、再生部34とを備える。テンプレート記憶部33及び再生部34は実施例3の環境音合成装置3における同一番号のテンプレート記憶部33及び再生部34と同じであるから説明を省略する。また、データ受信部71は実施例7の環境音合成装置7における同一番号のデータ受信部71と同じであるから説明を省略する。

【0111】

<残響付加音源合成部82>

残響付加音源合成部82は、入力された環境音量残響パラメタ $P_j$ で特定される音量に応じたテンプレートのうちいずれか1つをテンプレート記憶部33からランダムに選択する。つまり、 $P_j = E_i$ を満たす $E_i$ に対応づけられているテンプレート $T_i$ のうち、いずれか1つをランダムに選択する。残響付加音源合成部82は、選択したテンプレートに環境残響パラメタ $R P_j$ で特定される残響を加え、残響を加えたテンプレートを、必要に応じて前のフレームと補間をして、1フレーム分の音響信号を合成して環境音（この動作例では拍手音）を生成する（S82）。例えば、環境残響パラメタ $R P_j$ が直接音と残響音との平均エネルギー比を量子化したものである場合、平均エネルギー比が大きいほど残響が長くなるため、所定の閾値よりも環境残響パラメタ $R P_j$ の値が小さい場合にはテンプレートに短い残響 $H_s$ を畳み込み、環境残響パラメタ $R P_j$ の値が閾値以上の場合にはテンプレートに短い残響 $H_s$ とともに長い残響 $H_l$ を畳み込む。

【0112】

このような構成により、実施例7の場合に比べ、テンプレート記憶部に記憶するテンプレートの量を減らすことができる。なお、本実施例と実施例4とを組合せてもよい。

【0113】

<その他の変形例>

実施例7では、環境音分析装置により分析された部屋の広さ・大きさ（容積）に関連する情報である環境残響パラメタ $R P_j$ を用いて残響畳み込み済のテンプレートを選択したが、実施例4のように残響が含まれないテンプレートを用いて、環境残響パラメタ $R P_j$ に応じて特定される長さ、分布の残響を畳み込むこととしても良い。具体的には、環境残響パラメタ $R P_j$ に基づき、残響の最大値を求め、所定の残響の最小値から求めた残響の最大値までの範囲の中から、ランダムに、環境音量パラメタ $P_j$ に基づく個数の残響の長さを決定し、決定した各長さに応じてテンプレートに残響を畳み込むこととしても良い。

【0114】

残響付加音源合成部で環境音を収音する部屋の広さが、環境音合成装置で環境音を合成する部屋に比べて大きな場合、本来は部屋の中で聞こえない領域の音が部屋の中で発生しているような合成音となるため、合成音を聞いた人が違和感を感じる場合がある。この問題を低減するために、たとえば、収音を行う部屋の収容人数で環境音量パラメタ $P_j$ を正規化した値をパラメタとして伝送し、環境音合成装置において、予め定めた合成を行う部屋の収容人数と受信したパラメタとを掛けて得られる値を環境音量パラメタ $P_j$ に替えて用いて、畳み込む残響の長さを求めたり残響畳み込み済のDBからテンプレートを選択することとしても良い。

【0115】

別の方法としては、たとえば、環境音を合成する部屋に応じて予め、環境音量パラメタ $P_j$ の上限値や環境残響パラメタ $R P_j$ の上限値を設けておき、受信したパラメタが上限以上の場合には、受信したパラメタに替えて、部屋に応じた所定の上限値を用いることとしても良い。もしくは、環境音を合成する部屋に応じて予め、環境音量パラメタ $P_j$ の下限値や環境残響パラメタ $R P_j$ の下限値を設けておき、受信したパラメタが下限以下の場合には、受信したパラメタに替えて、部屋に応じた所定の下限値を用いることとしても良い。

10

20

30

40

50

## 【 0 1 1 6 】

なお、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

## 【 0 1 1 7 】

また、上述の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。

## 【 0 1 1 8 】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。

## 【 0 1 1 9 】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

## 【 0 1 2 0 】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。

## 【 0 1 2 1 】

なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

【 図 1 】

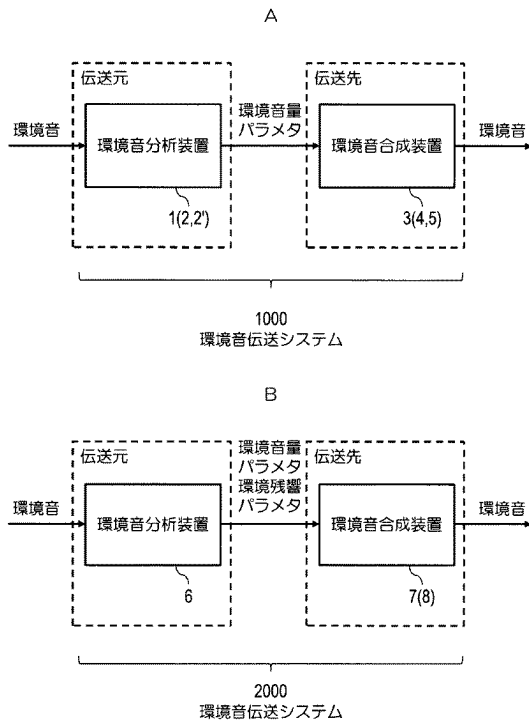


図1

【 図 2 】

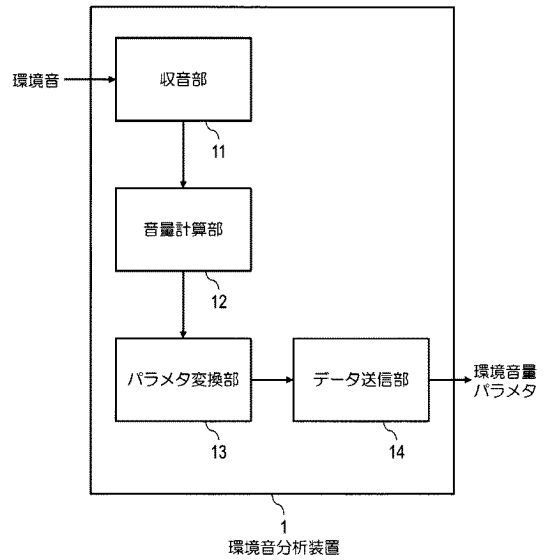


図2

【 図 3 】

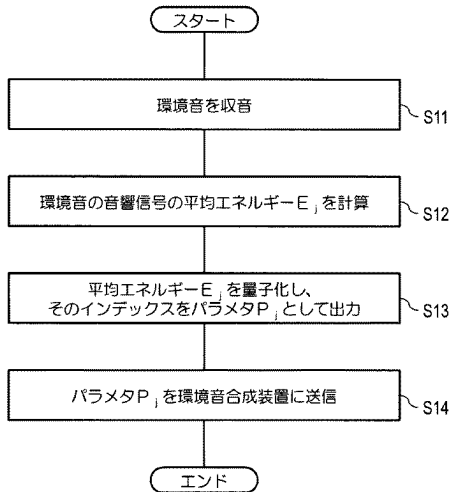


図3

【 図 4 】

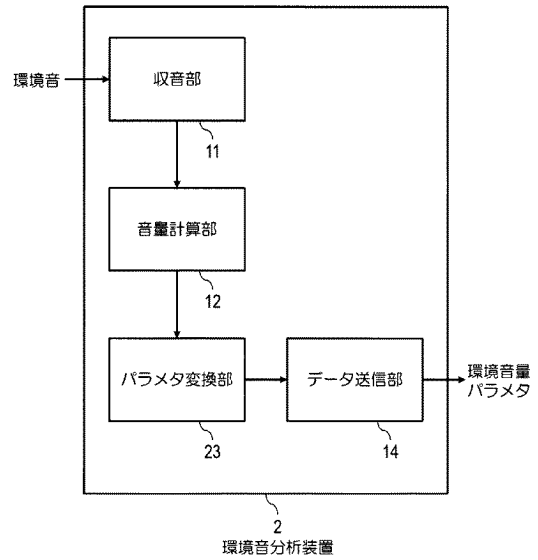


図4

【 図 5 】

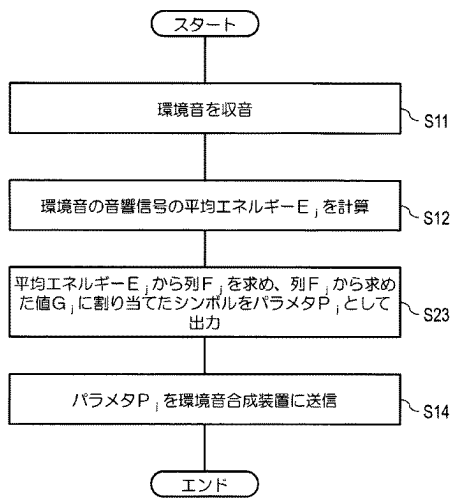


図5

【 図 6 】

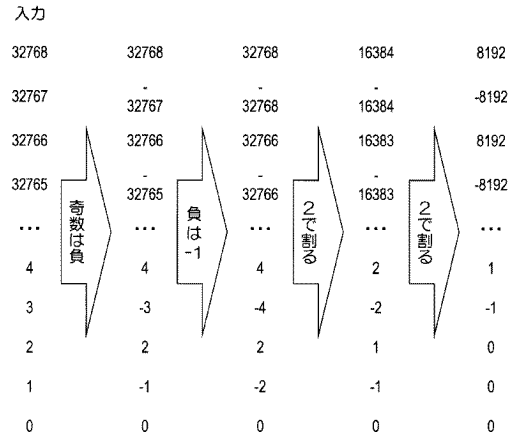


図6

【 図 7 】

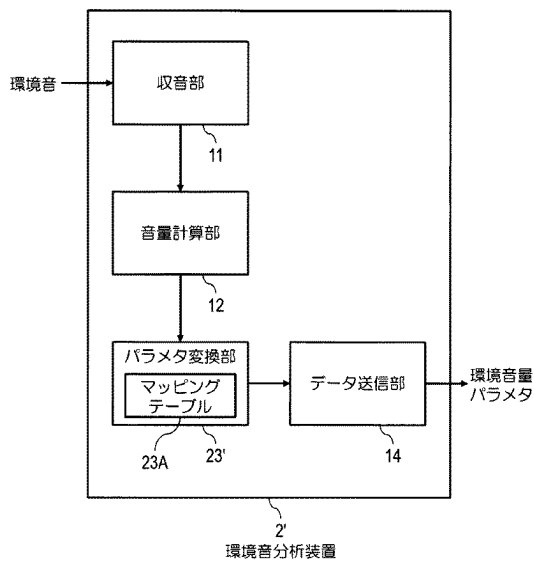


図7

【 図 8 】

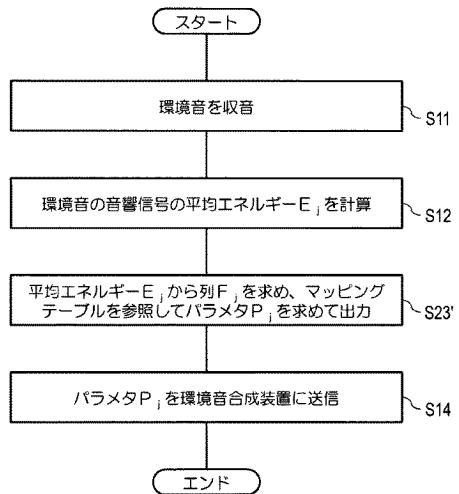


図8

【 図 9 】

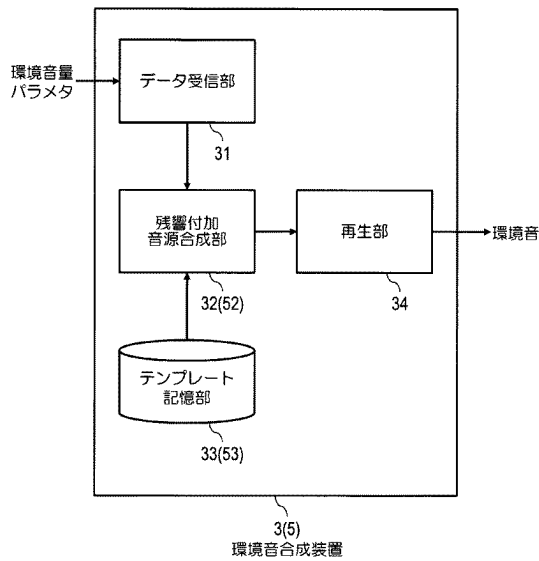


図9

【 図 1 0 】

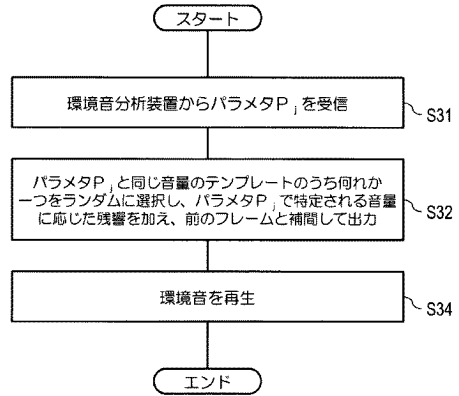


図10

【 図 1 1 】

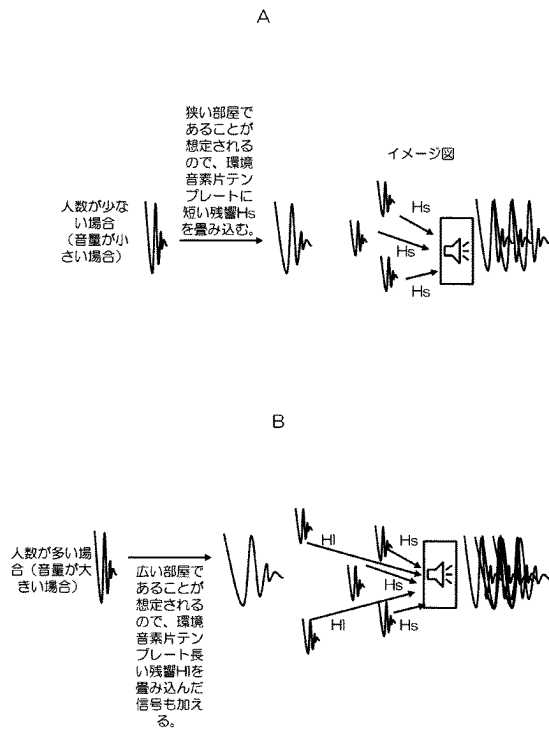


図11

【 図 1 2 】

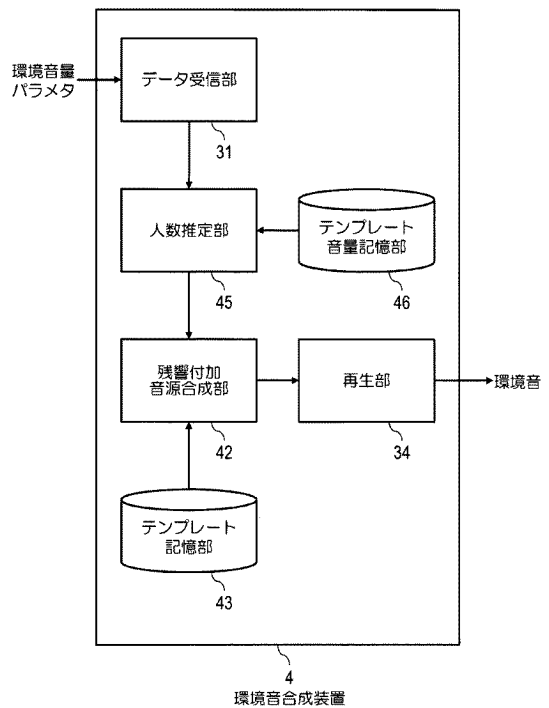


図12

【 図 1 3 】

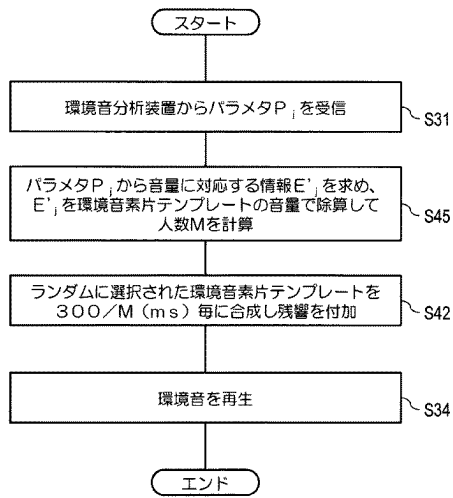


図13

【 図 1 4 】

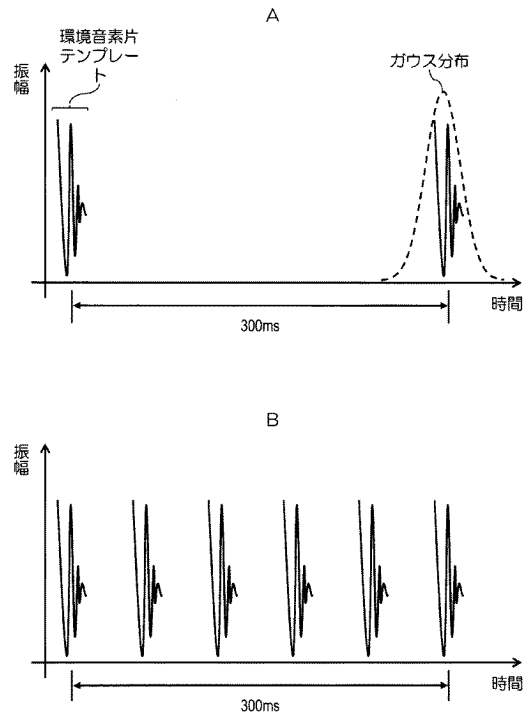


図14

【 図 1 5 】

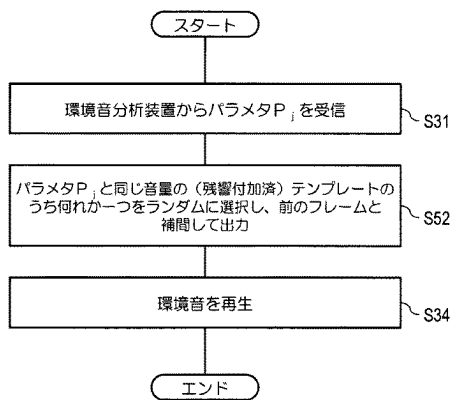


図15

【 図 1 6 】

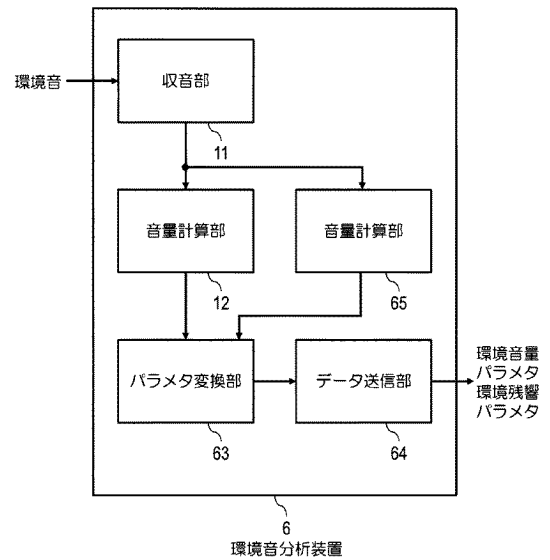


図16



【 図 1 7 】

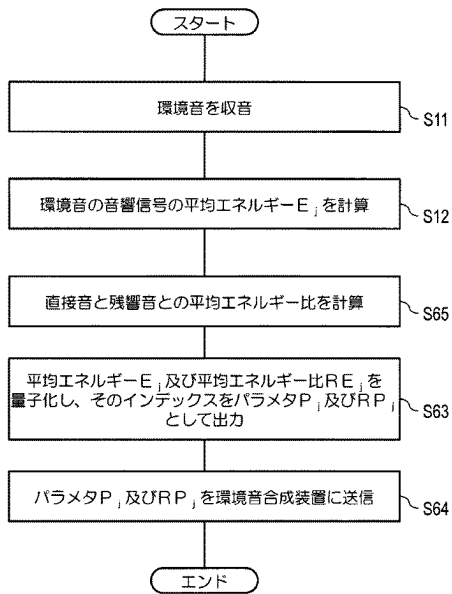


図17

【 図 1 8 】

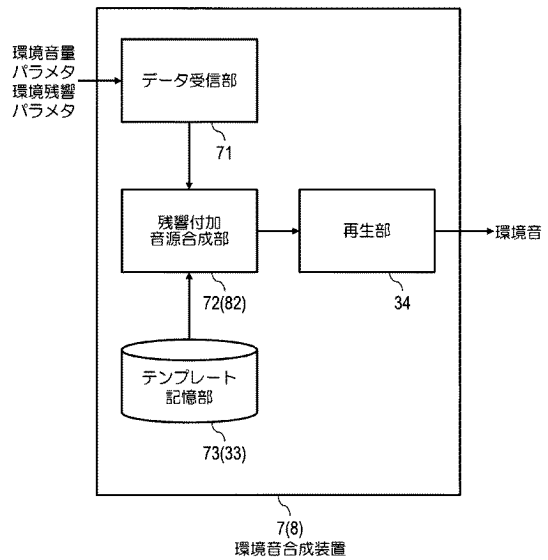


図18

【 図 1 9 】

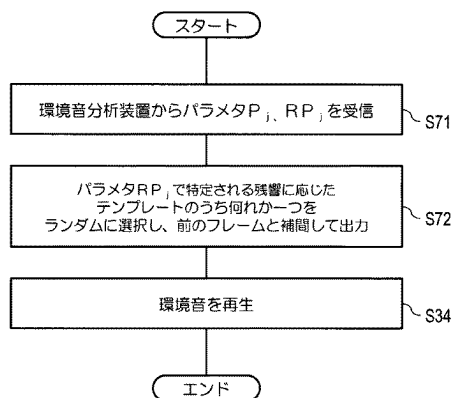


図19

【 図 2 0 】

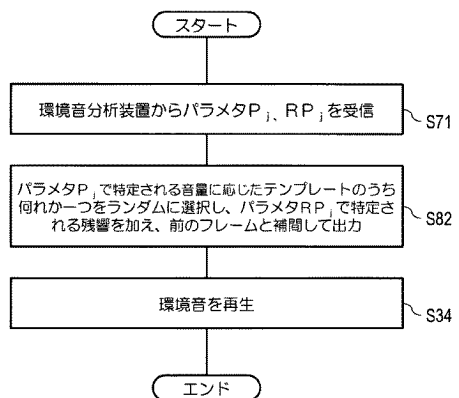


図20

## フロントページの続き

- (72)発明者 守谷 健弘  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 佐藤 尚  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 杉浦 亮介  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 白木 善史  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 川西 隆仁  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 野口 賢一  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 堤 公孝  
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
- (72)発明者 河原 一彦  
福岡県福岡市東区箱崎六丁目10番1号 国立大学法人九州大学内
- (72)発明者 藤森 朗穂  
福岡県福岡市東区箱崎六丁目10番1号 国立大学法人九州大学内
- (72)発明者 尾本 章  
福岡県福岡市東区箱崎六丁目10番1号 国立大学法人九州大学内

審査官 菊池 智紀

- (56)参考文献 特表2008-513845(JP,A)  
特開2014-63145(JP,A)  
特開平9-127983(JP,A)  
特開平10-340080(JP,A)  
特表2008-503793(JP,A)  
特開2003-322560(JP,A)  
特開2004-294712(JP,A)  
特開2000-148139(JP,A)  
特開平6-59669(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00 - 19/26  
G10K 15/00 - 15/12