

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6443841号
(P6443841)

(45) 発行日 平成30年12月26日 (2018.12.26)

(24) 登録日 平成30年12月7日 (2018.12.7)

(51) Int.Cl. F I
G 1 O L 25/57 (2013.01) G 1 O L 25/57
G 1 O L 19/00 (2013.01) G 1 O L 19/00 3 1 2 F

請求項の数 15 (全 31 頁)

(21) 出願番号	特願2015-86567 (P2015-86567)	(73) 特許権者	000004226
(22) 出願日	平成27年4月21日 (2015.4.21)		日本電信電話株式会社
(65) 公開番号	特開2016-206372 (P2016-206372A)		東京都千代田区大手町一丁目5番1号
(43) 公開日	平成28年12月8日 (2016.12.8)	(73) 特許権者	504145342
審査請求日	平成29年7月26日 (2017.7.26)		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】 環境音伝送システム及び環境音伝送方法

(57) 【特許請求の範囲】

【請求項1】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含み、

前記映像特徴量取得部は、前記映像特徴量として、会場の明るさを表す情報を得、
 前記環境音合成部は、前記会場の明るさを表す情報に対応する情報に基づき、会場が明るいほど、再生時の音量が大きくなるように、環境音を合成する、
 環境音伝送システム。

【請求項2】

請求項1の環境音伝送システムであって、
 前記映像特徴量取得部は、会場の観客を撮像した映像の信号の平均エネルギー E_j (j はフレームのインデックスを表す1以上の整数) を、前記会場の明るさを表す情報として得、
 前記平均エネルギー E_j から

【数 1 2】

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

で表される列 F_j を計算し、前記列 F_j を 2 回、2 で割って値 G_j を求め、前記値 G_j を符号に変換して前記値 G_j の各々にシンボルを割り当て、前記値 G_j に割り当てられたシンボルを j 番目のフレームの前記会場の明るさを表す情報に対応する情報として求めるパラメタ変換部をさらに含む、

環境音伝送システム。

10

【請求項 3】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含み、

P は 2 以上の整数とし、 $p=1, 2, \dots, P$ とし、前記映像特徴量取得部は、会場の観客を撮像した映像を P 個のブロック p に分割し、ブロック p 毎に前記映像特徴量として、会場の明るさを表す情報を得、

ブロック p 毎の前記会場の明るさを表す情報と所定の閾値とを比較し、所定の閾値を超えたブロックの個数に基づき、観客の人数を推定する人数推定部と、

20

一人の人間による拍手及び手拍子の少なくとも何れかの一拍分の環境音のテンプレートである環境音素片テンプレートを記憶するテンプレート記憶部とをさらに含み、

前記環境音合成部は、前記テンプレート記憶部から前記環境音素片テンプレートを選択し、当該選択された環境音素片テンプレートを、前記人数が増えるに従って小さくなるように定められた時間間隔ごとに配置して環境音を合成する、

環境音伝送システム。

【請求項 4】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、

30

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含み、

Q は 2 以上の整数とし、 $q=1, 2, \dots, Q$ とし、前記映像特徴量取得部は、会場の観客を撮像した映像を Q 個のブロック q に分割し、ブロック q 毎に前記映像特徴量として、会場の明るさを表す情報を得、

所定の時間区間毎に、ブロック q 毎の前記会場の明るさを表す情報から、会場の明るさの時間的変化の周期を得る周期取得部をさらに含み、

前記環境音合成部は、前記周期が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、前記周期が予め定めた閾値よりも大きい場合には、前記周期で拍子をとる手拍子に対応する環境音を合成する、

40

環境音伝送システム。

【請求項 5】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含み、

前記映像特徴量取得部は、前記映像特徴量として、会場の観客を撮像した映像を圧縮符号化する際に用いる動き補償ベクトルを得、

50

前記環境音合成部は、前記動き補償ベクトルに対応する情報に基づき、拍手又は手拍子に対応する環境音を合成する、
環境音伝送システム。

【請求項 6】

請求項 5 の環境音伝送システムであって、
前記動き補償ベクトルの方向の相関を取得する相関取得部をさらに含み、
前記環境音合成部は、前記相関が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、前記相関が予め定めた閾値よりも大きい場合には、手拍子に対応する環境音を合成する、
環境音伝送システム。

10

【請求項 7】

請求項 5 の環境音伝送システムであって、
所定の時間区間毎に、前記動き補償ベクトルの大きさの時間的変化の周期を得る周期取得部をさらに含み、
前記環境音合成部は、前記周期が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、前記周期が予め定めた閾値よりも大きい場合には、前記周期で拍子をとる手拍子に対応する環境音を合成する、
環境音伝送システム。

【請求項 8】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、

20

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含み、

前記映像特徴量取得部は、前記映像特徴量として、会場の観客を撮像した映像を符号化して得られる符号列から、当該符号列の所定時間毎の符号量を得、
前記環境音合成部は、前記符号量に基づき、環境音を合成する、
環境音伝送システム。

【請求項 9】

請求項 8 の環境音伝送システムであって、
所定の時間区間毎に、前記符号量の時間的変化の周期を得る周期取得部をさらに含み、
前記環境音合成部は、前記周期が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、前記周期が予め定めた閾値よりも大きい場合には、前記周期で拍子をとる手拍子に対応する環境音を合成する、
環境音伝送システム。

30

【請求項 10】

請求項 8 の環境音伝送システムであって、
所定の時間区間毎に、前記符号量の平均を得る平均取得部をさらに含み、
前記環境音合成部は、前記平均が予め定めた閾値以下の場合には手拍子に対応する環境音を合成し、前記平均が予め定めた閾値よりも大きい場合には、拍手に対応する環境音を合成する、
環境音伝送システム。

40

【請求項 11】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含み、

前記映像特徴量取得ステップは、前記映像特徴量として、会場の明るさを表す情報を得

50

前記環境音合成ステップは、前記会場の明るさを表す情報に対応する情報に基づき、会場が明るいほど、再生時の音量が大きくなるように、環境音を合成する、環境音伝送方法。

【請求項 1 2】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含み、

Pは2以上の整数とし、 $p=1,2,\dots,P$ とし、前記映像特徴量取得ステップは、会場の観客を撮像した映像をP個のブロックpに分割し、ブロックp毎に前記映像特徴量として、会場の明るさを表す情報を得、

ブロックp毎の前記会場の明るさを表す情報と所定の閾値とを比較し、所定の閾値を超えたブロックの個数に基づき、観客の人数を推定する人数推定ステップと、

テンプレート記憶部には、一人の人間による拍手及び手拍子の少なくとも何れかの一拍分の環境音のテンプレートである環境音素片テンプレートが記憶されるものとし、

前記環境音合成ステップは、前記テンプレート記憶部から前記環境音素片テンプレートを選択し、当該選択された環境音素片テンプレートを、前記人数が増えるに従って小さくなるように定められた時間間隔ごとに配置して環境音を合成する、

環境音伝送方法。

【請求項 1 3】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含み、

Qは2以上の整数とし、 $q=1,2,\dots,Q$ とし、前記映像特徴量取得ステップは、会場の観客を撮像した映像をQ個のブロックqに分割し、ブロックq毎に前記映像特徴量として、会場の明るさを表す情報を得、

所定の時間区間毎に、ブロックq毎の前記会場の明るさを表す情報から、会場の明るさの時間的変化の周期を得る周期取得ステップをさらに含み、

前記環境音合成ステップは、前記周期が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、前記周期が予め定めた閾値よりも大きい場合には、前記周期で拍子をとる手拍子に対応する環境音を合成する、

環境音伝送方法。

【請求項 1 4】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含み、

前記映像特徴量取得ステップは、前記映像特徴量として、会場の観客を撮像した映像を圧縮符号化する際に用いる動き補償ベクトルを得、

前記環境音合成ステップは、前記動き補償ベクトルに対応する情報に基づき、拍手又は手拍子に対応する環境音を合成する、

環境音伝送方法。

【請求項 1 5】

会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、

前記映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音

10

20

30

40

50

響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含み、

前記映像特徴量取得ステップは、前記映像特徴量として、会場の観客を撮像した映像を符号化して得られる符号列から、当該符号列の所定時間毎の符号量を得、

前記環境音合成ステップは、前記符号量に基づき、環境音を合成する、環境音伝送方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、伝送元である会場の環境音を、伝送先で再生する技術に関する。

10

【背景技術】

【0002】

伝送元において收音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再現する従来技術として特許文献1が知られている。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2014-63145号公報

【発明の概要】

【発明が解決しようとする課題】

20

【0004】

しかしながら、従来技術では、伝送元において環境音を收音する必要がある。

【0005】

本発明は、必ずしも伝送元において環境音を收音しなくとも、伝送先で伝送元の場の雰囲気再現することができる環境音伝送システム及び環境音伝送方法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】

【0006】

上記の課題を解決するために、本発明の一態様によれば、環境音伝送システムは、会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得部と、映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成部と、を含む。

30

【0007】

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、環境音伝送方法は、会場の観客を撮像した映像からその映像の特徴量である映像特徴量を得る映像特徴量取得ステップと、映像特徴量に対応する情報に基づき、事前に記憶された、観客の行動に対応する音響信号を選択、及び/又は、加工し、遠隔地側で再生するための環境音を合成する環境音合成ステップと、を含む。

【発明の効果】

40

【0008】

本発明によれば、必ずしも伝送元において環境音を收音しなくとも、伝送先で伝送元の場の雰囲気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】

【0009】

【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。

50

- 【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4の環境音分析装置の構成を示すブロック図。
- 【図12】実施例4の環境音分析装置の動作を示すフローチャート。
- 【図13】1フレームの映像信号をP個のブロックpに分割した例を示す図。
- 【図14】実施例5及びその変形例の環境音分析装置の構成を示すブロック図。
- 【図15】実施例5及びその変形例の環境音分析装置の動作を示すフローチャート。 10
- 【図16】実施例5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17】実施例6の環境音分析装置の構成を示すブロック図。
- 【図18】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19】1フレームの映像信号をQ個のブロックqに分割した例を示す図。
- 【図20】実施例7の環境音合成装置の構成を示すブロック図。
- 【図21】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22】実施例8の環境音分析装置の構成を示すブロック図。
- 【図23】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。
- 【図24】実施例9の環境音合成装置の構成を示すブロック図。
- 【図25】実施例9の環境音合成装置の動作を示すフローチャート。 20
- 【図26】実施例10の環境音分析装置の構成を示すブロック図。
- 【図27】実施例10の環境音分析装置の動作を示すフローチャート。
- 【図28】実施例11の環境音合成装置の構成を示すブロック図。
- 【図29】実施例11の環境音合成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】

【0010】

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。

【0011】

<本実施形態のポイント>

本実施形態の環境音分析装置は、伝送元で収録された映像信号に基づき、伝送元の環境、雰囲気、盛り上がりの度合いまたは熱気（以下、「環境等」ともいう）を推定し、推定した環境等に基づき、伝送先で環境音に変換し、再生する。なお、映像信号そのものを送るのではなく、環境等に対応する環境パラメタを伝送することで、伝送効率を高めながら、伝送先で伝送元の環境等を再現することができる。

【0012】

本実施形態の環境音合成装置は、環境音分析装置から伝送元の環境等に対応する環境パラメタを取得して環境音を合成する。

【0013】

ペンライト及び発光するウチワなどの発光体の輝度は利用者が多いほど大きくなる。本実施形態では、ペンライト等による観客の動きそのものを伝送するのではなく、映像の特徴量である映像特徴量に対応する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、映像特徴量に対応する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）を合成し、再生する。

【0014】

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、拍手の場合、一拍と一拍の時間間隔（以下、拍手間隔ともいう）は、個人差に応じたゆらぎを持ち、200ms～400ms程度（1秒間に2～5回手をたたく）である。一方、手拍子の場合、拍手間隔は、手拍子の対象に応じて変化し、個人差に応じたゆらぎは少ない。例えば、曲のテンポが148BPM（Beats Per Minute）の場合、 50

各観客の手拍子は405ms前後となり、個人差は小さい。また、手拍子の場合、手拍子の対象に応じて手を叩き、その時間間隔が揃うため、手拍子の時間間隔の標準偏差は、拍手の時間間隔の標準偏差よりも小さい。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、拍手の場合、拍手間隔を200ms~400msの範囲にし、ある程度ゆらぎ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。また、手拍子の場合、拍手間隔を手拍子の対象の変化(例えば曲のテンポ)に合わせて、ほぼ等間隔で、ゆらぎをあまり持たせずに繰り返し再生することにより、別の人の手拍子音に類似した音を構成することができる。

【0015】

<環境音伝送システム>

以下、図1を参照して本実施形態の環境音伝送システムについて説明する。図1は本実施形態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、本実施形態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。環境音分析装置は、入力された映像信号から映像特徴量に対応する情報(環境パラメタ P_j 、以下単にパラメタともいう)を抽出し、出力する。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、入力された環境パラメタ P_j に合わせてテンプレートを変換することにより環境音を合成し、出力する。

【実施例1】

【0016】

本実施例では、撮像の対象となる会場として、発光体が発光することで明るくなるような会場(客席の照明を落とした屋内会場や夜間の野外コンサート会場等)を想定する。そのような会場の観客(客席)を撮像した映像を用いて、多くの観客が発光体を提示すると(例えばペンライトを頭上で振ったり、電源をONにすると)、会場が盛り上がっていると推定し、発光体の提示が少ないと(ペンライトを下げてしまったり、電源をOFFにすると)、会場が盛り上がっていないと推定する。また、会場が盛り上がっているほど、拍手、手拍子、声援及び掛け声等の環境音の音量は大きくなると想定する。そこで、本実施例では、伝送元の会場が明るいほど、伝送先で再生する環境音の音量が大きくなるようにする。

【0017】

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、映像変換部11と、画像変化量計算部12と、パラメタ変換部13と、データ送信部14とを備える。

【0018】

<映像変換部11>

映像変換部11は、伝送元の風景を映像信号に変換し、映像信号を収録し、出力する(S11)。ここでは、映像変換部11には伝送元の客席風景が入力されるものとし、映像信号は会場の観客を撮像したものであるとする。例えば、映像変換部11はビデオカメラ等からなり、映像信号はRGB等の画素値からなる。

【0019】

<画像変化量計算部12>

画像変化量計算部12は、映像信号を受け取り、画素値を取得する。画像変化量計算部12が取得する画素値は、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フレームの映像信号とし、 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ (Nはフレームあたりのサンプル数)とする。例えば30fpsサンプリングのときに1フレーム500msとすると、 $N = 15$ 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。画像変化量計算部12は、フレーム毎に、入力された画素値の輝度に対応する値(以下、「盛り上がり度合いに対応する値」ともいう)を求めて出力する。具体的には、

10

20

30

40

50

画像変化量計算部 1 2 は、フレーム毎に、入力された映像信号 $X_j = (x_j(1), x_j(2), \dots, x_j(N))$ の平均エネルギー

【0020】

【数1】

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

【0021】

を計算する (S12)。撮像した客席が明るければこの値は大きくなり、暗ければこの値は小さくなる。よって、この値 E_j は、映像信号の特徴量(以下「映像特徴量」ともいう)と言え、会場の明るさを表す情報である。画像変化量計算部 1 2 は映像信号から映像特徴量を得るため、画像変化量計算部のことを映像特徴量取得部ともいう。

10

【0022】

<パラメタ変換部 1 3 >

パラメタ変換部 1 3 は、画像変化量計算部 1 2 から出力された盛り上がり度合いに対応する値を取得する。パラメタ変換部 1 3 は、取得した盛り上がり度合いに対応する値を量子化し、環境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パラメタ P_j として出力する (S13)。

20

【0023】

<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 が出力した環境パラメタ P_j を伝送先の環境音合成装置 3 に送信する (S14)。環境音合成装置 3 については実施例 3 に記載する。

【0024】

このように、本実施例の環境音分析装置 1 によれば、伝送元において推定された環境等を効率よく低遅延に伝送することができる。

【実施例 2】

30

【0025】

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、映像変換部 1 1 と、画像変化量計算部 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。映像変換部 1 1、画像変化量計算部 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。

【0026】

40

<画像変化量計算部 1 2 >

画像変化量計算部 1 2 は、30 fps サンプルングでサンプル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。画像変化量計算部 1 2 は、フレーム毎に、入力された拍手映像信号 $X_j = (x_j(1), x_j(2), \dots, x_j(6))$ から、平均エネルギー

【0027】

【数2】

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

50

【0028】

を計算する(S12)。

【0029】

<パラメタ変換部23>

パラメタ変換部23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。

【0030】

【数3】

$$F_j = \begin{cases} -\sqrt{E_j} - 1 & \left(\sqrt{E_j} \text{が奇数の場合} \right) \\ \sqrt{E_j} & \left(\sqrt{E_j} \text{が偶数の場合} \right) \end{cases} \dots (1) \quad 10$$

【0031】

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(0~32768)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_j 。711準拠の範囲に収めるために、 $\mu-law$ 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_j に対して、ITU-T G.711の符号化処理を行い、 G_j をG.711の符号(番号)に変換する。30fpsサンプリング6サンプル分を一塊(1フレーム)にすると5fpsの1サンプル分に相当するので、上記の G_j 毎にG.711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(S23)。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。

【0032】

【数4】

$$F_j = \begin{cases} -\log E_j - 1 & \left(\log E_j \text{が奇数の場合} \right) \\ \log E_j & \left(\log E_j \text{が偶数の場合} \right) \end{cases} \dots (2) \quad 30$$

【0033】

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。よって、パラメタ P_j (割り当てられたシンボル列)は、会場の明るさを表す情報(平均パワー E_j)に対応する情報である。

【0034】

[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23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、映像変換部11と、画像変化量計算部12と、パラメタ変換部23'と、データ送信部14とを備える。映像変換部11、画像変化量計算部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。

【0035】

<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8ビット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。

【0036】

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。画素値を用いた場合、収録された映像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 16 bit 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 16 bit 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。

10

【実施例 3】

【0037】

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、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') から伝送元の会場の明るさを表す情報に対応する環境パラメタを取得して環境音を合成する装置である。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。

20

【0038】

<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パラメタ P_j を受信する (S31)。

【0039】

<テンプレート記憶部 33>

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 の画像変化量計算部 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。

30

【0040】

<音源合成部 32>

音源合成部 32 は、入力された環境パラメタ P_j と同じ音量大きさの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。音源合成部 32 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 (S32)。例えば、20ms のフレームあたり環境パラメタに 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。

40

【0041】

言い換えると、音源合成部 32 は、会場の明るさを表す情報 (映像特徴量) に対応する情報である環境パラメタ P_j に基づき、テンプレート記憶部 33 に事前に記憶された、観客の行動 (本実施例では「拍手」) に対応する音響信号を選択し、観客の行動に対応する環境音を合成する。なお、音源合成部を環境音合成部ともいう。本実施例では、環境パラ

50

メタ P_j に基づき、会場が明るいほど、再生時の音量が大きくなるように、環境音を合成する。

【0042】

<再生部34>

再生部34は、音源合成部32が合成した拍手音を再生する(S34)。例えば、再生部34としてスピーカを用いる。

【0043】

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、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。

10

【0044】

[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、伝送元で収録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、伝送元の環境パラメタが入力され、伝送先で環境音を合成してもよい。

【0045】

20

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、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。

【0046】

<実施例1~3の変形例>

実施例1~3では、パラメタ P_j を送信しているが、他の情報を伝送する構成としてもよい。これらの実施例のポイントは、会場の観客を撮像した映像から、遠隔地側で再生するための環境音を合成する点である。例えば、環境音分析装置は、映像変換部及びデータ送信部のみを含み、映像信号を環境音合成装置に送信する。その場合、環境音合成装置はデータ受信部、画像変化量計算部、パラメタ変換部、音源合成部、テンプレート記憶部及び再生部とを含み、映像信号から環境音を合成し、再生する。また、例えば、環境音分析装置は、映像変換部、画像変化量計算部、パラメタ変換部、音源合成部、テンプレート記憶部及びデータ送信部を含み、合成した環境音を環境音合成装置に送信する。環境音合成装置はデータ受信部及び再生部のみを含み、環境音を再生する。つまり、どの時点のデータを伝送するかは、適宜設定すればよく、伝送するデータに応じて、画像変化量計算部、パラメタ変換部、音源合成部及びテンプレート記憶部を、伝送元(環境音分析装置)及び伝送先(環境音合成装置)の何れに設ければよい。ただし、伝送量を減らすという観点からみれば、実施例1~3のようにパラメタ P_j を伝送するのがよい。

30

【実施例4】

【0047】

40

本実施形態では、発光体の数に合わせて環境音を合成する。実施例1~3では、映像信号全体の明るさを表す情報から提示されている発光体の数を推定していたが、本実施形態では、映像信号を複数のブロックに分割し、ブロック毎に明るさを表す情報を求め、発光体の数を推定する。そして、本実施形態では、会場が盛り上がっているほど、発光体を提示する人の数が増え、映像中の発光体の個数が増えると推定し、会場が盛り上がっているほど、拍手や手拍子、掛け声等の環境音の発生回数が増えると想定する。

【0048】

以下、図11、図12を参照して本発明の実施例4の環境音分析装置について説明する。図11は本実施例の環境音分析装置4の構成を示すブロック図である。図12は本実施例の環境音分析装置4の動作を示すフローチャートである。

50

【 0 0 4 9 】

図 1 1 に示すように、本実施例の環境音分析装置 4 は、映像変換部 1 1 と、画像変化量計算部 4 2 と、人数推定部 4 3 と、データ送信部 1 4 とを備える。映像変換部 1 1 とデータ送信部 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省略する。

【 0 0 5 0 】

< 画像変化量計算部 4 2 >

画像変化量計算部 4 2 は、会場の観客を撮像した映像信号を受け取り、1 フレームの映像信号を P 個のブロック p に分割し、ブロック p 毎に映像特徴量として、会場の明るさを表す情報を得、出力する。ただし、P は 2 以上の整数とし、 $p=1, 2, \dots, P$ とする。例えば、図 1 3 は、1 フレームの映像信号を 16 個のブロックに分割した例を示す。具体的には、画像変化量計算部 4 2 は、フレーム j 毎、ブロック p 毎に、入力された映像信号 $X_{j,p} = (x_{j,p}(1), x_{j,p}(2), \dots, x_{j,p}(N))$ の平均エネルギー

【 0 0 5 1 】

【数 5】

$$E_{j,p} = \frac{x_{j,p}(1)^2 + x_{j,p}(2)^2 + \dots + x_{j,p}(N)^2}{N}$$

【 0 0 5 2 】

を計算する (S 4 2)。ブロック p が明るければこの値は大きくなり、暗ければこの値は小さくなる。よって、この値 $E_{j,p}$ は、映像特徴量と言え、ブロック p 毎の会場の明るさを表す情報である。

【 0 0 5 3 】

< 人数推定部 4 3 >

人数推定部 4 3 は、ブロック p 毎の平均エネルギー $E_{j,p}$ を受け取り、平均エネルギー $E_{j,p}$ と所定の閾値とを比較し、所定の閾値を超えた個数に基づき、観客の人数 M を推定し (S 4 3)、その人数 M をパラメタ P_j として出力する。例えば、一つのブロックに y 人の観客が映るように映像信号を分割した場合には、同様に映像を分割した場合にあるブロックにおいて y 個の発光体が発光したときに得られる平均エネルギーの代表値 (平均値、最頻値、最小値及び最大値等の平均エネルギーを代表する何らかの値) を予め求めておき、その代表値に対応する値を閾値として設定し、(閾値を超えた個数 $\times y$) を観客の人数 M として推定する。

【実施例 5】

【 0 0 5 4 】

以下、図 1 4、図 1 5 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 1 4 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 1 5 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 1 6 は本実施例の音源合成部 5 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 4 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 3 1 と、音源合成部 5 2 と、テンプレート記憶部 5 3 と、再生部 3 4 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。

【 0 0 5 5 】

< テンプレート記憶部 5 3 >

テンプレート記憶部 5 3 には、一人の人間による一拍分の拍手及び手拍子の少なくとも何れか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。例えば、拍手の場合、テンプレートの長さは 200 ms ~ 400 ms 程度とし、手拍子の場合、手拍子の対象に応じて変化し、例えば、曲のテンポが 148 BPM の場合、手拍子のテンプレートの長さは 405 ms 前後とする。本実施例では環境音の例として拍手音及び手拍子音の少なくとも何れかを扱うため、拍手音及び手拍子音の少なくとも何れか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、以下では拍手音又は手拍子音

10

20

30

40

50

の少なくとも何れか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及び手拍子音の何れか（環境音）のテンプレートを指すものとする。なお、テンプレート記憶部 5 3 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をあらかじめ正規化しておく。

【 0 0 5 6 】

< 音源合成部 5 2 >

音源合成部 5 2 は、テンプレート記憶部 5 3 から環境音素片テンプレートをランダムに選択して合成する（S 5 2）。環境パラメタにより一人分の拍手を合成する場合（人数推定部 4 3 において $M = 1$ となった場合）は、図 1 6 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を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、± 数 1 0 m s の揺らぎを持たせればよい。例えば音源合成部 5 2 は

【 0 0 5 7 】

【 数 6 】

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

【 0 0 5 8 】

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 5 2）。別の表現方法で書くと、時系列テンプレート信号 $T_i = (t_i[1] \quad t_i[2] \quad \dots \quad t_i[P])$ と拍手タイミングを表すインパルス $(i \cdot \tau + \sigma_i)$ の畳み込み演算で Y_i を出力とする。

【 0 0 5 9 】

【 数 7 】

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

【 0 0 6 0 】

ここで * は畳み込み演算を表す。ここで、 $\tau = 3 0 0$ m s であり、 σ_i は $- 1 0$ m s $\sigma_i + 1 0$ m s の範囲で生成した乱数である。環境パラメタにより一人分の手拍子を合成する場合、時間間隔の揺らぎは、拍手の場合よりも手拍子の場合のほうが小さく設定したほうがよく、例えば、手拍子の場合の σ_i の範囲が拍手の場合の σ_i の範囲よりも小さくなるように設定する。

【 0 0 6 1 】

環境パラメタにより M 人分の拍手を合成する場合は、図 1 6 B のように、時間間隔を約 3 0 0 / M (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を用いて合成された波形を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 3 0 0 / M (m s) 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。例えば音源合成部 5 2 は、

【 0 0 6 2 】

【 数 8 】

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

【 0 0 6 3 】

10

20

30

40

50

によりテンプレートを変換した環境音 Y_i ($i = 0, 1, 2, \dots$) を出力する (S 52)。環境パラメタにより M 人分の手拍子を合成する場合、M 人の拍手のタイミングはほぼ同じなので、例えば、

【0064】
【数9】

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

【0065】

によりテンプレートを変換した環境音 Y_i ($i = 0, 1, 2, \dots$) を出力する。なお、この場合も時間間隔の揺らぎは、拍手の場合よりも手拍子の場合のほうが小さく設定したほうがよい。

【0066】

このように、実施例4の環境音分析装置4、実施例5の環境音合成装置5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53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5のメモリ量を削減することができる。なお、<実施例1~3の変形例>の場合と同様に、どの時点のデータを伝送するかは、適宜設定すればよく、伝送するデータに応じて、画像変化量計算部、人数類推部、音源合成部及びテンプレート記憶部を、伝送元(環境音分析装置)及び伝送先(環境音合成装置)の何れに設ければよい。ただし、伝送量を減らすという観点からみれば、本実施形態のように人数 M からなるパラメタ P_j を伝送するのがよい。

【0067】

[実施例5の動作例2]

実施例5は、環境音の例として拍手音及び手拍子音の少なくとも何れかを対象とし、ブロック毎の会場の明るさを表す情報を取得して、伝送先で拍手音及び手拍子音の少なくとも何れかを生成する環境音合成装置5を説明したが、これに限らず拍手音及び手拍子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)及び手拍子音の少なくとも何れか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。

【0068】

実施例5の動作例2では、実施例5の環境音合成装置5のデータ受信部31と、音源合成部52と、テンプレート記憶部53と、再生部34において取り扱われるデータが拍手音又は手拍子音から他の環境音(例えば、声援、掛け声)に置き換わる点を除いては、上述の動作例と同じである。

【0069】

なお、音源合成部52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 ($m \cdot \tau + \sigma_m$) の畳み込み演算で Y_i を出力としても良い。

【0070】

【数10】

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

【0071】

ここで*は畳み込み演算を表す。

【実施例6】

【0072】

本実施例では、発光体の時間変化に合わせて環境音を合成する。発光体の時間変化に周

10

20

30

40

50

期がなく、または、周期が短い場合には、拍手が起こっていると判断し、所定の閾値以上の周期で変動する場合には、手拍子が起こっていると判断する。このようにして発光体の時間変化に合わせて環境音を合成する。

【0073】

以下、図17、図18を参照して本発明の実施例6及びその変形例の環境音分析装置について説明する。図17は本実施例及びその変形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18は本実施例及びその変形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17に示すように、本実施例の環境音分析装置6は、映像変換部11と、画像変化量計算部62と、人数推定部43と、周期取得部63と、データ送信部64とを備える。映像変換部11、人数推定部43は実施例1、実施例4の環境音分析装置1、4における同一番号の各構成部と同じであるから適宜説明を略する。

10

【0074】

<画像変化量計算部62>

画像変化量計算部62は、会場の観客を撮像した映像信号を受け取り、実施例4の画像変化量計算部42と同様の処理により、ブロックp毎に映像特徴量として、会場の明るさを表す情報を得、人数推定部43に出力する。また、画像変化量計算部62は、1フレームの映像信号をQ個のブロックqに分割し、ブロックq毎に映像特徴量として、会場の明るさを表す情報を得、周期取得部63に出力する。ただし、Qは2以上の整数とし、 $q=1,2,\dots,Q$ とする。例えば、図19は、1フレームの映像信号を128個のブロックに分割した例を示す。具体的には、画像変化量計算部42は、フレームj毎、ブロックq毎に、入力された映像信号 $X_{j,q} = (x_{j,q}(1), x_{j,q}(2), \dots, x_{j,q}(N))$ の平均エネルギー

20

【0075】

【数11】

$$E_{j,q} = \frac{x_{j,q}(1)^2 + x_{j,q}(2)^2 + \dots + x_{j,q}(N)^2}{N}$$

【0076】

を計算する(S62)。ブロックqが明るければこの値は大きくなり、暗ければこの値は小さくなる。よって、この値 $E_{j,q}$ は、映像特徴量と言え、ブロックq毎の会場の明るさを表す情報である。

30

【0077】

ブロックの大きさは、観客が発光体を振ったときに、発光体がブロック内からブロック外に出たり、ブロック外からブロック内に入ったりする大きさとする。例えば、あるブロックに対して、発光体を右に振ったときには発光体の位置がブロック外となり、発光体を左に振ったときにはブロック内になるようにブロックの大きさを設定する。例えば、ブロックは、対象(観客)に対して、(1)肩幅の半分以下、(2)発光体を振るときの振れ幅の半分以下、(3)一人一人の幅の半分以下等とすればよく、例えば、20~30cm四方以下とする。映像変換部11(例えばビデオカメラ)から対象までの距離と(例えばビデオカメラの)ズーム率から自動的に上述の大きさとなるようにブロックの大きさを決定してもよい。

【0078】

40

<周期取得部63>

周期取得部63は、ブロックq毎の平均エネルギー $E_{j,q}$ を受け取り、所定の時間区間毎に、平均エネルギー $E_{j,q}$ から、ブロックq毎の会場の明るさの時間的変化の周期 T_q を得る。例えば、1秒から数秒程度(例えば90フレーム:3秒)の窓幅で分析を行う。分析したフレームが一定の周期性がある場合には、周期取得部63は、自己相関関数のピーク間隔を周期 T_q とする。さらに、周期取得部63は、ブロックq毎の周期 T_q から分割前の映像信号全体の周期 T を得、パラメタ P_j として出力する(S63)。例えば、Q個の周期 T_q の代表値(平均値、最頻値、最小値及び最大値等のQ個の周期 T_q を代表する何らかの値)を映像信号全体の周期 T として取得する。なお、Q個のブロックの中には、全く発光体が出入りしないブロックが存在しうる。そのようなブロックの周期 T_q に基づき代表値を求

50

めると、発光体の動きに基づく映像信号全体の周期 T_q を得ることができない。そのため、例えば、1秒から数秒程度（例えば90フレーム：3秒）、平均エネルギー $E_{j,q}$ が所定の閾値（発光体がない状態の平均エネルギーに対応する値）より小さい場合には、そのブロックには発光体が入りしていないと判断し、対応する周期 T_q を代表値を算出する際の値から除く。よって、1以上Q個以下のブロックq毎の周期 T_q から映像信号全体の周期 T を得てもよい。

【0079】

<データ送信部64>

データ送信部14は、人数推定部43が出力した環境パラメタ P_j （人数M）と周期取得部63が出力した環境パラメタ P_j （周期 T_j ）とを併せて、伝送先の環境音合成装置7に送信する（S64）。

10

【実施例7】

【0080】

以下、図20、図21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。実施例7の環境音合成装置7は、実施例6の環境音分析装置6に対応する装置である。図20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21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20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31と、音源合成部72と、テンプレート記憶部53と、再生部34とを備える。音源合成部72以外の各構成部は実施例5の環境音合成装置5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。

20

【0081】

<音源合成部72>

音源合成部72は、環境パラメタ P_j （人数Mと周期 T_j ）を受け取り、周期 T_j が予め定めた閾値以下の場合には拍手に対応する環境音を合成し、周期 T_j が予め定めた閾値よりも大きい場合には、周期で拍子をとる手拍子に対応する環境音を合成し（S72）、出力する。例えば、周期 T_j が予め定めた閾値以下の場合（S7）、当該フレームの映像信号が拍手に対応する映像信号であると判定して、実施例5の環境音合成装置5の音源合成部52と同じ処理を行い、拍手音を出力する（S72A）。一方、音源合成部72は、周期 T_j が予め定めた閾値よりも大きい場合（S7）、当該フレームの映像信号が手拍子に対応する映像信号であると判定して、その周期（たとえば500ms）が中心となるような実施例5の時よりも分散の小さいガウス分布やラプラス分布に従う揺らぎを持たせた波形を合成し、拍手音（実際は手拍子音）として出力する（S72B）。音源合成部72は、例えば上記式（3）または（4）において、 T_j の値を周期情報（例えば $T_j = 500\text{ms}$ ）として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（S72B）。この例では、手拍子音における $T_j = 500\text{ms}$ （ms）として前述の拍手音の合成の際に設定された $T_j = 300\text{ms}$ （ms）より長い時間間隔に設定しているため、拍手音と判定された場合と比較して手拍子音と判定された場合の時間間隔が長くなるように環境音素片テンプレートが配置され合成される。

30

【0082】

このように、実施例6の環境音分析装置6と本実施例の環境音合成装置7により構成される環境音伝送システムによれば、周期取得部63がフレームを分析して周期を取得し、音源合成部72において、周期の大小を考慮して拍手、手拍子の何れかの合成音を生成するため、伝送元において収録された映像信号に対応する拍手や手拍子音を効率よく伝送する効果に加えて、拍手、手拍子の双方をより正確に合成することができる。

40

【0083】

<実施例6及び7の変形例>

なお、本実施例では、発光体の時間変化に合わせて環境音を合成する点がポイントである。そのため、必ずしも実施例5の環境音合成装置5の音源合成部52と同じ処理を行い、拍手音を出力する必要はない。例えば、本実施例の環境音分析装置6は、人数推定部43を備えずに、環境パラメタとして人数Mを出力しなくともよい。環境音合成装置7の音

50

源合成部 72 は、M に依らずに、手拍子の周期よりも短い間隔で拍手に対応する環境音を合成すればよい。

【0084】

また、＜実施例 1～3 の変形例＞の場合と同様に、どの時点のデータを伝送するかは、適宜設定すればよく、伝送するデータに応じて、画像変化量計算部、人数類推部、周期取得部、音源合成部及びテンプレート記憶部を、伝送元（環境音分析装置）及び伝送先（環境音合成装置）の何れに設ければよい。ただし、伝送量を減らすという観点からみれば、本実施形態のように人数 M 及び周期 からなるパラメタ P_j を伝送するのがよい。

【実施例 8】

【0085】

本実施例では、映像特徴量として、会場の明るさを表す情報の代わりに、動画圧縮用の特徴量である動き補償ベクトル（参考文献 1 参照）を利用する。

（参考文献 1）村上篤道、浅井光太郎、関口俊一、「高効率映像符号化技術 HEVC/H.265 とその応用」、オーム社、2013 年、p.20-28,125-132

【0086】

一般的な映像信号では、前のフレームと現在のフレームとは全く同じわけではなく、フレーム中の物体が動いたり、ビデオカメラがパンをしたりすることにより、フレーム間で動きがある。動き補償とは、フレーム間予測においてこれらの動きを補うものである。具体的には、現在のフレームを予測する場合に、動きの分だけずらした位置の画像を用いるものである。動き補償を行うためには、画像の動き量を推定する動き補償ベクトル探索が必要になる。動画圧縮符号化する場合には、この動き補償ベクトルも同時に符号化する必要がある。本実施形態では、動き補償ベクトルを利用することで、拍手及び手拍子識別のための追加的計算量を抑える。また、動き補償ベクトルから、集団のまとまった動き（マクロな動作）を推定することができる。

【0087】

なお、実施例 8 及び実施例 9 では、会場の明るさを表す情報に変えて動き補償ベクトルを利用しているため、対象は発光体に限らず、会場は発光体が発光することで明るくなるような会場に限らない。例えば、観客が拍手、手拍子、手を振る映像信号から動き補償ベクトルを取得してもよいし、そのような動きが確認できる昼間の野外コンサート会場であってもよい。

【0088】

以下、図 22、図 23 を参照して本発明の実施例 8 の環境音分析装置について説明する。図 22 は本実施例の環境音分析装置 8 の構成を示すブロック図である。図 23 は本実施例の環境音分析装置 8 の動作を示すフローチャートである。

【0089】

図 22 に示すように、本実施例の環境音分析装置 8 は、映像変換部 11 と、符号化部 82 とデータ送信部 14 とを備える。映像変換部 11、データ送信部 14 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適宜説明を略する。

【0090】

<符号化部 82>

符号化部 82 は、映像信号を受け取り、圧縮符号化し、圧縮符号化する際に用いる動き補償ベクトルを得、動き補償ベクトルを環境パラメタ P_j として出力する（S82）。なお、圧縮符号化する場合、1 フレームの映像信号を複数のブロックに区切って、ブロック毎に動き保障ベクトルを得る。このとき、動体の輪郭部分では動きを細かくみるために、ブロックの大きさを小さくして詳細な情報を取得し、動体以外の背景等ではブロックの大きさを大きくして大まかな情報を取得することで圧縮効率を高める。ブロックの大きさは圧縮符号化の中で自動的に決定される。

【実施例 9】

【0091】

以下、図 24、図 25 を参照して本発明の実施例 9 の環境音合成装置について説明する

10

20

30

40

50

。実施例 9 の環境音合成装置 9 は、実施例 8 の環境音分析装置 8 に対応する装置である。図 2 4 は本実施例の環境音合成装置 9 の構成を示すブロック図である。図 2 5 は本実施例の環境音合成装置 9 の動作を示すフローチャートである。図 2 4 に示すように、本実施例の環境音合成装置 9 は、データ受信部 3 1 と、音源合成部 9 2 と、テンプレート記憶部 3 3 と、再生部 3 4 とを備える。さらに、環境音合成装置 1 1 A は、その合成方法に応じて相関取得部 9 5 または周期取得部 9 6 を備える。

【 0 0 9 2 】

音源合成部 9 2、相関取得部 9 5、周期取得部 9 6 以外の各構成部は実施例 3 の環境音合成装置 7 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。

【 0 0 9 3 】

< 相関取得部 9 5 >

相関取得部 9 5 は、環境パラメタ P_j (動き補償ベクトル) を受け取り、動き補償ベクトルの方向の相関を取得し (s_{95})、出力する。動き補償ベクトルの方向が、映像内でほぼ同じ向きときは、観客が発光体を曲の拍子に合わせて振っていると考えられるため、手拍子に対応する環境音を合成する。その際、相関は大きくなる。一方、動き補償ベクトルの方向が、映像内で方向、長さがバラバラのときは、発光体を曲の拍子とは関係なく、各観客は異なるタイミングで発光体を振っていると考えられるため、拍手に対応する環境音を合成する。その際、相関は小さくなる。以下、相関の取得方法について説明する。

【 0 0 9 4 】

(相関の取得方法 1)

動き補償ベクトルあるいは動き補償ベクトルをいくつかのブロックにわたって平均化したもの (以下まとめて動き補償ベクトルとする) の (2次元の) 和をとりその絶対値 (あるいは長さ) を利用する。絶対値が大きい場合は動画中の物体の動きの量が多くかつ向きがそろっていると推定でき、相関が大きいと判断できる。逆に、小さい場合は物体の動きが小さいかまたは向きがそろっていないと推定でき、相関が小さいと判断できる。

【 0 0 9 5 】

(相関の取得方法 2)

(相関の取得方法 1) における和の計算の際に、動き補償ベクトルの長さを正規化する。これによって動きが小さい場合でも識別が可能になる。

【 0 0 9 6 】

(相関の取得方法 3)

任意の 2 つの動き補償ベクトルあるいは動き補償ベクトルをいくつかのブロックにわたって平均化したもの (以下まとめて動き補償ベクトルとする) の内積を任意のペアの数だけ計算し、その和を利用する。和が大きい場合は動画中の物体の動きの量が多くかつ向きがそろっていると推定でき、相関が大きいと判断できる。逆に、小さい場合または負の場合は物体の動きが小さいかまたは向きがそろっていないと推定でき、相関が小さいと判断できる。

【 0 0 9 7 】

(相関の取得方法 4)

(相関の取得方法 3) における内積の計算の際に、動き補償ベクトルの長さを正規化する。これによって動きが小さい場合でも識別が可能になる。

【 0 0 9 8 】

< 周期取得部 9 6 >

周期取得部 9 6 は、環境パラメタ P_j (動き補償ベクトル) を受け取り、所定の時間区間毎に、動き補償ベクトルの大きさの時間的変化の周期を取得し (S_{96})、出力する。例えば、全ての動き補償ベクトルの大きさを加算し、その総和について、1 秒から数秒程度 (例えば 90 フレーム : 3 秒) の窓幅で分析を行う。総和はある程度大きな周期をもつ場合は、観客が発光体を曲の拍子に合わせて振っている (発光体の移動のタイミングと停止のタイミングが一致する) と考えられるため、手拍子に対応する環境音を合成する。また、総和は周期を持たないか、または、小さな周期となる場合には、各観客が任意のタイ

10

20

30

40

50

ミングで発光体を曲の拍子に合わせずに振っていると考えられるため、拍手に対応する環境音を合成する。なお、必ずしも全ての動き補償ベクトルの大きさを加算する必要はなく、所定の個数以上の発光体の動きが分かればよい。対応する範囲の動き補償ベクトルの大きさを加算したものをいれればよい。

【0099】

<音源合成部92>

音源合成部92は、環境パラメタ P_j （動き補償ベクトル）に対応する情報（動き補償ベクトルの方向の相関、または、動き補償ベクトルの大きさの時間的変化の周期）を受け取り、この情報に基づき、拍手又は手拍子に対応する環境音を合成し（S92）、出力する。例えば、音源合成部92は、動き補償ベクトルの方向の相関、または、動き補償ベクトルの大きさの時間的変化の周期が予め定めた閾値 T_1 以下の場合には拍手に対応する環境音を合成し、動き補償ベクトルの方向の相関、または、動き補償ベクトルの大きさの時間的変化の周期が予め定めた閾値 T_1 よりも大きい場合には、手拍子に対応する環境音を合成する。なお、拍手及び手拍子に対応する環境音の合成方法は、<実施例6及び7の変形例>の部分で説明した方法により、人数 M を用いずに、音源合成部72と同様の方法で行えばよい。

10

【0100】

例えば、閾値 T_1 は以下のようにして設定する。まず、拍手している観客が映っている映像信号から動き補償ベクトルを求め、動き補償ベクトルの方向の相関、及び、動き補償ベクトルの大きさの時間的変化の周期を求める。同様に、手拍子している観客が映っている映像信号から相関及び周期を求める。拍手している観客が映っている映像信号から求めた相関及び周期と、手拍子している観客が映っている映像信号から求めた相関及び周期とから、それぞれこれらの値を区分できる値（境界となる値、例えば平均値）を求め、閾値 T_1 として設定する。

20

【0101】

なお、全ての動き補償ベクトルの大きさを加算したものの（前述の総和）が、閾値 T_2 より小さい場合には、動いておらず、観客は拍手及び手拍子をしていないと考えられるため、拍手及び手拍子をしていない状況に対応する環境音を合成する構成としてもよい。この場合、周期取得部96から周期を求める過程で利用した全ての動き補償ベクトルの大きさを加算したものの（総和）を受け取る構成としてもよいし、音源合成部92が動き補償ベクトルを受け取り総和を求める構成としてもよい。

30

【0102】

「動き補償ベクトルの方向の相関」を見るということは空間方向の違いを見ること意味し、「動き補償ベクトルの大きさの時間的変化」を見るということは時間方向の動きの見ることを意味する。

【0103】

なお、<実施例1~3の変形例>の場合と同様に、どの時点のデータを伝送するかは、適宜設定すればよく、伝送するデータに応じて、画像変化量計算部、相関取得部、周期取得部、音源合成部及びテンプレート記憶部を、伝送元（環境音分析装置）及び伝送先（環境音合成装置）の何れに設ければよい。ただし、伝送量を減らすという観点からみれば、本実施形態のように動き補償ベクトルからなるパラメタ P_j を伝送するのがよい。ただし、従来の動画配信サービス等では、動画の符号を送る際に、動き補償ベクトルと一緒に伝送する。そこで、画像変化量計算部を伝送先（環境音合成装置）に設ければ、従来の動画配信サーバを環境音分析装置として利用することができる。

40

【実施例10】

【0104】

本実施例では、映像特徴量として、会場の明るさを表す情報の代わりに、動画を符号化したときの符号量を利用する。実施例8及び実施例9で述べたように動き補償ベクトルと拍手及び手拍子の識別には相関があると考えられる。また、動き補償ベクトルと符号量にも相関がある。そこで本実施例では、動き補償ベクトルの符号量の変化から、集団のまと

50

まった動き（マクロな動作）を推定する。

【0105】

なお、本実施例では、映像変換部11の向きが変化すると符号量が増加するため、固定とする。ただし、映像変換部11の向きが変化した際の符号量の増加分を差し引くことで本実施例の構成を適用することができるため、必ずしも映像変換部11の向きを固定する必要はない。

【0106】

なお、本実施例では、映像特徴量として、会場の明るさを表す情報の代わりに、動画を符号化したときの符号量を利用するため、実施例8及び実施例9と同様、対象は発光体に限らず、会場は発光体が発光することで明るくなるような会場に限らない。

10

【0107】

以下、図26、図27を参照して本発明の実施例10の環境音分析装置について説明する。図26は本実施例の環境音分析装置10の構成を示すブロック図である。図27は本実施例の環境音分析装置10の動作を示すフローチャートである。

【0108】

図26に示すように、本実施例の環境音分析装置8は、映像変換部11と、符号化部102とデータ送信部14とを備える。

【0109】

<符号化部102>

符号化部102は、映像信号を受け取り、圧縮符号化し、その結果、得られる符号列から所定時間毎の符号量を得、符号量を環境パラメタ P_j として出力する(S102)。なお、あるフレーム内ですべての動き補償ベクトルが同じ向きで同じ大きさであるとき、符号量が小さくなると期待できる。よって、符号量が小さいときは手拍子に対応する環境音を合成する。逆に、動き補償ベクトルが向き及び大きさが揃っていないとき、符号量が大きくなると期待できる。よって、符号量が大きいときは拍手に対応する環境音を合成する。

20

【実施例11】

【0110】

以下、図28、図29を参照して本発明の実施例11の環境音合成装置について説明する。実施例11の環境音合成装置11Aは、実施例10の環境音分析装置10に対応する装置である。図28は本実施例の環境音合成装置11Aの構成を示すブロック図である。図29は本実施例の環境音合成装置11Aの動作を示すフローチャートである。図29に示すように、本実施例の環境音合成装置11Aは、データ受信部31と、音源合成部112と、テンプレート記憶部33と、再生部34とを備える。さらに、環境音合成装置11Aは、その合成方法に応じて平均取得部115または周期取得部116を備える。

30

【0111】

音源合成部112、平均取得部115、周期取得部116以外の各構成部は実施例3の環境音合成装置7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。

【0112】

<平均取得部115>

平均取得部115は、環境パラメタ P_j （符号量）を受け取り、所定の時間区間毎に、符号量の平均を得(S115)、出力する。この場合、中期的な符号量の増減を利用する。あるフレーム内ですべての動き補償ベクトルが同じ向きで同じ大きさであるとき、符号量が小さくなると期待できる。すなわち、動画内の物体がそろって動く際は符号量が小さくなる。中期的な（1秒～数秒）符号量の平均値が小さい場合は、動きはそろっており、すなわち手拍子だと判断できる。逆に、符号量の平均値が大きい場合は、動きはそろっておらず、すなわち拍手であると判断できる。1秒から数秒程度（例えば90フレーム：3秒）単位で、符号量の平均値を求める。

40

【0113】

<周期取得部116>

50

周期取得部 1 1 6 は、所定の時間区間毎に、環境パラメタ P_j (符号量) を受け取り、符号量の時間的変化の周期を得 (S 1 1 6)、出力する。この場合、短期的な符号量の増減を利用する。短期的 (~500 msec) で符号量を観察したときに、手拍子のようなそろった動きであれば、その動きの大きさに合わせて周期的な増減がみられると期待できる。逆に、周期的な増減がない場合、または、周期が小さい場合は、拍手であると識別できる。そこで、例えば500 msec以下の窓幅で分析を行い、周期を得る。

【0 1 1 4】

<音源合成部 1 1 2 >

音源合成部 1 1 2 は、環境パラメタ P_j (符号量) に対応する情報 (所定の時間区間毎の符号量の平均、または、符号量の時間的変化の周期) を受け取り、この情報に基づき、拍手又は手拍子に対応する環境音を合成し (S 1 1 2)、出力する。以下、二つの合成方法を説明する。

【0 1 1 5】

(合成方法 1)

この場合、音源合成部 1 1 2 は、環境パラメタ P_j (符号量) に対応する情報として、所定の時間区間毎の符号量の平均を受け取る。音源合成部 1 1 2 は、平均が予め定めた閾値 T_1 以下の場合には手拍子に対応する環境音を合成し、平均が予め定めた閾値 T_1 よりも大きい場合には、拍手に対応する環境音を合成する。

【0 1 1 6】

(合成方法 2)

この場合、音源合成部 1 1 2 は、環境パラメタ P_j (符号量) に対応する情報として、符号量の時間的変化の周期を受け取る。音源合成部 1 1 2 は、周期が予め定めた閾値 T_1 以下の場合には拍手に対応する環境音を合成し、周期が予め定めた閾値 T_1 よりも大きい場合には、周期で拍子をとる手拍子に対応する環境音を合成する。

【0 1 1 7】

なお、拍手及び手拍子に対応する環境音の合成方法は、音源合成部 9 2 と同様の方法で行えばよい。

【0 1 1 8】

例えば、閾値 T_1 は以下のように設定する。拍手している観客が映っている映像信号を符号化して、その符号量を求め、所定の時間区間毎の符号量の平均、または、符号量の時間的変化の周期を求める。同様に、手拍子している観客が映っている映像信号から平均及び周期を求める。拍手している観客が映っている映像信号から求めた平均及び周期と、手拍子している観客が映っている映像信号から求めた平均及び周期とから、それぞれこれらの値を区分できる値 (境界となる値、例えば平均値) を求め、閾値 T_1 として設定する。

【0 1 1 9】

なお、符号量が、極端に小さい場合には、映像に動きが見られず、観客は拍手及び手拍子をしていないと考えられるため、拍手及び手拍子をしていない状況に対応する環境音を合成する構成としてもよい。例えば、閾値 T_1 よりも小さい閾値 T_2 を設定する。例えば、閾値 T_2 は以下のように設定する。観客は拍手及び手拍子をしていない映像信号を符号化して、その符号量を求め、その符号量に対応する値を閾値 T_2 に設定すればよい。音源合成部 1 1 2 が符号量を受け取り判断する構成としてもよい。

【0 1 2 0】

「符号量の中長期間の平均」を見るということは空間方向の違いを見ること意味し、「符号量の時間的変化」を見るということは時間方向の動きの見ることを意味する。

【0 1 2 1】

符号量は、観客の動きの変化量に対応する値なので、観客の動きが大きい場合に会場が盛り上がっていると判断できる場合には、実施例 1 ~ 3 において、会場の明るさに変えて、符号量を用いて、会場の環境等を推定し、環境音を合成してもよい。

【0 1 2 2】

<その他の変形例 >

10

20

30

40

50

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。

【0123】

<プログラム及び記録媒体>

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。

【0124】

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。

【0125】

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。

【0126】

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。

【0127】

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。

10

20

30

【図1】

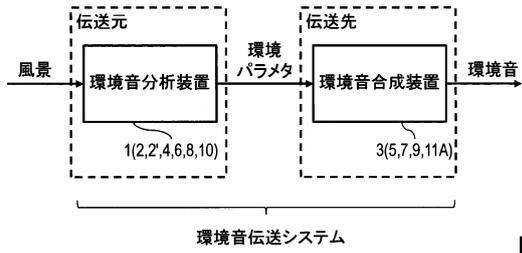


図1

【図2】

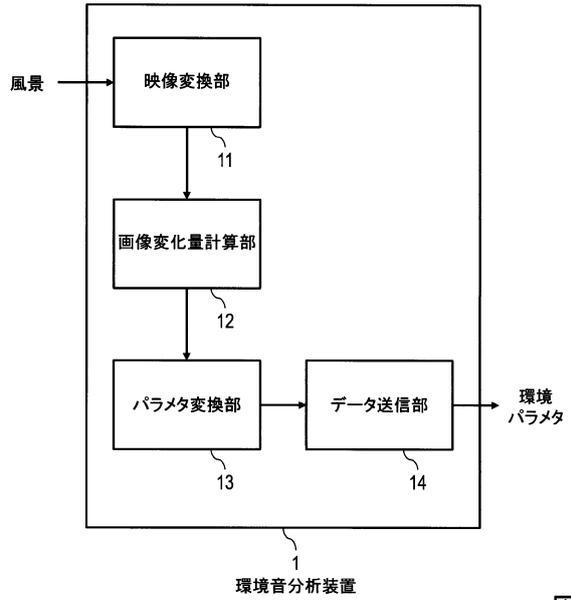


図2

【図3】

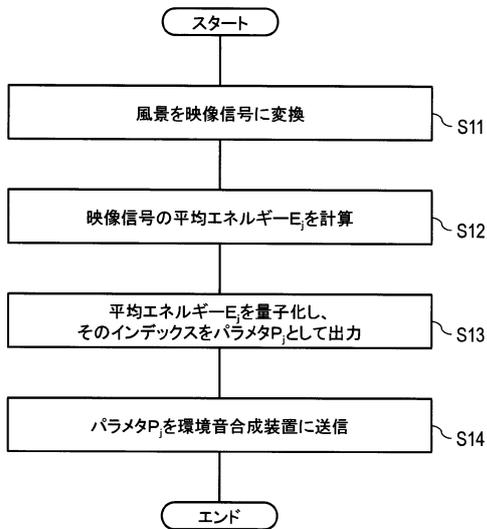


図3

【図4】

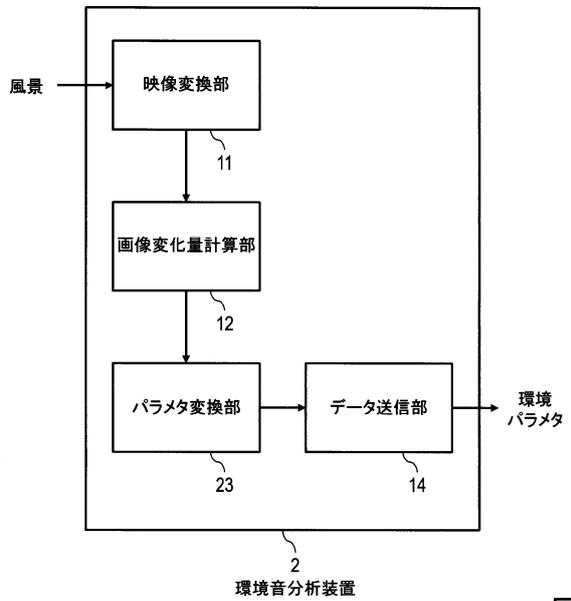


図4

【 図 5 】

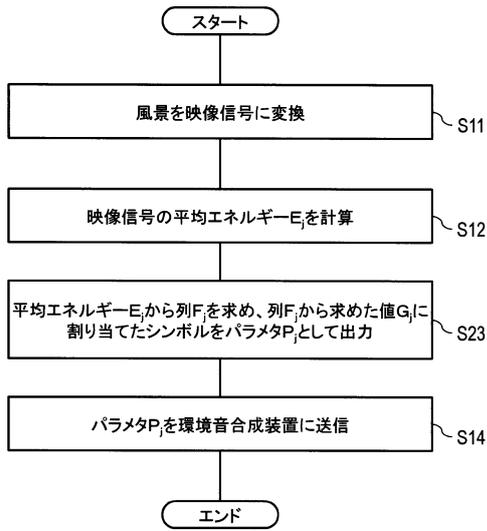


図5

【 図 6 】

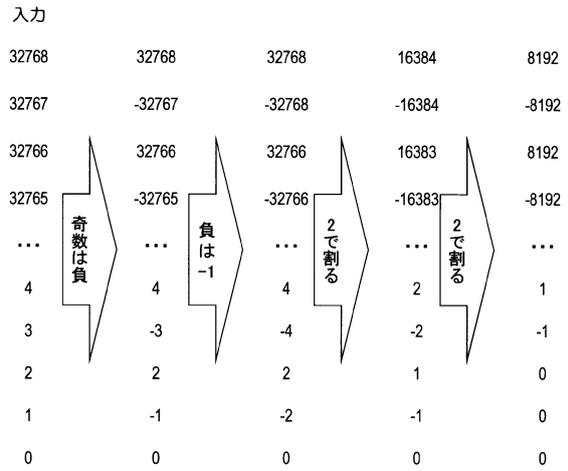


図6

【 図 7 】

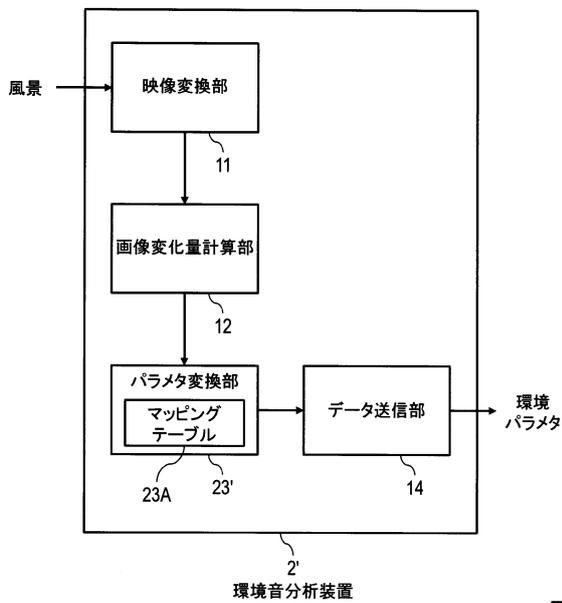


図7

【 図 8 】

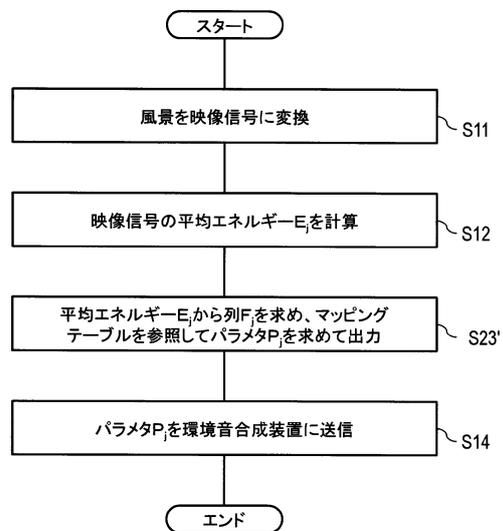


図8

【図9】

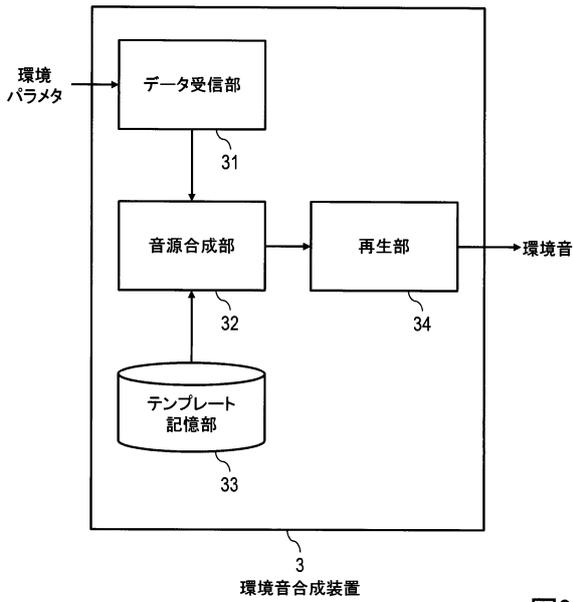


図9

【図10】

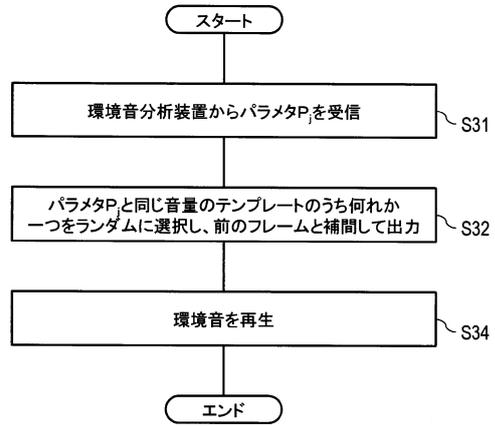


図10

【図11】

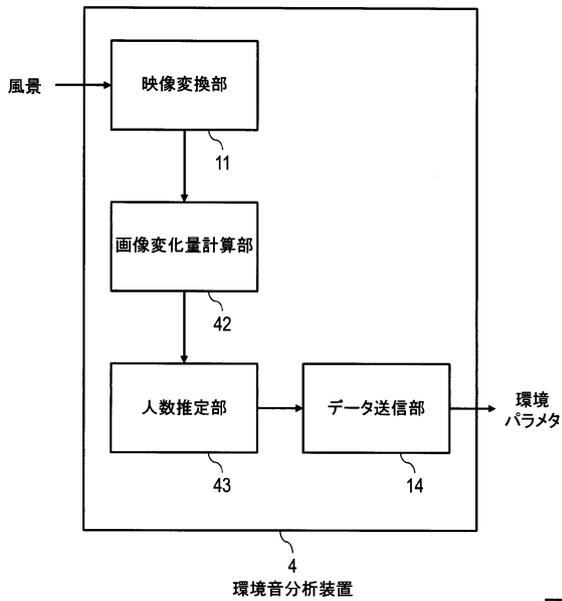


図11

【図12】

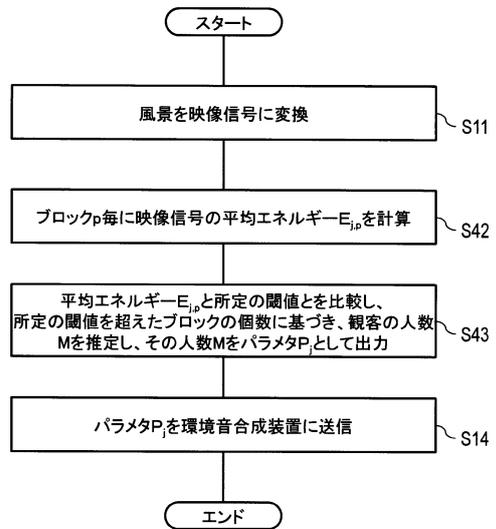


図12

【図13】

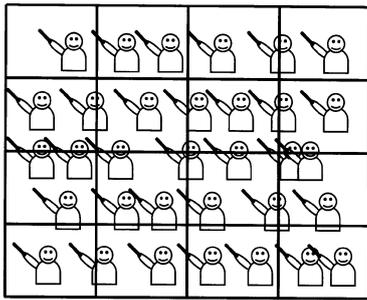


図13

【図14】

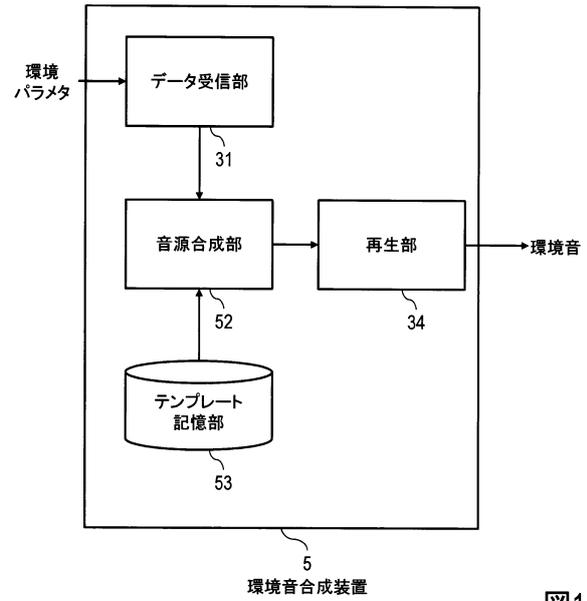


図14

【図15】

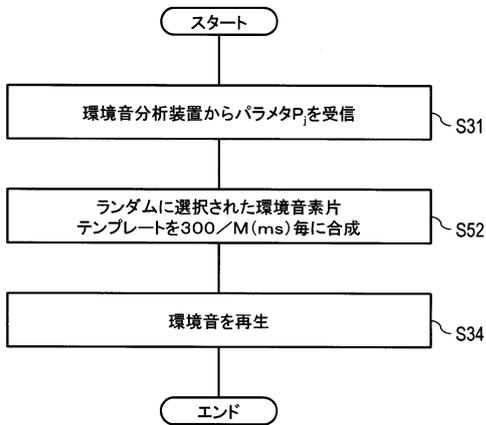


図15

【図16】

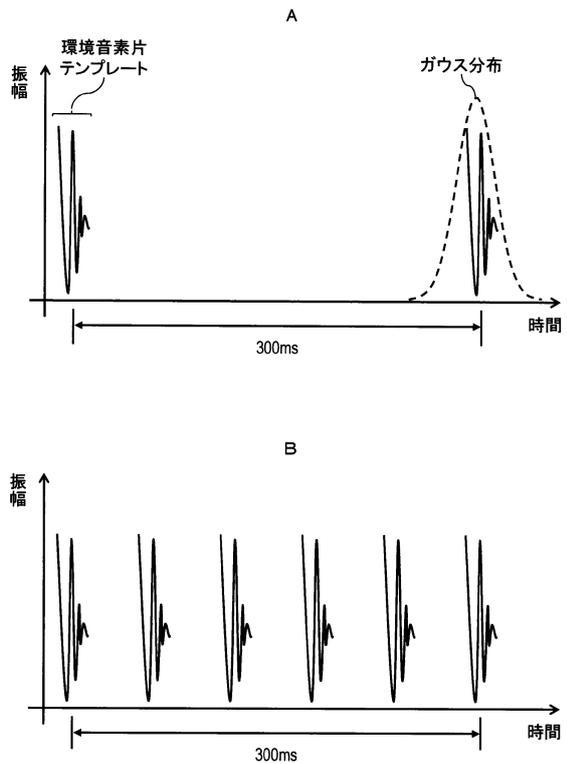


図16

【図17】

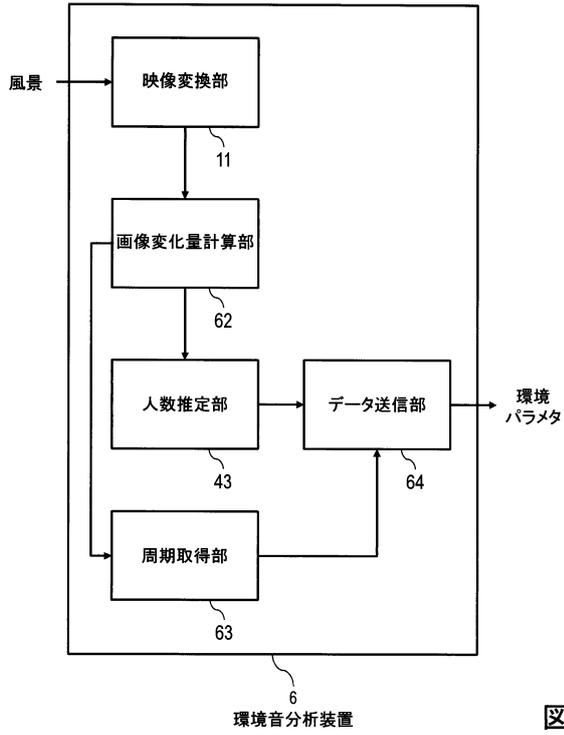


図17

【図18】

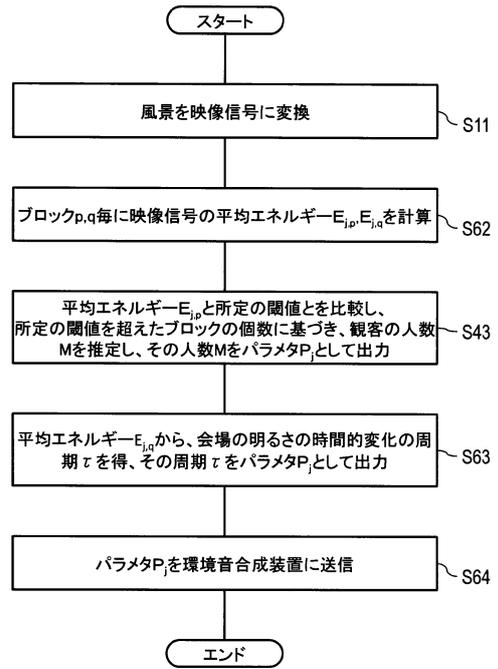


図18

【図19】

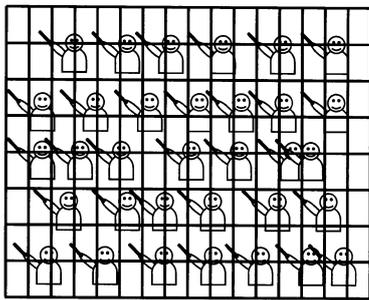


図19

【図20】

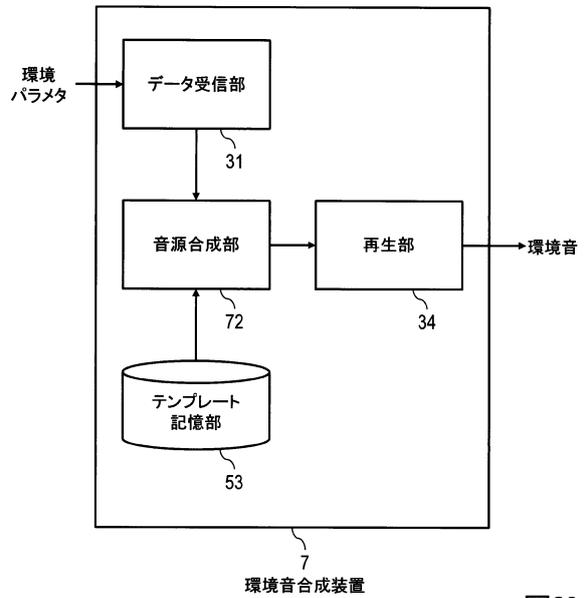


図20

【図21】

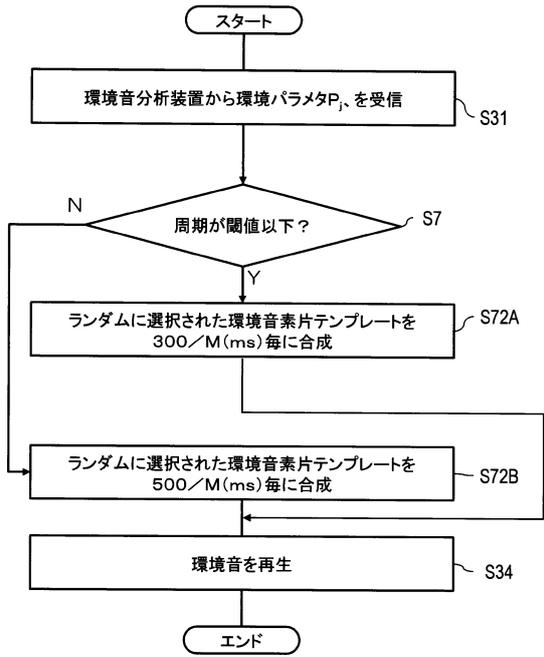


図21

【図22】

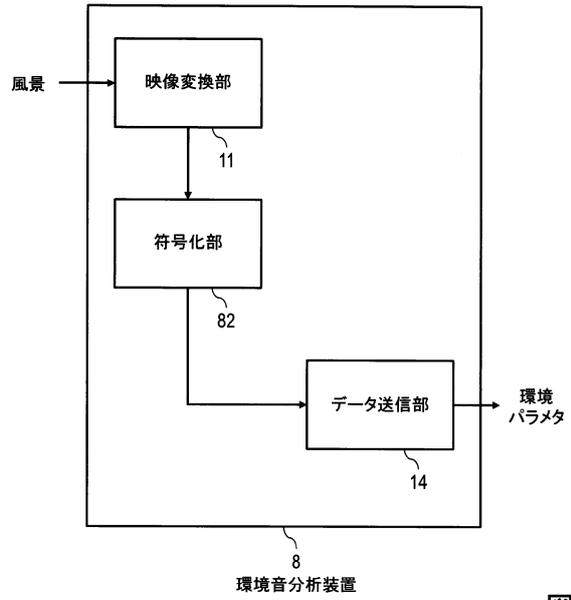


図22

【図23】

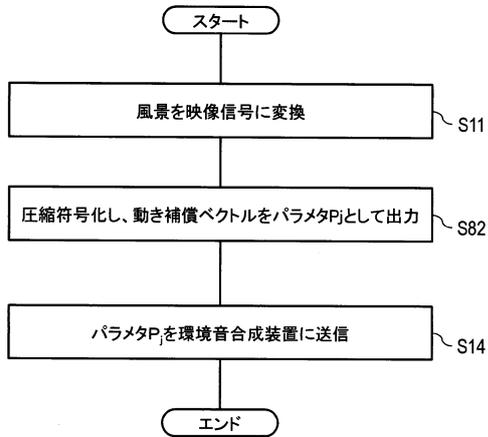


図23

【図24】

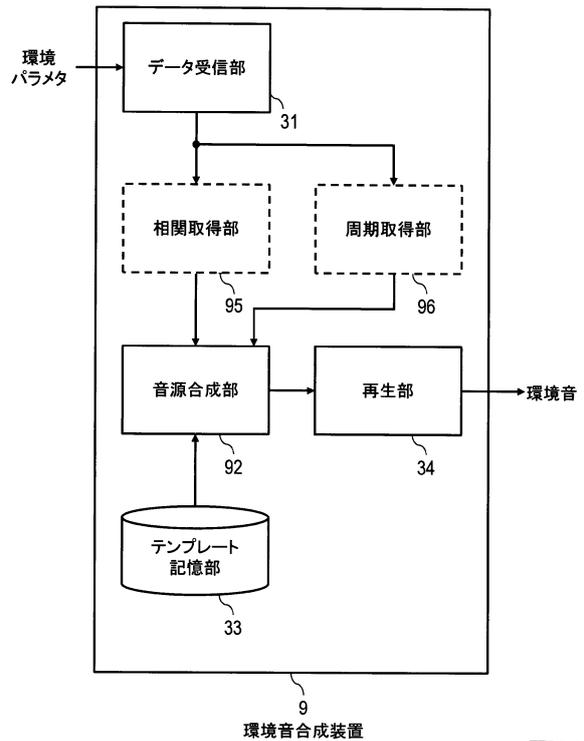


図24

【図25】

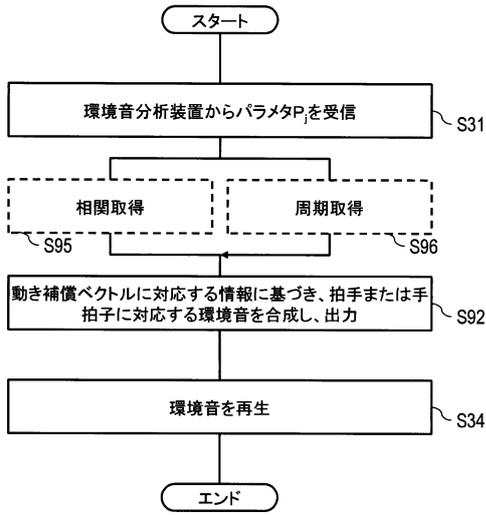


図25

【図26】

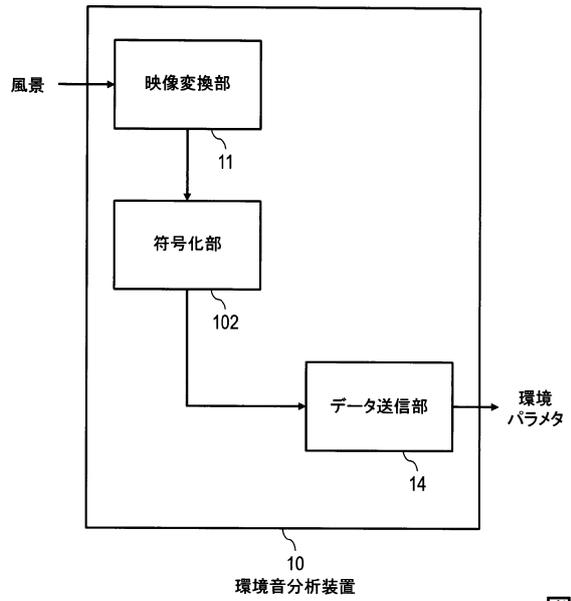


図26

【図27】

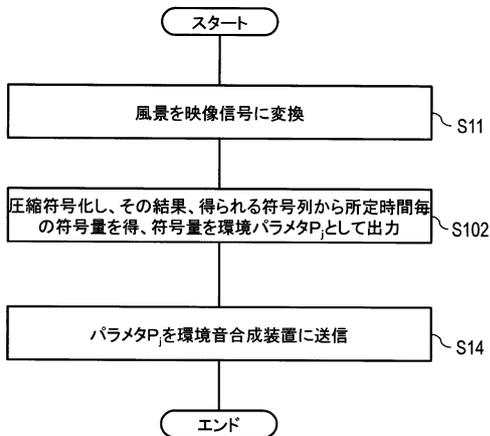


図27

【図28】

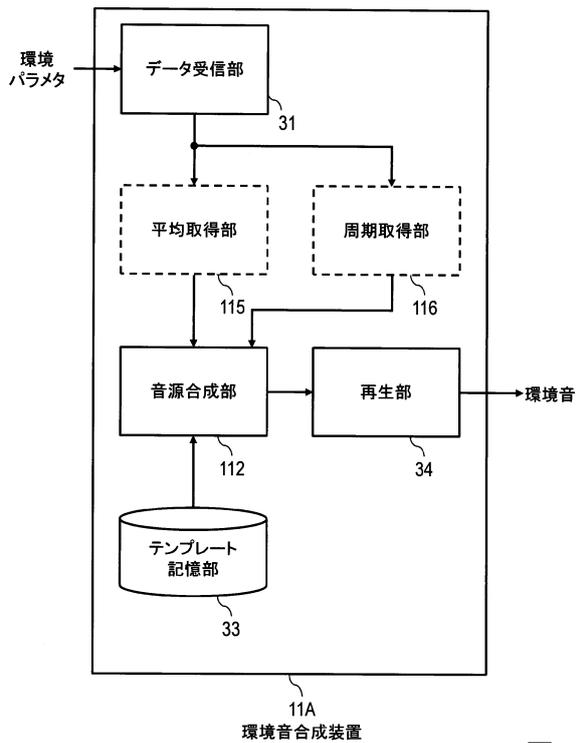


図28

【図29】

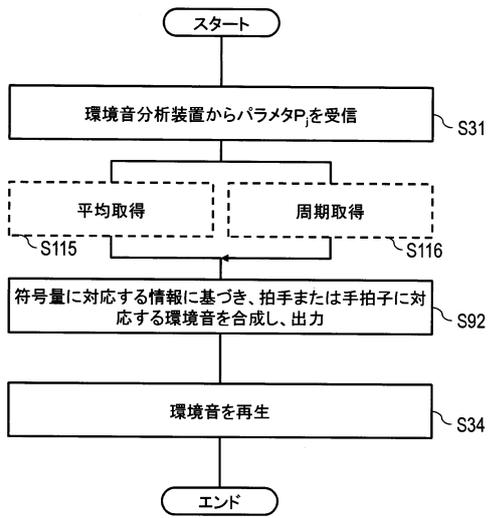


図29

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 原田 登
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 ガブリエル パプロ ナバ
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 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 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 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 上田 雄

- (56)参考文献 特開2004-015523(JP,A)
特開2014-063145(JP,A)
特開2002-109116(JP,A)
特開2010-239245(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 19/00-19/26
G10L 25/00-25/93