

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2016 年 9 月 1 日(01.09.2016)

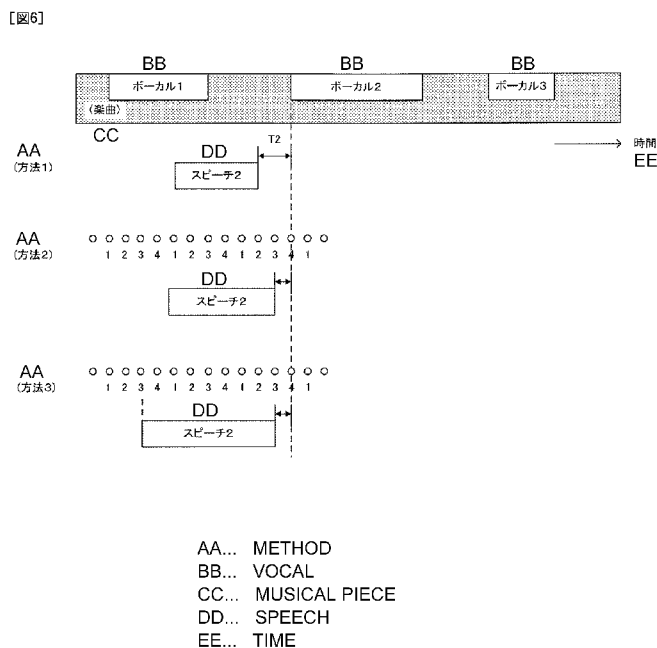


(10) 国際公開番号
WO 2016/135920 A1

- (51) 国際特許分類:
G10K 15/04 (2006.01)
- (21) 国際出願番号: PCT/JP2015/055609
- (22) 国際出願日: 2015 年 2 月 26 日(26.02.2015)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: パイオニア株式会社(PIONEER CORPORATION) [JP/JP]; 〒2120031 神奈川県川崎市幸区新小倉 1 番 1 号 Kanagawa (JP).
- (72) 発明者: 石岡 克仁 (ISHIOKA, Katsuhito); 〒2120031 神奈川県川崎市幸区新小倉 1 番 1 号 パイオニア株式会社内 Kanagawa (JP). 木村 知道 (KIMURA, Tomomichi); 〒2120031 神奈川県川崎市幸区新小倉 1 番 1 号 パイオニア株式会社内 Kanagawa (JP). 菅原 啓太郎(SUGAWARA, Keitaro); 〒2120031 神奈川県川崎市幸区新小倉 1 番 1 号 パイオニア株式会社内 Kanagawa (JP).
- (74) 代理人: 中村 聡延, 外(NAKAMURA, Toshinobu et al.); 〒1040031 東京都中央区京橋一丁目 1 6 番 10 号 オークビル京橋 3 階 東京セントラル特許事務所内 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).
- 添付公開書類:
— 国際調査報告 (条約第 21 条(3))

(54) Title: REPRODUCTION DEVICE, REPRODUCTION METHOD, AND PROGRAM

(54) 発明の名称: 再生装置、再生方法、及び、プログラム



(57) Abstract: This reproduction device outputs, along with a music signal, a lyric voice signal expressing lyrics included in the music signal. More specifically, the reproduction device outputs the lyric voice signal such that the output timing of the lyric voice signal is a predetermined time before the output start timing of a lyric section in the music signal that corresponds to the lyric voice signal. In this way, the lyric voice signal corresponding to a lyric is output a predetermined time before starting to output that lyric in the music signal, and thus, a user can sing along with the music signal after listening to the lyric voice signal and grasping the lyrics.

(57) 要約: 再生装置は、音楽信号に合わせて、音楽信号に含まれる歌詞を表す歌詞音声信号を出力する。具体的に、再生装置は、歌詞音声信号の出力タイミングが、音楽信号中のその歌詞音声信号対応する歌詞区間の出力開始タイミングより一定時間前となるように、歌詞音声信号を出力する。これにより、音楽信号中の歌詞の出力開始より一定時間前に、その歌詞に対応する歌詞音声信号が出力されるので、ユーザは歌詞音声信号を聞いて歌詞を把握してから音楽信号に合わせて歌唱することができる。

明 細 書

発明の名称：再生装置、再生方法、及び、プログラム

技術分野

[0001] 本発明は、楽曲の再生に伴って歌詞の情報を出力する手法に関する。

背景技術

[0002] カラオケの演奏曲に先行して歌詞データを音声合成して出力するカラオケ装置が知られている（例えば、特許文献１、２）。

先行技術文献

特許文献

[0003] 特許文献１：特開４－６７４６７号公報

特許文献２：特開１０－６３２７４号公報

発明の概要

発明が解決しようとする課題

[0004] カラオケ装置の場合、再生される楽曲に歌詞が含まれないため、先行技術により出力される歌詞音声聞き取りにくくなることはない。しかし、カラオケではなく通常の音楽を再生して聞いているような場合には、先行技術の手法により歌詞を音声出力すると、出力された歌詞音声は元の音楽に含まれる歌詞の部分と重なって聞き取りにくくなってしまうことがある。また、例えば車両の運転中に音楽を聞いている場合には、先行技術の手法により出力される歌詞音声は車載用ナビゲーション装置による道案内の音声メッセージなどと重なって聞き取りにくくなってしまうこともある。

[0005] 本発明の解決しようとする課題としては、上記のものが一例として挙げられる。本発明は、歌詞を含む音楽を再生している際に、ユーザがその曲を歌うための歌詞音声を聞き取り易く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] 請求項１に記載の発明は、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置であって、前記歌詞音声信号の出力終了タイミン

グが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段を備えることを特徴とする。

[0007] 請求項4に記載の発明は、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行される再生方法であって、前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力工程を備えることを特徴とする。

[0008] 請求項5に記載の発明は、コンピュータを備え、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行されるプログラムであって、前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段として前記コンピュータを機能させることを特徴とする。

図面の簡単な説明

[0009] [図1]アシストボーカルの概念を示す図である。

[図2]アシストボーカル処理のフローチャートである。

[図3]スピーチ情報生成処理のフローチャートである。

[図4]スピーチ情報生成処理の概要を示す。

[図5]歌詞ブロック化の例を示す。

[図6]スピーチ挿入方法の例を示す。

[図7]スピーチ強調処理の例を示す。

[図8]スピーチ強調処理の他の例に係る構成を示す。

[図9]スピーチ強調処理の他の例に係る構成を示す。

[図10]楽曲再生システムの全体構成を示すブロック図である。

[図11]端末装置の内部構成例を示すブロック図である。

[図12]第1実施例の楽曲再生システムによるアシストボーカル処理のフローチャートである。

[図13]第2実施例の楽曲再生システムによるアシストボーカル処理のフローチャートである。

[図14]スピーチのみを再生するアシストボーカル処理のフローチャートである。

[図15]外部ソースにより再生されている楽曲の特定方法を説明する図である。

発明を実施するための形態

[0010] 本発明の好適な実施形態では、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置は、前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段を備える。

[0011] 上記の再生装置は、音楽信号に合わせて、音楽信号に含まれる歌詞を表す歌詞音声信号を出力する。ここで、再生装置は、歌詞音声信号の出力タイミングが、音楽信号中のその歌詞音声信号対応する歌詞区間の出力開始タイミングより一定時間前となるように、歌詞音声信号を出力する。これにより、音楽信号中の歌詞の出力開始より一定時間前に、その歌詞に対応する歌詞音声信号が出力されるので、ユーザは歌詞音声信号を聞いて歌詞を把握してから音楽信号に合わせて歌唱することができる。

[0012] 上記の再生装置の一態様は、前記音楽信号の拍位置を検出する拍検出手段を備え、前記出力手段は、前記歌詞音声信号の出力終了タイミングを、前記音楽信号の拍位置と一致させる。この態様では、歌詞音声信号の出力が音楽信号の拍位置と一致したタイミングで終了するので、ユーザが音楽信号に合わせて歌唱しやすくなる。

[0013] 上記の再生装置の他の一態様では、前記出力手段は、前記歌詞音声信号の出力開始タイミングを、前記音楽信号の拍位置と一致させる。この態様では、歌詞音声信号の出力が音楽信号の拍位置と一致したタイミングで開始するので、ユーザが音楽信号に合わせて歌唱しやすくなる。

[0014] 本発明の他の好適な実施形態では、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行される再生方法は、前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力工程を備える。この方法によれば、音楽信号中の歌詞の出力開始より一定時間前に、その歌詞に対応する歌詞音声信号が出力されるので、ユーザは歌詞音声信号を聞いて歌詞を把握してから音楽信号に合わせて歌唱することができる。

[0015] 本発明の他の好適な実施形態では、コンピュータを備え、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行されるプログラムは、前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段として前記コンピュータを機能させる。このプログラムを実行することにより、音楽信号中の歌詞の出力開始より一定時間前に、その歌詞に対応する歌詞音声信号が出力されるので、ユーザは歌詞音声信号を聞いて歌詞を把握してから音楽信号に合わせて歌唱することができる。このプログラムは記憶媒体に記憶して取り扱うことができる。好適には、このプログラムは記憶媒体に記憶して取り扱うことができる。

実施例

[0016] 以下、図面を参照して本発明の好適な実施例について説明する。

[0017] [1] アシストボーカル

[1.1] アシストボーカルの概念

車両を運転しているユーザが車内で音楽を再生して聞いている際、聞いている曲を歌いたくなることがある。しかし、運転中は歌詞の情報を見ることができないため、ユーザはその曲の歌詞を記憶していないと歌うことはできない。

[0018] 本実施例では、歌詞を含む楽曲を再生している際に、その楽曲に含まれる

歌詞を音声信号として出力し、ユーザに教える。具体的には、端末装置のメモリなどに記憶されている楽曲を再生している際に、その楽曲に含まれる歌詞を、その歌詞がその楽曲中で再生される前に、音声として出力してユーザに伝える。これにより、ユーザは、運転中であっても、再生中の曲を歌うことができる。また、運転手以外のユーザも、歌詞集などを見ることなく曲を歌うことができる。

[0019] このように、楽曲中で歌詞が再生されるタイミングに先行して、その歌詞の内容を音声出力してユーザに伝える機能を「アシストボーカル」と呼ぶ。なお、本実施例では、再生される楽曲はカラオケではなく、歌詞を含む通常の曲であるものとする。

[0020] 図1は、アシストボーカルの概念を示す。図1は、1つの楽曲を模式的に示したものである。図1の横軸は時間を示す。1つの楽曲中には、複数のブロックに分けて歌詞の部分が含まれている。再生される楽曲に含まれる歌詞の部分を「ボーカル」と呼ぶ。また、楽曲中において、ボーカル以外の部分を「間奏」と呼ぶ。よって、通常1つの楽曲は、複数の間奏と複数のボーカルとにより構成される。

[0021] 図1の例では、楽曲は、3つのボーカル1～3と、複数の間奏とにより構成されている。ボーカル1の内容（歌詞）は「あいうえお」であり、ボーカル2の内容は「かきくけこ」であり、ボーカル3の内容は「さしすせそ」であるものとする。

[0022] このような楽曲が再生されている状況において、本実施例では、楽曲中のボーカル1が再生されるタイミングに先行して、ボーカル1に対応する歌詞「あいうえお」が音声出力される。なお、本明細書では、アシストボーカルにより音声出力される歌詞音声を「スピーチ」と呼んで、楽曲中に含まれる「ボーカル」と区別する。

[0023] 図1の例では、ボーカル1に先行して、ボーカル1に対応するスピーチ1が出力される。同様に、ボーカル2に先行してスピーチ2が出力され、ボーカル3に先行してスピーチ3が出力される。

[0024] スピーチは、曲に含まれるボーカルの歌詞のみを音声信号として出力するものであり、基本的に音程やリズムなどの要素を含まない。また、後述するように、スピーチは基本的に対応するボーカルの前の間奏に挿入されるので、必要に応じてその長さが調整され、通常は楽曲の再生中にボーカルとして再生される場合よりも短い時間とされる。典型的な例では、スピーチは対応するボーカルの歌詞を早口で話した音声となる。

[0025] [1. 2] アシストボーカル処理

次に、スピーチを出力するためのアシストボーカル処理について説明する。図2は、アシストボーカル処理のフローチャートである。なお、この処理は、車両に搭載された端末装置、典型的にはスマートフォンなどの携帯端末などにより実行されるが、その詳細については後述する。以下の説明では、端末装置が処理を実行するものとして説明する。

[0026] まず、端末装置は、アシストボーカルがオンになっているか否かを判定する（ステップS1）。ここで、アシストボーカルのオン／オフは、ユーザが手動で行う場合と、自動で行う場合とがある。手動で行う場合、ユーザはアシストボーカルによりスピーチの再生を行いたいときに所定のボタンなどを操作してアシストボーカルをオンに設定し、端末装置はこれを検出する。一方、自動で行う場合、端末装置は例えばマイクなどを利用してユーザの声を判定し、ユーザが歌唱している又は歌唱に準ずる行為を行っている場合に、自動的にアシストボーカルをオンに設定する。なお、アシストボーカルの自動設定方法についてはさらに後述する。

[0027] アシストボーカルがオンに設定されていない場合（ステップS1：No）、処理は終了する。一方、アシストボーカルがオンに設定されている場合（ステップS1：Yes）、端末装置は、再生中の楽曲を特定する（ステップS2）。この場合に、車内で再生されている楽曲は、サーバからダウンロードされるなどして端末装置の内部に記憶されている楽曲、CDや車載器のメモリなどの記憶媒体に記憶されている楽曲、ラジオなどから再生されている楽曲などを含む。端末装置の内部に記憶されている楽曲を再生している場合

、端末装置はその再生中の楽曲を容易に特定することができる。一方、CDなどの記憶媒体に記憶されている楽曲が再生されている場合やラジオから楽曲が再生されている場合には、端末装置は、車内のスピーカから再生されている楽曲をマイクで集音し、そのオーディオデータを外部の音楽検索サーバに送信する。音楽検索サーバは、多数の楽曲のデータをデータベース化して記憶しており、端末装置から受信したオーディオデータと一致する楽曲を特定してその楽曲を示す情報（例えば、曲名、アーティスト名など、以下、「楽曲特定情報」と呼ぶ。）を端末装置に送信する。こうして、端末装置は、現在再生されている楽曲の楽曲特定情報を取得する。

[0028] こうして、再生中の楽曲が特定されると、端末装置は、スピーチ情報生成処理を実行する（ステップS3）。図3は、スピーチ情報生成処理のフローチャートである。また、図4は、スピーチ情報生成処理の概要を示す。

[0029] 図3において、端末装置は、ステップS2で特定された楽曲の歌詞データを外部サーバなどから取得する（ステップS31）。ここで、「歌詞データ」とは、その楽曲において、どのタイミングにどのような歌詞が再生されるかを規定する情報であり、具体的には、楽曲に含まれる歌詞を示す歌詞テキストデータと、その歌詞が再生される再生時刻（曲の開始時刻からの経過時間）を示す再生時刻データとを対応付けた情報である。

[0030] 次に、端末装置は、楽曲解析データを取得する（ステップS32）。楽曲解析データとは、その楽曲における拍位置、小節位置などの音楽的特徴を示す情報であり、再生された楽曲のオーディオデータに基づいて生成される。具体的には、端末装置は内部に楽曲解析アプリケーションを内蔵しておき、車両のスピーカから再生された楽曲をマイクで集音してオーディオデータを取得し、そのオーディオデータを解析することにより拍位置などの楽曲解析データを取得する。なお、端末装置に楽曲解析アプリケーションを内蔵する代わりに、外部の楽曲解析装置やサーバなどを利用して楽曲解析データを取得してもよい。

[0031] 次に、端末装置は、歌詞ブロック化を行う（ステップS33）。歌詞ブロ

ック化とは、ステップS 3 1 で取得した歌詞データに含まれる歌詞テキストデータをブロック化する処理であり、1つのブロックは、1つのスピーチに対応する。即ち、歌詞ブロック化は、歌詞テキストデータを、スピーチの単位に分割する処理である。

[0032] 図4の例では、端末装置は、歌詞テキストデータとして「あいうえおかきくけこさしすせそ」を取得しており、端末装置は、これを「あいうえお」、「かきくけこ」、「さしすせそ」の3つのブロックに分割してブロック歌詞データを生成する。

[0033] 図5は、歌詞ブロック化の例を示す。図5（A）に第1の方法を示す。この方法では、楽曲に含まれる間奏と間奏との間を1つのブロックとする。なお、「間奏」は、楽曲のうち「ボーカル」以外の部分である。具体的には、端末装置は、ボーカル以外の区間（非ボーカル区間）の長さ l_t が所定長さ t_1 よりも長い場合に、その区間を間奏と判定する。

[0034] 但し、例外的に、間奏の長さとの関係で複数のブロックを1つのブロックにまとめる場合がある。図5（B）に示す例のように、ボーカル3の長さ V_t3 に対して、その直前の間奏2の長さ l_t2 が非常に短い（ $l_t2 < \alpha 1 \cdot V_t3$ ； $\alpha 1$ は任意の係数）場合、間奏2の間にボーカル3のスピーチを出力することは難しい。このような場合に、その1つの前の間奏1の長さ l_t1 が所定長より長ければ、端末装置は、ボーカル2とボーカル3を1つのブロックとする。これにより、ボーカル2とボーカル3に対応するスピーチは間奏1においてされる。

[0035] 図5（C）に第2の方法を示す。この方法では、端末装置は、歌詞データに含まれる区切りに基づいて各ブロックを決定する。即ち、歌詞データに含まれる歌詞テキストデータに予め区切りの情報が含まれている場合には、端末装置はその区切りに従って歌詞テキストデータをブロック化することができる。

[0036] 次に、端末装置は、歌詞スピーチ化を行う（ステップS 3 4）。歌詞ブロック化により得られたブロック歌詞データはあくまで歌詞を示すテキストデ

ータであり、歌詞スピーチ化はブロック歌詞データを音声データに変換する処理である。具体的には、端末装置は、テキスト音声変換（TTS: Text To Speech）ソフトウェアを内蔵し、ステップS33で得られた各ブロック歌詞データを音声データに変換する。これにより、図4に示すように、各ブロック歌詞データから、音声データであるスピーチ1～3が生成される。なお、端末装置にTTSソフトウェアを内蔵する代わりに、外部サーバなどによるTTS変換を利用してもよい。

[0037] 次に、端末装置は、スピーチ長変更を行う（ステップS35）。スピーチ長変更とは、歌詞スピーチ化により得られた各スピーチの時間的な長さを短縮して、短い時間で再生できるようにする処理である。既に述べたように、各スピーチは対応するボーカルに先行する間奏において再生されるが、間奏の時間的な長さには制限があるので、スピーチを短くして再生する必要がある。このため、スピーチ長変更が行われる。

[0038] 基本的には、人間により聞き取り可能な範囲で、各スピーチの再生時間を短く（再生速度を速く）する。例えば、ステップS34で得られた各スピーチの時間的な長さ（「オリジナルスピーチ長」と呼ぶ。）を「 S_t 」とし、スピーチ長変換係数を「 α_2 」とすると、スピーチ長変更による変更後の長さ「 S_{tv} 」は、

$$S_{tv} = S_t \cdot \alpha_2 \quad (\alpha_2 < 1.0) \quad (1)$$

で与えられる。例えば、 $\alpha_2 = 0.7$ とすれば、スピーチ長変更により各スピーチは元の3割増しの速さで再生されることになる。

[0039] また、上記のような一括変更に加えて、各スピーチ毎に対応する間奏の時間に応じてさらに再生時間を短くしてもよい。なお、この場合、同じ文字数のスピーチ、又は、同じ歌詞の言葉であっても、曲中の位置（先行する間奏の長さ）に応じて、再生時間が異なることになる。

[0040] 次に、端末装置は、スピーチ挿入タイミングを算出する（ステップS36）。端末装置は、あるボーカルに対応するスピーチを、そのボーカルの再生

タイミングに先行して挿入する。図4に示す例では、ボーカル1に対応するスピーチ1はボーカルの再生タイミングより前に挿入される。同様に、ボーカル2に対応するスピーチ2はボーカル2の再生タイミングより前に挿入され、ボーカル3に対応するスピーチ3はボーカル3の再生タイミングより前に挿入される。

[0041] スピーチを挿入する方法の具体例を図6に示す。図6は、ボーカル2に対応するスピーチ2を挿入するタイミングの例を示す。

[0042] 方法1では、スピーチは、対応するボーカルの開始タイミングよりも一定時間前に終了する。具体的に、図6に示すように、スピーチ2は、ボーカル2の再生開始タイミングより一定時間 T_2 前に終了するように挿入される。即ち、スピーチ2はボーカル2の再生開始より一定時間 T_2 前に終了する。この場合、スピーチ2の再生開始タイミングはスピーチ2の長さに応じて決まる。方法1では、スピーチの再生が終了してから、対応するボーカルが再生されるまでに一定時間が確保されるので、ユーザは余裕を持ってボーカル部分を歌うことができる。

[0043] 方法2では、スピーチの終了タイミングを楽曲の拍位置と一致させる。具体的に、図6の例では、スピーチ2は、ボーカル2の再生開始タイミングより N 拍前（ N は任意の整数；本例では $N=1$ ）に終了するように挿入される。この場合、スピーチ2の再生開始タイミングはスピーチ2の長さに応じて決まる。なお、楽曲の拍の位置は、前述の楽曲解析データから取得される。

[0044] 方法3では、スピーチの再生開始タイミングと再生終了タイミングの両方を楽曲の拍位置と一致させる。具体的に、図6の例では、スピーチ2の再生開始タイミング及び再生終了タイミングをともに4拍子の3拍目に一致させている。

[0045] 方法2、3のように、スピーチの終了タイミング、又は、開始／終了タイミングの両方を楽曲の拍位置と一致させると、スピーチが楽曲と連動するのでユーザが楽曲を歌いやすくなる。

[0046] 以上のようにして、端末装置は、スピーチの挿入タイミングを決定する。

具体的には、各スピーチについて、その再生開始タイミングと再生終了タイミングとを、楽曲の先頭からの経過時間により規定する。各スピーチの再生開始タイミングと再生終了タイミングは、スピーチ情報の一部として記憶される。即ち、スピーチ情報は、各スピーチに対応する音声信号（以下、「スピーチ信号」とも呼ぶ。）と、各スピーチの再生開始タイミング／再生終了タイミングとを含む。

[0047] 次に、処理は図2に示すメインルーチンに戻り、端末装置は、再生中の楽曲の現在の再生位置を取得する（ステップS4）。具体的には、端末装置は、再生中の楽曲の再生開始時刻からの経過時間をカウントすることにより、現在の再生位置を取得する。

[0048] 次に、端末装置は、スピーチ強調処理を行う（ステップS5）。スピーチ強調処理は、楽曲に含まれるボーカルと、スピーチとを区別して聞き取りやすくする処理であるが、その詳細は後述する。

[0049] 次に、端末装置は、スピーチ情報に含まれる各スピーチの再生開始タイミング／再生終了タイミングと、現在の再生位置とに基づいて、スピーチを再生する（ステップS6）。具体的には、スピーチの再生開始タイミングでスピーチの再生を開始し、スピーチの再生終了タイミングでスピーチの再生を終了する。これにより、楽曲中のボーカルに先行して、対応するスピーチが再生されることになる。

[0050] 次に、端末装置は、スピーチの再生を終了すべきか否かを判定する（ステップS7）。スピーチの再生を終了すべき場合とは、スピーチ情報が無くなった場合、楽曲の再生自体が終了した場合、ユーザの操作によりアシストボーカルがオフされた場合、などが挙げられる。スピーチの再生を終了すべきでない場合（ステップS7：No）、処理はステップS4に戻り、スピーチの再生を継続する。一方、スピーチの再生を終了すべきである場合（ステップS7：Yes）、アシストボーカル処理は終了する。

[0051] [1. 3] アシストボーカルの自動オン設定方法

次に、図2に示すアシストボーカル処理のステップS1においてアシスト

ボーカルを自動的にオンに設定する方法について説明する。

[0052] 基本的な方法としては、端末装置は、ユーザが発している音声をマイクで集音し、ユーザが楽曲に合わせて歌唱している（歌を歌っている）又は歌唱に準ずる行為を行っているとは判定される場合にアシストボーカルを自動的にオンにする。例えば、マイクにより集音した音声データを解析した結果、鼻歌を歌っている、断片的に曲を歌っている、ハミングしているなどと判定される場合には、アシストボーカルをオンにする。一方、音声データが歌唱しているのではなく、同乗者との会話である場合にはアシストボーカルをオンにしない。音声データが鼻歌を歌っている部分を含んでいるような場合でも、大部分が会話であるような場合にもアシストボーカルをオンにはしない。

[0053] なお、音声データに含まれるユーザの音声歌唱であるか否かは、音声データに含まれるリズムや音程の有無に基づいて判断することができる。例えばリズムが規則的である場合や音程の変化が大きい場合には歌唱であると判断し、リズムが不規則である場合は音程の変化が小さい場合に歌唱ではない（会話である）と判断することができる。また、前述の楽曲解析アプリケーションを利用し、音声データから拍や小節が抽出できた場合に歌唱であると判断し、抽出できない場合に歌唱ではないと判断してもよい。また、前述の音楽検索サーバ又は音楽検索機能を利用し、音声データから楽曲が特定できた場合に歌唱であると判断し、楽曲が特定できない場合に歌唱ではないと判断してもよい。

[0054] また、端末装置は、集音した音声データと、再生中の楽曲との相関を算出し、一定値以上の相関がある場合に、ユーザが歌唱していると判断してアシストボーカルをオンにしてもよい。また、端末装置が再生中の曲の歌詞データを既に取り得している場合には、マイクにより集音した音声データと歌詞データとの相関が一定値以上である場合に、ユーザが歌っていると判断してもよい。また、歌詞データに基づいて、歌詞が存在しないはずの楽曲の間奏位置においてもユーザの音声が出力されている場合には、それは会話であると判断してもよい。

- [0055] また、マイクで集音したリズムの情報を利用してもよい。例えば、ユーザが楽曲のリズムに合わせて手や指でステアリングなどを叩いているとか、足で床を踏んでリズムを取っていると判断される場合には、ユーザが歌唱に準ずる行為を行っているとは判定し、アシストボーカルをオンにしてもよい。この場合、マイクで集音したリズムと再生中の楽曲のリズムとの相関を算出し、相関が一定値以上である場合にアシストボーカルをオンにしてもよい。また、再生中の楽曲のリズムとの相関を算出しなくても、マイクで集音されたリズムが、一定のリズムの繰り返しになっているような場合には、アシストボーカルをオンにしてもよい。
- [0056] さらに、車内を撮影するカメラでユーザの状態を撮影し、ユーザが楽曲に合わせて首を振っているような場合に、アシストボーカルをオンにしてもよい。また、車内を撮影するカメラにより、助手席や後部座席に同乗者がいるか否かを検出し、同乗者の有無により、ユーザが歌っているのか会話しているのかの判定基準を変化させてもよい。
- [0057] また、上記の例では、ユーザが歌唱していると判断した場合に、アシストボーカルをオンにする例を説明したが、ユーザが歌唱していても、ユーザが歌詞を知っていてアシストボーカルを再生する必要がないと判断した場合には、アシストボーカルをオンにしなくてもよい。具体的には、例えば集音した音声データと、再生中の楽曲との相関が一定値以上であり、かつ歌詞データとの相関が一定値以上である場合には、ユーザが歌詞を知っていると判断し、歌唱していてもアシストボーカルをオンにしない。
- [0058] ただしこの場合、ユーザが途中から歌詞が分からなくなる可能性があるため、スピーチ情報を生成し、出力する準備をしておいてもよい。そのあとに、集音した音声データと、再生中の楽曲との相関が一定値未満であり、または歌詞データとの相関が一定値未満である場合には、ユーザは歌詞を知らないとは判断し、アシストボーカルを出力する。
- [0059] また、上記の例では、アシストボーカルの自動オン設定の方法について説明したが、アシストボーカルの自動オフ設定も行うことができる。アシスト

ボーカルをオンしている間に、ユーザが楽曲に合わせて歌唱していない（歌を歌っていない）又は歌唱に準ずる行為（鼻歌を歌っている、断片的に曲を歌っている、ハミングをしている等）を行っていないと判定された場合に、アシストボーカルを自動的にオフにしてもよい。同様に、会話が検出されたら、アシストボーカルを自動的にオフにしてもよいし、リズムをとっていないと判断されたり、ユーザが楽曲に合わせて頭を振っていないと判断された場合、アシストボーカルを自動的にオフにしてもよい。

[0060] また、上記の例では、ユーザが歌唱しているもしくは歌唱に準ずる行為をしているか否かに基づき、アシストボーカルの自動オン設定もしくは自動オフ設定を行うことを説明したが、再生されている楽曲の構成に基づき自動オン設定もしくは自動オフ設定してもよい。

例えば、楽曲のサビの部分だけ歌唱したいというユーザに対しては、楽曲のサビの部分再生時に、アシストボーカルを自動的にオン設定し、楽曲のサビ以外の部分を再生する際に、アシストボーカルを自動的にオフ設定してもよい。逆に、サビの部分は知っていてサビ以外の部分を練習したいというユーザに対しては、楽曲のサビ以外の部分を再生する際に、アシストボーカルを自動的にオン設定し、楽曲のサビの部分再生時に、アシストボーカルを自動的にオフ設定してもよい。

[0061] [1. 4] スピーチ強調処理

次に、図2に示すアシストボーカル処理のステップS5において実行されるスピーチ強調処理について説明する。スピーチ強調処理は、ユーザがスピーチとボーカルとを区別して聞き取り易くする方法であり、以下のいくつかの方法を示す。

[0062] [1. 4. 1] スピーチとボーカルが重なる場合の処理

スピーチは基本的に対応するボーカルの直前の間奏中に再生され、ボーカルとは時間的に重ならないことが好ましい。このために前述のスピーチ長変更処理（ステップS35）を行うのであるが、スピーチの長さと同様の長さによっては、スピーチ長を短縮してもスピーチを間奏中に再生しきれないこ

ともある。即ち、間奏の長さよりも、スピーチの長さの方が長い場合、スピーチとボーカルとが部分的に重なって再生される。このようにスピーチとボーカルとを重ねて再生することに代えて、以下のいずれかの処理を行ってもよい。

[0063] (1) ボーカルのレベルを調整する。

[0064] スピーチとボーカルとが重なってしまう場合、ボーカルの音量レベルを下げる方法がある。図7(A)は、スピーチの後方部分と、ボーカルの先頭部分とが重なり、重複部分Xが生じる場合を示す。この場合、重複部分Xにおいてボーカルの音量を調整する。具体的には、ボーカルの音量をスピーチが聞こえる程度まで低下させる、もしくはゼロにする。これにより、重複部分Xでは、スピーチの再生が優先され、スピーチが聞き取り易くなる。

[0065] 図7(B)は、逆にスピーチの先頭部分と、1つ前のボーカルの後方部分とが重なり、重複部分Xが生じる場合を示す。この場合にも、重複部分Xにおいて、ボーカルの音量を調整する。具体的には、ボーカルの音量をスピーチが聞こえる程度まで低下させる、もしくはゼロにする。また、重複部分Xにおいて、急にボーカルの音量レベルを下げるのではなく、ボーカルをフェードアウトさせて徐々に音量レベルを下げるようにしてもよい。これにより、重複部分Xでは、スピーチの再生が優先され、スピーチが聞き取り易くなる。

[0066] 具体的に上記のレベル調整は、楽曲信号においてボーカルの成分と楽器などの演奏の成分とが分離している場合には、ボーカルの成分の音量レベルを低下させればよい。一方、ボーカルの部分が楽器などの演奏の部分と合成されており、ボーカルのみの音量を調整できない場合には、楽曲信号全体の音量レベルを低下させてもよいし、又は、楽曲信号のうち一般的にボーカル（人間の声）に相当する周波数帯域の成分のみ音量レベルを低下させるようにしてもよい。

[0067] (2) スピーチのレベルを調整する。

[0068] スピーチとボーカルとが重なってしまう場合、逆にスピーチの音量レベル

を下げる方法もある。図7（C）は、スピーチの後方部分と、ボーカルの先頭部分とが重なり、重複部分Xが生じる場合を示す。この場合、重複部分Xにおいて、スピーチの音量を調整する。具体的には、スピーチの音量を低下させる、もしくはゼロにする。急にスピーチの音量を下げるのではなく、スピーチをフェードアウトさせて徐々に音量を下げるようにしてもよい。この場合、重複部分Xでは、スピーチが聞き取れなくなるが、一般的にユーザがある程度知っている楽曲を聞く場合には、歌詞の全てを覚えてはいないものの、歌詞の先頭部分がわかれば、その後は歌詞を思い出して歌うことができるということも多い。よって、図7（C）のように、スピーチの先頭部分が聞き取れれば、スピーチの後方部分が聞き取りにくくなっても構わないということも多い。この手法はそのような場合に有効である。

[0069] [1. 4. 2] スピーチとボーカルを異なる方向から聞かせる処理

人間には、同時に異なる方向から到来する音を聞き分ける能力がある（いわゆるカクテルパーティ効果）。これを利用し、ユーザがスピーチとボーカルとを聞き分けることができるようにする手法が考えられる。なお、この手法は、スピーチとボーカルとが時間的に重なるか否かに拘わらず実行される。

[0070] （1）左右のスピーカで位相を調整する方法

図8（A）は、左右のスピーカから出力されるスピーチの位相を反転させる構成を示す。左（L）チャンネルの楽曲信号は加算器32に供給され、右（R）チャンネルの楽曲信号は加算器33に供給される。一方、スピーチ信号は、そのまま加算器33に供給されるとともに、位相反転器31で位相が反転されて加算器32に供給される。加算器32の出力は左スピーカ30Lに供給され、加算器33の出力は右スピーカ30Rに供給される。

[0071] この構成によれば、ボーカルを含む楽曲の音像は左右スピーカの間に定位するのに対し、スピーチの音像はユーザの耳回りに定位することになり、ユーザはスピーチと楽曲中のボーカルとを聞き分けやすくなる。なお、図8（A）の例では、位相反転器31により左スピーカ30Lに供給されるスピー

チ信号の位相のみを反転しているが、逆に右スピーカ 30R に供給されるスピーチ信号の位相のみを反転させてもよい。また、左右のスピーカに供給されるスピーチ信号の間に一定の位相差があればスピーチの音像位置と楽曲の音像位置とを異ならせることができるので、一方のスピーカに供給されるスピーチ信号を必ずしも反転（ 180° 変化）させる必要はない。即ち、一方のスピーカに供給されるスピーチ信号と、他方のスピーカに供給されるスピーチ信号との間に一定の位相差を与えてやればよい。

[0072] (2) 音像の定位を制御する方法

図8 (B) は、スピーチの音像を任意の位置に設定可能な構成を示す。左 (L) チャンネルの楽曲信号は加算器 32 に供給され、右 (R) チャンネルの楽曲信号は、加算器 33 に供給される。一方、スピーチ信号は、音像定位制御演算部 34、クロストークキャンセル部 35 を経由して加算器 32、33 に供給される。音像定位制御演算部 34 は、目標のスピーカ位置と聴取位置 (ユーザの位置) との間の伝達関数をスピーチ信号に畳み込み、クロストークキャンセル部 35 は楽曲を出力しているスピーカと聴取位置との間の伝達関数をキャンセルする処理を行う。これにより、楽曲の音像は左右のスピーカ 30L、30R の間に定位させるとともに、スピーチの音像を目標のスピーカ位置に定位させることができるので、ユーザはスピーチとボーカルとを聞き分けやすくなる。

[0073] (3) ヘッドレストスピーカを利用する方法

車両のスピーカに加えて車両のシートにヘッドレストスピーカが搭載されている場合、車両のスピーカからボーカルを含む楽曲を出力し、ヘッドレストスピーカからスピーチを出力することができる。この場合の構成例を図9 に示す。

[0074] 左右チャンネルの楽曲信号はそれぞれ車両のスピーカ 30L、30R に供給される。また、スピーチ信号は、そのまま右のヘッドレストスピーカ 35R に供給されるとともに、位相反転器 31 で位相が反転されて左のヘッドレストスピーカ 35L に供給される。この場合も、2つのヘッドレストスピー

カ 3 5 L、3 5 R に供給されるスピーチ信号に位相差が与えられているため、スピーチの音像は楽曲の音像と異なる位置に定位し、ユーザはスピーチと楽曲中のボーカルとを聞き分けやすくなる。なお、この例においても、図 8 (A) の例と同様に、一方のヘッドレストスピーカに供給されるスピーチ信号と、他方のヘッドレストスピーカに供給されるスピーチ信号との間に一定の位相差を与えてやればよい。

[0075] ヘッドレストスピーカを利用する場合には、運転席のヘッドレストスピーカの代わりに、助手席のヘッドレストスピーカを利用してスピーチを再生してもよい。また、車両の複数の座席にヘッドレストスピーカが搭載されている場合には、各座席毎にスピーチの再生の可否を選択して設定できるようにしてもよい。こうすると、スピーチを聞いて楽曲を歌いたい搭乗者の座席のヘッドレストスピーカのみからスピーチが再生されるように設定することができる。

[0076] また、位相差を与えることに代えて、図 8 (B) で説明した処理と同様に、音像定位制御演算部 3 4 と、クロストークキャンセル部 3 5 とを用いることで、スピーチの音像を任意の位置に定位させてもよい。これにより、ユーザがスピーチとボーカルとを聞き分けやすくなることができる。

[0077] [2] システム構成

次に、上述のアシストボーカルを実現する楽曲再生システムの構成例を説明する。

[0078] [2. 1] 第 1 実施例

第 1 実施例では、アシストボーカル処理を主として端末装置側で実行する。第 1 実施例による楽曲再生システムの全体構成を図 10 に示す。第 1 実施例の楽曲再生システムでは、複数の車両 1 と、コンテンツプロバイダ 2 と、ゲートサーバ 3 とがネットワーク 4 を介して通信可能とされる。なお、複数の車両 1 は、無線通信によりネットワーク 4 を介してコンテンツサーバ 2、ゲートサーバ 3 と通信可能となっている。

[0079] コンテンツプロバイダ 2 は、音楽配信業者などのサーバであり、楽曲デー

タ、楽曲のメタデータ、歌詞データなどを提供する。ゲートサーバ3は、本実施例によるアシストボーカルを実現するために機能するサーバであり、コンテンツプロバイダ2から必要な楽曲の楽曲データ、メタデータ、歌詞データなどを取得して、図示しないデータベースに記憶している。

[0080] 車両1の内部構成の一例を図11(A)に示す。車両1は、端末装置10と、音楽再生装置20と、スピーカ30とを備える。

[0081] 端末装置10は、典型的にはスマートフォンなどの携帯端末であり、通信部11と、制御部12と、記憶部13と、マイク14と、操作部15とを備える。通信部11は、ネットワーク4を通じてゲートサーバ3と通信する。制御部12は、CPUなどからなり、端末装置10の全体を制御する。

[0082] 記憶部13は、ROM、RAMなどのメモリであり、制御部12が各種の処理を実行するためのプログラムを記憶するとともに、ワークメモリとしても機能する。記憶部13に記憶されたプログラムを制御部12が実行することにより、アシストボーカル処理を含む処理が実行される。また、記憶部13は、ユーザが保存した楽曲の楽曲データを記憶していてもよい。

[0083] マイク14は、車内で再生されている楽曲、ユーザによる歌唱、会話などの音声を集音して音声データを生成する。操作部15は、典型的にはタッチパネルなどであり、ユーザによる操作、選択の入力を受け付ける。

[0084] 音楽再生装置20は、例えばカーオーディオなどであり、アンプなどを含む。スピーカ30は、車両に搭載されたスピーカである。音楽再生装置20は、端末装置10から供給される楽曲データに基づいて楽曲をスピーカ30から再生する。

[0085] 車両1の内部構成の他の例を図11(B)に示す。この例では、車両1は端末装置10xを備える。端末装置10xは、図11(A)に示す携帯端末などの端末装置10とカーオーディオなどの音楽再生装置20の機能を併せ持つ装置である。端末装置10xは、端末装置10と同様に通信部11、制御部12、記憶部13、マイク14、操作部15を備えるとともに、音楽再生装置20に相当する音楽再生部16を備える。端末装置10xはスピーカ

30に接続され、楽曲データに基づいて楽曲をスピーカ30から再生する。

[0086] 次に、第1実施例の楽曲再生システムによるアシストボーカル処理について説明する。図12は、第1実施例に係るアシストボーカル処理のフローチャートである。この処理では、アシストボーカル処理を主として端末装置10又は10x（以下、代表して単に「端末装置10」と記す。）により実行する。

[0087] まず、ゲートサーバ3は、ネットワーク4を介してコンテンツプロバイダ2に接続し、複数の楽曲について、楽曲データ及び歌詞データを取得し、内部のデータベースに保存しておく（ステップS101）。

[0088] 端末装置10は、ユーザによる操作部15の操作により、再生すべき楽曲の指定を受け取り（ステップS102）、その楽曲を指定する楽曲指定情報をゲートサーバ3へ送信する（ステップS103）。ゲートサーバ3は、受け取った楽曲指定情報に対応する楽曲の楽曲データ及び歌詞データをデータベースから取得し、端末装置10へ送信する（ステップS104）。

[0089] 次に、端末装置10は、受信した楽曲データ及び歌詞データを利用して、ステップS105～S109の処理を行う。ここで、ステップS105～S109の処理は、図2におけるステップS3～S7と同様であるので、説明を省略する。

[0090] こうして、第1実施例の楽曲再生システムにおいては、車両1に搭載された端末装置10が主としてアシストボーカル処理を実行する。

[0091] 上記の例では、ステップS101でゲートサーバ3はコンテンツプロバイダから楽曲データを取得しているが、楽曲データが端末装置10に保存されている場合には、ゲートサーバ3は端末装置10から楽曲データを取得してもよい。また、ゲートサーバ3内のデータベースに楽曲データが保存されている場合には、そこから楽曲データを取得してもよい。

[0092] [2.2] 第2実施例

第2実施例では、アシストボーカル処理の一部をゲートサーバ3側で実行する。第2実施例による楽曲再生システムの全体構成は、図10に示す第1

実施例と同様であるので、説明を省略する。

[0093] 次に、第2実施例の楽曲再生システムによるアシストボーカル処理について説明する。図13は、第2実施例に係るアシストボーカル処理のフローチャートである。この処理では、ゲートサーバ3がスピーチ情報を生成し、さらにスピーチ付楽曲データを生成して端末装置10へ送信する。端末装置10は、スピーチ付楽曲データを受信して再生する。以下、詳しく説明する。

[0094] まず、ゲートサーバ3は、ネットワーク4を介してコンテンツプロバイダ2に接続し、複数の楽曲について、楽曲データ及び歌詞データを取得し、内部のデータベースに保存する（ステップS201）。そして、ゲートサーバ3は、各楽曲について、取得した楽曲データと歌詞データとに基づいてスピーチ情報を生成する（ステップS202）。このスピーチ情報生成処理は、図2のステップS3と同一であるので、説明を省略する。

[0095] スピーチ情報を生成すると、ゲートサーバ3は、楽曲データにスピーチを付加してスピーチ付楽曲データを生成する（ステップS203）。具体的に、ゲートサーバ3は、生成したスピーチ情報に基づいて、各スピーチに対応するスピーチ信号を、図3のステップS36の処理により算出したタイミングで楽曲データに合成し、スピーチ付楽曲データを生成してデータベースに記憶する。つまり、スピーチ付楽曲データは、そのまま再生することにより、楽曲に加えてスピーチが再生されるデータである。

[0096] 端末装置10は、ユーザによる操作部15の操作により、再生すべき楽曲の指定を受け取り（ステップS204）、その楽曲を指定する楽曲指定情報をゲートサーバ3へ送信する（ステップS205）。ゲートサーバ3は、受け取った楽曲指定情報に対応する楽曲のスピーチ付楽曲データを端末装置10へ送信する（ステップS206）。

[0097] 次に、端末装置10は、受信したスピーチ付楽曲データを再生する（ステップS207）。これにより、楽曲の再生中の適切なタイミングで、スピーチが再生される。次に、端末装置10は、楽曲の再生を終了すべきか否かを判定する（ステップS208）。その楽曲が最後まで再生された場合、又は

、ユーザが再生を中止した場合など、再生を終了すべき場合には（ステップ S 2 0 8 : Y e s）、端末装置 1 0 は再生を終了する。一方、楽曲の再生を終了すべきではない場合（ステップ S 2 0 8 : N o）、処理はステップ S 2 0 7 へ戻り、スピーチ付楽曲データの再生が継続される。

[0098] こうして、第 2 実施例の楽曲再生システムにおいては、ゲートサーバ 3 側でスピーチ付楽曲データが生成され、端末装置 1 0 へ提供される。端末装置 1 0 は、受信したスピーチ付楽曲データを再生することにより、スピーチを含む楽曲を聞くことができる。

[0099] 上記の例では、ステップ S 2 0 1 でゲートサーバ 3 はコンテンツプロバイダから楽曲データを取得しているが、楽曲データが端末装置 1 0 に保存されている場合には、ゲートサーバ 3 は端末装置 1 0 から楽曲データを取得してもよい。また、ゲートサーバ 3 内のデータベースに楽曲データが保存されている場合には、そこから楽曲データを取得してもよい。

[0100] [3] スピーチのみを再生するアシストボーカル

上述のアシストボーカル処理では、端末装置 1 0 により再生している楽曲に対して、スピーチを付加して再生している。しかし、端末装置 1 0 以外のソース、例えば車内のラジオ、CD など(以下、「外部ソース」と呼ぶ。)から再生されている楽曲に対してスピーチを付加することができれば便利である。この場合、端末装置 1 0 は、基本的に上述の方法でスピーチ情報を生成し、外部ソースから再生されている楽曲の再生位置に応じたタイミングでスピーチのみを再生すればよい。

[0101] この場合のアシストボーカル処理のフローチャートを図 1 4 に示す。まず、端末装置 1 0 は、外部ソースから再生されている楽曲をマイク 1 4 により集音して再生楽曲データを取得し（ステップ S 1 5 1）、これをゲートサーバ 3 へ送信する（ステップ S 1 5 2）。

[0102] ゲートサーバ 1 5 3 は、端末装置 1 0 から再生楽曲データを受信し、対応する楽曲及びその再生位置を特定する（ステップ S 1 5 3）。具体的には、ゲートサーバ 3 は、前述の音楽検索サーバの機能を有する音楽検索部を備え

、再生楽曲データに基づいて、その楽曲を特定するとともに、その再生楽曲データの部分に対応する再生位置を特定する。そして、ゲートサーバ3は、特定した楽曲の楽曲名やアーティスト名とともに、歌詞データと、再生位置情報とを端末装置10へ送信する（ステップS154）。

[0103] 端末装置10は、受信した歌詞データを利用して、スピーチ情報を生成する（ステップS155）。なお、スピーチ情報の生成は、図3を参照して説明したのと同様の方法で行われる。なお、端末装置10は、マイク14で取得した再生楽曲データを解析することにより、楽曲解析データを取得することができる（図3のステップS32の処理）。

[0104] 次に、端末装置10は、ゲートサーバ3から取得した再生位置情報に基づいて、その楽曲における現在の再生位置を算出する（ステップS156）。この方法については後述する。次に、端末装置10は、スピーチ強調処理を行い（ステップS157）、外部ソースにより再生されている楽曲に合わせて適切なタイミングでスピーチを再生する（ステップS158）。これにより、外部ソースから再生されている楽曲に合わせて、スピーチが再生される。

[0105] そして、端末装置10は、スピーチの再生を終了すべきか否かを判定し（ステップS159）、終了させるべきでない場合には、ステップS156へ戻って処理を継続する。一方、外部ソースからの楽曲の再生が終了した場合、再生されている楽曲が別の楽曲に変わった場合、再生すべきスピーチが無くなった場合など、スピーチの再生を終了すべき場合には（ステップS159：Yes）、処理を終了する。

[0106] 次に、図15を参照して、ステップS156において楽曲の現在の再生位置を特定する方法を説明する。端末装置10からゲートサーバ3へ送信される再生楽曲データは、実際には複数のオーディオフレームのデータとなる。即ち、端末装置10は、外部ソースにより再生されている楽曲をマイク14で集音し、複数のオーディオフレームとして順次ゲートサーバ3へ送信する。

[0107] 図15の例では、端末装置10は、外部ソースにより再生されている楽曲のオーディオフレーム n 、 $(n+1)$ 、 $(n+2)$ 、...を、再生楽曲データとして順次ゲートサーバ3へ送信する。この際、端末装置10は、最初に再生楽曲データを送信した時刻、図15の例ではオーディオフレーム n を送信した時刻（以下、「基準時刻 t_0 」と呼ぶ。）を記憶しておく。

[0108] ゲートサーバ3の音楽検索部は、データベースに記憶された多数の楽曲の情報を参照し、受信した複数のオーディオフレームに基づいて楽曲を特定する。図15の例では、ゲートサーバ3の音楽検索部は、オーディオフレーム $n \sim (n+4)$ に基づいて楽曲を特定できたものとする。この場合、ゲートサーバ3は、楽曲判定結果として、楽曲名、アーティスト名などに加えて、端末装置10から最初に受信したオーディオフレーム n の曲先頭からの再生時間（ t_n ）を再生位置情報として端末装置10へ送信する。即ち、図14のステップS154でゲートサーバ3から端末装置10へ送信される再生位置情報は、端末装置10がゲートサーバ3へ最初に送信したオーディオフレーム n の、その楽曲の先頭からの経過時間となっている。そこで、ステップS156において、端末装置10は、予め記憶していた基準時刻 t_0 から現在までの経過時間の経過時間 Δt を算出し、これを再生時間 t_n に加算する。即ち、ゲートサーバ3から送信される再生時間 t_n は、その楽曲の先頭からオーディオフレーム n までの時間であり、経過時間 Δt はオーディオフレーム n から現在までの時間である。よって、現在の再生位置（再生時間） T_c は、以下の式で算出される。

$$[0109] \quad T_c = t_n + \Delta t \quad (2)$$

以上のように、ゲートサーバ3に音楽検索機能を設け、再生楽曲データに基づいて楽曲及びその再生位置を特定することにより、外部ソースから再生されている楽曲に合わせてスピーチを再生することができる。また、ゲートサーバ3に音楽検索機能を設ける代わりに、外部の音楽検索サーバを利用しても良い。

[0110] なお、ステップS159では、1つの楽曲が終了したときに再生を終了し

てもよいが、1つの楽曲が終了した後で別の楽曲が再生されているような場合には、処理を継続してもよい。即ち、端末装置10からゲートサーバ3への楽曲再生データの送信を継続している間は、スピーチの再生を継続することとしてもよい。これにより、外部ソースから再生される曲が変わっても、それに追従してスピーチの再生を継続することが可能となる。

産業上の利用可能性

[0111] 本発明は、音楽を再生する装置に利用することができる。

符号の説明

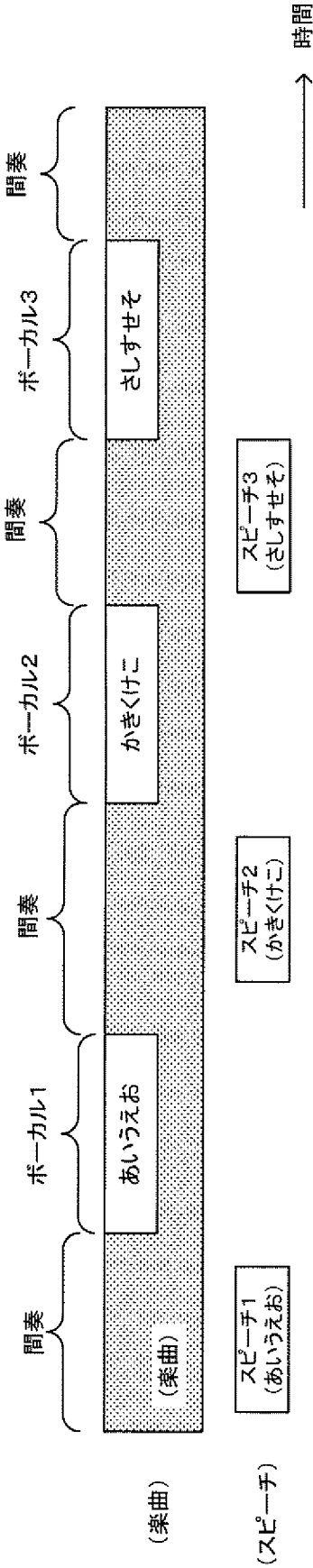
- [0112]
- 1 車両
 - 2 コンテンツプロバイダ
 - 3 ゲートサーバ
 - 4 ネットワーク
 - 10、10x 端末装置
 - 12 制御部
 - 13 記憶部
 - 14 マイク
 - 20 音楽再生装置
 - 30 スピーカ

請求の範囲

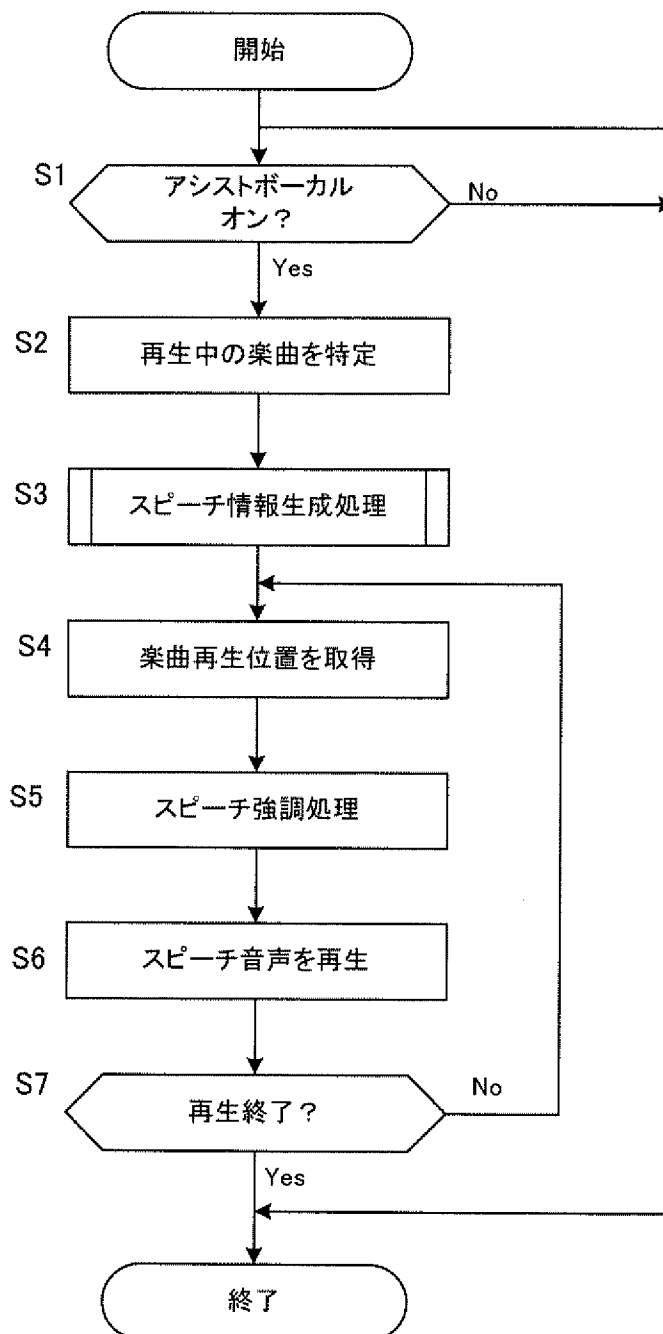
- [請求項1] 音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置であって、
- 前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段を備えることを特徴とする再生装置。
- [請求項2] 前記音楽信号の拍位置を検出する拍検出手段を備え、
- 前記出力手段は、前記歌詞音声信号の出力終了タイミングを、前記音楽信号の拍位置と一致させることを特徴とする請求項1に記載の再生装置。
- [請求項3] 前記出力手段は、前記歌詞音声信号の出力開始タイミングを、前記音楽信号の拍位置と一致させることを特徴とする請求項2に記載の再生装置。
- [請求項4] 音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行される再生方法であって、
- 前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力工程を備えることを特徴とする再生方法。
- [請求項5] コンピュータを備え、音楽信号と当該音楽信号の歌詞を表す歌詞音声信号とを出力する再生装置により実行されるプログラムであって、
- 前記歌詞音声信号の出力終了タイミングが、前記音楽信号中の前記歌詞音声信号に対応する歌詞区間の出力開始タイミングより一定時間前となるように、前記歌詞音声信号を出力する出力手段として前記コンピュータを機能させることを特徴とするプログラム。
- [請求項6] 請求項5に記載のプログラムを記憶したことを特徴とする記憶媒体。

[図1]

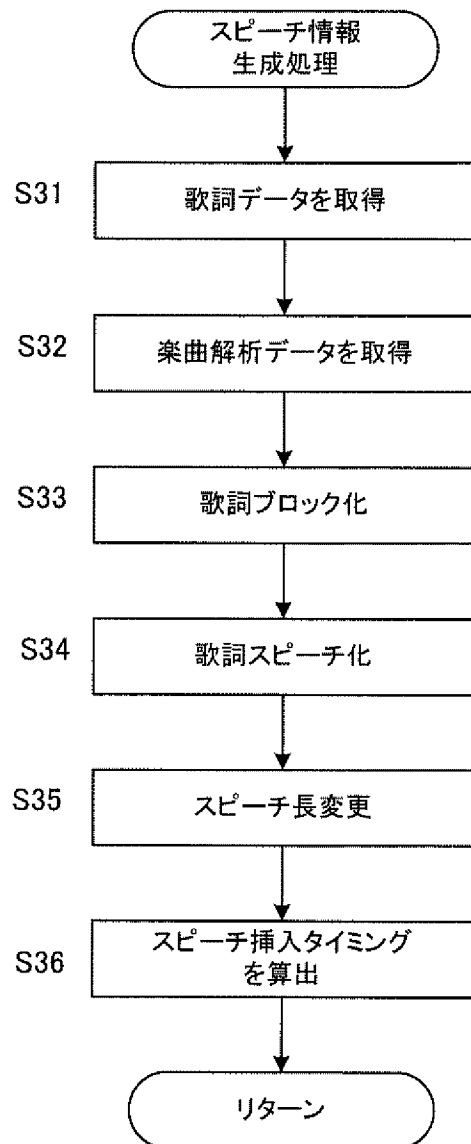
<アシストボーカルの概念>



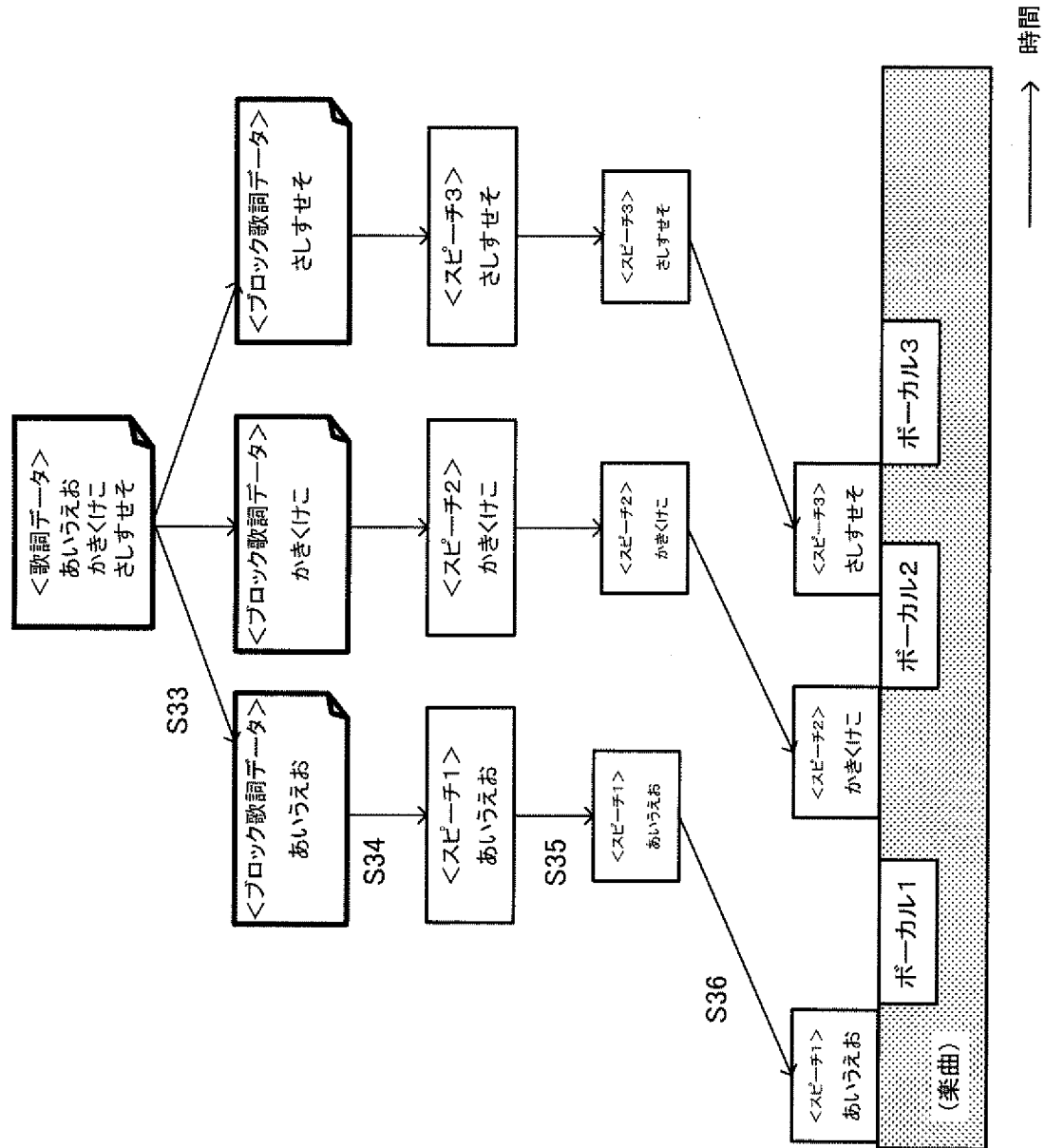
[図2]



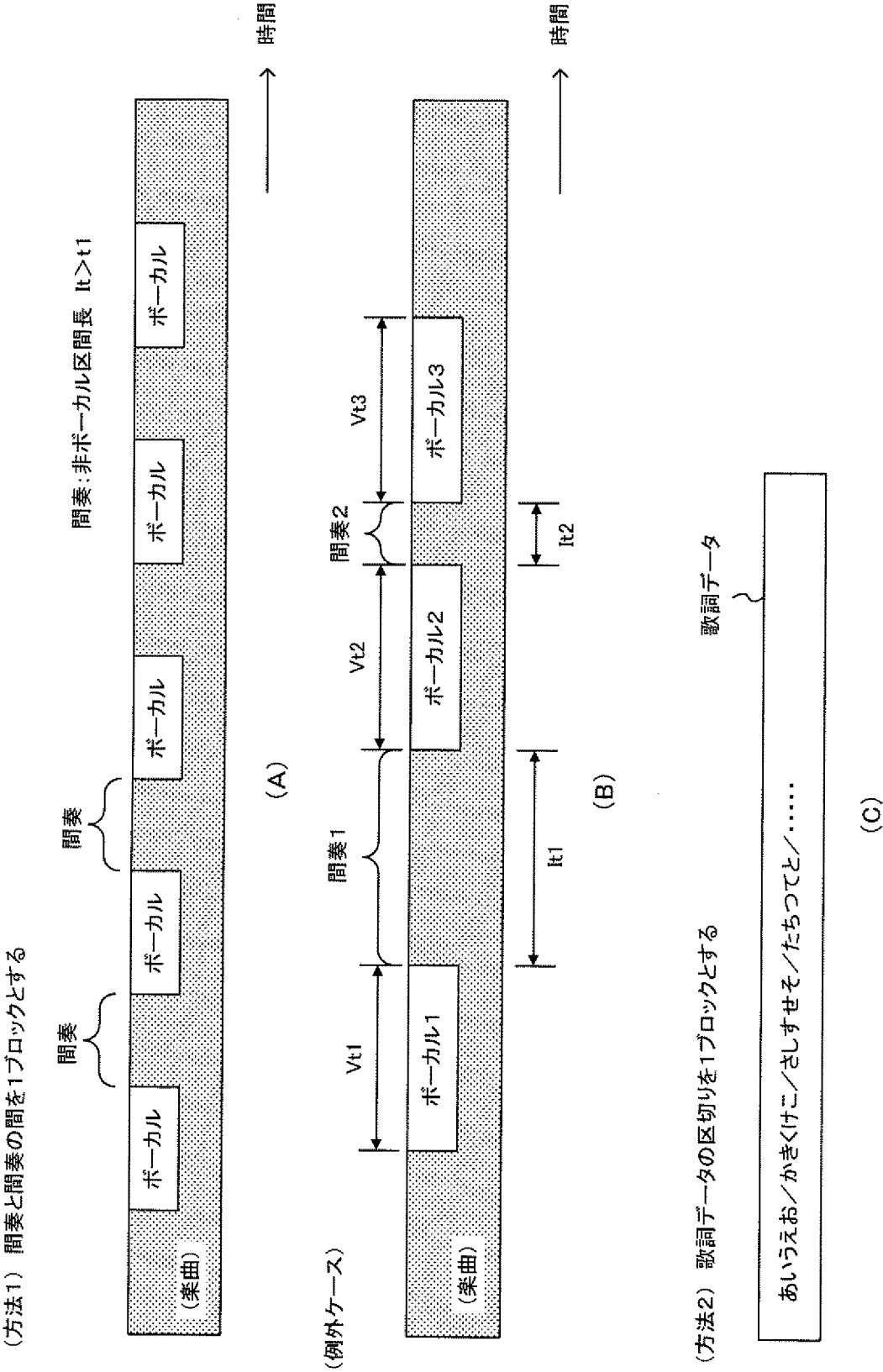
[図3]



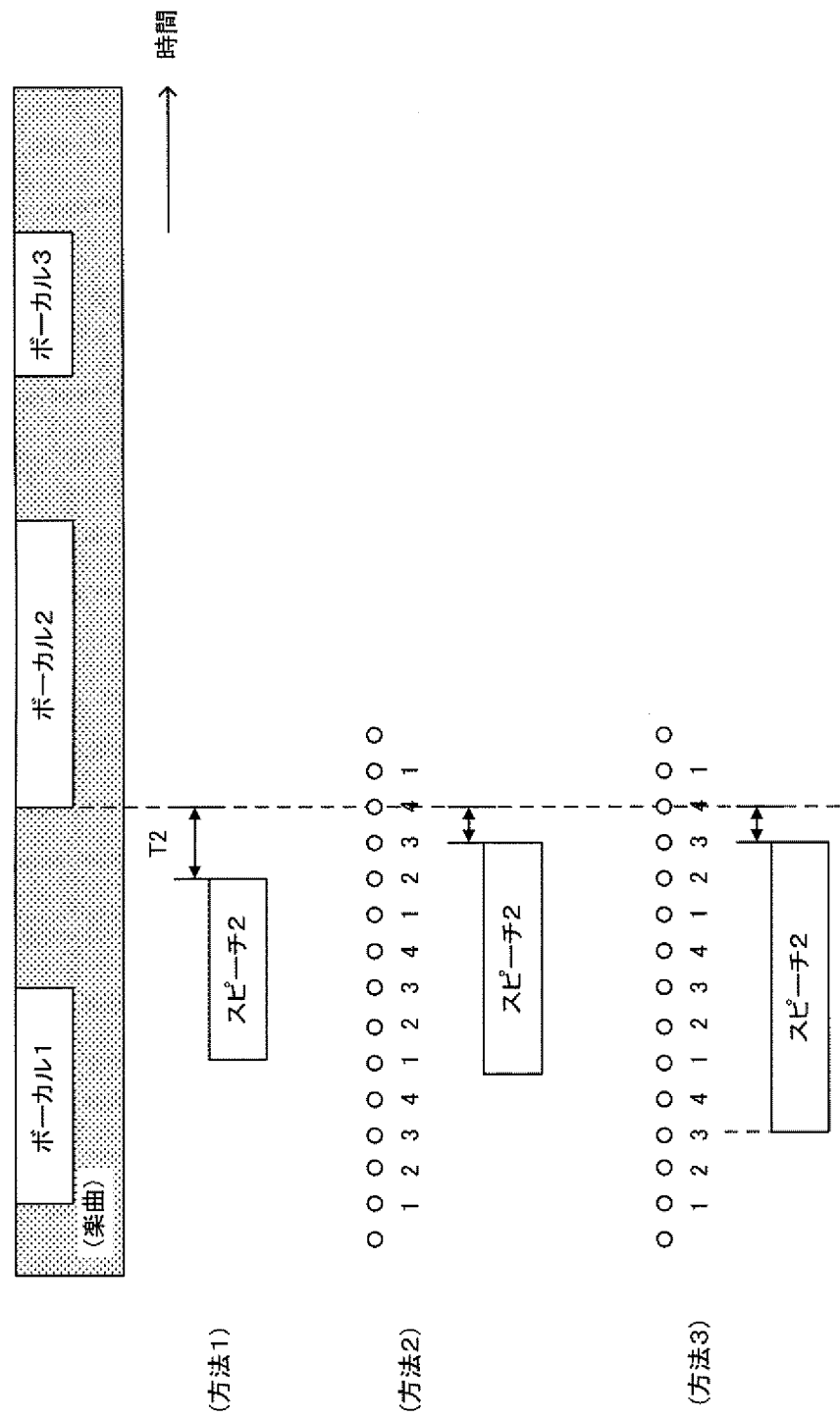
[図4]



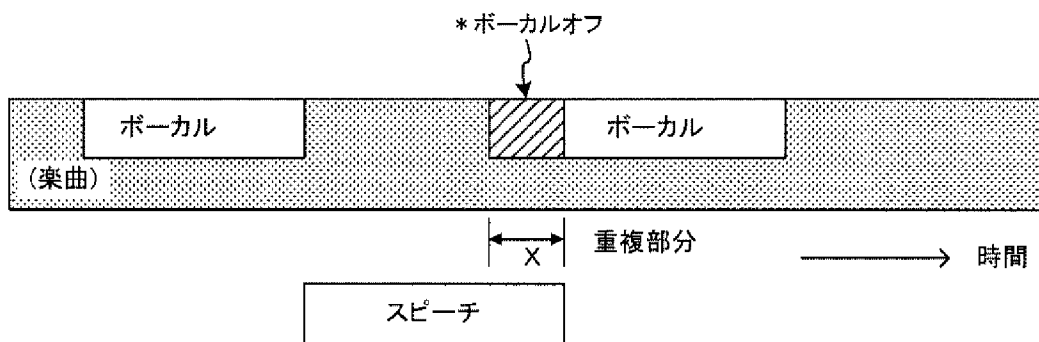
[図5]



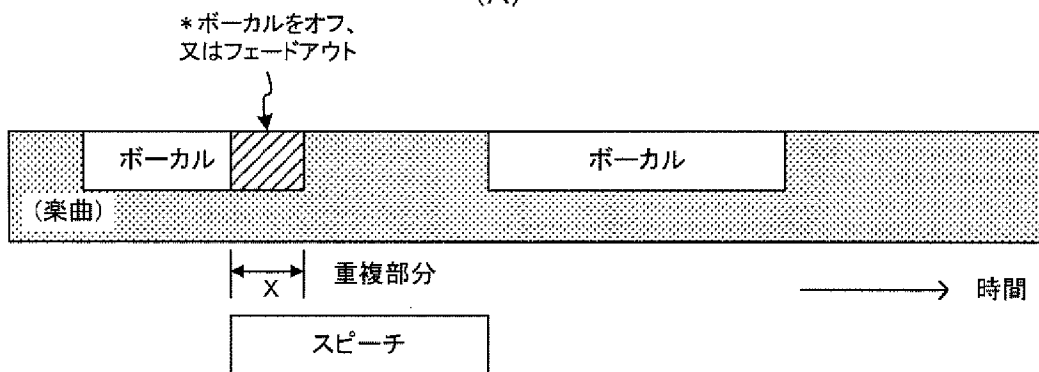
[図6]



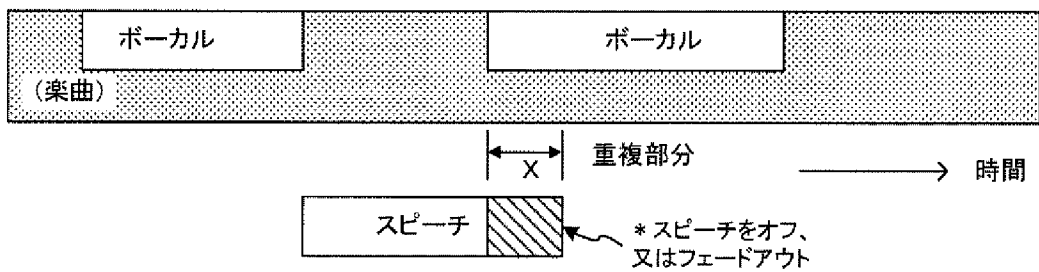
[図7]



(A)

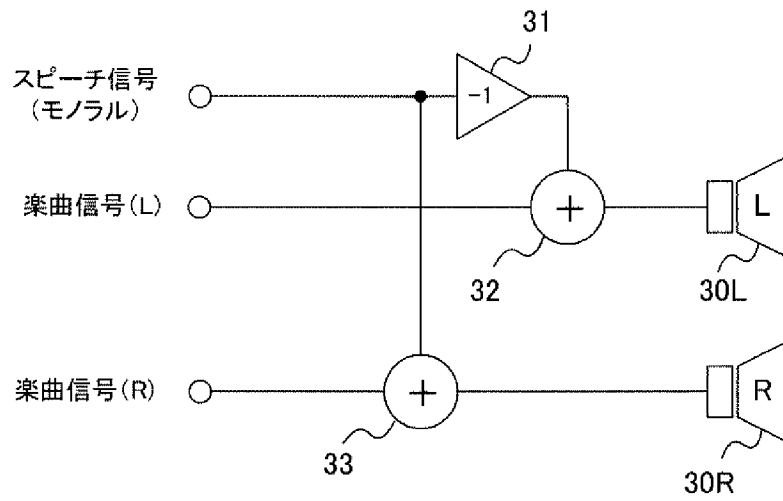


(B)

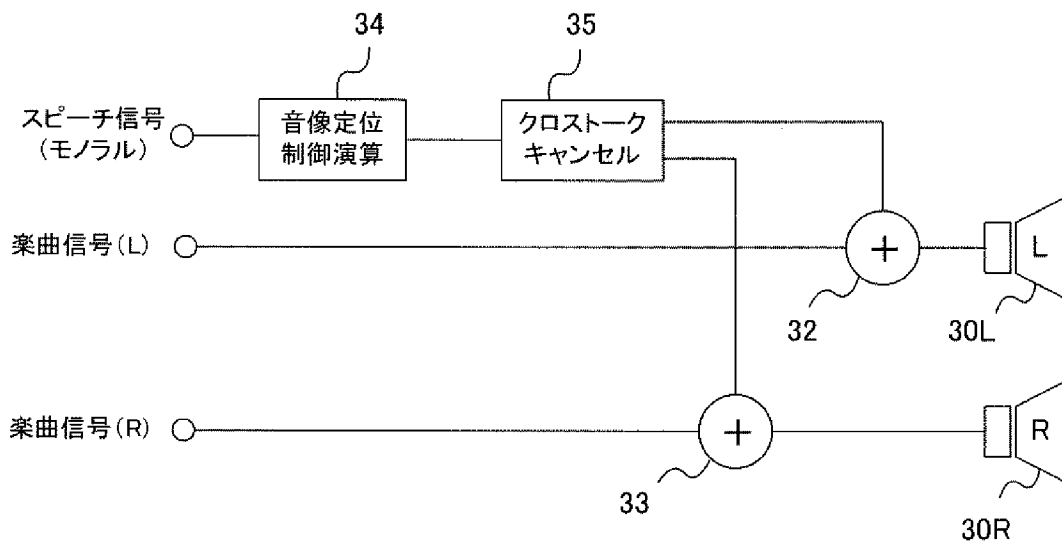


(C)

[図8]

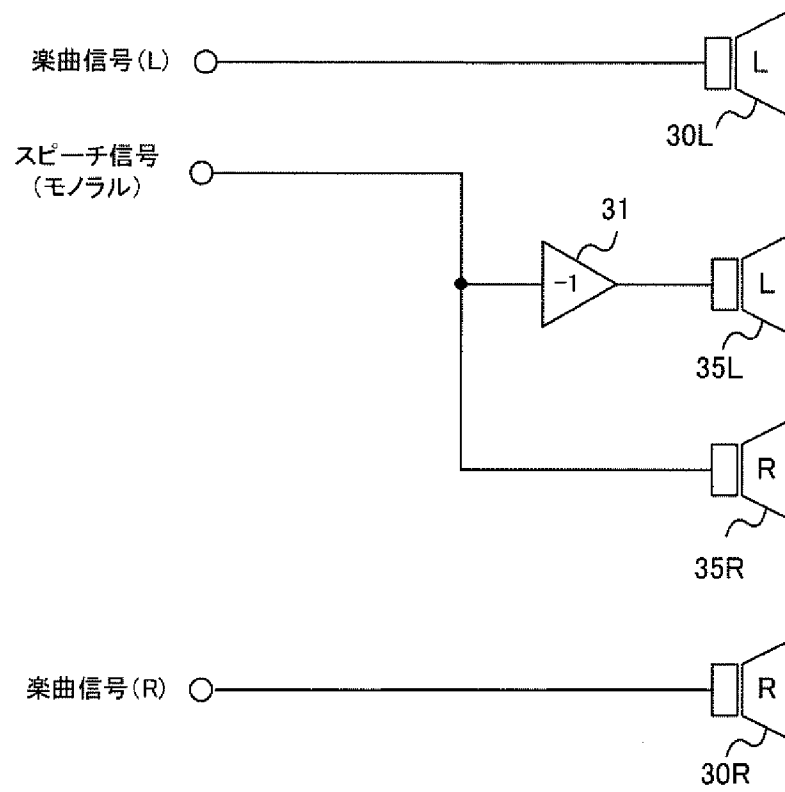


(A)

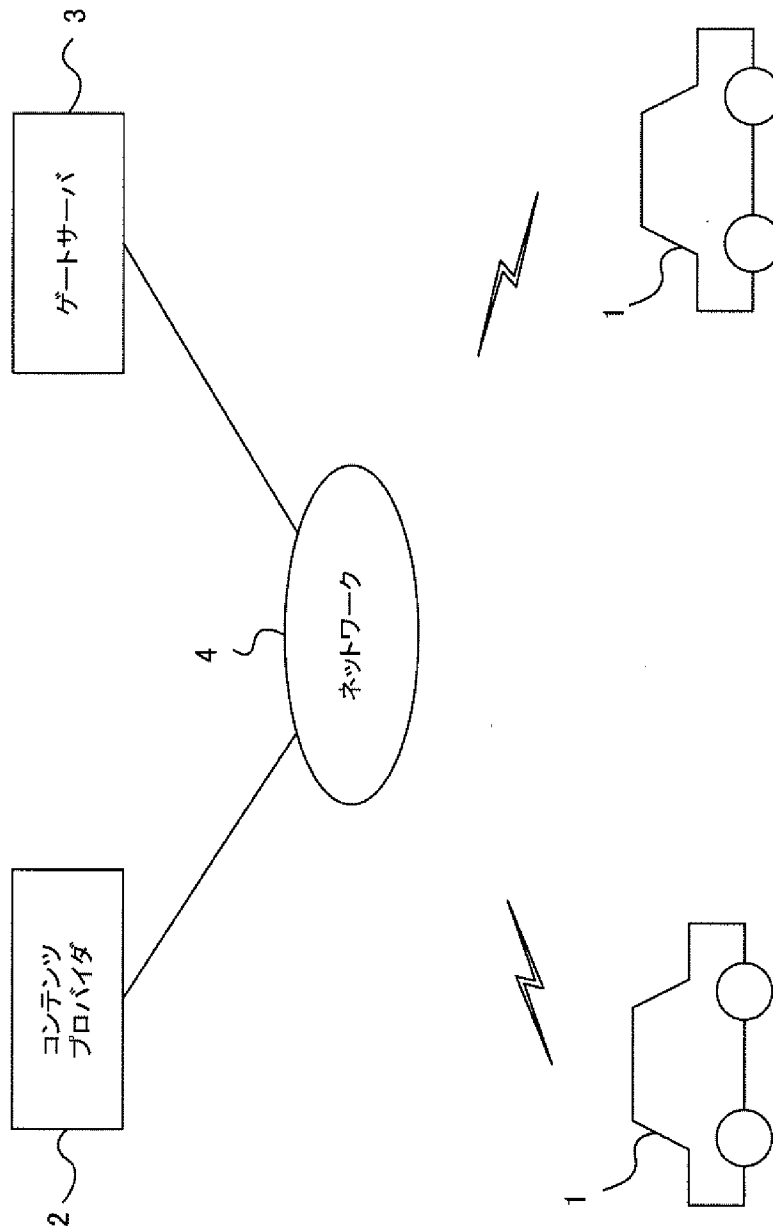


(B)

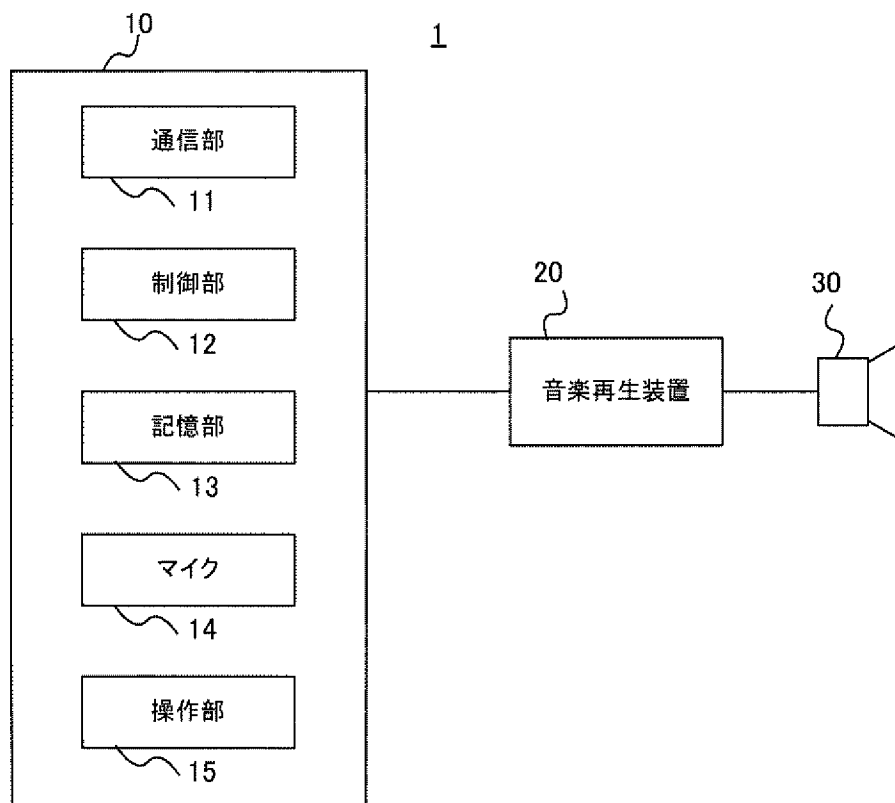
[図9]



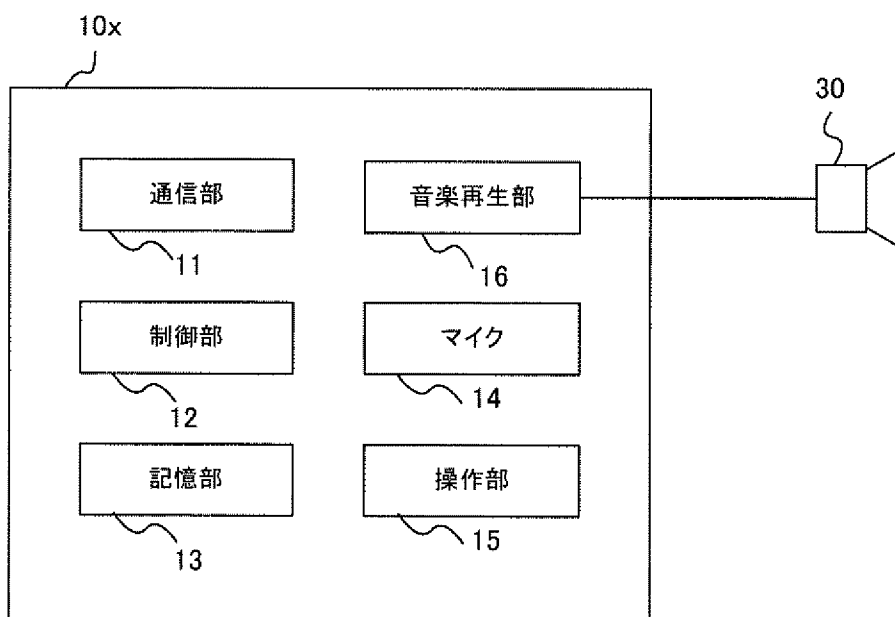
[図10]



[図11]

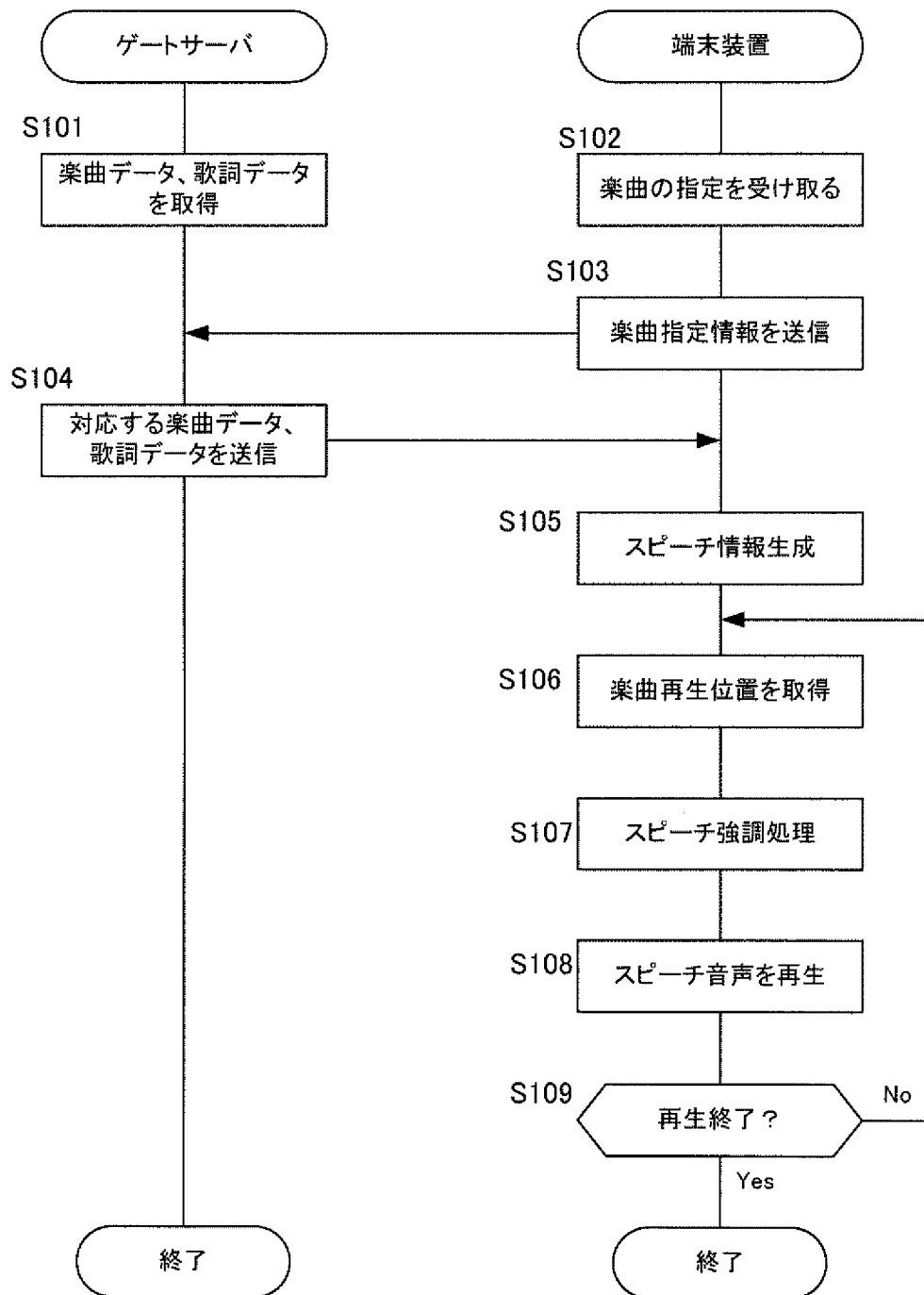


(A)

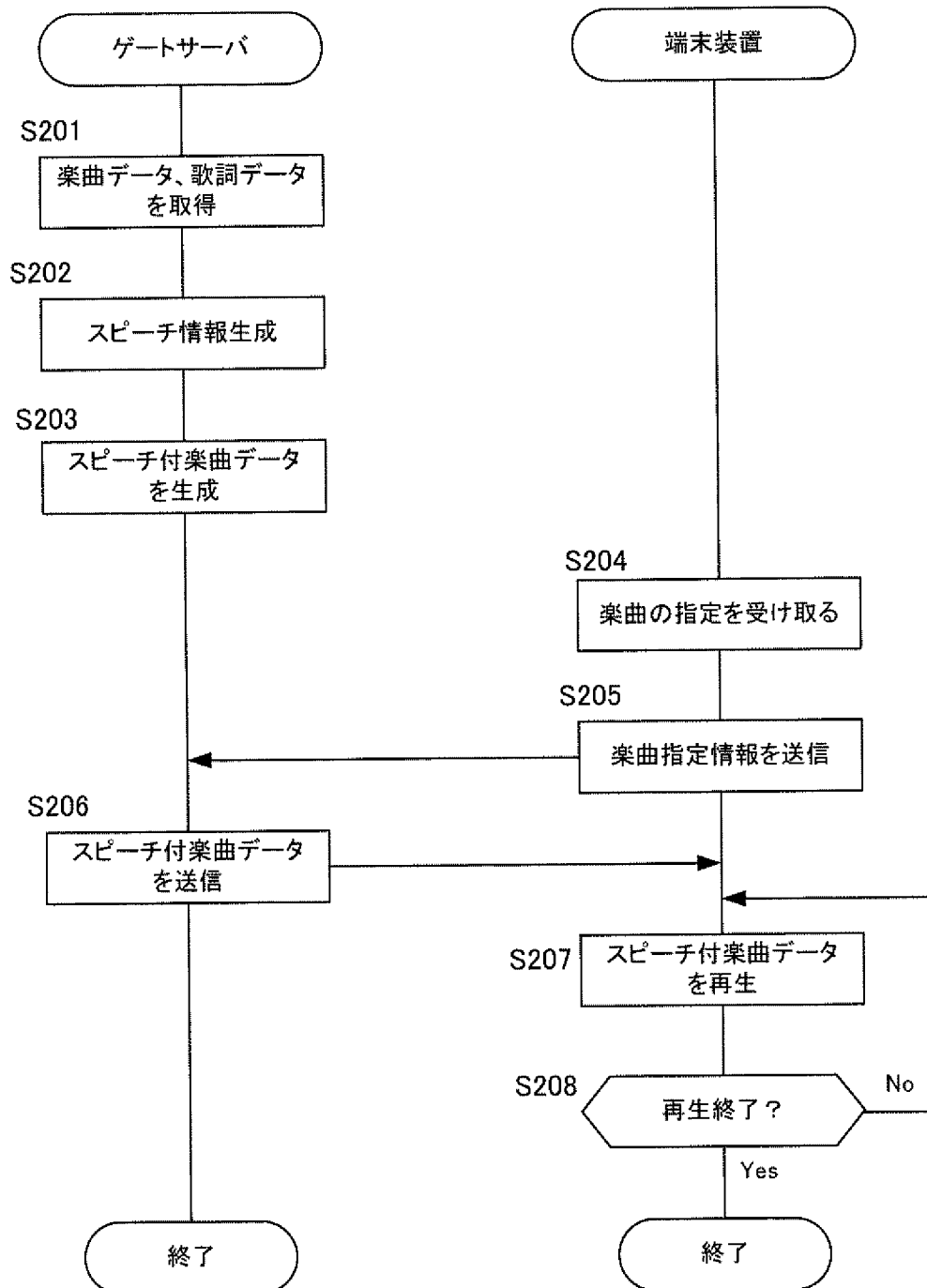


(B)

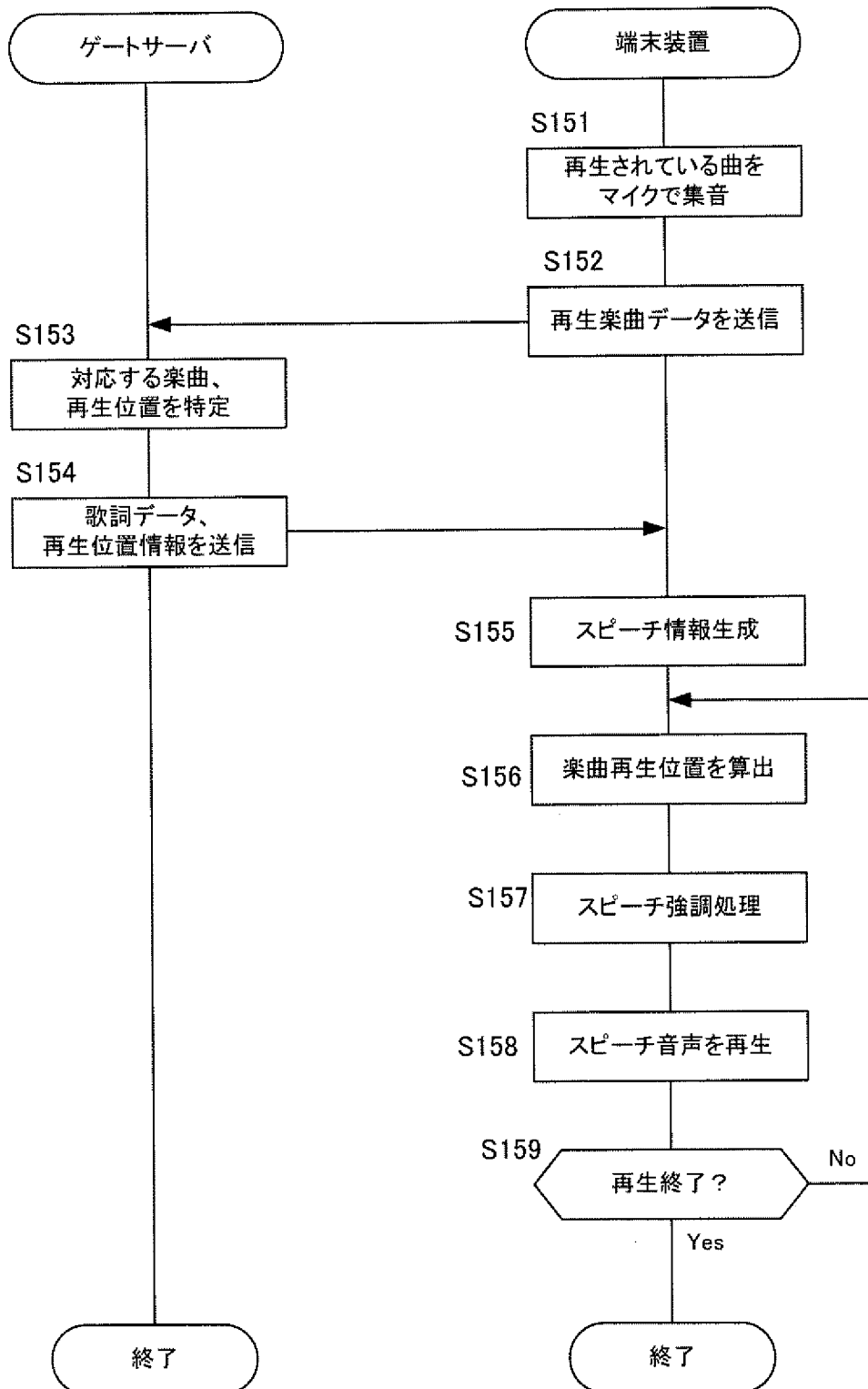
[図12]



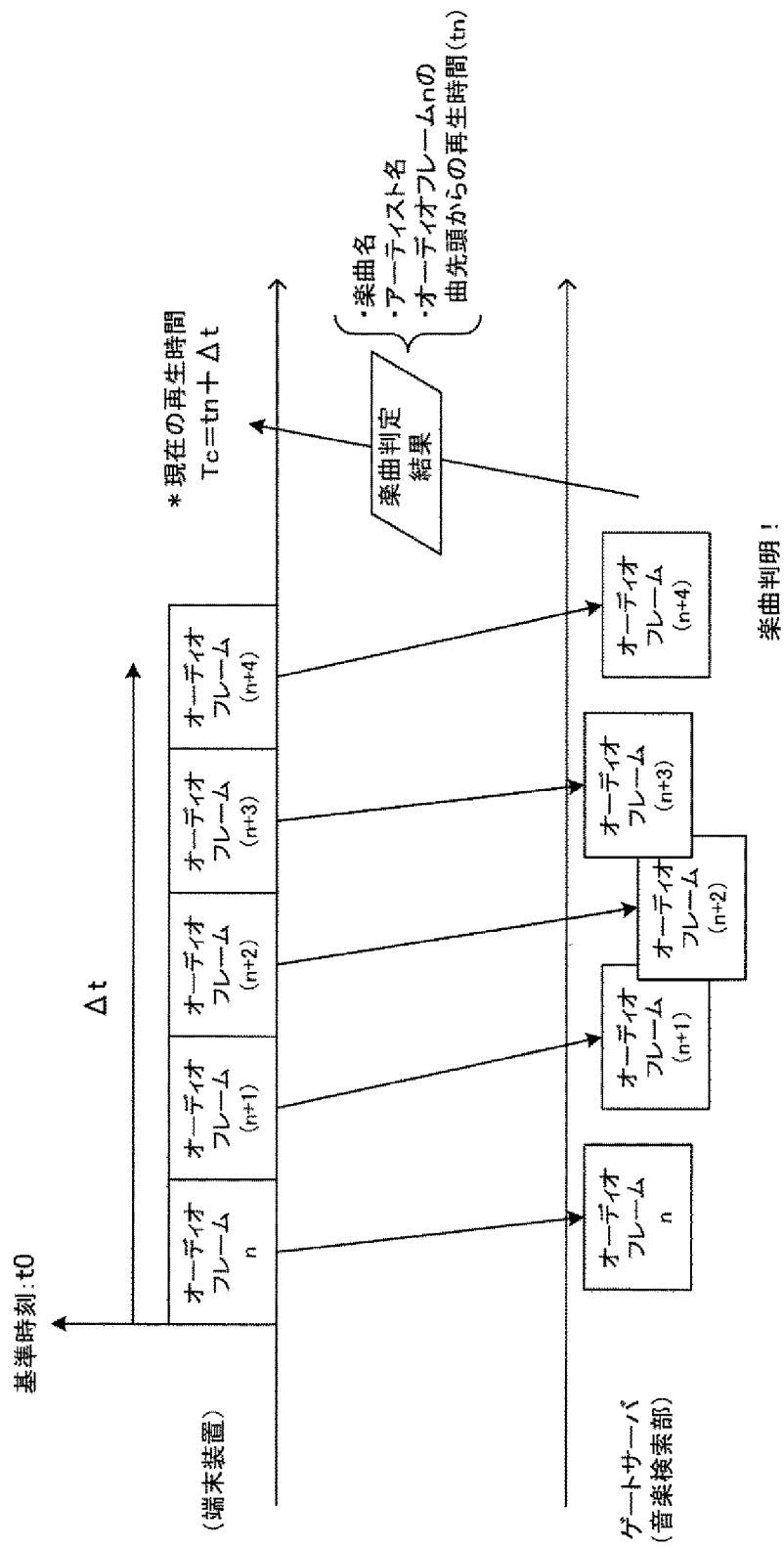
[図13]



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/055609

A. CLASSIFICATION OF SUBJECT MATTER

G10K15/04(2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10K15/04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2015
Kokai Jitsuyo Shinan Koho	1971-2015	Toroku Jitsuyo Shinan Koho	1994-2015

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2005-037846 A (Xing Inc.), 10 February 2005 (10.02.2005), claim 1 (Family: none)	1-6
A	JP 11-024676 A (Yamaha Corp.), 29 January 1999 (29.01.1999), paragraph [0038] & US 6036498 A	2, 3
A	JP 10-240272 A (Taito Corp.), 11 September 1998 (11.09.1998), paragraph [0019] (Family: none)	2, 3

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
30 March 2015 (30.03.15)

Date of mailing of the international search report
07 April 2015 (07.04.15)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2015/055609

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2002-244677 A (Alpine Electronics, Inc.), 30 August 2002 (30.08.2002), paragraph [0011] (Family: none)	1-6
A	JP 2008-216398 A (Daiichikosho Co., Ltd.), 18 September 2008 (18.09.2008), claim 1 (Family: none)	1-6

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G10K15/04(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. G10K15/04

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2015年
日本国実用新案登録公報	1996-2015年
日本国登録実用新案公報	1994-2015年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2005-037846 A（株式会社エクシング）2005.02.10, 【請求項1】 （ファミリーなし）	1-6
A	JP 11-024676 A（ヤマハ株式会社）1999.01.29, 【0038】 & US 6036498 A	2,3
A	JP 10-240272 A（株式会社タイトー）1998.09.11, 【0019】（フ ァミリーなし）	2,3
A	JP 2002-244677 A（アルパイン株式会社）2002.08.30, 【0011】 （ファミリーなし）	1-6

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

30.03.2015

国際調査報告の発送日

07.04.2015

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

富澤 直樹

電話番号 03-3581-1101 内線 3591

5Z

4188

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2008-216398 A (株式会社第一興商) 2008.09.18, 【請求項1】 (ファミリーなし)	1-6