(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
2014 年 7 月 31 日(31.07.2014)

WIPO | PCT

(10) 国際公開番号
WO 2014/115290 A1

(51) 国際特許分類:
H04B 3/23 (2006.01) *H04M 3/56* (2006.01)
G10L 21/0224 (2013.01) *H04R 3/02* (2006.01)
H04M 1/60 (2006.01)

(21) 国際出願番号: PCT/JP2013/051520

(22) 国際出願日: 2013 年 1 月 25 日(25.01.2013)

(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: 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目６番６号 Tokyo (JP).

(72) 発明者: 戸上 真人(TOGAMI Masahito); 〒1858601 東京都国分寺市東恋ヶ窪一丁目２８０番地 株式会社日立製作所 中央研究所内 Tokyo (JP).

(74) 代理人: 井上 学, 外(INOUE Manabu et al.); 〒1008220 東京都千代田区丸の内一丁目６番１号 株式会社日立製作所内 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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.

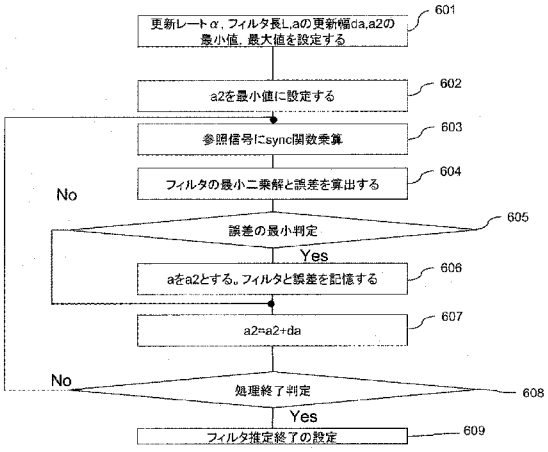
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[続葉有]

(54) **Title**: SIGNAL PROCESSING DEVICE/ACOUSTIC PROCESSING SYSTEM

(54) 発明の名称 : 信号処理装置・音響処理システム

図6



601... SET MINIMUM AND MAXIMUM VALUES FOR UPDATE RATE (α), FILTER LENGTH (L), UPDATE WIDTH (da) FOR (a), AND (a2)
602... SET a2 TO MINIMUM VALUE
603... sync FUNCTION MULTIPLICATION FOR REFERENCE SIGNAL
604... CALCULATE LEAST SQUARE AND ERROR FOR FILTER
605... DETERMINE MINIMUM ERROR
606... SET A TO a2; STORE FILTER AND ERROR
608... DETERMINE THAT PROCESS IS COMPLETE
609... SET FILTER ESTIMATED END

(57) **Abstract**: With prior acoustic echo cancellers, removing an echo component with high accuracy is difficult in cases when a time lag occurs in proportion to time, caused by a lag in the sampling rate between an A/D converter and a D/A converter. In the present invention, an acoustic echo component can be removed with high accuracy, even when a time lag occurs in proportion to time, by using an acoustic echo canceller characterized in having a time-lag coefficient estimation unit for estimating a proportional coefficient for the time-proportional time lag, and further characterized in using the time lag coefficient output by the estimation unit to remove the acoustic echo component.

(57) 要約: 従来の音響エコーキャンセラでは, A/D 変換機と D/A 変換機の間でサンプリングレートがずれていることに起因する時間に比例した時間ずれが起こる場合に、高精度にエコー成分を除去することが困難であった。 本発明では, 時間に比例して起こる時間ずれの比例係数を推定する時間ずれ係数推定部を有し, 更に本推定部が出力する時間ずれ係数を使用して音響エコー成分を除去することを特徴とする音響エコーキャンセラにより, 時間に比例した時間ずれが起こる場合でも, 高精度に音響エコー成分を除去できる。

WO 2014/115290 A1

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

明　細　書

発明の名称 ： 信号処理装置・音響処理システム

技術分野

[0001] A/D変換機でデジタル変換した後の信号から，D/A変換機で出力した信号成分が，何らかの要因によりA/D変換機でデジタル変換した後の信号に混入するような信号処理系において，該混入成分を除去する信号処理除去装置，もしくは音響エコーキャンセラ技術に属する。

背景技術

[0002] D/A変換機及びスピーカ経由で空間に放射した音がマイクロホン及びA/D変換機経由で得られるデジタル信号に混入する音響エコー成分を除去する音響エコーキャンセラが広く用いられてきている。一般的な音響エコーキャンセラは，A/D変換機とD/A変換機に同一のクロックが供給されていることを前提としている。D/A変換機を通して空中に放射する音の原信号をd(t)と記載する。ここで，tは時間インデックスとする。更にA/D変換機を通して得られる音をx(t)と記載する。ここで，x(t)は，一般に，x(t)=h*d(t)+s(t)と記述することができる。ここで，s(t)は環境内にいる話者や雑音からなる信号とする。*は畳み込み混合を表す演算子とする。hは，スピーカから出た音がマイクロホンで受音されるまでに環境内の壁や床などを伝わる伝達関数とする。一般にhは環境に依存するため事前に知ることはできない。一方，d(t)は、D/A変換機に出す前にシステムが保持している信号であり，既知であると考えることができる。

[0003] 遠隔会議システムでは，x(t)の中に含まれるs(t)を抽出し，s(t)を接続先に転送し，接続先のスピーカより放射することで相手に環境内の話者の声を届けることができる。したがってs(t)を如何にx(t)から抽出するかがポイントとなるが，hがもし分かっていれば，s(t)はx(t)からh*d(t)を引き算により抽出することができる。

[0004] 音響エコーキャンセラは、x(t)とd(t)から何らかの方法でhを推定し，得ら

れたhを使ってs(t)を推定する構成を取る。hの推定法としては，e(t)=x(t)-h*d(t)で定義されるe(t)の分散値の期待値E[e(t)*e(t)]ができる限り小さくなるように推定する最小二乗法に基づくLMS(Least Mean Square)法などが一般に使われている。

先行技術文献

特許文献

[0005] 特許文献1：特開２００７－２１４９７６

発明の概要

発明が解決しようとする課題

[0006] 従来の音響エコーキャンセラでは，A/D変換機とD/A変換機において，信号が同じタイミングでサンプリングされている環境で動作することを前提に設計されている。したがって，A/D変換機とD/A変換機のサンプリングのタイミングがずれる場合には，音響エコーキャンセラの音響エコー成分除去性能が低下することが問題となっていた。A/D変換機とD/A変換機のサンプリングのタイミングのずれとして，例えばテレビ内蔵のD/A変換機とスピーカとを接続しているような場合においては，遠隔会議システムにおいて，スピーカ再生用の信号を出力した後，該信号が数百ms遅れで出力されることがよく起こる。このような時間タイミングがスピーカで遅れる場合には，遅れ時間分のオフセットを加味した音響エコーキャンセラにより，遅れの影響無く音響エコー成分が除去できる構成が示されている（例えば，特許文献１参照）。

[0007] しかし，D/A変換機とA/D変換機が同じクロックで駆動されていない場合においては，D/A変換機とA/D変換機の間の時間ずれは時間毎に変化することになる。例えば，D/A変換機とA/D変換機のサンプリングレートが同じ8kHzに設定されている場合であっても，同じクロックを供給されていない限り，同じサンプリングレートとなることは非常にまれである。通常、一方がサンプリングレート8000Hzである場合でも，他方がサンプリングレート8001Hzといったように微小にサンプリングレートがずれることになる。

[0008] 一方，サンプリングレートは，クロックを生成するデバイス（例えば水晶など）に固有であると考えられるため， D/A変換機とA/D変換機毎に，時間毎に一定のレートとなると考えられる。したがって，D/A変換機とA/D変換機は時間に比例して時間がずれることになると考えられる。

[0009] 図４にA/D変換機で収録した音とD/A変換機で収録した音との間の相関係数を時間毎に算出した時の値を示す。横軸は時間軸であり，縦軸はD/A変換機とA/D変換機との間の時間ずれを示している。時間に比例してD/A変換機とA/D変換機との間の時間ずれが大きくなっている様子が見て取れる。よって、特許文献１では、このような時間に比例したD/A変換機とA/D変換機毎の間の時間ずれにより、音響エコー成分を精度よく除去することが出来ない。

[0010] 本発明は、時間に比例したD/A変換機とA/D変換機毎の間の時間ずれがある場合における音響エコー成分の除去を課題とする。

[0011] 従来のLMSアルゴリズムなどの適応フィルタに基づく適応型の音響エコーキャンセラでは，D/A変換後の信号がA/D変換機に届くまでの伝達関数を適応的に推定することができる。理論上は、時間に比例する時間ずれは，この伝達関数に含まれると考えられるが，一般に時間に比例する時間ずれは数秒の間に数サンプルもずれるような非常に早いずれであるため，一般の適応型フィルタでは追従が追いつかない（図４参照）。

課題を解決するための手段

[0012] 本発明では，時間に比例して起こる時間ずれに限定し，時間ずれの時間比例係数を推定することを特徴とする時間ずれ係数推定部と適応型フィルタによる音響エコーキャンセラとを併用する。

発明の効果

[0013] 本発明により広い部屋同士をつなぐビデオ会議システムにおいて，音響エコーの影響が少ないクリアな音声で快適な音声通話が可能となる。また、一般の適応型フィルタを用いた音響エコーキャンセラだけを用いた場合では，理論的に追従ができないような速い時間比例型の時間ずれが起こる場合でも，高精度に追従することを可能とする。

図面の簡単な説明

[0014] [図1]本発明を実施するための最良の形態

[図2]遠隔会議システムにおいて各拠点の会議システムを接続する構成の例

[図3]中央演算装置１０２内で実行する音響処理のブロック構成

[図4]A/D変換機とD/A変換機の間の時間ずれの一例を示した図

[図5]音響処理３０１の具体的なブロック構成

[図6]フィルタ学習部５０５の時間領域処理における処理フロー一例

[図7]フィルタ学習部５０５の周波数領域処理における処理フロー一例

[図8]音響処理３０１の処理フロー一例

[図9]遠隔会議システムにおけるA/D変換機とD/A変換機の接続構成が変化した際に，時間に比例した時間ずれの比例係数を再学習するような構成の一例

[図10]A/D変換機とD/A変換機の種類ごとに時間ずれの比例係数を情報として格納する際のテーブル構造

[図11]D/A変換機付きのスピーカとA/D変換機つきのマイクロホンを用いた遠隔会議システムのシステム構成

発明を実施するための形態

[0015] 本発明は，例えば、広い部屋で使われる遠隔会議システムなどで使用されることを想定している。遠隔会議において，本発明を実施するための一例を図１に示す。

[0016] 図１は，遠隔会議の各拠点毎に設置された会議システムのハードウェア構成の一例を示している。

[0017] 各拠点毎の会議システム１００では，各会議室の中の音声波形をマイクロホン１０５で集音する。マイクロホン１０５は，単一のマイクロホンで無くとも，例えば複数のマイクロホン素子からなったマイクロホンアレイであっても良い。

[0018] 集音したアナログの音声波形は，A/D変換機１０４でアナログ信号からデジタル信号に変換される。変換されたデジタル音声波形は，中央演算装置１０２で残響除去処理を施された後，ＨＵＢ１０８を介してパケットに変換され

ネットワークに放出される。

[0019] 中央演算装置１０２では，不揮発性メモリ１０１に記憶されているプログラム，及びプログラムで用いるパラメータを読み込み，該プログラムを実行する。また，プログラム実行時に用いるワークメモリは，揮発性メモリ１０３上に確保され，音響処理に必要な各種パラメータの記憶領域が定義される。

[0020] 中央演算装置１０２では，遠隔会議における，他拠点（遠端）の音声波形を，ネットワーク越しに，HUB１０８から受け取る。受け取った遠端音声波形（デジタル音声波形）は，中央演算装置１０２経由で，Ｄ／Ａ変換機１０６に送られて，デジタル信号からアナログ信号に変換された後，変換されたアナログの音声波形は、スピーカ１０７から放出される。スピーカ１０７は単一のスピーカ素子で無くとも，例えば複数のスピーカ素子からなるものであっても良い。また，各拠点毎の映像情報は，一般的なカメラ１０９で撮像されＨＵＢ１０８を経由して他拠点に送信される。他拠点の映像情報は，ネットワーク経由でＨＵＢ１０８に送られ，更に中央演算装置１０２を経由して，各拠点毎に設置されたディスプレイ１１０上で表示される。カメラ１０９を複数台設置したり，ディスプレイ１１０を複数台設置するような構成を、取っても良い。

[0021] また，遠隔会議システムの接続や切断をコントロールするユーザーインターフェース１１１を保持する。ユーザーインターフェース１１１は，マウスやキーボード、もしくは赤外線のリモコンなど，当業者が一般的に用いるものを想定する。

[0022] 図２に，遠隔会議システムにおいて各拠点の会議システムを接続する構成の例を示している。拠点数をＮとしＮ個の各拠点毎会議システム（１００－１，１００－２，・・・１００－Ｎ）は，ネットワークでつながれており，音声や映像の流れを制御するＭＣＵ２０２と接続され，各拠点毎の音声や映像の流れが制御される。ＭＣＵは，当業者であれば，既知のシステムであるため，説明は割愛する。

[0023] 図３は，中央演算装置１０２内で実行する音響処理プログラムのブロック構成を示している。マイクロホン１０５からＡ／Ｄ変換機１０４経由で得られたデジタル音声波形は，音響処理３０１で処理される。音響処理３０１は、デジタル音声波形から、音響エコー成分を除去し，各会議室の近端音声発話をピックアップし出力する。この際，音響エコー成分とは，スピーカ１０７から出力された音声波形が各拠点の壁や天井などで反射した後，マイクロホン１０５に混入する成分を指す。音響処理３０１では音響エコー成分を除去するためにＨＵＢ１０８経由で得られる遠端音声波形を用いる。遠隔会議制御処理３０２は，遠隔会議の開始や終了，また接続先の選択といった遠隔会議システムの制御に関する処理を行うブロックである。遠隔会議制御処理３０２は，当業者であれば，容易に類推可能な遠隔会議用の制御処理であるため，詳細構成の説明は割愛する。

[0024] 図５に音響処理３０１の詳細なブロック構成を示す。

[0025] 音声取り込み部５０１は，各時間のマイクロホン入力信号x(t)を取り込む。ここで，tはA/D変換のサンプル番号とする。また音声受信部５０２は，遠隔会議システムが接続されている接続先から送信されてくるスピーカ再生用の信号d(t)をHUB１０８から受け取る。スピーカ再生用の信号d(t)はD/A変換機１０４を経由してスピーカ１０７から放射される。

[0026] バッファリング部５０３は，音声取り込み部５０１で取り込んだマイクロホン入力信号x(t)と音声受信部５０２で取り込んだスピーカ再生用の信号d(t)を一定量，揮発性メモリ１０３か，不揮発性メモリ１０１内に貯めておく。

[0027] 短時間フーリエ変換５０４では，バッファリング部５０３がバッファリングしたマイクロホン入力信号x(t)，及びスピーカ再生用の信号d(t)に対して，に対して、短時間フーリエ変換を施し，時間周波数領域信号を得る。ここで，得られる時間周波数領域信号をx(f,τ)及びd(f,τ)と表記する。ここで，x(f,τ)はマイクロホン入力信号x(t)の短時間フーリエ変換後の信号であり，d(f,τ)はスピーカ再生用の信号d(t) の短時間フーリエ変換後の信号である。ここで，fは周波数インデックス，τは短時間フーリエ変換のフレームイ

ンデックスとする。

[0028] フィルタ学習部５０５は、時間周波数領域信号かまたは時間領域信号から、音響エコー除去用のフィルタhと，D/A変換機とA/D変換機との間の時間に比例した時間ずれの比例係数aとを学習する。ここで，音響エコー除去用のフィルタhとD/A変換機とA/D変換機との間の時間に比例した時間ずれの比例係数 aの学習方法は，時間領域信号から学習する場合と，時間周波数領域信号から学習するかで異なる。学習した音響エコー除去用のフィルタhとD/A変換機とA/D変換機との間の時間に比例した時間ずれの比例係数aは，揮発性メモリ１０３か不揮発性メモリ１０１上に確保するフィルタDB５０６に保存される。

[0029] 音響エコーキャンセラ部５０７は，フィルタ学習部５０５で学習した音響エコー除去用のフィルタhと時間に比例した時間ずれの比例係数aを用いてマイクロホン入力信号x(t)中に含まれるスピーカからの放射成分を除去し，近端音声を抽出する。音響エコーキャンセラ部５０７においても，時間領域信号から音響エコー成分を除去するか，時間周波数領域信号から音響エコー成分を除去するかで処理が分かれる。いずれの場合においても，音響エコーキャンセラ部５０７は，スピーカ再生用の信号d(t)もしくはd(f,τ)に対して音響エコー除去用のフィルタhを畳込むことで、マイクロホン入力信号x(t)内の音響エコー成分の推定値e(t)もしくはe(f,τ)を推定し，これをマイクロホン入力信号x(t)もしくはx(f,τ)から引き算することで，音響エコー成分を除去する。

[0030] 逆フーリエ変換５０８では，音響エコーキャンセラ部において音響エコー成分を除去した後の時間周波数領域信号を、時間領域信号に変換する。音声送信部５０９では，音響エコー成分除去後の時間領域信号を接続先に送信する。

[0031] 図６は，フィルタ学習部５０５において，時間領域信号から音響エコー除去用のフィルタhとD/A変換機とA/D変換機との間の時間に比例した時間ずれの比例係数aを学習する際の処理フローを示した図である。本フィルタ部は，時間領域信号を１サンプル得る毎に動作させるような構成を取っても良いし，

複数サンプル得る毎に動作させるような構成を取っても良い。本構成は，従来の音響エコーキャンセラの適応フィルタの一種であるＬＭＳ（Least Mean Square）アルゴリズムや，ＲＬＳ（Recursive Least Square）アルゴリズムの拡張構成となっている。

[0032] まず，ＬＭＳアルゴリズムやＲＬＳアルゴリズムと同様に，フィルタ学習部５０５は、フィルタを時間毎に変化させていく際の変化率である更新レートαと，音響エコー除去用フィルタhのフィルタ長Ｌを定める。また，比例係数aを探索的に最適化する際の探索範囲の最小値と最大値をそれぞれamin，amaxとして定める。また比例係数aを探索する際の刻み幅をdaとする（６０１）。本処理フローでは，比例係数aを最小値aminから最大値amaxまで変化させた時の最も良い値を選択する。探索に用いる仮の比例係数aをa2とする。仮の比例係数a2をaminとして設定する（６０２）。

[0033] 次に，参照信号（スピーカ再生信号）d(t)を時間比例係数a2分だけ時間に比例して変化させる。つまり，dd(t)=d(t-a2t)となるような信号dd(t)を生成する。ここで，dd(t)は，元のd(t)にsync関数を掛けたり，サンプリングレート変換を行うことで生成できる（６０３）。これは当業者であればよく知られて方法であるため，詳細な説明は割愛する。

[0034] 次に，求めたdd(t)を用いて，フィルタの最小二乗解と誤差を算出する。フィルタhtempは，数1に示すような重み付き最小二乗法により求める。当業者であれば，本重み付き最小二乗解は容易に算出可能である。またその時の誤差を数2で示すerrとする。

[0035] [数1]

【数１】
$$h_{temp} = \arg\min_{h} \sum_{t_2=0}^{t} \alpha^{t-t_2} \left\| x(t_2) - h * dd(t_2) \right\|^2$$

[0036] [数2]

【数２】
$$err = \sum_{t_2=0}^{t} \alpha^{t-t_2} \left\| x(t_2) - h_{temp} * dd(t_2) \right\|^2$$

[0037] 比例係数a2が最小値aminの場合には誤差errをmin_errとし，音響エコー除

去用フィルタhをhtempとする。それ以外の場合には，誤差errがmin_errよりも小さかった場合に，音響エコー除去用フィルタhをhtempとし，誤差errをmin_errとする（６０５，６０６）。次に，比例係数a2に刻み幅daを加える（６０７）。比例係数a2が最大値amaxよりも大きい場合には，フィルタ推定処理を終了する（６０９）。逆に小さい場合には，６０３の前に戻る（６０８）。以上により、フィルタ学習部５０５は，時間領域信号から音響エコー除去用のフィルタhとD/A変換機とA/D変換機との間の時間に比例した時間ずれの比例係数aとを学習する。

[0038] 図７は、フィルタ学習部５０５において，時間周波数領域の信号から時間に比例した時間ずれの比例係数a及び音響エコー除去用フィルタhを求める処理フローを示した図である。時間領域の場合と同様に，フィルタを時間毎に変化させていく際の変化率である更新レートαと，フィルタ長Ｌ，短時間フーリエ変換のフレーム幅Ｌｆ，フレームシフトＬｓを定める（７０１）。次に短時間フーリエ変換により時間周波数領域のマイクロホン入力信号x(f,τ)と，参照信号d(f,τ)を算出する（７０２）。次に，比例係数aに暫定的な値を入れる（例えば，0など）（７０３）。算出した時間周波数領域のマイクロホン入力信号x(f,τ)と，参照信号d(f,τ)を用いてフィルタの最小二乗解を数３で求める（７０４）。

[0039] [数3]

【数3】

$$h_{temp}(f) = \arg\min_{h(f)} \sum_{\tau_2=0}^{L} \alpha^{t[\tau]-t[\tau_2]} \left\| x(f,\tau_2) - h(f) * d(f,\tau_2) \exp(-j2\pi fat[\tau_2]) \right\|^2$$

[0040] ここで、t[τ]は，フレームτの開始サンプルに相当する時間インデックスと定義する。次に，音響エコー成分の推定値e(f,τ)を数４で定義する。

[0041] [数4]

【数4】

$$e(f,\tau) = h_{temp}(f) * d(f,\tau)$$

[0042] 次に，数５でe(f,τ)とx(f,τ)の正規化した位相差を算出する（７０５）。

[0043]

[数5]

【数5】 $$c(f,\tau_2) = \frac{-1}{2\pi f}\arg\left(e(f,\tau_2)^H x(f,\tau_2)\right)$$

[0044] ここで，Hは共役転置を取るための演算子であり，argは複素数の位相差を算出するための演算子とする。C(f,τ)より数６で時間に比例した時間ずれの比例係数aの推定値であるa^を求め，a^をaに代入する（７０６）。

[0045] [数6]

【数6】 $$\hat{a} = \frac{\sum_f \sum_{\tau_2=0}^{\tau} \alpha^{t[\tau]-t[\tau_2]} c(f,\tau_2) t[\tau_2]}{\sum_f \sum_{\tau_2=0}^{\tau} \alpha^{t[\tau]-t[\tau_2]} t[\tau_2]^2}$$

[0046] ７０６の処理を所定回数行っていれば，求めたhtemp(f)を音響エコーキャンセラ用のフィルタhとして，フィルタ推定を終了する（７０８）。それ以外の場合，７０４に戻る（７０７）。以上により、フィルタ学習部５０５は，時間周波数領域の信号から時間に比例した時間ずれの比例係数a及び音響エコー除去用フィルタhを求める。

[0047] 図８に，本発明において，時間に比例した時間ずれ比例係数を遠隔会議システムの立ち上げ時に学習する場合の遠隔会議システムの動作フローチャートを示す。A/D変換機１０４及びD/A変換機１０６はサンプリングをまず開始する（８０１）。次に，中央演算装置１０２は、前述した時間領域もしくは，時間周波数領域におけるフィルタ学習により，時間に比例した時間ずれ比例係数aを算出する（８０２）。算出した比例係数aはフィルタDBに格納される。時間に比例した時間ずれ比例係数aの算出にあたってD/A変換機から音を出す必要があるが，これは必ずしも遠隔地の音である必要はなく，遠隔会議システム内に予め用意してある音を再生するような構成をとっても良い。そして，遠隔地と接続し遠隔会議を開始する（８０３）。８０２において音響処理３０１のフィルタ学習部５０５が時間に比例した時間ずれ比例係数aを算出すると、遠隔会議制御処理３０２は、遠隔会議開始制御を開始する。

[0048] 遠隔会議中は，遠隔地から送られてきたスピーカ再生用信号d(t)（参照信号）を受信（８０４），及び、マイクロホン入力信号x(t)を受信（８０５）

するたびに，音響処理３０１の音響エコーキャンセラ部５０７は、まず、受信した参照信号d(t)を予めステップ８０２で算出した時間に比例した時間ずれ係数aを用いてd(t-at)のように補正する（８０６）。参照信号d(t)の補正は，sync関数を波形に重畳したり，サンプリングレート変換など，当業者であれば，よく知られた方法により行う。補正後の参照信号をdd(t)=d(t-at)と記載する。また，時間周波数領域における音響エコーキャンセラ用フィルタを用いる場合は，dd(t)を時間周波数領域信号dd(f，τ)に変換する。

[0049] 次に，フィルタ学習部５０５は、参照信号としてdの代わりにddを用いて，ＬＭＳやＲＬＳなどの方法で，逐次音響エコーキャンセラフィルタhを更新する。音響エコーキャンセラフィルタhの更新は，時間領域かまたは時間周波数領域で行う（８０７）。

[0050] 次に，音響エコーキャンセラ部５０７は、求めたフィルタhを参照信号dd(t)に重畳し，音響エコーレプリカを作成し，作成した音響エコーレプリカをマイクロホン入力信号x(t)から引き算することにより，音響エコー成分が除去された信号を得る。時間周波数領域における音響エコーキャンセラの場合は、逆フーリエ変換５０８において、逆フーリエ変換により時間領域のエコー成分除去後の信号を得る（８０８）。

[0051] そして，音声送信部５０８は、音響エコー成分除去後の信号をＨＵＢを介して，接続先に送信する（８０９）。ユーザーインターフェースを介したユーザー制御により遠隔会議システムが終了される場合は，遠隔会議制御処理３０２は、処理を終了し，それ以外の場合は，参照信号受信（８０４）に戻る。

[0052] 図９は，遠隔会議システムで用いるマイクロホンまたはスピーカが切り替わる場合に，接続構成毎の時間に比例した時間ずれ係数を再設定するように，本発明を使用する際のフローチャートである。

[0053] A/D変換機１０４及びD/A変換機１０６はサンプリングをまず開始する（９０１）。次に，フィルタ学習部５０５は、前述した時間領域もしくは，時間周波数領域におけるフィルタ学習により，時間に比例した時間ずれ係数aを算

出する（９０２）。このaは，初期値として用いる。そして，遠隔地と接続し遠隔会議を開始する（９０３）。次に，フィルタ学習部５０５は、A/D変換機１０４及びD/A変換機１０６の接続が変化したかどうかを判定する。変化している場合は，aの再設定(９０５)に移る。

[0054] ここで，接続されているA/D変換機１０４、D/A変換機１０６及びそれらのサンプリングレート毎にそれに対応した時間に比例した時間ずれ比例係数aをフィルタDB上に，図１０に示すようなデータ構造で，有しておくものとする。

[0055] ここで，切り替わった後のA/D変換機とD/A変換機の構成が図１０のテーブル上に存在する場合，フィルタ学習部５０５は、図１０のテーブルを参照し，接続されているA/D変換機１０４及びD/A変換機１０６及びそれらのサンプリングレートに対応する時間に比例した時間ずれ比例係数aを新しい時間に比例した時間ずれ比例係数aとして用いる。図１０のテーブル上に対応するA/D変換機、D/A変換機、及びそれらのサンプリングレートが無い場合は，フィルタ学習部５０５は、前述した時間領域もしくは時間周波数領域におけるフィルタ学習によりaを求める。A/D変換機とD/A変換機の接続関係が変化していない場合は，通常の遠隔会議の音声通信状態に移る。

[0056] 遠隔会議中は，遠隔地から送られてきたスピーカ再生用信号（参照信号）の受信と（９０６），マイク入力信号の取得（９０７）を行うたびに，音響エコーキャンセラ部は、まず参照信号補正（９０８）にて、得られた参照信号を予め算出した時間に比例した時間ずれ係数aを用いてd(t-at)のように補正する。補正は，sync関数を波形に重畳したり，サンプリングレート変換など，当業者であれば，よく知られた方法により行う。補正後の信号をdd(t)=d(t-at)と記載する。また，時間周波数領域における音響エコーキャンセラ用フィルタを用いる場合は，dd(t)を時間周波数領域信号dd(f, τ)に変換する。

[0057] 次に，フィルタ学習部は、参照信号としてdの代わりにddを用いて，ＬＭＳやＲＬＳなどの方法で，逐次音響エコーキャンセラフィルタhを更新する。更

新は，時間領域かまたは時間周波数領域で行う（９０９）。次に音響エコーキャンセラ部は、求めたフィルタhを参照信号に重畳し，音響エコーレプリカを作成し，作成した音響エコーレプリカをマイク入力信号から引き算することにより，音響エコー成分が除去された信号を得る。時間周波数領域における音響エコーキャンセラの場合は逆フーリエ変換により時間領域の音響エコー成分除去後の信号を得る（９１０）。

[0058] そして，音声送信部５０９は、音響エコー成分除去後の信号をＨＵＢを介して，接続先に送信する（９１１）。ユーザーインターフェースを介したユーザー制御により遠隔会議システムが終了される場合は，処理を終了し，それ以外の場合は，接続切り替え判定（８０４）に戻る。

[0059] 図１１に，D/A変換機付きのスピーカとA/D変換機つきのマイクロホンを用いた遠隔会議システムの実施例を示す。D/A変換機能が内在されたD/A変換機能付きスピーカ1101と，A/D変換機能付き無線マイクロホン1102には，アナログ信号ではなく，デジタル信号が入力もしくは出力される。したがってこれらは異なるクロックでサンプリングされることとなる。

[0060] 音声受信部１１０６は，接続先の遠隔会議システムから送信された音声信号をＨＵＢ１０８経由で受け取る。受け取った音声信号は，スピーカ再生１１０５により，D/A変換機能が内在されたD/A変換機能付きスピーカ1101に送られて再生される。無線受信機１１０３は，A/D変換機能付き無線マイクロホン1102で収録したマイク入力信号を受信する。音響エコーキャンセラ１１０４は，マイク入力信号中に含まれる，D/A変換機能が内在されたD/A変換機能付きスピーカ1101から再生した音声の音響エコー成分を除去する。除去は，本発明の第一の実施例で示した構成により，時間に比例した時間ずれを補正すると共に行う。音声送信部１１０７は，音響エコーキャンセラ後の信号をＨＵＢ１０８経由で遠隔地に送信する。音響エコーキャンセラ１１０４は，音響処理３０１と同様の処理を行うものとする。

符号の説明

[0061] 100…拠点毎会議システム、101…不揮発性メモリ、102…中央演算装置、10

3…揮発性メモリ、104…A/D変換機、105…マイクロホン、106…D/A変換機、107…スピーカ、108…ＨＵＢ、109…カメラ、110…ディスプレイ、110…ユーザーインターフェース、201…ＭＣＵ、301…音響処理、302…遠隔会議制御処理、501…音声取り込み部、502…音声受信部、503…バッファリング、504…短時間フーリエ変換、505…フィルタ学習部、506…フィルタDB、507…音響エコーキャンセラ、508…短時間フーリエ変換、509…音声送信部、1101…D/A変換機能付きスピーカ、1102…A/D変換機能付き無線マイクロホン、1103…無線受信機、 1104…音響エコーキャンセラ、 1105…スピーカ再生、1106…音声受信部、1107…音声送信部

請求の範囲

[請求項1] 音声を入力するマイクロホンと、

前記マイクロホンからの信号をデジタル変換するＡＤ変換器と、

前記ＡＤ変換器からのデジタル信号を処理し音響エコー成分を抑圧する情報処理装置と、

前記情報処理装置からの信号をネットワークに送出する出力インターフェースと、

前記ネットワークからの信号を受信する入力インターフェースと、

前記入力インターフェースからの信号をアナログ変換するＤＡ変換器と、

前記ＤＡ変換器からの信号を音声として出力するスピーカと、備えた音声システムであって、

前記情報処理装置は、

前記A/D変換機と前記D/A変換機との間の時間に比例した時間ずれ係数を推定するフィルタ学習部と、

前記推定した時間に比例した時間ずれ係数を用いて、前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去する音響エコーキャンセラ部と、を備えることを特徴とする音声システム。

[請求項2] 前記フィルタ学習部は、

前記ネットワークから入力されるデジタル音声波形と、前記ＡＤ変換器で得られたデジタル音声波形と、に基づいて前記時間に比例した時間ずれ係数を推定することを特徴とする請求項１に記載の音声システム。

[請求項3] 前記音響エコーキャンセラ部は、前記ネットワークから入力されるデジタル音声波形を前記推定した時間に比例した時間ずれ係数を用いて補正し、

前記フィルタ学習部は、補正したデジタル音声波形を用いて音響エコーキャンセラフィルタを更新し、

前記音響エコーキャンセラ部は、更新した音響エコーキャンセラフィルタを前記補正したデジタル音声波形に重畳して音響エコーレプリカを作成し，作成した前記音響エコーレプリカを前記ＡＤ変換器で得られたデジタル音声波形から引き算することで音響エコー成分を除去することを特徴とする請求項２に記載の音声システム。

［請求項4］ 前記情報処理装置は、

A/D変換機を識別する識別子、D/A変換機を識別する識別子、該A/D変換機のサンプリングレート、及び、該D/A変換機のサンプリングレートの組み合わせに対応した時間に比例した時間ずれ係数を保持するデータベースを備え、

接続されるA/D変換機またはD/A変換機が変更されたことを検出すると、前記データベースを参照して、接続されるA/D変換機、D/A変換機、及び、該A/D変換機と該D/A変換機のサンプリングレートに対応する時間に比例した時間ずれ係数を取得し、取得した時間に比例した時間ずれ係数を用いて前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去することを特徴とする請求項３に記載の音声システム。

［請求項5］ 音響エコー成分を除去する方法であって、

マイクロホンで音声を受信するステップと、

前記マイクロホンからの信号をＡＤ変換器でデジタル変換するステップと、

ネットワークからの信号を受信するステップと、

前記ネットワークから受信した信号をＤＡ変換器でアナログ変換するステップと、

前記ＤＡ変換器からの信号を音声としてスピーカから出力するステップと、

前記A/D変換機と前記D/A変換機との間の時間に比例した時間ずれ係数を推定するステップと、

前記推定した時間に比例した時間ずれ係数を用いて、前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去するステップと、

前記音響エコー成分を抑圧したデジタル信号をネットワークに送出するステップと、を含む音響エコー成分を除去する方法。

[請求項6] 前記A/D変換機と前記D/A変換機との間の時間に比例した時間ずれ係数を推定するステップは。

前記ネットワークから入力されるデジタル音声波形と、前記ＡＤ変換器で得られたデジタル音声波形と、に基づいて前記時間に比例した時間ずれ係数を推定することを特徴とする請求項５に記載の音響エコー成分を除去する方法。

[請求項7] 前記ネットワークから入力されるデジタル音声波形を前記推定した時間に比例した時間ずれ係数を用いて補正するステップと、

前記補正したデジタル音声波形を用いて音響エコーキャンセラフィルタを更新するステップと、

前記更新した音響エコーキャンセラフィルタを前記補正したデジタル音声波形に重畳して音響エコーレプリカを作成し，作成した前記音響エコーレプリカを前記ＡＤ変換器で得られたデジタル音声波形から引き算することで音響エコー成分を除去するステップと、を含むことを特徴とする請求項６に記載の音響エコー成分を除去する方法。

[請求項8] A/D変換機を識別する識別子、D/A変換機を識別する識別子、該A/D変換機のサンプリングレート、及び、該D/A変換機のサンプリングレートの組み合わせに対応した時間に比例した時間ずれ係数を保持するデータベースを備え、

接続されるA/D変換機またはD/A変換機が変更されたことを検出するステップと、

前記データベースを参照して、接続されるA/D変換機、D/A変換機、及び、該A/D変換機と該D/A変換機のサンプリングレートに対応する時

間に比例した時間ずれ係数を取得するステップと、

取得した時間に比例した時間ずれ係数を用いて前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去するステップと、を含むことを特徴とする請求項７に記載の音響エコー成分を除去する方法。

[請求項9] マイクロホンからの音声信号をＡＤ変換器でデジタル変換したデジタル音声波形を取り込む音声取り込み部と、

ＤＡ変換器で変換してスピーカで再生するためのスピーカ再生用のデジタル音声信号を受信する音声受信部と、

前記A/D変換機と前記D/A変換機との間の時間に比例した時間ずれ係数を推定するフィルタ学習部と、

前記推定した時間に比例した時間ずれ係数を用いて、前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去する音響エコーキャンセラ部と、

前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去したデジタル音声波形を出力する音声送信部と、を備えることを特徴とする音響エコーキャンセラ。

[請求項10] 前記フィルタ学習部は、

前記音声受信部で受信するデジタル音声波形と、前記ＡＤ変換器で得られたデジタル音声波形と、に基づいて前記時間に比例した時間ずれ係数を推定することを特徴とする請求項９に記載の音響エコーキャンセラ。

[請求項11] 前記音響エコーキャンセラ部は、前記音声受信部で受信するデジタル音声波形を前記推定した時間に比例した時間ずれ係数を用いて補正し、

前記フィルタ学習部は、補正したデジタル音声波形を用いて音響エコーキャンセラフィルタを更新し、

前記音響エコーキャンセラ部は、更新した音響エコーキャンセラフ

ィルタを前記補正したデジタル音声波形に重畳して音響エコーレプリカを作成し，作成した前記音響エコーレプリカを前記ＡＤ変換器で得られたデジタル音声波形から引き算することで音響エコー成分を除去することを特徴とする請求項１０に記載の音響エコーキャンセラ。

［請求項12］ A/D変換機を識別する識別子、D/A変換機を識別する識別子、該A/D変換機のサンプリングレート、及び、該D/A変換機のサンプリングレートの組み合わせに対応した時間に比例した時間ずれ係数を保持するデータベースを備え、

接続されるA/D変換機またはD/A変換機が変更されたことを検出すると、前記データベースを参照して、接続されるA/D変換機、D/A変換機、及び、該A/D変換機と該D/A変換機のサンプリングレートに対応する時間に比例した時間ずれ係数を取得し、取得した時間に比例した時間ずれ係数を用いて前記ＡＤ変換器で得られたデジタル音声波形から音響エコー成分を除去することを特徴とする請求項１１に記載の音響エコーキャンセラ。

[図1]

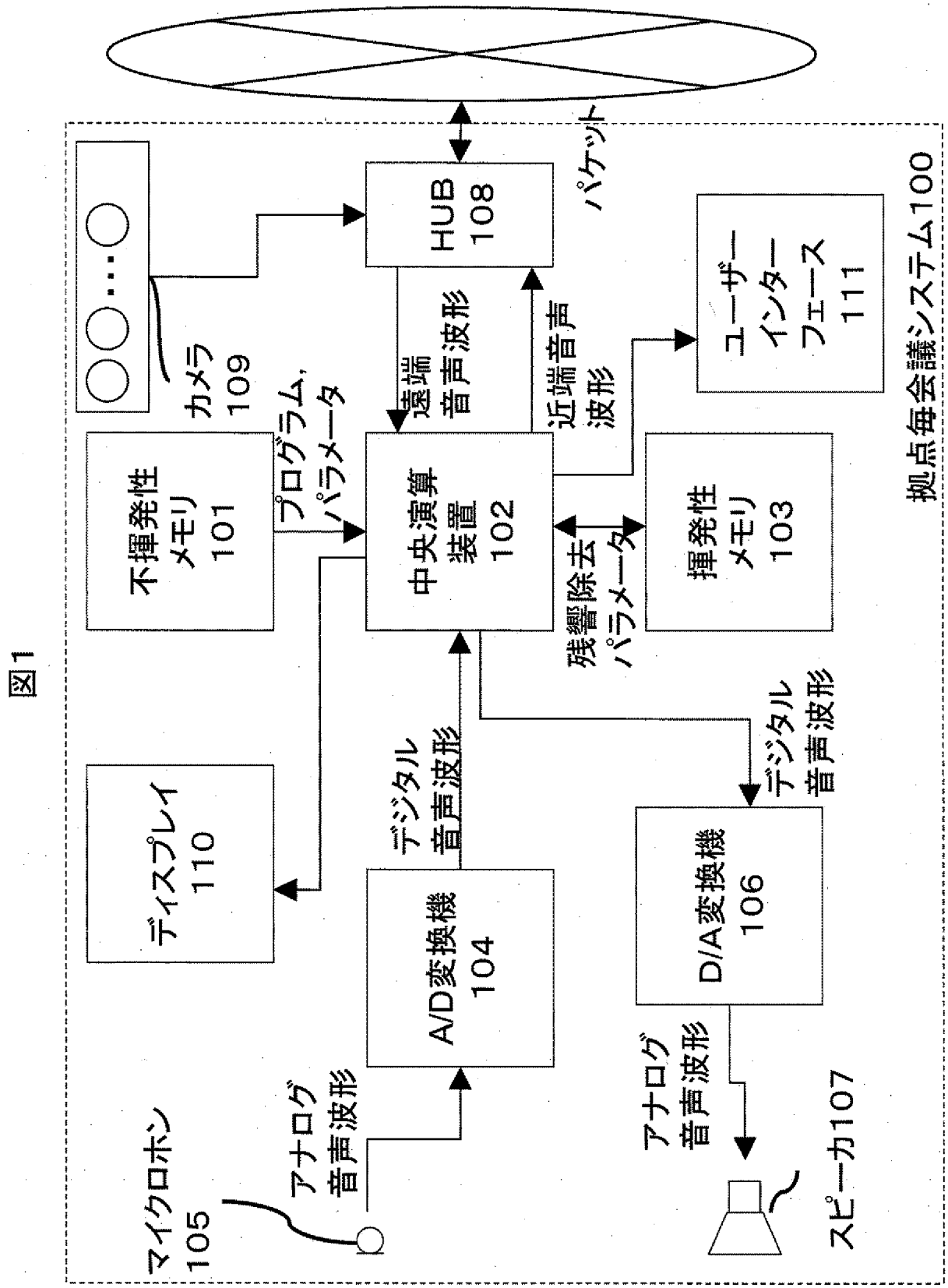
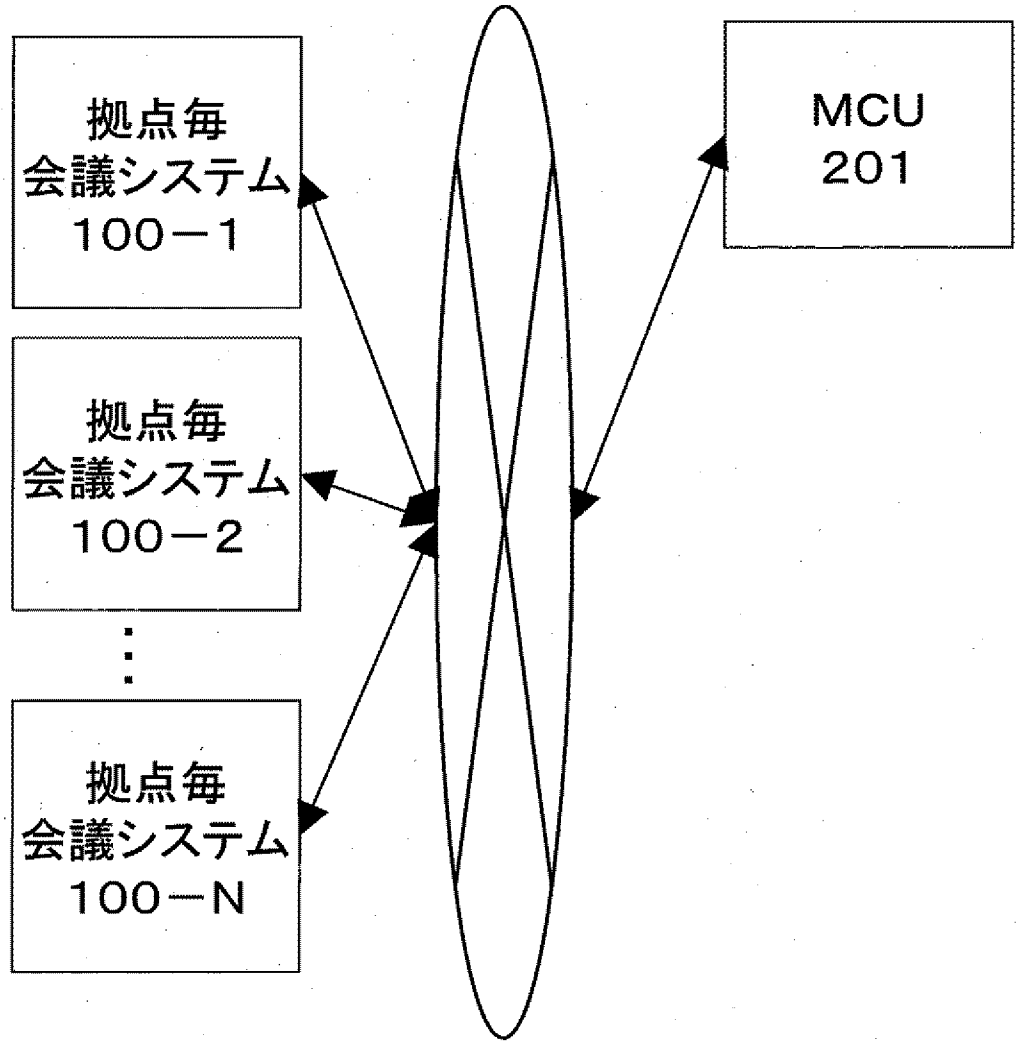

拠点毎会議システム100
パケット
HUB
108
カメラ
109
遠端
音声波形
近端音声
波形
ユーザー
インター
フェース
111
不揮発性
メモリ
101
プログラム,
パラメータ
中央演算
装置
102
揮発性
メモリ
103
残響除去
パラメータ
ディスプレイ
110
デジタル
音声波形
デジタル
音声波形
A/D変換機
104
D/A変換機
106
マイクロホン
105
アナログ
音声波形
アナログ
音声波形
スピーカ107

図1

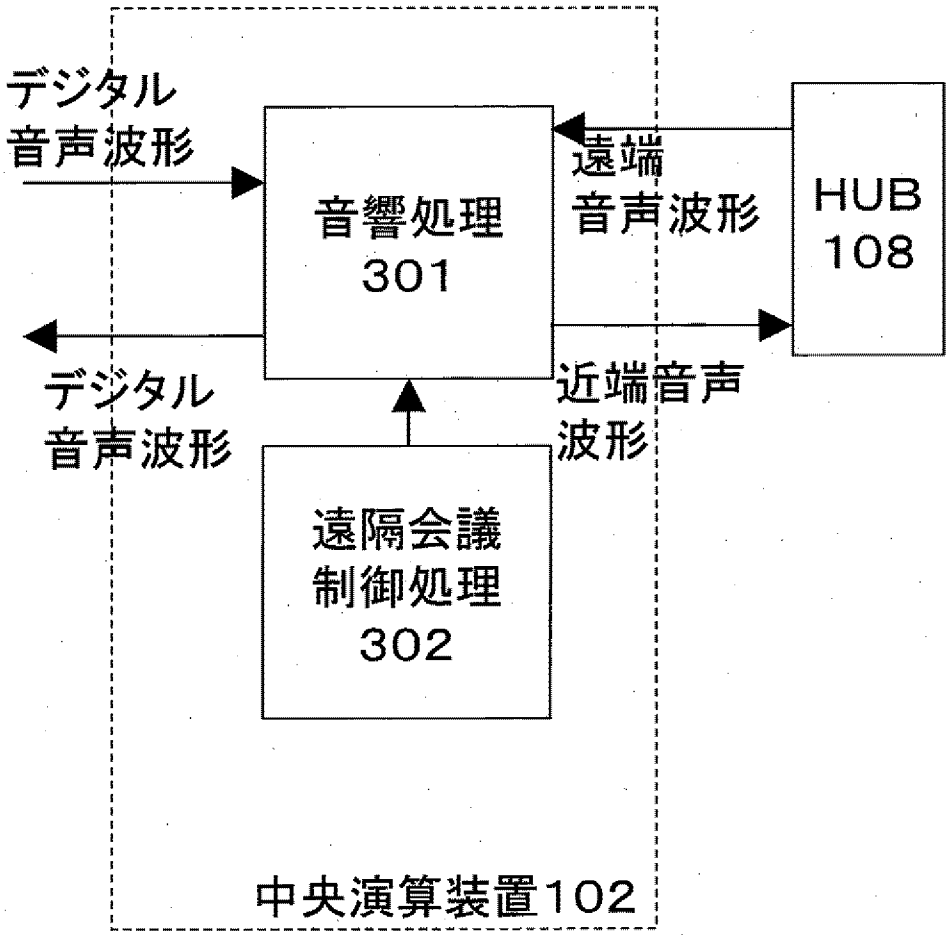
[図2]

図2



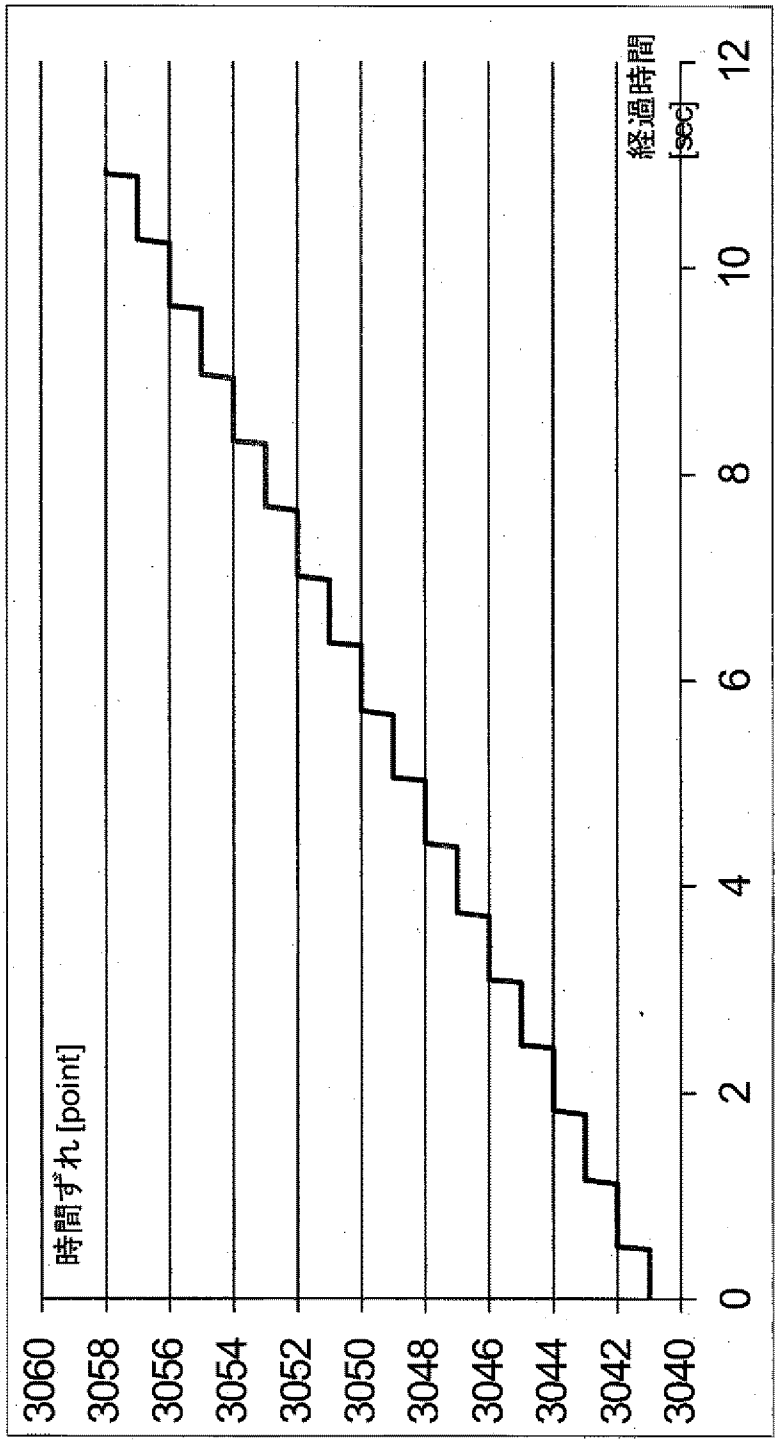
[図3]

図3


デジタル
音声波形
音響処理
301
遠端
音声波形
HUB
108
デジタル
音声波形
近端音声
波形
遠隔会議
制御処理
302
中央演算装置102

[図4]

図4


A/D変換機とD/A変換機の時間ずれの一例
時間ずれ [point]
経過時間 [sec]
3060
3058
3056
3054
3052
3050
3048
3046
3044
3042
3040
0
2
4
6
8
10
12

[図5]

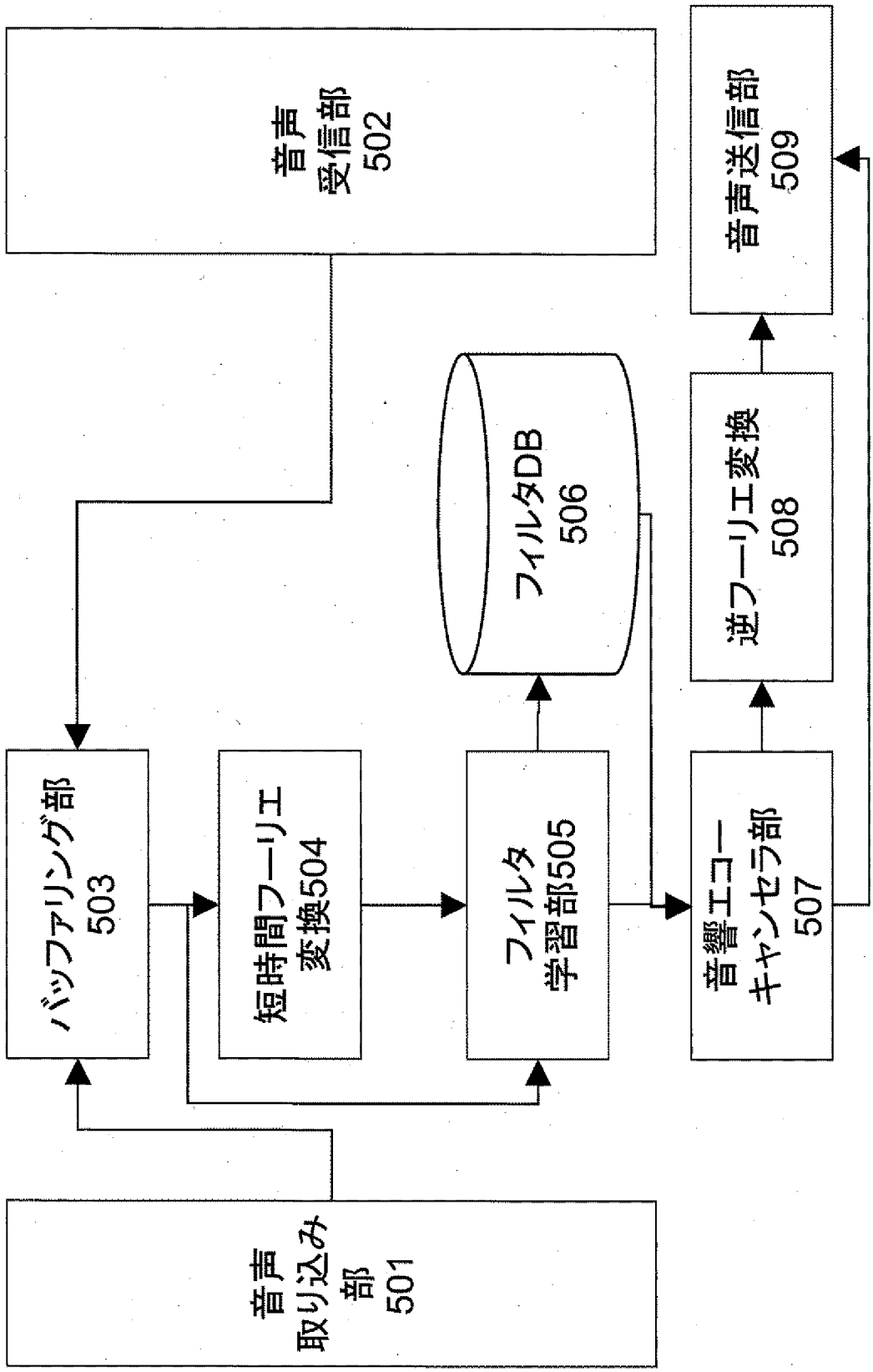
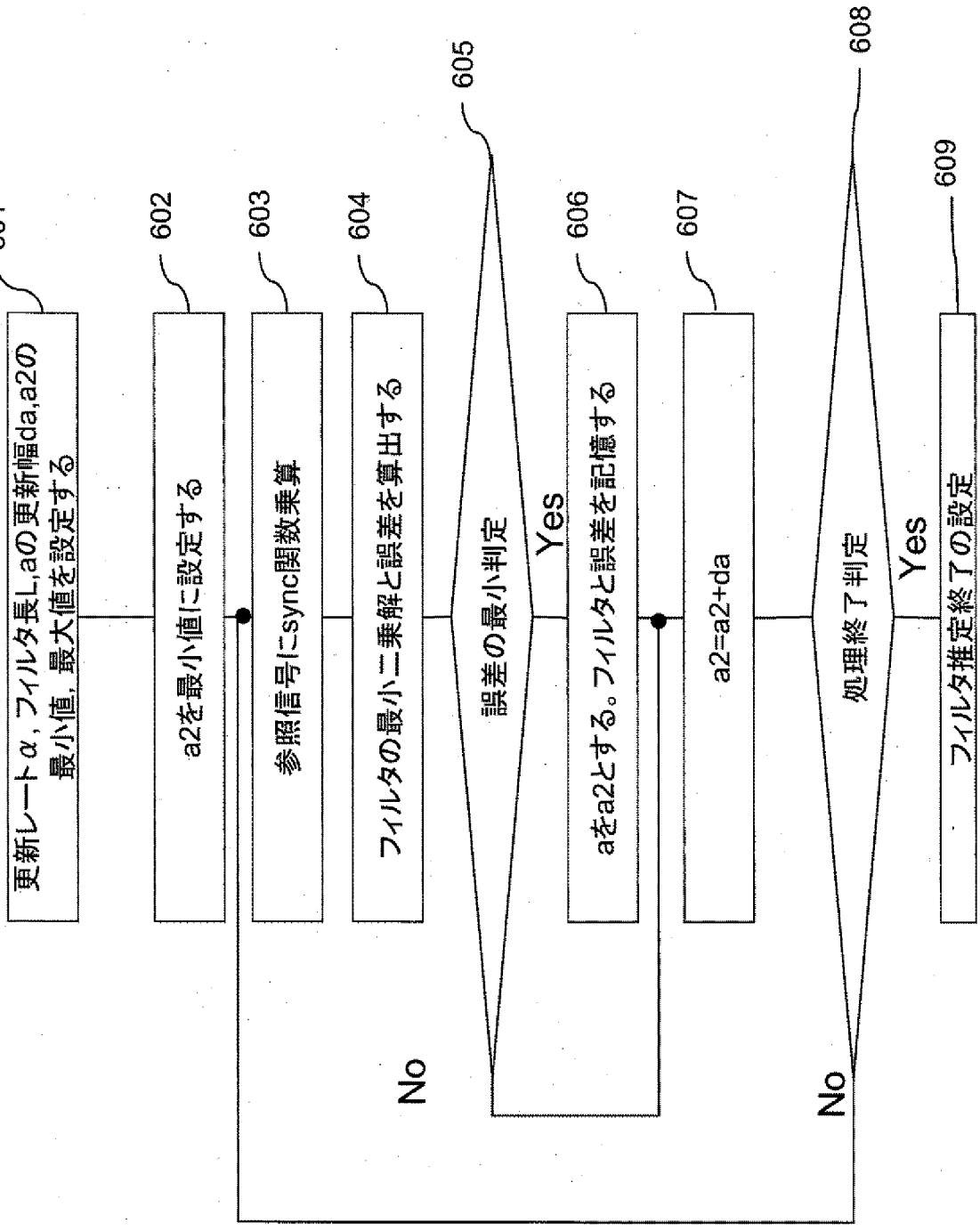

音声
取り込み
部
501
バッファリング部
503
短時間フーリエ
変換504
フィルタ
学習部505
フィルタDB
506
音響エコー
キャンセラ部
507
逆フーリエ変換
508
音声送信部
509
音声
受信部
502

図5

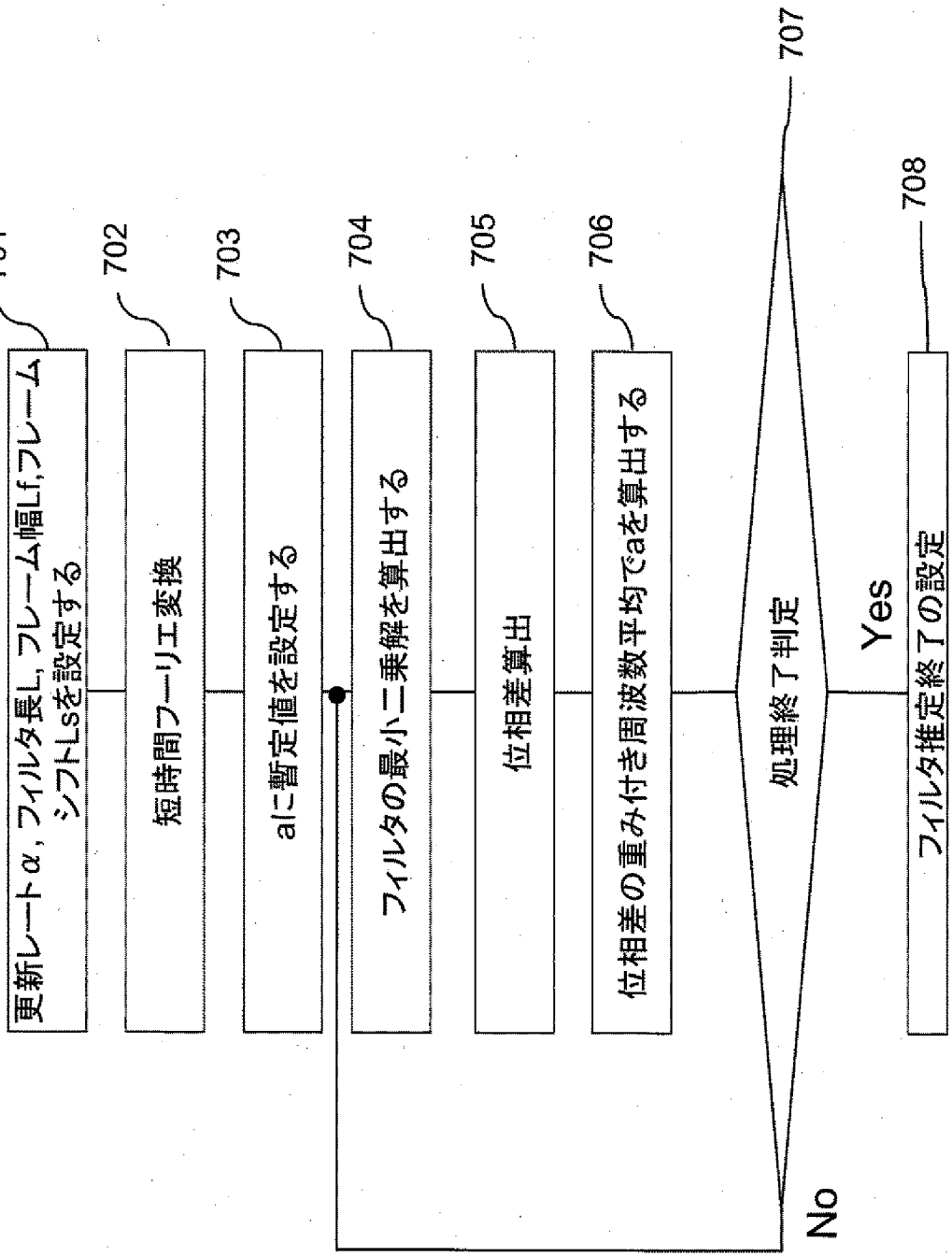
[図6]

図6


601
更新レートα, フィルタ長L,aの更新幅da,a2の最小値, 最大値を設定する
602
a2を最小値に設定する
603
参照信号にsync関数乗算
604
フィルタの最小二乗解と誤差を算出する
605
誤差の最小判定
Yes
606
aをa2とする。フィルタと誤差を記憶する
No
607
a2=a2+da
608
処理終了判定
Yes
609
フィルタ推定終了の設定
No

[図7]

図7


701
更新レートα, フィルタ長L, フレーム幅Lf,フレームシフトLsを設定する
702
短時間フーリエ変換
703
aに暫定値を設定する
704
フィルタの最小二乗解を算出する
705
位相差算出
706
位相差の重み付き周波数平均でaを算出する
707
処理終了判定
No
Yes
708
フィルタ推定終了の設定

[図8]

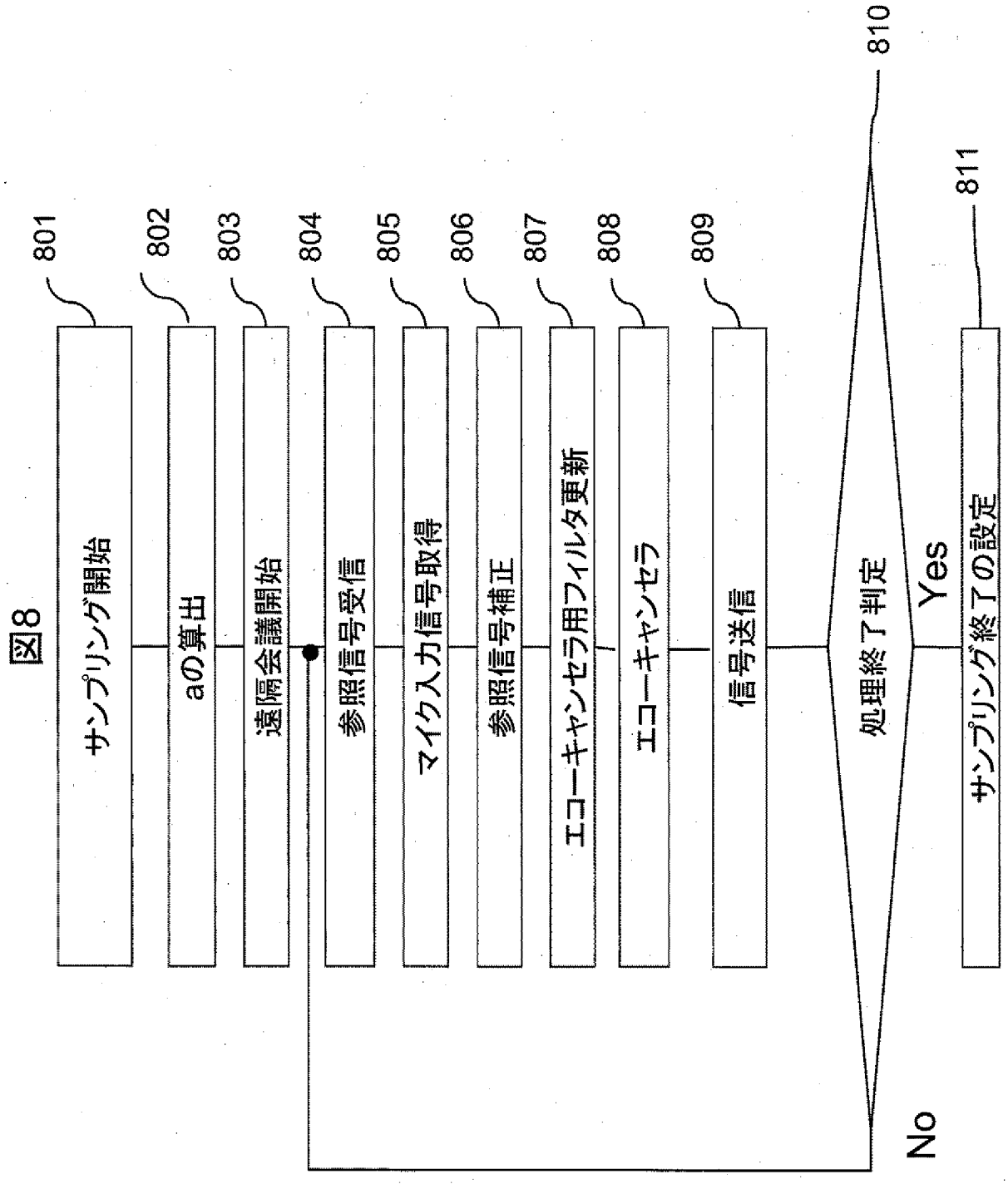
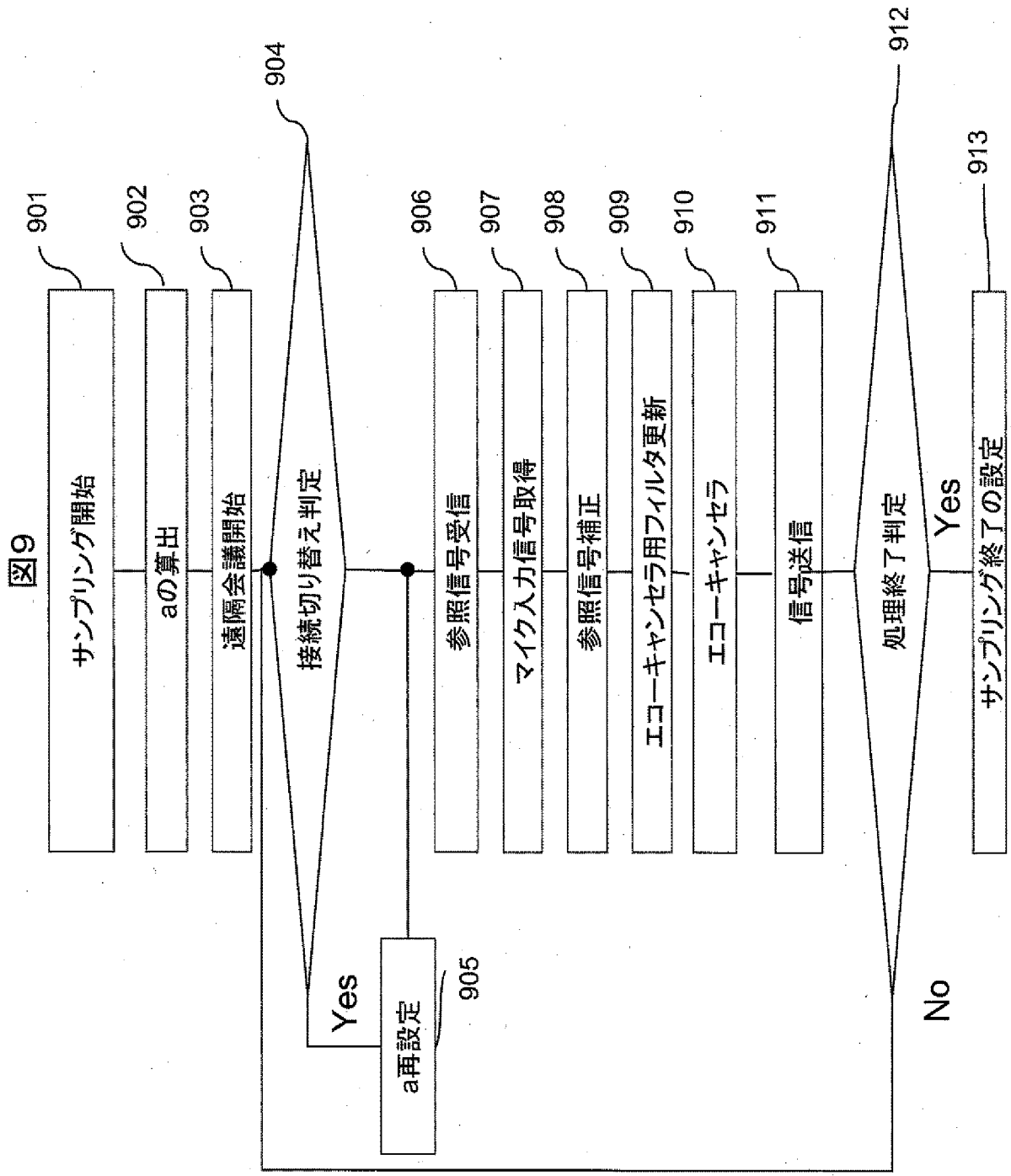

801
サンプリング開始
802
aの算出
803
遠隔会議開始
804
参照信号受信
805
マイク入力信号取得
806
参照信号補正
807
エコーキャンセラ用フィルタ更新
808
エコーキャンセラ
809
信号送信
810
処理終了判定
Yes
No
811
サンプリング終了の設定

図8

[図9]


図9
サンプリング開始
901
aの算出
902
遠隔会議開始
903
接続切り替え判定
904
Yes
a再設定
905
参照信号受信
906
マイク入力信号取得
907
参照信号補正
908
エコーキャンセラ用フィルタ更新
909
エコーキャンセラ
910
信号送信
911
処理終了判定
912
Yes
No
サンプリング終了の設定
913

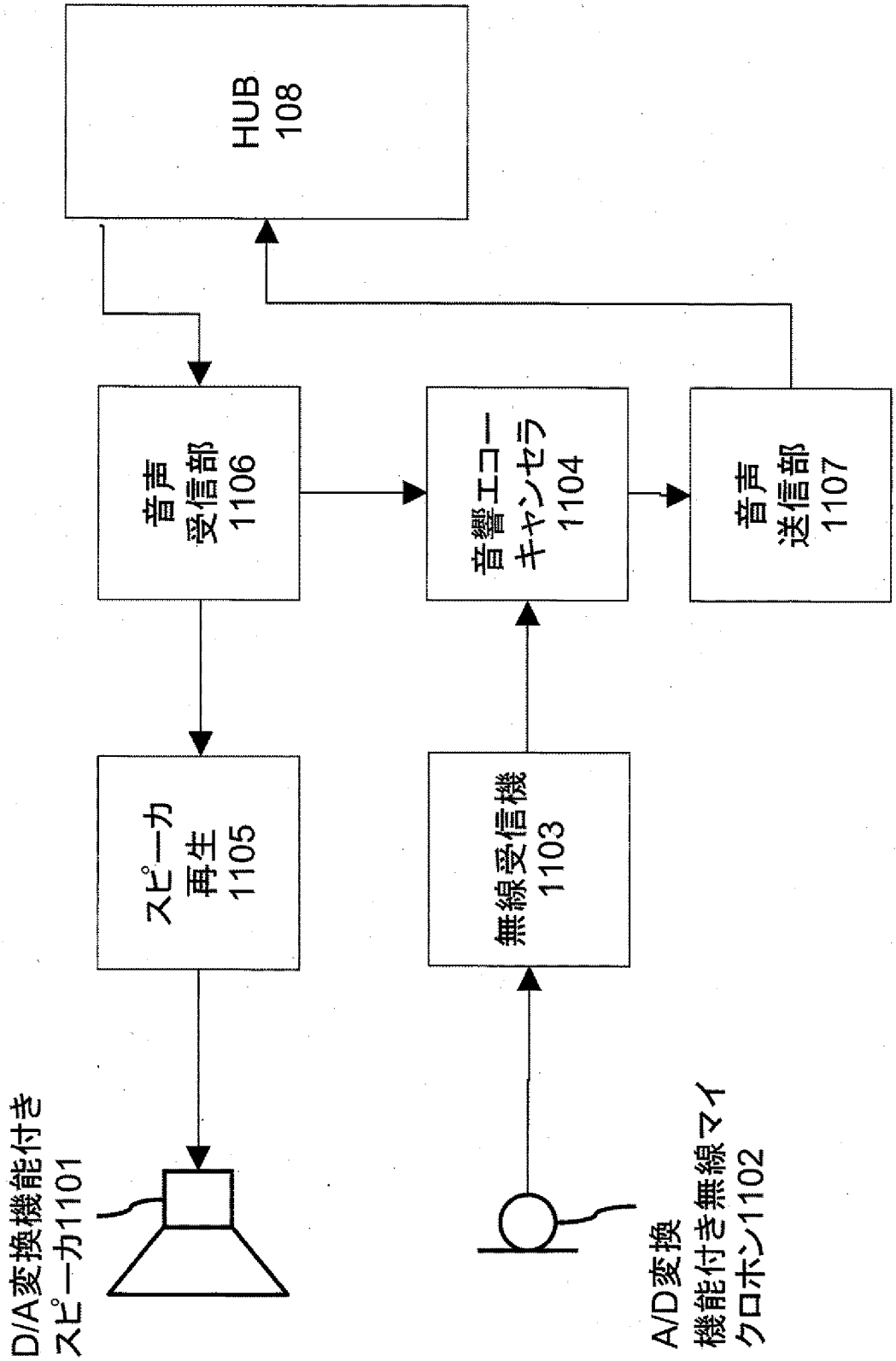
[図10]

図10

No.	A/D	D/A	a
1	Product 1 8 kHz	Product 2 8 kHz	0.1
2	Product 1 8 kHz	Product 3 8 kHz	0.2
3	Product 1 16 kHz	Product 3 16 kHz	0.1

[図11]

図11


HUB
108
音声
受信部
1106
音響エコー
キャンセラ
1104
音声
送信部
1107
スピーカ
再生
1105
無線受信機
1103
D/A変換機能付き
スピーカ1101
A/D変換
機能付き無線マイ
クロホン1102

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2013/051520

A. CLASSIFICATION OF SUBJECT MATTER
H04B3/23(2006.01)i, *G10L21/0224*(2013.01)i, *H04M1/60*(2006.01)i, *H04M3/56*(2006.01)i, *H04R3/02*(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)
H04B3/23, G10L21/0224, H04M1/60, H04M3/56, H04R3/02

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	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

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 Y	JP 2010-56778 A (Nippon Telegraph and Telephone Corp.), 11 March 2010 (11.03.2010), paragraphs [0006], [0017], [0020], [0027], [0031], [0032]; fig. 3 (Family: none)	1-3,5-7,9-11 4,8,12
Y	JP 2011-80868 A (Hitachi, Ltd.), 21 April 2011 (21.04.2011), paragraph [0029]; fig. 11 & US 2011/0082690 A1 paragraph [0052]; fig. 11 & CN 102036158 A	4,8,12

☐ 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	Date of mailing of the international search report
08 February, 2013 (08.02.13)	19 February, 2013 (19.02.13)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１３／０５１５２０

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　H04B3/23(2006.01)i, G10L21/0224(2013.01)i, H04M1/60(2006.01)i, H04M3/56(2006.01)i, H04R3/02(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.　H04B3/23, G10L21/0224, H04M1/60, H04M3/56, H04R3/02

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	１９７１－２０１３年
日本国実用新案登録公報	１９９６－２０１３年
日本国登録実用新案公報	１９９４－２０１３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 ２０１０－５６７７８ Ａ（日本電信電話株式会社）２０１０．０３．１１，第０００６段落，第００１７段落，第００２０段落，第００２７段落，第００３１段落，第００３２段落，第３図（ファミリー無し）	1-3, 5-7, 9-11
Y		4, 8, 12
Y	ＪＰ ２０１１－８０８６８ Ａ（株式会社日立製作所）２０１１．０４．２１，第００２９段落，第１１図 ＆ ＵＳ ２０１１／００８２６９０ Ａ１，第００５２段落，第１１図 ＆ ＣＮ １０２０３６１５８ Ａ	4, 8, 12

☐　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 ０８．０２．２０１３	国際調査報告の発送日 １９．０２．２０１３
国際調査機関の名称及びあて先 日本国特許庁（ＩＳＡ／ＪＰ） 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）　５Ｊ　３０５５ 川口　貴裕 電話番号　０３－３５８１－１１０１　内線　３５３４