

(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
2017年9月8日(08.09.2017)



(10) 国際公開番号
WO 2017/150103 A1

- (51) 国際特許分類:
G10L 25/63 (2013.01) H04R 1/40 (2006.01)
G10L 25/57 (2013.01) H04R 3/00 (2006.01)
G10L 25/84 (2013.01)
- (21) 国際出願番号: PCT/JP2017/004483
- (22) 国際出願日: 2017年2月8日(08.02.2017)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2016-038227 2016年2月29日(29.02.2016) JP
- (71) 出願人: パナソニックIPマネジメント株式会社 (PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) [JP/JP]; 〒5406207 大阪府大阪市中央区城見2丁目1番61号 Osaka (JP).
- (72) 発明者: 辻 寿嗣(TSUJI Hisashi), 藤井 亮太(FUJII Ryota), 田中 久裕(TANAKA Hisahiro).
- (74) 代理人: 鎌田 健司, 外(KAMATA Kenji et al.); 〒5406207 大阪府大阪市中央区城見2丁目1番61号 パナソニックIPマネジメント株式会社内 Osaka (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, DJ, 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, KH, KN, KP, KR, KW, 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).

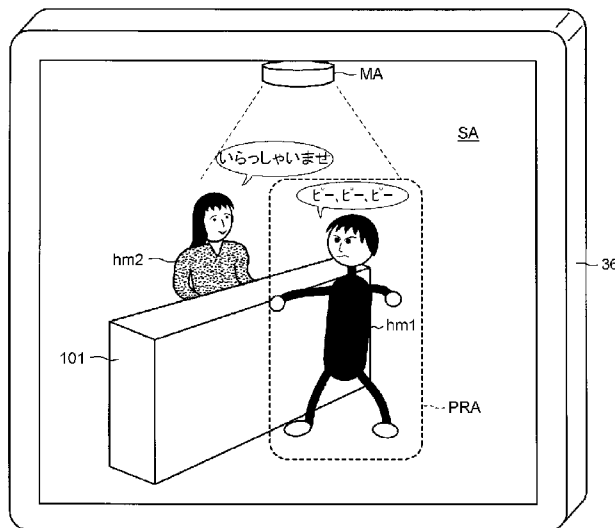
添付公開書類:

- 国際調査報告 (条約第21条(3))

(54) Title: AUDIO PROCESSING DEVICE, IMAGE PROCESSING DEVICE, MICROPHONE ARRAY SYSTEM, AND AUDIO PROCESSING METHOD

(54) 発明の名称: 音声処理装置、画像処理装置、マイクアレイシステム、及び音声処理方法

【図5】



(57) Abstract: This audio processing device comprises: an acquisition unit for acquiring audio collected by an audio collection unit; a detection unit for detecting the position of the audio; a determination unit that determines whether or not the audio contains an utterance when the audio position is within a privacy protection area; an analysis unit that analyzes the utterance and acquires an emotion value; a converting unit for converting the utterance into an alternative sound corresponding to the emotion value; and an output control unit that causes an audio output unit for outputting audio to output the alternative sound.

(57) 要約: 音声処理装置は、收音部により收音された音声を取得する取得部と、音声の音声位置を検出する検出部と、音声位置がプライバシー保護エリア内である場合に、音声が発話の音声であるか否かを判定する判定部と、発話の音声进行分析して感情値を取得する分析部と、発話の音声を感情値に対応する代替音に変換する変換部と、音声を出力する音声出力部に、代替音を出力させる出力制御部と、備える。

WO 2017/150103 A1

明 細 書

発明の名称：

音声処理装置、画像処理装置、マイクアレイシステム、及び音声処理方法

技術分野

[0001] 本開示は、音声処理装置、画像処理装置、マイクアレイシステム、及び音声処理方法に関する。

背景技術

[0002] 近年、カメラやマイクを用いて収録されたデータを扱う機会が増えている。防犯・証左の用途で店舗の窓口等に設置されるネットワークカメラシステムの台数は増加傾向にある。例えば、窓口でお客様と従業員の会話を録音する場合、お客様のプライバシー保護を考慮して、録音並びに再生を行う必要がある。また、録画を行う場合も同様である。

[0003] このようなシステムでは、マイクアレイ装置から指定された音声位置に向かう指向方向に、收音された音声に対する指向性を形成する。そして、このシステムは、音声位置がプライバシー保護領域である場合、收音された音声の出力を制御（ミュート処理、マスキング処理、又はボイスチェンジ処理）し、又は音声の收音を休止する（特許文献1参照）。

[0004] 本開示は、プライバシー保護を図りつつ、発話者の感情を察知することを目的とする。

先行技術文献

特許文献

[0005] 特許文献1：特開2015-29241号公報

発明の概要

[0006] 本開示の音声処理装置は、收音部により收音された音声を取得する取得部と、音声の音声位置を検出する検出部と、音声位置がプライバシー保護エリア内である場合に、音声が発話の音声であるか否かを判定する判定部と、発話の音声进行分析して感情値を取得する分析部と、発話の音声を感情値に対応

する代替出力に変換する変換部と、音声を出力する音声出力部に、代替出力を出力させる出力制御部と、備える。

[0007] 本開示によれば、プライバシー保護を図りつつ、発話者の感情を察知できる。

図面の簡単な説明

[0008] [図1]図1は、第1の実施形態におけるマイクアレイシステムの構成を示すブロック図である。

[図2A]図2Aは、ピッチの変化に対応する感情値が登録された感情値テーブルの登録内容を示す図である。

[図2B]図2Bは、話速に対応する感情値が登録された感情値テーブルの登録内容を示す図である。

[図2C]図2Cは、音量に対応する感情値が登録された感情値テーブルの登録内容を示す図である。

[図2D]図2Dは、滑舌に対応する感情値が登録された感情値テーブルの登録内容を示す図である。

[図3]図3は、感情値に対応する対応する代替音に登録された代替音テーブルの登録内容を示す図である。

[図4]図4は、マイクアレイ装置により收音された音声に対して所定の方向に指向性を形成する原理の一例の説明図である。

[図5]図5は、店舗の窓口に設置されたマイクアレイ装置によって受付係とお客様との会話が收音される状況を表す映像を示す図である。

[図6]図6は、マイクアレイ装置で收音された音声の出力手順を示すフローチャートである。

[図7]図7は、第2の実施形態におけるマイクアレイシステムの構成を示すブロック図である。

[図8]図8は、代替画像テーブルの登録内容を示す図である。

[図9]図9は、店舗の窓口に設置されたマイクアレイ装置によって受付係とお客様との会話が收音される状況を表す映像を示す図である。

[図10]図10は、マイクアレイ装置で収録された音声に基づく顔アイコンを含む映像の出力手順を示すフローチャートである。

[図11]図11は、第3の実施形態におけるマイクアレイシステムの構成を示すブロック図である。

[図12]図12は、店舗の窓口に設置されたマイクアレイ装置によって受付係とお客様との会話が収録される状況を表す映像を示す図である。

発明を実施するための形態

[0009] 以下、適宜図面を参照しながら、実施形態を詳細に説明する。但し、必要以上に詳細な説明は省略する場合がある。例えば、既によく知られた事項の詳細説明や実質的に同一の構成に対する重複説明を省略する場合がある。これは、以下の説明が不必要に冗長になることを避け、当業者の理解を容易にするためである。尚、添付図面及び以下の説明は、当業者が本開示を十分に理解するために提供されるものであり、これらにより特許請求の範囲に記載の主題を限定することは意図されていない。

[0010] (本開示の一形態を得るに至った経緯)

録音された従業員とお客様の会話録を、トラブル事案としてクレーム発生時の振り返りや社内研修資料に使用するとする。この会話録に対してプライバシー保護の必要がある場合、会話録の音声出力の制御等がされる。そのため、お客様の発話内容を把握することが困難であり、どのような経緯が存在するかを理解し難い。また、従業員と対面しているお客様の感情の変化を察することが困難である。

[0011] 以下、プライバシー保護を図りつつ、発話者の感情を察知できる音声処理装置、画像処理装置、マイクアレイシステム、及び音声処理方法について説明する。

[0012] (第1の実施形態)

[構成等]

図1は、第1の実施形態におけるマイクアレイシステム10の構成を示すブロック図である。マイクアレイシステム10は、カメラ装置CAと、マイ

クアレイ装置MAと、レコーダRCと、指向性制御装置30とを含む構成を有する。

[0013] カメラ装置CA、マイクアレイ装置MA、レコーダRC及び指向性制御装置30は、ネットワークNWを介して相互にデータ通信可能に接続されている。ネットワークNWは、有線ネットワーク（例えばイントラネット、インターネット）でもよいし、無線ネットワーク（例えば無線LAN（Local Area Network））でもよい。

[0014] カメラ装置CAは、例えば、室内の天井や壁等に設置された、画角が固定された固定カメラである。カメラ装置CAは、自装置が設置された撮像空間である撮像エリアSA（図5参照）を撮像可能な監視カメラとして機能する。

[0015] なお、カメラ装置CAは、固定カメラに限られず、全方位カメラ、パン・チルト・ズーム動作自在なPTZカメラであってもよい。カメラ装置CAは、映像を撮像した時刻（撮像時刻）を映像データと対応付けて記憶し、ネットワークNWを介して指向性制御装置30に送信する。

[0016] マイクアレイ装置MAは、例えば室内の天井に設置された全方位マイクアレイ装置である。マイクアレイ装置MAは、自装置が設置された收音空間（收音エリア）における全方位の音声を收音する。

[0017] マイクアレイ装置MAは、中央に開口部が形成された筐体、及びこの開口部の周囲に円周方向に沿って同心円状に配置された複数のマイクロホンユニットを有する。マイクロホンユニット（以下、単にマイクロホンと称する）には、例えば高音質小型エレクトレットコンデンサーマイクロホン（ECM：Electret Condenser Microphone）が用いられる。

[0018] 尚、カメラ装置CAが、例えばマイクアレイ装置MAの筐体に形成された開口部に收容される全方位カメラである場合、撮像エリアと收音エリアは略同一となる。

[0019] マイクアレイ装置MAは、收音した音声データを、收音した時刻（收音時刻）と対応付けて記憶するとともに、記憶した音声データ及び收音時刻のデ

ータを、ネットワークNWを介して、指向性制御装置30に送信する。

[0020] 指向性制御装置30は、例えばマイクアレイ装置MA及びカメラ装置CAが設置された室内の外に設置される。指向性制御装置30は、例えば、据置型のPC (Personal Computer) である。

[0021] 指向性制御装置30は、マイクアレイ装置MAで収録された全方位の音声に対し指向性を形成し、その指向方向の音声を強調する。指向性制御装置30は、撮像エリア内の音源の位置（音声位置ともいう）を推定し、推定された音源の位置がプライバシー保護エリアの範囲内である場合、所定のマスク処理を行う。マスク処理の詳細については、後述する。

[0022] 尚、指向性制御装置30は、PCの代わりに、携帯電話機、タブレット端末、スマートフォン等の通信端末でもよい。

[0023] 指向性制御装置30は、通信部31と、操作部32と、信号処理部33と、ディスプレイ装置36と、スピーカ装置37と、メモリ38と、設定管理部39と、音声分析部45と、を少なくとも含む構成である。信号処理部33は、指向性制御部41、プライバシー判断部42、発話判定部34及び出力制御部35を含む。

[0024] 設定管理部39は、初期設定として、ディスプレイ装置36に表示された、カメラ装置CAで撮像された映像に対し、ユーザによって指定されたプライバシー保護エリアの座標を、マイクアレイ装置MAからプライバシー保護エリアに対応する音声エリアに向かう指向方向を示す角度に変換する。

[0025] この変換処理では、設定管理部39は、プライバシー保護エリアの指定に応じて、マイクアレイ装置MAからプライバシー保護エリアに対応する音声エリアに向かう指向角 ($\theta MA h$, $\theta MA v$) を算出する。この算出処理の詳細については、例えば特許文献1に記載されている。

[0026] $\theta MA h$ は、マイクアレイ装置MAから音声位置に向かう指向方向の水平角を表す。 $\theta MA v$ は、マイクアレイ装置MAから音声位置に向かう指向方向の垂直角を表す。音声位置は、操作部32がディスプレイ装置36に表示された映像データにおいてユーザの指又はスタイラスペンによって指定され

た指定位置に対応する実際の位置である。なお、この変換処理は、信号処理部33が行ってもよい。

[0027] また、設定管理部39は、メモリ39zを有する。設定管理部39は、カメラ装置CAで撮像された映像に対し、ユーザによって指定されたプライバシー保護エリアの座標、及び、変換されたプライバシー保護エリアに対応する音声エリアに向かう指向方向を示す座標、をメモリ39zに記憶する。

[0028] 通信部31は、カメラ装置が送信した撮像時刻を含む映像データ、及びマイクアレイ装置MAが送信した収録時刻を含む音声データ、を受信して、信号処理部33に出力する。

[0029] 操作部32は、ユーザの入力操作の内容を信号処理部33に通知するためのユーザインターフェース（UI：User Interface）であり、例えばマウス、キーボード等のポインティングデバイスを含んで構成される。また、操作部32は、例えばディスプレイ装置36の画面に対応して配置され、ユーザの指やスタイラスペンによって入力操作が可能なタッチパネル又はタッチパッドを用いて構成されてもよい。

[0030] 操作部32は、ディスプレイ装置36に表示されたカメラ装置CAの映像データ（図5参照）において、ユーザがプライバシー保護を希望するエリアであるプライバシー保護エリアPRAを指定する。そして、操作部32は、指定されたプライバシー保護エリアの位置を表す座標データを取得して、信号処理部33に出力する。

[0031] メモリ38は、例えばRAM（Random Access Memory）を用いて構成され、指向性制御装置30が動作する際、プログラムメモリ、データメモリ、ワークメモリとして機能する。メモリ38は、マイクアレイ装置MAで収録される音声の音声データを収録時刻とともに記憶する。

[0032] 信号処理部33は、機能的構成として、発話判定部34、指向性制御部41、プライバシー判断部42及び出力制御部35を有する。信号処理部33は、ハードウェアとして、例えばCPU（Central Processing Unit）、MPU（Micro Processing Unit）又はDSP（Digital Signal Processor）を用

いて構成される。信号処理部33は、指向性制御装置30の各部の動作を全体的に統括するための制御処理、他の各部との間のデータの入出力処理、データの演算（計算）処理及びデータの記憶処理を行う。

[0033] 発話判定部34は、收音された音声を分析し、音声が発話であるか否かを認識する。ここでの音声は、可聴周波数帯（例えば20Hz－23kHz）の周波数を有する音であり、人が話す音声以外を含んでもよい。また、発話は、人が話す音声であり、可聴周波数帯に比べて狭い帯域（例えば300Hz－4kHz）の周波数を有する音である。例えば、入力音から音声が発話された区間を検出する技術であるVAD（Voice Activity Detectors）によって、発話が認識される。

[0034] プライバシー判断部42は、メモリ38に記憶された音声データを用いて、マイクアレイ装置MAで收音された音声プライバシー保護エリア内で検出されたものであるか否かを判定する。

[0035] プライバシー判断部42は、マイクアレイ装置MAで音声收音された場合、音源の方向がプライバシー保護エリアの範囲内であるか否かを判定する。この場合、プライバシー判断部42は、例えば、撮像エリアを複数のブロックに分割し、ブロック毎に音声の指向性を形成し、その指向方向に閾値を超える音声があるか否かを判定し、撮像エリア内の音声位置を推定する。

[0036] 音声位置の推定方法として、公知の方法を用いてよく、例えば、『論文「マイクロホンアレイを用いたCSP法に基づく複数音源位置推定」西浦敬信等、電子情報通信学会論文誌 D-11 Vol. J83-D-11 No. 8 pp. 1713-1721 2000年 8月』の文献に記載されている方法を用いてもよい。

[0037] また、プライバシー判断部42は、マイクアレイ装置MAが收音した音声データに対し、プライバシー保護エリア内の位置に指向性を形成し、その指向方向に音声検出されているか否かを判定してもよい。この場合、音声位置がプライバシー保護エリアの範囲内にあるか否かを判定できるが、プライバシー保護エリアの外側に音声位置があっても、その位置は特定されない。

- [0038] 出力制御部35は、カメラ装置CA、マイクアレイ装置MA、ディスプレイ装置36及びスピーカ装置37の動作を制御する。出力制御部35は、カメラ装置CAから送信された映像データをディスプレイ装置36に出力させ、マイクアレイ装置MAから送信された音声データをスピーカ装置37に音声出力させる。
- [0039] 指向性制御部41は、マイクアレイ装置MAが收音して指向性制御装置30に送信した音声データを用いて指向性の形成処理を行う。ここでは、指向性制御部41は、設定管理部39により算出された指向角 (θMAh , θMAv) の方向に、音声データの指向性を形成する。
- [0040] プライバシー判断部42は、算出された指向方向を示す座標データを基に、音声位置が予め指定されたプライバシー保護エリアPRA (図5参照) 内に含まれるか否かを判定してもよい。
- [0041] 出力制御部35は、プライバシー保護エリアPRA内に音声位置が含まれると判定された場合、マイクアレイ装置MAにより收音された音声を制御し、例えば、この音声に代えて代替音を再生して出力する。代替音は、例えば、プライバシー音の一例としての通称「ピー音」を含む。
- [0042] なお、出力制御部35は、マイクアレイ装置MAにより收音されたプライバシー保護エリアPRA内の音声の音圧を算出し、この算出された音声の音圧閾値を超える場合に、代替音を出力してもよい。
- [0043] 出力制御部35は、代替音を出力する際、マイクアレイ装置MAにより收音されたプライバシー保護エリアPRA内の音声を音声分析部45に送る。出力制御部35は、音声分析部45によって音声分析が行われた結果に基づく代替音の音声データを、音声分析部45から取得する。
- [0044] 音声分析部45は、マイクアレイ装置MAにより收音されたプライバシー保護エリアPRA内の音声を受けると、この音声を分析し、音声を発した人物の感情を感情値として取得する。この音声分析では、音声分析部45は、プライバシー保護エリアPRA内の音声のうち、例えば、話者が発する発話の音声のピッチ (周波数) の変化を分析し、声が上がった、下がった、上が

った等の感情値を得る。感情値として、例えば「高」、「中」、「低」の3段階に分けられる。なお、感情値を任意の段数に分けてもよい。

[0045] 音声分析部45のプライバシー音データベース(DB)48には、4つの感情値テーブル47A、47B、47C、47Dが保持されている(図2A~図2D参照)。特にこれらのテーブルを区別する必要が無い場合、感情値テーブル47と総称する。感情値テーブル47は、プライバシー音DB48に記憶される。

[0046] 図2Aは、ピッチの変化に対応する感情値が登録された感情値テーブル47Aの登録内容を示す模式図である。

[0047] 感情値テーブル47Aでは、例えば、ピッチの変化が「大」の場合、声が上がっている等として、感情値に「高」が設定される。例えば、ピッチの変化が「中」の場合、声が僅かに上がっている等として、感情値に「中」が設定される。例えば、ピッチの変化が「小」の場合、声下がって落ち着いている等として、感情値に「小」が設定される。

[0048] 図2Bは、話速に対応する感情値が登録された感情値テーブル47Bの登録内容を示す模式図である。話速は、例えば、所定時間内に話者が発した単語数により表される。

[0049] 感情値テーブル47Bでは、例えば、話速が早い場合、早口になっている等として、感情値に「高」が設定される。例えば、話速が普通(中程度)の場合、話が少し早い等として、感情値に「中」が設定される。例えば、話速が遅い場合、気分が落ち着いている等として、感情値に「小」が設定される。

[0050] 図2Cは、音量に対応する感情値が登録された感情値テーブル47Cの登録内容を示す模式図である。

[0051] 感情値テーブル47Cでは、例えば、話者が発する音声の音量が大きい場合、気分が高揚している等として、感情値に「高」が設定される。例えば、音量が普通(中程度)の場合、通常気分である等として、感情値に「中」が設定される。例えば、音量が小さい場合、気分が落ち着いている等として

、感情値に「小」が設定される。

[0052] 図2Dは滑舌に対応する感情値が登録された感情値テーブル47Dの登録内容を示す模式図である。

[0053] 滑舌の善し悪しは、例えば、音声認識による認識率の高低で判断される。感情値テーブル47Cでは、例えば、音声の認識率が低く、滑舌が悪い場合、怒っている等として、感情値に「大」が設定される。例えば、音声の認識率が中で滑舌が普通（中程度）の場合、平静である等として、感情値に「中」が設定される。例えば、音声の認識率が高く、滑舌が良い場合、冷静である等として、感情値に「小」が設定される。

[0054] 音声分析部45は、いずれの感情値テーブル47を用いてもよく、また、複数の感情値テーブル47を用いて感情値を導出してもよい。ここでは、一例として、音声分析部45が感情値テーブル47Aにおけるピッチの変化から感情値を取得する場合を示す。

[0055] 音声分析部45は、プライバシー音変換部46、及び、プライバシー音DB48を有する。

[0056] プライバシー音変換部46は、プライバシー保護エリアPRA内の発話の音声を、感情値に対応する代替音に変換する。

[0057] プライバシー音DB48には、例えば、プライバシー音としてピー音を表す正弦波（サイン波）の音声データが1つ登録されている。プライバシー音変換部46は、プライバシー音DB48に登録されている正弦波の音声データを読み出し、発話の音声が出力されている期間、読み出した音声データを基に、感情値に対応する周波数の正弦波の音声データを出力する。

[0058] 例えば、プライバシー音変換部46は、感情値が「高」である場合、1kHzのピー音を出力し、感情値が「中」である場合、500Hzのピー音を出力し、感情値が「低」である場合、200Hzのピー音を出力してもよい。尚、この周波数は、一例であり、他の高さでもよい。

[0059] 尚、プライバシー音変換部46は、1つの正弦波の音声データを基に、複数の周波数の音声データを生成する代わりに、予め感情値に対応する音声デ

ータを、例えばプライバシー音DB48に登録しておき、この音声データを読み出してもよい。

[0060] 図3は、感情値に対応する代替音が登録された代替音テーブル49の登録内容を示す模式図である。代替音テーブル49は、プライバシー音DB48に記憶される。

[0061] 代替音テーブル49には、感情値に対応する代替音として、前述した3つの異なる周波数のプライバシー音が登録されている。尚、これに限らず、プライバシー音DB48には、感情値が「高」である場合に怒りを表す大砲の音データ、感情値が「中」である場合に怒っていないことを表す豆鉄砲の音データ、感情値が「低」である場合に喜びを表すメロディ音の音データ、等が登録されてもよい。

[0062] ディスプレイ装置36は、カメラ装置CAが撮像した映像データを画面に表示する。

[0063] スピーカ装置37は、マイクアレイ装置MAが收音した音声データ、又は指向角(θMAh , θMAV)に指向性が形成されたマイクアレイ装置MAが收音した音声データ、を音声出力する。尚、ディスプレイ装置36及びスピーカ装置37は、指向性制御装置30とは別体の装置として構成されてもよい。

[0064] 図4は、マイクアレイ装置MAにより收音された音声に対して所定の方向に指向性を形成する原理の一例の説明図である。

[0065] 指向性制御装置30は、マイクアレイ装置MAから送信された音声データを用いて、音声データの指向性制御処理によって、各々のマイクロホンMA1~MANにより收音された各音声データを加算する。そして、指向性制御装置30は、マイクアレイ装置MAの各マイクロホンMA1~MANの位置から特定方向への音声(音量レベル)を強調(増幅)するために、特定方向への指向性を形成した音声データを生成する。特定方向とは、マイクアレイ装置MAから操作部32で指定された音声位置に向かう方向である。

[0066] 尚、マイクアレイ装置MAによって收音される音声の指向性を形成するた

めの音声データの指向性制御処理に関する技術は、例えば特開2014-143678号公報や特開2015-029241号公報（特許文献1）等に表示されるように、公知の技術である。

[0067] 図4では、説明を分かり易くするため、マイクロホンMA1～MA n は直線上に一次元配列されている。この場合、指向性は面内の二次元空間になる。更に、三次元空間で指向性を形成するためには、マイクロホンMA1～MA n を二次元配列し、同様な処理を実施されればよい。

[0068] 音源80から発した音波は、マイクアレイ装置MAに内蔵される各マイクロホンMA1, MA2, MA3, ..., MA($n-1$), MA n に対し、ある一定の角度（入射角＝ $(90-\theta)$ [度]）で入射する。入射角 θ は、マイクアレイ装置MAから音声位置に向かう指向方向の水平角 θ_{MAh} でも垂直角 θ_{MAv} でもよい。

[0069] 音源80は、例えば、マイクアレイ装置MAが收音する收音方向に存在するカメラ装置CAの被写体である人物の会話である。音源80は、マイクアレイ装置MAの筐体21の面上に対し、所定角度 θ の方向に存在する。また、各マイクロホンMA1, MA2, MA3, ..., MA($n-1$), MA n 間の間隔 d は、一定とする。

[0070] 音源80から発した音波は、例えば、最初にマイクロホンMA1に到達して收音され、次にマイクロホンMA2に到達して收音され、同様に次々に收音され、最後にマイクロホンMA n に到達して收音される。

[0071] マイクアレイ装置MAは、各マイクロホンMA1, MA2, MA3, ..., MA($n-1$), MA n が收音したアナログの音声データを、A/D変換器241, 242, 243, ..., 24($n-1$), 24 n においてデジタルの音声データにAD変換する。

[0072] 更に、マイクアレイ装置MAは、遅延器251, 252, 253, ..., 25($n-1$), 25 n において、各々のマイクロホンMA1, MA2, MA3, ..., MA($n-1$), MA n における到達時間差に対応する遅延時間を与えて、全ての音波の位相を揃えた後、加算器26において遅延処理後の音

声データを加算する。

[0073] これにより、マイクアレイ装置MAは、各マイクロホンMA 1, MA 2, MA 3, ..., MA (n - 1), MA nに、所定角度 θ の方向に音声データの指向性を形成する。

[0074] このように、マイクアレイ装置MAは、遅延器251, 252, 253, ..., 25(n - 1), 25nに設定される遅延時間D 1, D 2, D 3, ..., D n - 1, D nを変更することで、收音した音声データの指向性を簡易に形成できる。

[0075] [動作等]

次にマイクアレイシステム10の動作について説明する。ここでは、店舗に来店したお客様と受付係との会話を收音して音声出力する場合を一例として示す。

[0076] 図5は、店舗の窓口に設置されたマイクアレイ装置MAによって、受付係hm 2とお客様hm 1との会話が收音される状況を表す映像を示す模式図である。

[0077] 図5の映像では、店舗内の天井に設置された固定カメラであるカメラ装置CAによって撮像された撮像エリアSAが、ディスプレイ装置36に映し出されている。例えば、受付係hm 2（従業員の一例）がお客様hm 1と対面するカウンタ101の真上に、マイクアレイ装置MAが設置される。マイクアレイ装置MAは、受付係hm 2とお客様hm 1との会話を含む、店舗内の音声を收音する。

[0078] お客様hm 1が位置するカウンタ101は、プライバシー保護エリアPRAに設定されている。プライバシー保護エリアPRAは、例えば、ユーザが予めディスプレイ装置36に表示された映像に対して、タッチ操作等で範囲を指定することで設定される。

[0079] 図5の映像では、撮像エリアSAにおいて、お客様hm 1が来店し、カウンタ101の前に設置されたプライバシー保護エリアPRAに入っている状況が示されている。例えば、受付係hm 2が「いらっしゃいませ」と挨拶す

ると、その音声はスピーカ装置37から出力される。また、例えば、お客様hm1は険しい表情で話しかけているが、その音声はスピーカ装置37から「ピー、ピー、ピー」とプライバシー音が出力される。

[0080] これにより、発話内容の秘匿性が担保される。また、マイクアレイシステム10のユーザは、スピーカ装置37から出力されるプライバシー音のピッチの変化等から、お客様hm1の感情を察知できる。

[0081] 尚、受付係hm2とお客様hm1が発した発話の音声を表す吹き出しは、説明を分かり易くするために付加されたものである。

[0082] 図6は、マイクアレイ装置MAで收音された音声の出力手順を示すフローチャートである。この音声出力動作は、例えば、マイクアレイ装置MAで收音された音声の音声データをレコーダRCに一旦記憶させた後に行われる。

[0083] 通信部31は、ネットワークNWを介してレコーダRCに記録された、所定時間の音声データ及び映像データを取得する(S1)。

[0084] 指向性制御部41は、マイクアレイ装置MAで收音された音声データに対し、指向性を形成し、店舗内等の所定方向を指向方向とする音声データを取得する(S2)。

[0085] プライバシー判断部42は、指向性制御部41によって指向性が形成される音声位置がプライバシー保護エリアPRA内であるか否かを判別する(S3)。

[0086] 音声位置がプライバシー保護エリアPRA内でない場合、出力制御部35は、指向性形成済みの音声データをそのままスピーカ装置37に出力する(S4)。また、この場合、出力制御部35は、映像データをディスプレイ装置36に出力する。この後、信号処理部33は本動作を終了する。

[0087] S3で、指向性制御部41によって指向性が形成される音声位置がプライバシー保護エリアPRA内である場合、発話判定部34は、指向性形成済みの音声が発話の音声であるか否かを判別する(S5)。

[0088] S5では、例えば、発話判定部34は、指向性形成済みの音声が発話係hm2とお客様hm1との会話のような人が話す音声であり、可聴周波数帯に

比べて狭い帯域（例えば300Hz－4kHz）の周波数を有する音であるか否かを判別する。

[0089] 尚、ここでは、発話の音声を音声分析の対象としたが、プライバシー保護エリアPRAで発せられる全ての音声を音声分析の対象としてもよい。

[0090] S5において、指向性形成済みの音声が発話の音声でない場合、信号処理部33は、前述したS4の処理に進む。

[0091] S5において、指向性形成済みの音声が発話の音声である場合、音声分析部45は、指向性形成済みの音声データに対し、音声分析する（S6）。

[0092] 音声分析の結果、音声分析部45は、プライバシー音DB48に登録されている感情値テーブル47を用いて、発話の音声の感情値が「高」か、「中」か、「低」か、を判別する（S7）。

[0093] S7で、発話の音声の感情値が「高」である場合、プライバシー音変換部46は、代替音テーブル49を用いて、正弦波の音声データを読み出し、高域の周波数（例えば1kHz）の音声データに変換する（S8）。

[0094] 出力制御部35は、高い周波数の音声データをプライバシー音としてスピーカ装置37に出力する（S11）。スピーカ装置37は、プライバシー音である「ピー音」を出力する。この後、信号処理部33は本動作を終了する。

[0095] S7で、発話の音声の感情値が「中」である場合、プライバシー音変換部46は、代替音テーブル49を用いて、正弦波の音声データを読み出し、中域の周波数（例えば500Hz）の音声データに変換する（S9）。

[0096] 出力制御部35は、S11で、中域の周波数の音声データをプライバシー音としてスピーカ装置37に出力する。スピーカ装置37は、プライバシー音である「ピー音」を出力する。この後、信号処理部33は本動作を終了する。

[0097] S7で、発話の音声の感情値が「低」である場合、プライバシー音変換部46は、代替音テーブル49を用いて、正弦波の音声データを読み出し、低域の周波数（例えば200Hz）の音声データに変換する（S10）。

[0098] 出力制御部35は、S11で、低域の周波数の音声データをプライバシー音としてスピーカ装置37に出力する。スピーカ装置37は、プライバシー音である「ピー音」を出力する。この後、信号処理部33は本動作を終了する。

[0099] マイクアレイシステム10では、ユーザは、例えばスピーカ装置37から出力されるお客様hm1の発話の内容が分からなくても、プライバシー音として発せられる「ピー音」の音の高さから、お客様hm1が怒っている等の感情を察することができる。

[0100] 従って、例えば、受付係hm2とお客様hm1の会話録をトラブル事案として、振り返りや社内研修に使用したとしても、ユーザは、お客様hm1の発話の内容が秘匿された状態で、お客様hm1の感情の変化を理解できる。

[0101] [効果等]

このように、音声処理装置は、收音部により收音された音声を取得する取得部と、音声の音声位置を検出する検出部と、音声位置がプライバシー保護エリアPRA内である場合に、音声が発話の音声であるか否かを判定する判定部と、発話の音声进行分析して感情値を取得する分析部と、発話の音声感情値に対応する代替音に変換する変換部と、音声を出力する音声出力部に、代替音を出力させる出力制御部35と、備える。

[0102] 音声処理装置は、例えば指向性制御装置30である。收音部は、例えばマイクアレイ装置MAである。取得部は、例えば通信部31である。検出部は、例えば指向性制御部41である。判定部は、例えば発話判定部34である。分析部は、例えば音声分析部45である。音声出力部は、例えばスピーカ装置37である。変換部は、例えばプライバシー音変換部46である。代替音は、例えばプライバシー音である。

[0103] これにより、音声処理装置は、プライバシー保護を図りつつ、発話者の感情を把握できる。例えば、発話の音声代替音によって秘匿化でき、お客様hm1のプライバシー保護が担保される。また、音声処理装置は、発話された音声を一律にマスキングするのではなく、発話された音声に応じて代替音

を使い分けるので、発話者の感情に応じた代替音を出力できる。よって、また、受付係 h m 2 とお客様 h m 1 の会話録を、トラブル事案としてクレーム発生時の振り返りや社内研修資料に使用しても、ユーザは、お客様 h m 1 の感情の変化を推察できる。つまり、ユーザは、例えば、トラブル時にお客様 h m 1 に対して従業員 h m 2 がどのような対応をすると、お客様 h m 1 が落ち着くのかを把握できる。

[0104] また、分析部は、発話の音声に対し、ピッチの変化、話速、音量及び滑舌の少なくとも1つ（複数の組み合わせを含む）を分析して、感情値を取得してもよい。

[0105] これにより、音声処理装置は、発話の音声に対し、様々な方法で音声分析できる。従って、ユーザは、お客様 h m 1 の感情を適切に把握できる。

[0106] また、変換部は、感情値に応じて代替音の周波数を変更してもよい。

[0107] これにより、音声処理装置は、感情値に応じて異なる周波数のプライバシー音を出力できる。よって、ユーザは、お客様 h m 1 の感情を適切に把握できる。

[0108] （第2の実施形態）

第1の実施形態では、音声分析部45で音声分析を行った結果得られる感情値に対応する代替音を、プライバシー音として出力することを示した。第2の実施形態では、感情値に対応する顔アイコンを、カメラ装置CAによって撮像される音声位置の映像の代わりに出力することを示す。

[0109] [構成等]

図7は、第2の実施形態におけるマイクアレイシステム10Aの構成を示すブロック図である。第2の実施形態のマイクアレイシステムは、第1の実施形態とほぼ同一の構成を有する。第1の実施形態と同一の構成要素については、同一の符号を用いることで、その説明を省略又は簡略化する。

[0110] マイクアレイシステム10Aは、第1の実施形態のマイクアレイシステム10と同様の構成の他、音声分析部45A及び映像変換部65を有する。

[0111] 音声分析部45Aは、プライバシー音変換部46を省き、プライバシー音

DB48Aを有する。音声分析部45Aは、マイクアレイ装置MAにより収録されたプライバシー保護エリアPRA内の音声を受けると、この音声を分析し、音声を発した人物の感情を感情値として取得する。この音声の分析では、プライバシー音DB48Aに登録された感情値テーブル47が用いられる。

[0112] 映像変換部65は、顔アイコン変換部66及び顔アイコンDB68を有する。映像変換部65は、カメラ装置CAによって撮像される音声位置の映像を、感情値に対応する代替画像（例えば顔アイコン）に変換する。顔アイコンDB68には、代替画像テーブル67が記憶されている。

[0113] 図8は代替画像テーブル67の登録内容を示す模式図である。

[0114] 代替画像テーブル67には、感情値に対応する顔アイコンfm（fm1, fm2, fm3, …）が登録されている。例えば、感情値が高くて「高」である場合、怒っているような表情を持つ顔アイコンfm1に変換される。例えば、感情値が普通（中程度）で「中」である場合、穏やかな表情を持つ顔アイコンfm2に変換される。例えば、感情値が低くて「低」である場合、笑っているような表情を持つ顔アイコンfm3に変換される。

[0115] 尚、図8では3つの登録例を示したが、任意の数の顔アイコンが感情値に対応するように登録されていてもよい。

[0116] 顔アイコン変換部66は、音声分析部45Aによる音声分析の結果、得られる感情値に対応する顔アイコンfmを、顔アイコンDB68内の代替画像テーブル67から取得する。顔アイコン変換部66は、カメラ装置CAによって撮像される音声位置の映像に、取得された顔アイコンfmを重ねる。映像変換部65は、顔アイコン変換後の画像データを出力制御部35に送る。出力制御部35は、顔アイコン変換後の画像データをディスプレイ装置36に表示させる。

[0117] [動作等]

次に、マイクアレイシステム10Aの動作について説明する。ここでは、来店したお客様と受付係との会話を収録して音声出力する場合を一例として

示す。

- [0118] 図9は、店舗の窓口に設置されたマイクアレイ装置MAによって、受付係hm2とお客さまhm1との会話が収録される状況を表す映像を示す模式図である。
- [0119] 図9の映像は、店舗内の天井に設置された固定カメラであるカメラ装置CAによって撮像された撮像エリアSAが、ディスプレイ装置36に映し出されている。例えば、受付係hm2がお客さまhm1と対面するカウンタ101の真上に、マイクアレイ装置MAが設置される。マイクアレイ装置MAは、受付係hm2とお客さまhm1との会話を含む、店舗内の音声を収録する。
- [0120] お客さまhm1が位置するカウンタ101は、プライバシー保護エリアPRAに設定されている。プライバシー保護エリアPRAは、例えば、ユーザが予めディスプレイ装置36に表示された映像に対して、タッチ操作等で範囲を指定することで設定される。
- [0121] 図9の映像では、撮像エリアSAにおいて、お客さまhm1が来店し、カウンタ101の前に設置されたプライバシー保護エリアPRAに入っている状況が示されている。例えば、受付係hm2が「いらっしゃいませ」と挨拶をすると、その音声はスピーカ装置37から出力される。また、例えば、お客さまhm1が発した音声はスピーカ装置37から「先日のトラブルの件」と出力される。発音内容は認識可能である。
- [0122] 一方、プライバシー保護エリアPRA内に立つ、お客さまhm1の顔近傍（音声位置）には、怒っているような表情を持つ顔アイコンfm1が描画されている。
- [0123] これにより、ユーザは、発話内容を察知でき、顔アイコンfm1からお客さまhm1の感情を察知できる。一方、顔アイコンfm1によってお客さまhm1の顔は秘匿化（マスク）され、お客さまhm1のプライバシー保護が担保される。
- [0124] 尚、受付係hm2とお客さまhm1が発した発話の音声を表す吹き出しは、説明を分かり易くするために付加されたものである。

- [0125] 図10は、マイクアレイ装置MAで收音された音声に基づく顔アイコンを含む映像の出力手順を示すフローチャートである。この映像出力動作は、例えば、マイクアレイ装置MAで收音された音声の音声データ及び画像データをレコーダRCに一旦記憶させた後に行われる。
- [0126] 尚、第1の実施形態と同一のステップ処理については、同一のステップ番号を付すことで、その説明を省略又は簡略化する。
- [0127] S3では、音声位置がプライバシー保護エリアPRA内でない場合、出力制御部35は、カメラ装置CAで撮像された、顔画像を含む映像データをディスプレイ装置36に出力する(S4A)。この場合、出力制御部35は、指向性形成済みの音声データをそのままスピーカ装置37に出力する。この後、信号処理部33は本動作を終了する。
- [0128] S7で、発話の音声の感情値が「高」である場合、顔アイコン変換部66は、代替画像テーブル67に登録されている、感情値が「高」に対応する顔アイコンfm1を読み出す。顔アイコン変換部66は、読み出された顔アイコンfm1を、カメラ装置CAで撮像された映像データの顔画像(音声位置)に重畳することで、映像データを変換する(S8A)。
- [0129] 尚、顔アイコン変換部66は、カメラ装置CAで撮像された映像データの顔画像(音声位置)を、読み出された顔アイコンfm1に置換することで、映像データを変換してもよい(S8A)。
- [0130] 出力制御部35は、変換後の映像データをディスプレイ装置36に出力する(S11A)。ディスプレイ装置36は、顔アイコンfm1を含む映像データを表示する。また、この場合、出力制御部35は、指向性形成済みの音声データをそのままスピーカ装置37に出力する。この後、信号処理部33は本動作を終了する。
- [0131] S7で、発話の音声の感情値が「中」である場合、顔アイコン変換部66は、代替画像テーブル67に登録されている、感情値が「中」に対応する顔アイコンfm2を読み出す。顔アイコン変換部66は、読み出された顔アイコンfm2を、カメラ装置CAで撮像された映像データの顔画像(音声位置)

)に重畳することで、映像データを変換する(S9A)。

[0132] 尚、顔アイコン変換部66は、カメラ装置CAで撮像された映像データの顔画像(音声位置)を、読み出された顔アイコンfm2に置換することで、映像データを変換してもよい(S9A)。

[0133] 出力制御部35は、S11Aで、変換後の映像データをディスプレイ装置36に出力する。ディスプレイ装置36は、顔アイコンfm2を含む映像データを表示する。また、この場合、出力制御部35は、指向性形成済みの音声データをそのままスピーカ装置37に出力する。この後、信号処理部33は本動作を終了する。

[0134] S7で、発話の音声の感情値が「低」である場合、顔アイコン変換部66は、代替画像テーブル67に登録されている、感情値が「低」に対応する顔アイコンfm3を読み出す。顔アイコン変換部66は、読み出された顔アイコンfm3を、カメラ装置CAで撮像された映像データの顔画像(音声位置)に重畳することで、映像データを変換する(S10A)。

[0135] 尚、顔アイコン変換部66は、カメラ装置CAで撮像された映像データの顔画像(音声位置)を、読み出された顔アイコンfm3に置換することで、映像データを変換してもよい(S10A)。

[0136] 出力制御部35は、S11Aで、変換後の映像データをディスプレイ装置36に出力する。ディスプレイ装置36は、顔アイコンfm3を含む映像データを表示する。また、この場合、出力制御部35は、指向性形成済みの音声データをそのままスピーカ装置37に出力する。この後、信号処理部33は本動作を終了する。

[0137] マイクアレイシステム10Aでは、ユーザは、例えばディスプレイ装置36に表示されるお客様hm1の顔画像を視認し難くても、表示された顔アイコンfmの種類に基づいて、お客様hm1が怒っている等の感情を察することができる。

[0138] 従って、例えば、受付係hm2とお客様hm1の会話録をトラブル事案として、振り返りや社内研修に使用したとしても、ユーザは、お客様hm1の

顔画像が秘匿された状態で、お客様 h m 1 の感情の変化を理解できる。

[0139] [効果等]

このように、音声処理装置では、取得部は、撮像部により撮像された撮像エリア S A の映像を取得し、收音部により收音された撮像エリア S A の音声を取得する。変換部は、音声位置の映像を感情値に対応する代替画像に変換する。出力制御部 3 5 は、映像を表示する表示部に、代替画像を表示させる。

[0140] 撮像部は、例えばカメラ装置 C A である。変換部は、例えば顔アイコン変換部 6 6 である。代替画像は、例えば顔アイコン f m である。表示部は、例えばディスプレイ装置 3 6 である。

[0141] また、本実施形態の画像処理装置は、撮像部により撮像された撮像エリア S A の映像と、收音部により收音された撮像エリア S A の音声を取得する取得部と、音声の音声位置を検出する検出部と、音声位置が前記プライバシー保護エリア P R A 内である場合に、音声が発話の音声であるか否かを判定する判定部と、発話の音声を分析して感情値を取得する分析部と、音声位置の映像を感情値に対応する代替画像に変換する変換部と、映像を表示する表示部に、代替画像を表示させる出力制御部 3 5 と、を備える。尚、画像処理装置は、例えば指向性制御装置 3 0 である。

[0142] これにより、ユーザは、顔アイコン f m からお客様 h m 1 の感情を察知できる。また、顔アイコンによってお客様 h m 1 の顔を秘匿化（マスク）でき、お客様 h m 1 のプライバシー保護が担保される。よって、音声処理装置は、プライバシー保護を図りつつ、発話者の感情を視覚的に把握できる。

[0143] また、変換部は、感情値に応じて、感情を示す異なる代替画像を表示させてもよい。

[0144] これにより、音声処理装置は、感情値に応じて異なる表情の顔アイコン f m 等を出力できる。よって、ユーザは、お客様 h m 1 の感情を適切に把握できる。

[0145] （第 3 の実施形態）

第3の実施形態では、第1の実施形態におけるプライバシー音に変換する処理と、第2の実施形態における顔アイコンに変換する処理と、を組み合わせた場合を示す。

[0146] 図11は、第3の実施形態におけるマイクアレイシステム10Bの構成を示すブロック図である。第1及び第2の実施形態と同一の構成要素については、同一の符号を用いることで、その説明を省略又は簡略化する。

[0147] マイクアレイシステム10Bは、第1及び第2の実施形態と同様の構成を有し、音声分析部45及び映像変換部65の両方を有する。音声分析部45及び映像変換部65の構成及び動作は前述した通りである。

[0148] マイクアレイシステム10Bでは、第1の実施形態及び第2の実施形態と同様、例えば、来店したお客様と受付係との会話を収録して音声出力し、お客様と受付係とが所在する撮像エリアを録画する場合を想定する。

[0149] 図12は、店舗の窓口に設置されたマイクアレイ装置MAによって、受付係hm2とお客様hm1との会話が収録される状況を表す映像を示す模式図である。

[0150] 図12に示すディスプレイ装置36に表示される映像では、お客様hm1が来店し、カウンタ101の前に設置されたプライバシー保護エリアPRAに入っている状況が示される。例えば、受付係hm2が「いらっしゃいませ」と挨拶すると、その音声はスピーカ装置37から出力される。また、例えば、お客様hm1も受付係hm2に話しかけるが、スピーカ装置37から「ピー、ピー、ピー」とプライバシー音が出力される。

[0151] これにより、発話内容の秘匿性が担保される。また、マイクアレイシステム10Bのユーザは、スピーカ装置37から出力されるプライバシー音のピッチの変化等から、お客様hm1の感情を察知できる。

[0152] 図12の映像では、プライバシー保護エリアPRA内に立つ、お客様hm1の顔近傍（音声位置）には、怒っているような表情を持つ顔アイコンfm1が配置される。

[0153] これにより、ユーザは、顔アイコンfm1からお客様hm1の感情を察知

できる。また、顔アイコン f m 1 によってお客様 h m 1 の顔が秘匿化（マスク）され、お客様 h m 1 のプライバシー保護が担保される。

[0154] [効果等]

このように、マイクアレイシステム 1 0 B は、撮像エリア S A の映像を撮像する撮像部と、撮像エリアの音声を收音する收音部と、收音部により收音された音声の音声位置を検出する検出部と、音声位置がプライバシー保護エリア P R A 内である場合に、音声が発話の音声であるか否かを判定する判定部と、発話の音声を分析して感情値を取得する分析部と、感情値に対応する変換処理を行う変換部と、変換処理の結果を出力させる出力制御部 3 5 と、を備える。変換処理は、例えば、プライバシー音に変換する音声処理と、顔アイコン f m に変換する画像変換処理と、の少なくとも一方を含む。

[0155] これにより、マイクアレイシステム 1 0 B は、例えば、プライバシー音によってお客様 h m 1 の発話内容が秘匿化され、顔アイコン f m によってお客様 h m 1 の顔が秘匿化されるので、プライバシーを更に保護できる。上記の発話内容の秘匿化と顔の秘匿化とは、少なくとも一方が実施される。また、ユーザは、プライバシー音のピッチの変化や顔アイコンの種類によって、お客様 h m 1 の感情を更に察知し易くなる。

[0156] (他の実施形態)

以上のように、本開示における技術の例示として、第 1 ～ 第 3 の実施形態を説明した。しかし、本開示における技術は、これに限定されず、変更、置き換え、付加、省略などを行った実施形態にも適用できる。また、各実施形態を組み合わせてもよい。

[0157] 第 1, 第 3 の実施形態では、マイクアレイ装置 M A で検出される音声の音声位置がプライバシー保護エリア P R A 内である場合、撮像エリア S A で検出された音声をユーザに依存せずにプライバシー音に変換する処理を行うことを示した。この代わりに、プライバシー音への変換処理が、ユーザに依存して行われてもよい。プライバシー音への変換処理に限らず、顔アイコンの変換処理についても同様である。

- [0158] 例えば、指向性制御装置30を操作するユーザが、一般ユーザである場合、プライバシー音への変換処理を行い、管理者等の権限のあるユーザである場合、プライバシー音への変換処理をしなくてもよい。いずれのユーザであるかは、例えば、指向性制御装置30にログインする際のユーザID等によって判断されてもよい。
- [0159] 第1, 第3の実施形態では、プライバシー音変換部46は、感情値に対応するプライバシー音として、マイクアレイ装置MAにより収録された音声の音声データに対してボイスチェンジ処理（加工処理）を施してもよい。
- [0160] プライバシー音変換部46は、ボイスチェンジ処理の一例として、例えば、マイクアレイ装置MAにより収録された音声の音声データの周波数（ピッチ）の高低を変化させてもよい。つまり、プライバシー音変換部46は、スピーカ装置37から出力される音声の周波数を音声の内容が分かり難くなるような他の周波数に変更してもよい。
- [0161] これにより、プライバシー保護エリアPRA内の音声の内容を認識し難くしつつ、ユーザは話者の感情を察することができる。また、プライバシー音DB48に予めプライバシー音を複数保持することが不要となる。
- [0162] このように、出力制御部35は、マイクアレイ装置MAにより収録され、加工処理された音声をスピーカ装置37から出力させてもよい。これにより、プライバシー保護エリアPRA内に存在する被写体（例えば人物）のプライバシーを効果的に保護できる。
- [0163] 第1～第3の実施形態では、出力制御部35は、ユーザの指又はスタイラスペンによって画面上で指定された指定位置に対応する音声位置がプライバシー保護エリアPRAに含まれる旨を、画面上でユーザに対して明示的に通知してもよい。
- [0164] 第1～第3の実施形態では、音源位置や音源位置の方向がプライバシー保護エリアの範囲や方向にある場合、感情値に応じて音声や映像の少なくとも一部が、代替される別の音声、映像又は画像（代替出力又は変換処理の結果）に変換されることを例示した。この代わりに、プライバシー判断部42は

、収録された時間帯がプライバシー保護を必要とする時間帯（プライバシー保護時間）に含まれるか否かを判断してもよい。プライバシー保護時間に収録時間が含まれる場合に、プライバシー音変換部46や顔アイコン変換部66により、感情値に応じて音声や映像の少なくとも一部が変換されてもよい。

[0165] また、本開示の実施形態では、お客様hm1をプライバシー保護エリアPRAに設定し、お客様hm1の発話から検出される感情値に応じて音声や映像の少なくとも一部が、代替される別の音声、映像又は画像に変換される例を示したが、逆に受付係hm2をプライバシー保護エリアに設定し、受付係hm2の発話から検出される感情値に応じて音声や映像の少なくとも一部が、代替される別の音声、映像又は画像に変換されてもよい。これにより、例えばトラブル事案としてクレーム発生時の振り返りや社内研修資料に使用する際に、受付係の顔をアイコンに変更することで、社員の特定を困難にするという効果が期待出来る。

[0166] さらに、本開示の実施形態では、マイクアレイ装置MA及び指向性制御装置30を用いて、お客様hm1及び受付係hm2の発話を収録しているが、これらの代わりにお客様hm1及び受付係hm2それぞれの近傍に設置された複数のマイク（例えば指向性マイクなど）を用いて、それぞれの発話を収録してもよい。

産業上の利用可能性

[0167] 本開示は、プライバシー保護を図りつつ、発話者の感情を察知できる音声処理装置、画像処理装置、マイクアレイシステム、及び音声処理方法等に有用である。

符号の説明

[0168] 10, 10A, 10B マイクアレイシステム
21 筐体
26 加算器
30 指向性制御装置

- 3 1 通信部
- 3 2 操作部
- 3 3 信号処理部
- 3 4 発話判定部
- 3 5 出力制御部
- 3 6 ディスプレイ装置
- 3 7 スピーカ装置
- 3 8 メモリ
- 3 9 設定管理部
- 3 9 z メモリ
- 4 1 指向性制御部
- 4 2 プライバシー判断部
- 4 5, 4 5 A 音声分析部
- 4 6 プライバシー音変換部
- 4 7, 4 7 A, 4 7 B, 4 7 C, 4 7 D 感情値テーブル
- 4 8, 4 8 A プライバシー音データベース (D B)
- 4 9 代替音テーブル
- 6 5 映像変換部
- 6 6 顔アイコン変換部
- 6 7 代替画像テーブル
- 6 8 顔アイコンデータベース (D B)
- 8 0 音源
- 1 0 1 カウンタ
- 2 4 1, 2 4 2, 2 4 3, ..., 2 4 n A / D 変換器
- 2 5 1, 2 5 2, 2 5 3, ..., 2 5 n 遅延器
- C A カメラ装置
- f m, f m 1, f m 2, f m 3 顔アイコン
- h m 1 お客様

h m 2 受付係

NW ネットワーク

MA マイクアレイ装置

MA 1, MA 2, ..., MA n, MB 1, MB 2, ..., MB n マイクロ

ホン

RC レコーダ

SA 撮像エリア

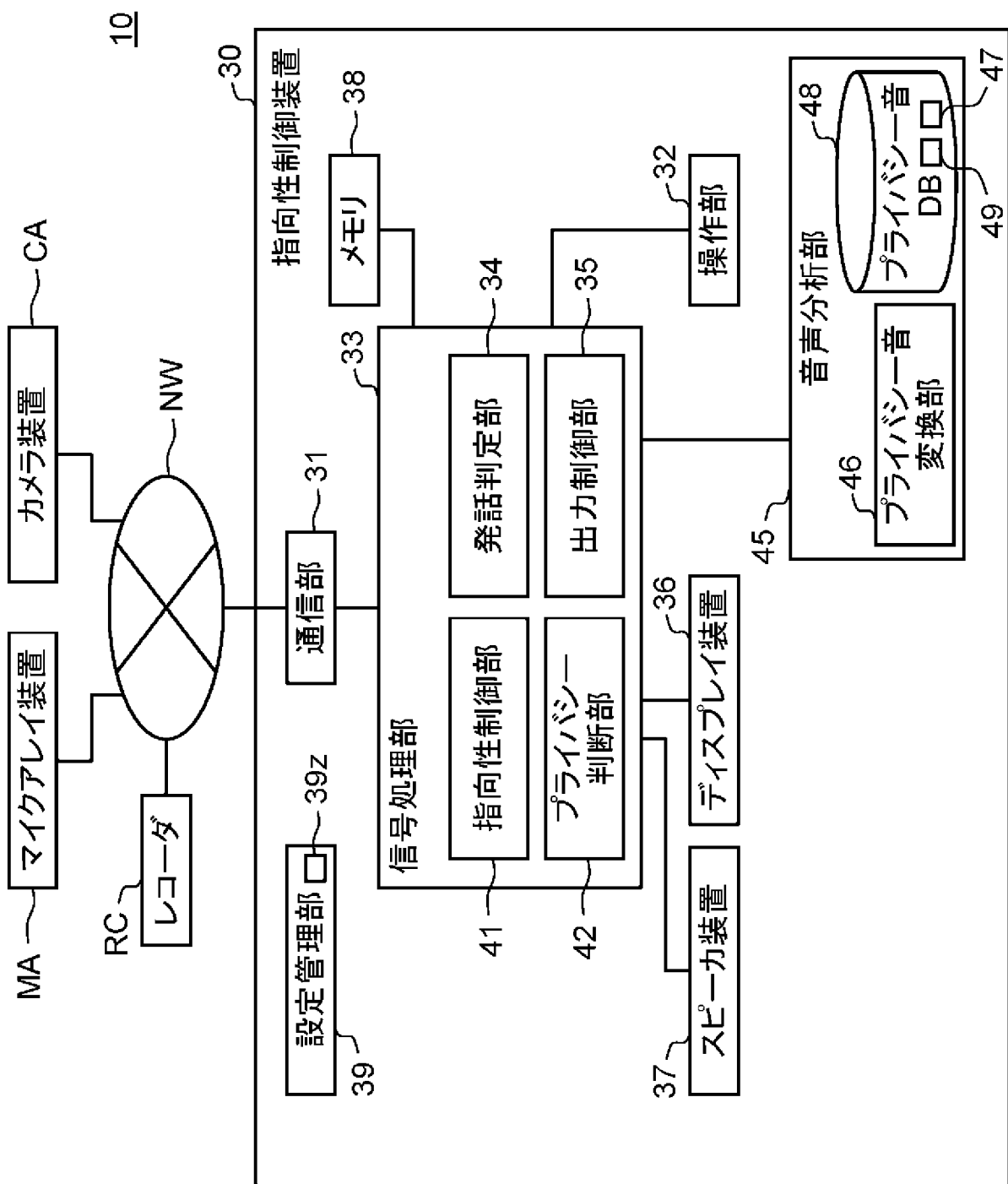
請求の範囲

- [請求項1] 收音部により收音された音声を取得する取得部と、
前記音声の音声位置を検出する検出部と、
前記音声位置がプライバシー保護エリア内である場合に、前記音声が発話の音声であるか否かを判定する判定部と、
前記発話の音声を分析して感情値を取得する分析部と、
前記発話の音声を前記感情値に対応する代替音に変換する変換部と、
、
前記音声を出力する音声出力部に、前記代替音を出力させる出力制御部と、
備える音声処理装置。
- [請求項2] 請求項1に記載の音声処理装置であって、
前記分析部は、前記発話の音声に対し、ピッチの変化、話速、音量及び滑舌の少なくとも1つを分析して、前記感情値を取得する、音声処理装置。
- [請求項3] 請求項1に記載の音声処理装置であって、
前記変換部は、前記感情値に応じて前記代替音の周波数を変更する、音声処理装置。
- [請求項4] 請求項1に記載の音声処理装置であって、
前記取得部は、撮像部により撮像された撮像エリアの映像を取得し、
前記收音部により收音された前記撮像エリアの音声を取得し、
前記変換部は、前記音声位置の前記映像を前記感情値に対応する代替画像に変換し、
前記出力制御部は、前記映像を表示する表示部に、前記代替画像を表示させる、音声処理装置。
- [請求項5] 請求項4に記載の音声処理装置であって、
前記変換部は、前記感情値に応じて、感情を示す異なる代替画像を表示させる、音声処理装置。

- [請求項6] 撮像部により撮像された撮像エリアの映像と、收音部により收音された前記撮像エリアの音声を取得する取得部と、
前記音声の音声位置を検出する検出部と、
前記音声位置がプライバシー保護エリア内である場合に、前記音声が発話の音声であるか否かを判定する判定部と、
前記発話の音声进行分析して感情値を取得する分析部と、
前記音声位置の映像を前記感情値に対応する代替画像に変換する変換部と、
前記映像を表示する表示部に、前記代替画像を表示させる出力制御部と、
備える画像処理装置。
- [請求項7] 撮像エリアの映像を撮像する撮像部と、
前記撮像エリアの音声を收音する收音部と、
前記收音部により收音された前記音声の音声位置を検出する検出部と、
前記音声位置がプライバシー保護エリア内である場合に、前記音声が発話の音声であるか否かを判定する判定部と、
前記発話の音声进行分析して感情値を取得する分析部と、
前記感情値に対応する変換処理を行う変換部と、
前記変換処理の結果を出力させる出力制御部と、
備える、マイクアレイシステム。
- [請求項8] 音声処理装置における音声処理方法であって、
收音部により收音された音声を取得し、
前記音声の音声位置を検出し、
前記音声位置がプライバシー保護エリア内である場合に、前記音声が発話の音声であるか否かを判定し、
前記発話の音声进行分析して感情値を取得し、
前記発話の音声を前記感情値に対応する代替音に変換し、

前記音声を出力する音声出力部に、前記代替音を出力させる、音声処理方法。

[図1]



[図2A]

47A

ピッチの変化	感情値
大	高
中	中
小	低

[図2B]

47B

話速	感情値
早い	高
普通	中
遅い	低

[図2C]

47C

音量	感情値
大	高
中	中
小	低

[図2D]

47D

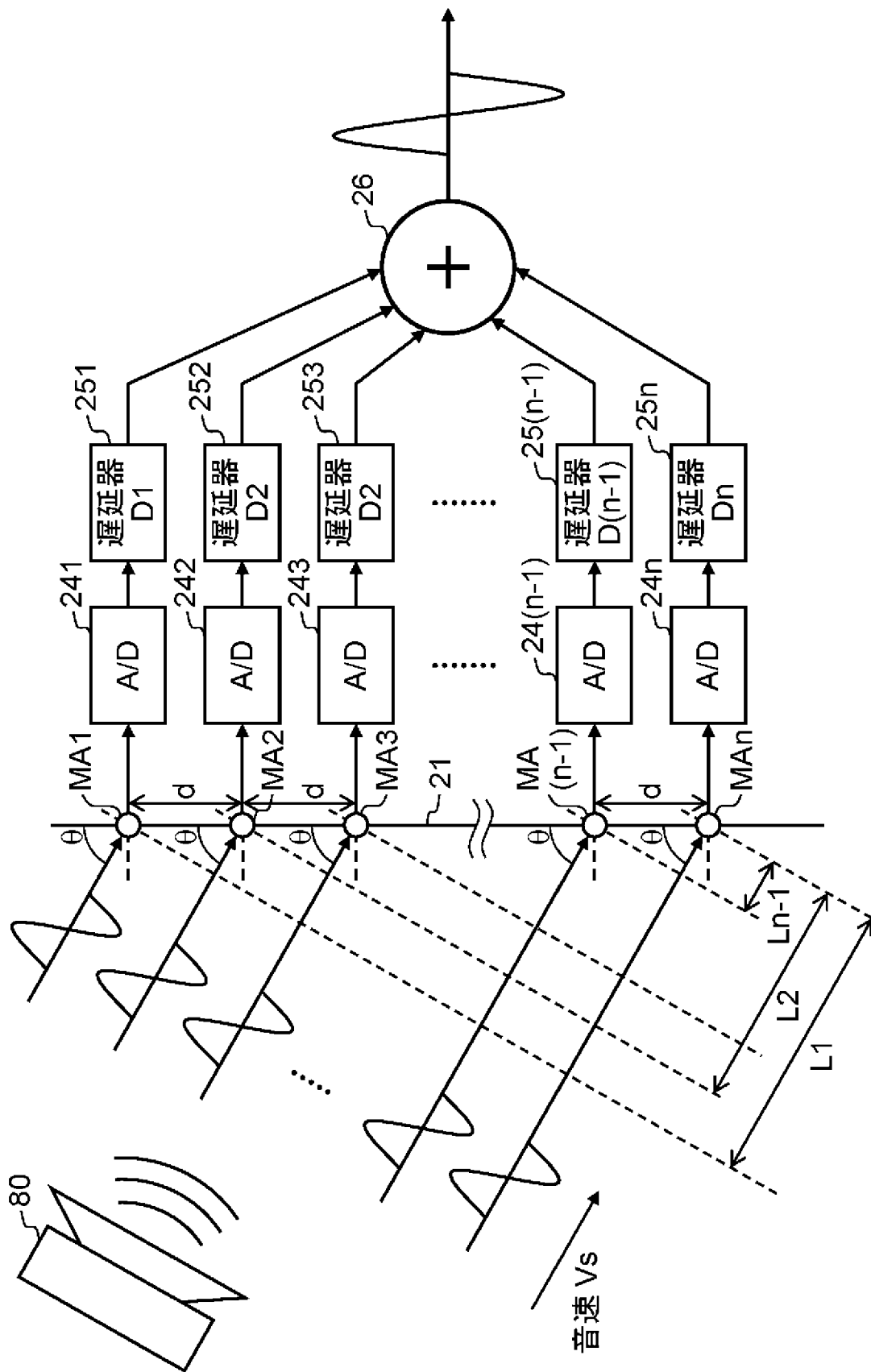
滑舌(音声認識率)	感情値
悪い	高
普通	中
良い	低

[図3]

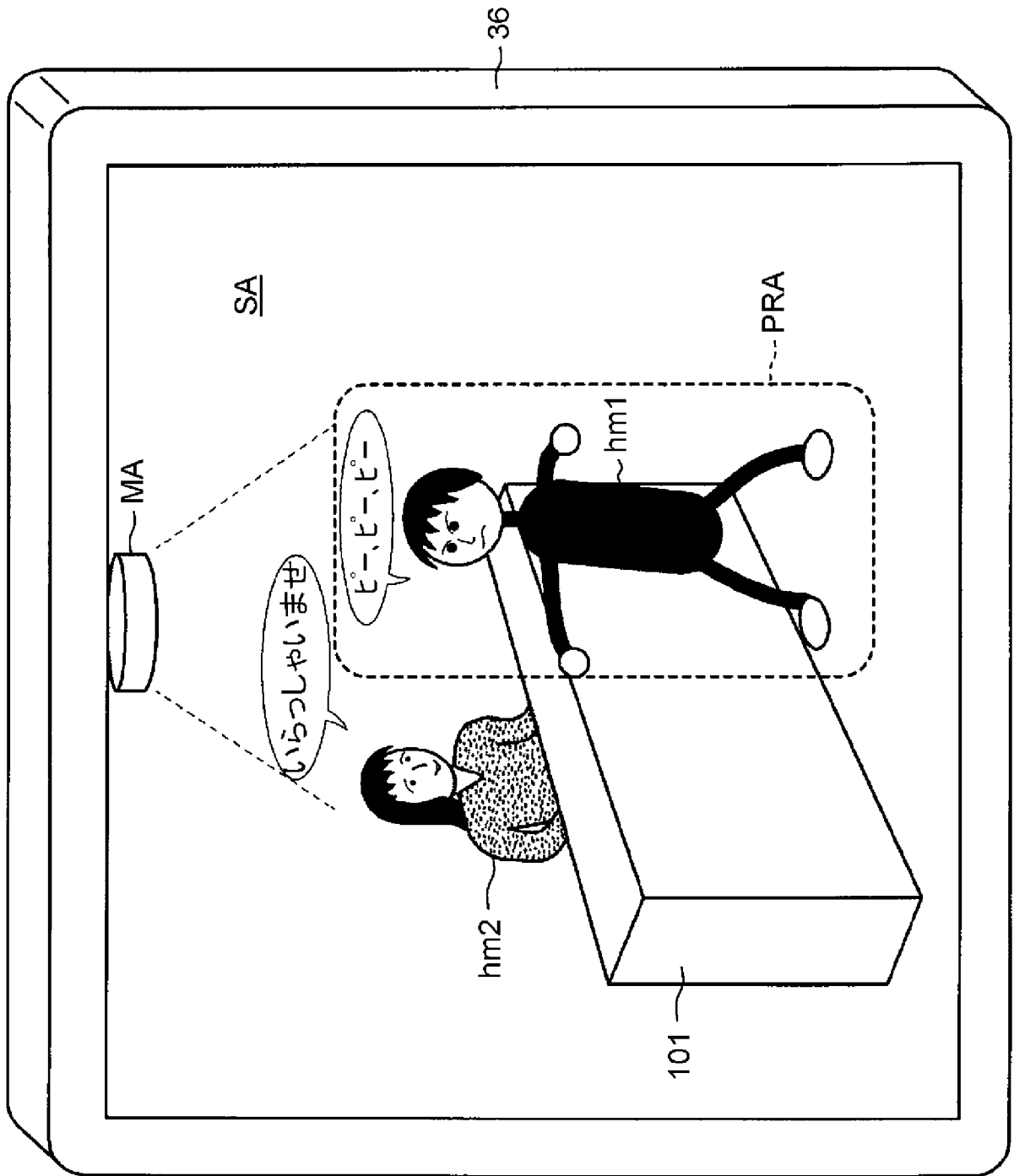
49

感情値	ピ一音(正弦波)
高	高域(1kHz)
中	中域(500Hz)
低	低域(200Hz)

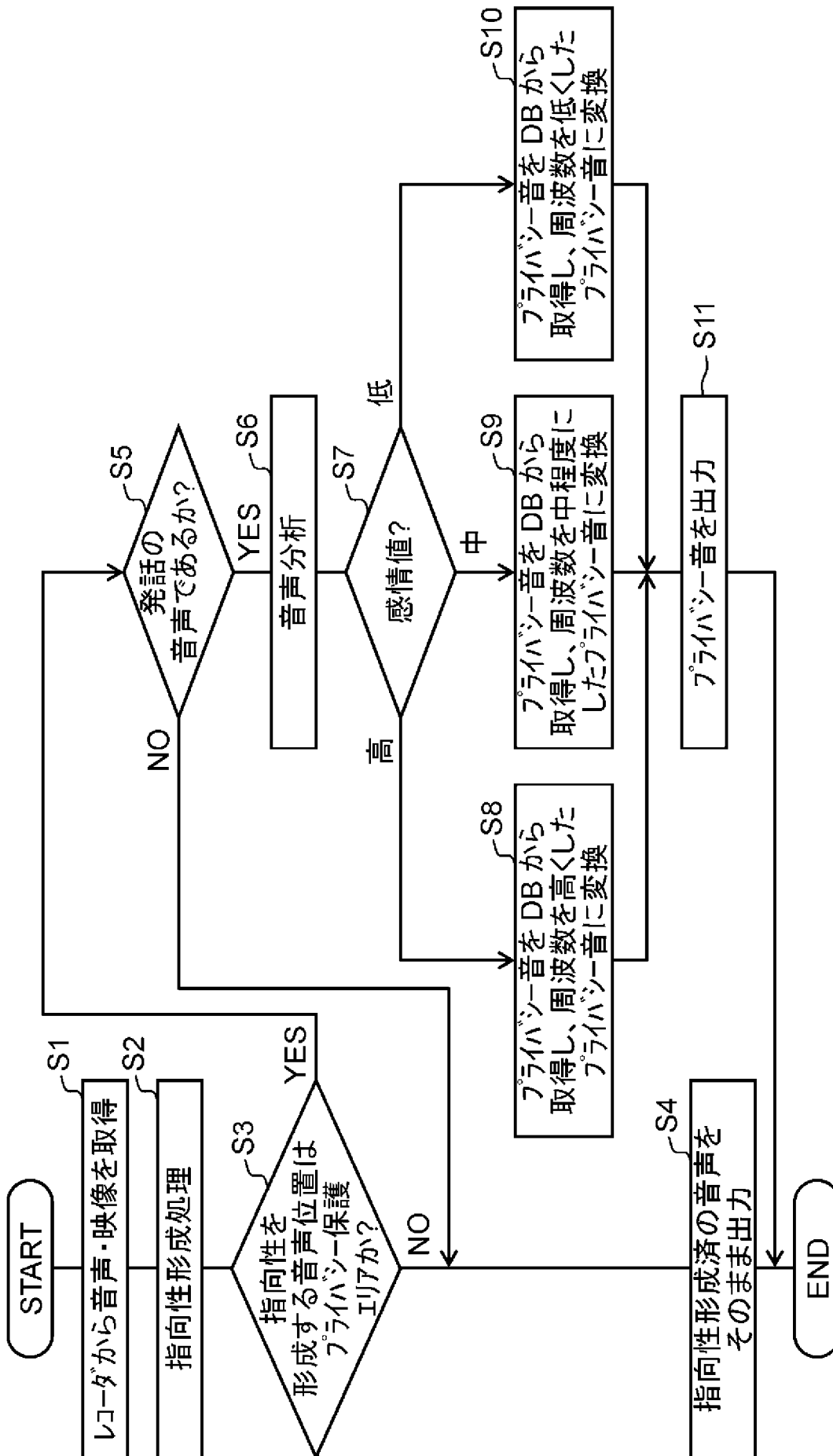
[図4]



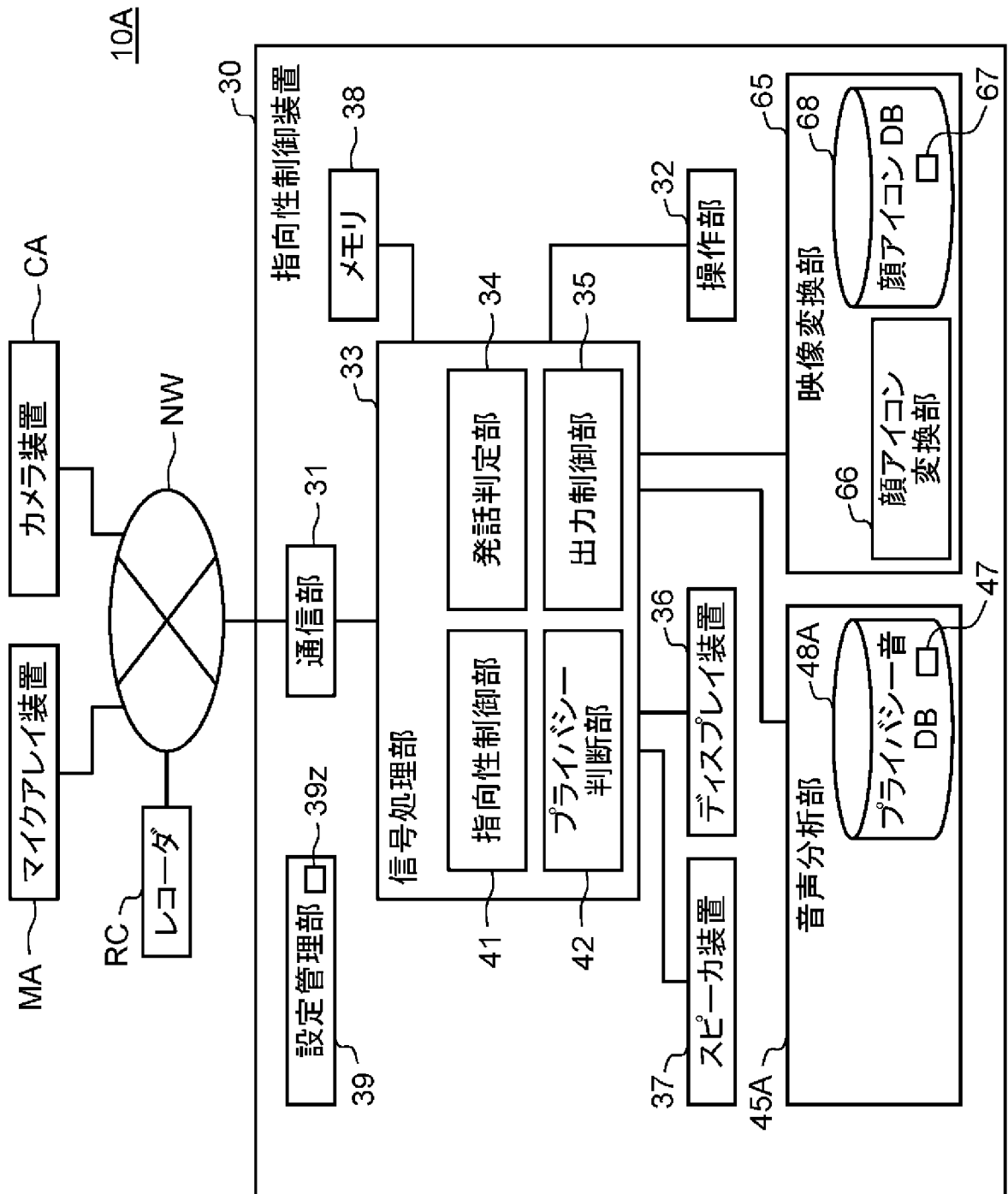
[図5]



[図6]






[図7]

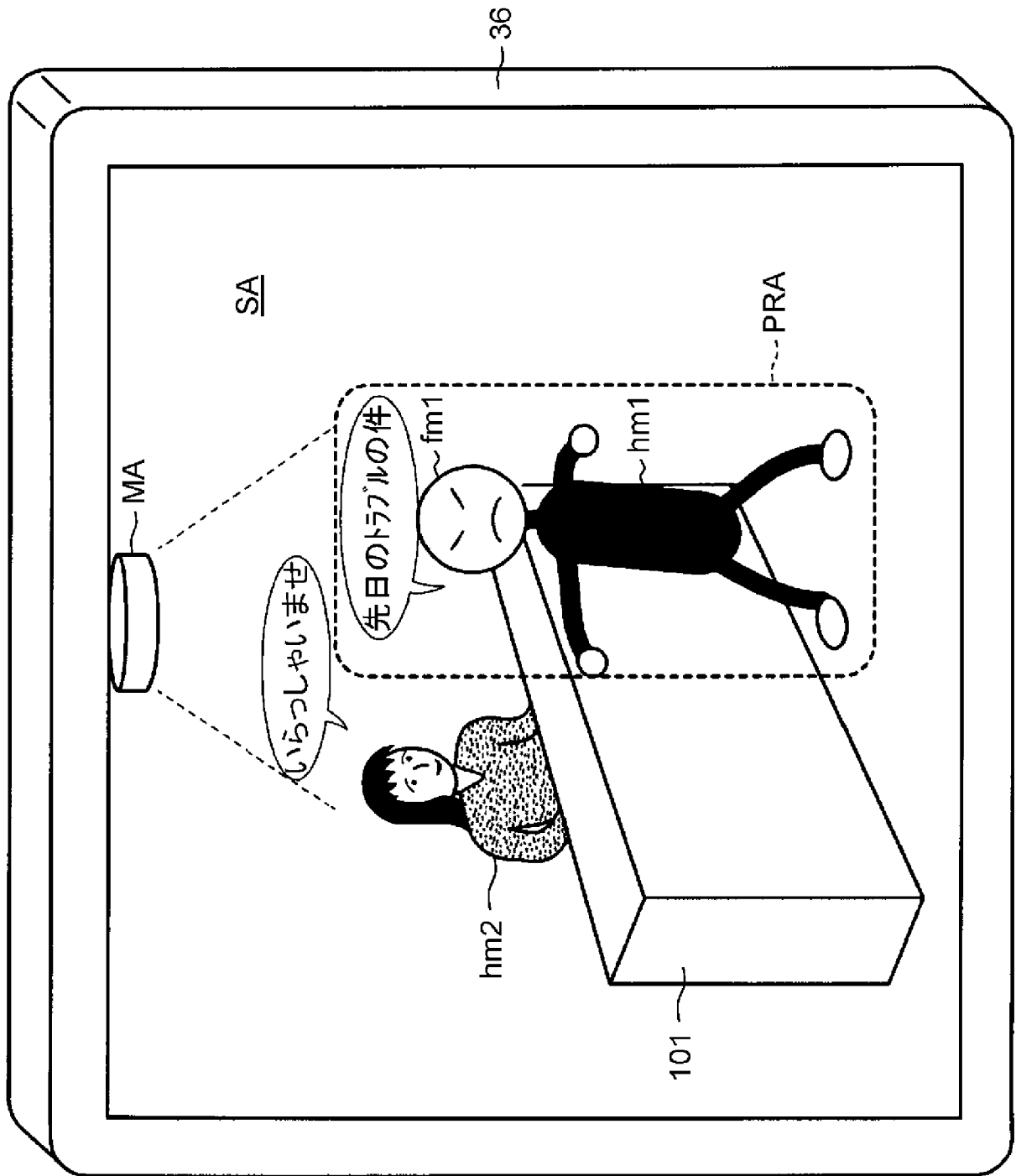


[図8]

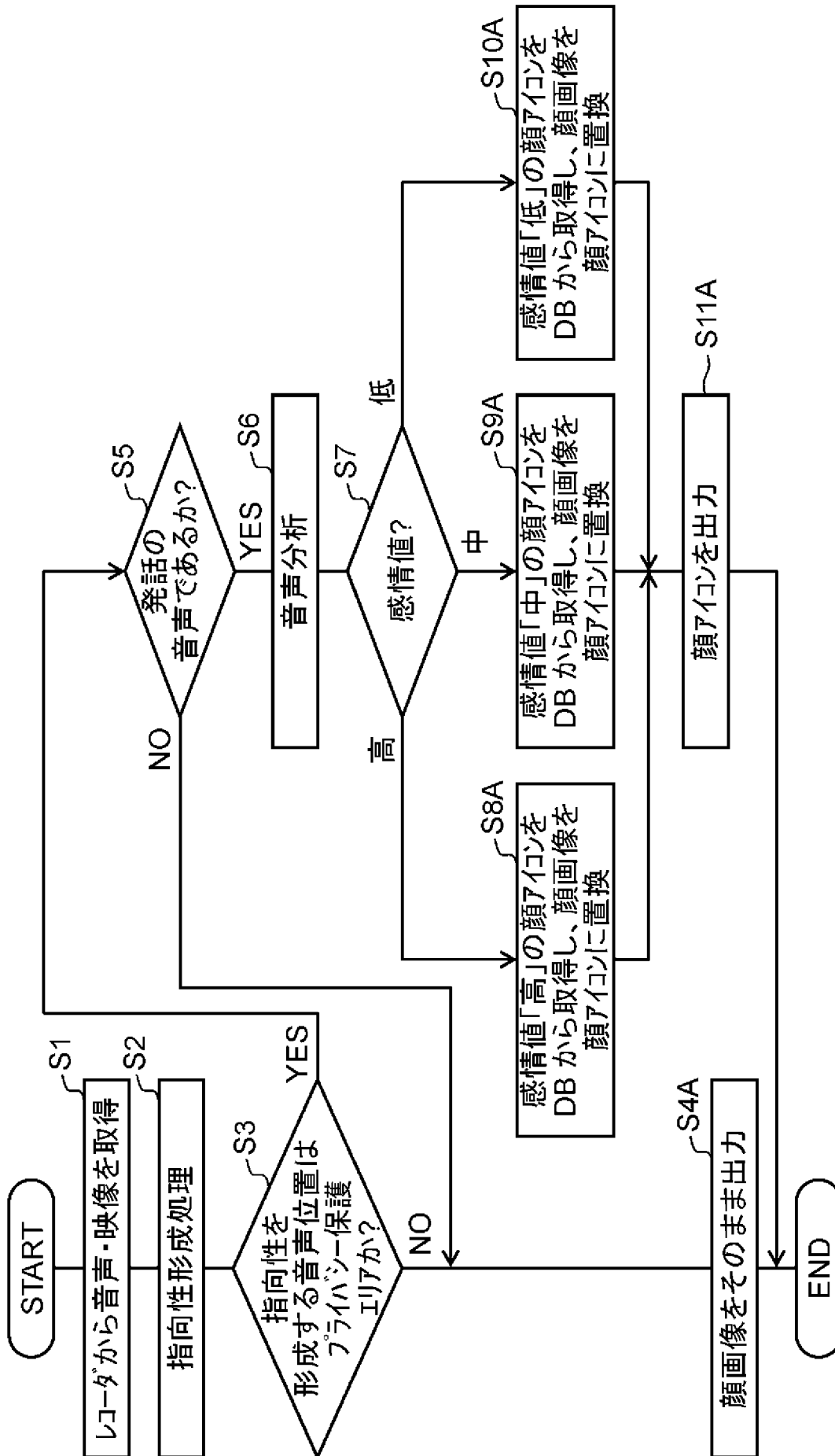
67

感情値	顔アイコン
高	 fm1
中	 fm2
低	 fm3

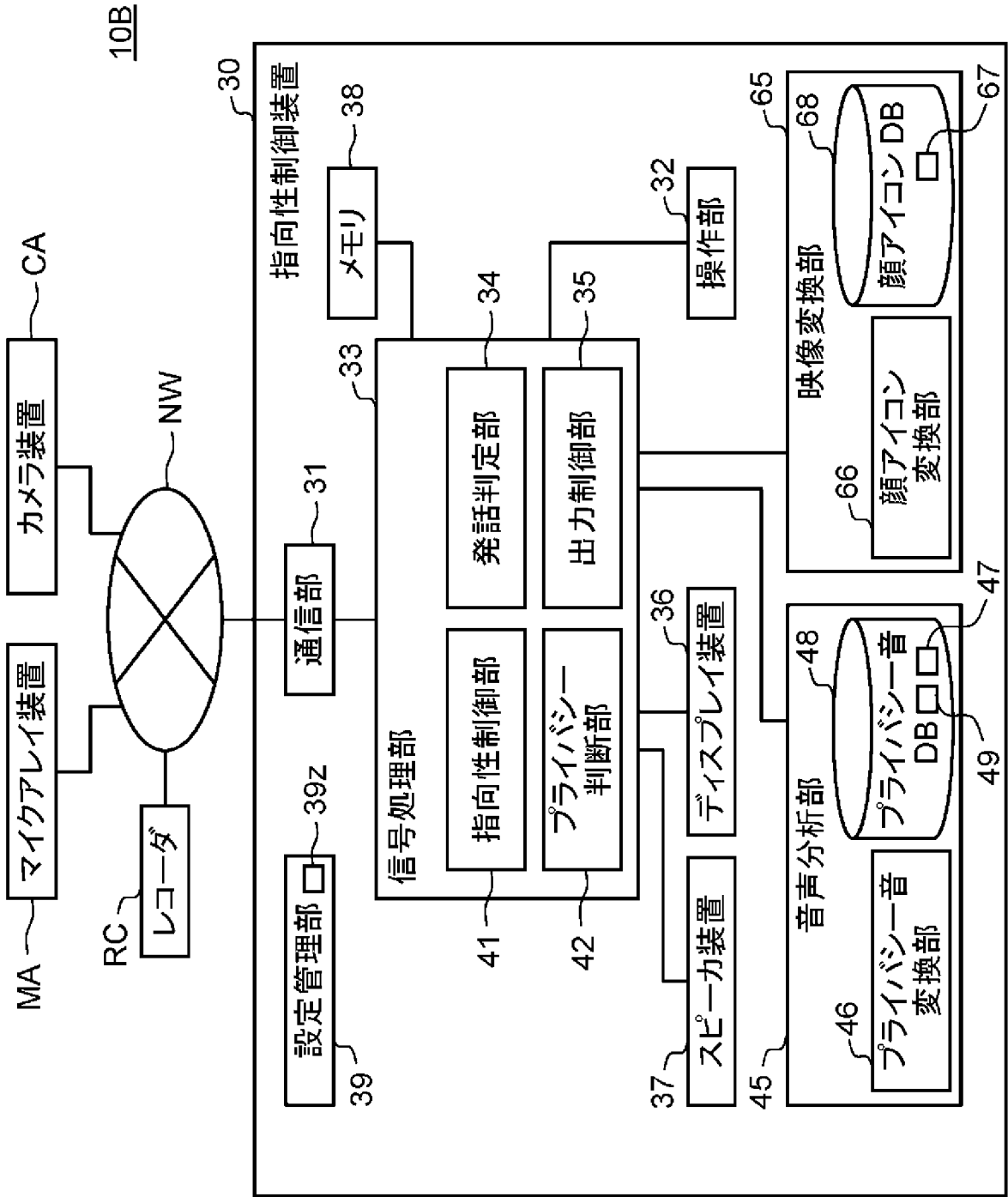
[図9]



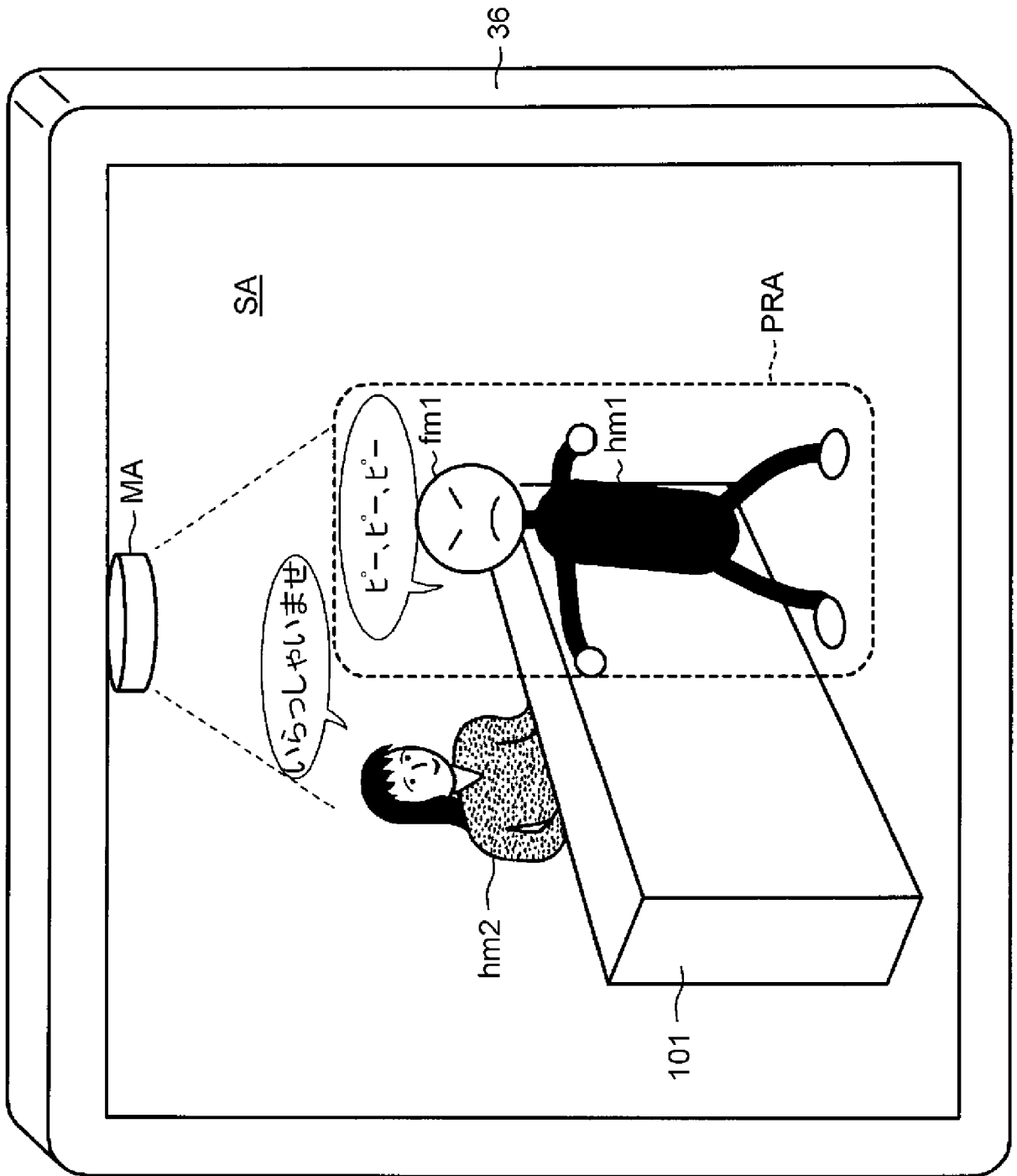
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2017/004483

A. CLASSIFICATION OF SUBJECT MATTER
G10L25/63(2013.01)i, G10L25/57(2013.01)i, G10L25/84(2013.01)i, H04R1/40(2006.01)i, H04R3/00(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)
G10L25/63, G10L25/57, G10L25/84, H04R1/40, H04R3/00

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	1996-2017
Kokai Jitsuyo Shinan Koho	1971-2017	Toroku Jitsuyo Shinan Koho	1994-2017

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 A	JP 2015-029241 A (Panasonic Intellectual Property Management Co., Ltd.), 12 February 2015 (12.02.2015), paragraphs [0010], [0049]; fig. 6 & US 2014/0376740 A1 paragraphs [0011], [0076]; fig. 6 & CN 104243779 A	1-2, 4-8 3
A	JP 2011-002704 A (Nippon Telegraph and Telephone Corp.), 06 January 2011 (06.01.2011), paragraphs [0032] to [0056] (Family: none)	1-8

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 29 March 2017 (29.03.17)	Date of mailing of the international search report 11 April 2017 (11.04.17)
---	--

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/JP2017/004483

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 2003-248837 A (Mega Chips Corp.), 05 September 2003 (05.09.2003), paragraphs [0173] to [0226]; fig. 4 (Family: none)	1-8
A	WO 2014/192457 A1 (Sony Corp.), 04 December 2014 (04.12.2014), paragraph [0049] & US 2016/0142767 A1 paragraph [0071] & CN 105247879 A	1-8
A	JP 2010-169925 A (Konami Digital Entertainment Co., Ltd.), 05 August 2010 (05.08.2010), paragraph [0020] & WO 2010/084830 A1 & TW 201040940 A	1-8
A	JP 2004-248145 A (Mega Chips System Solutions Inc.), 02 September 2004 (02.09.2004), paragraphs [0087] to [0103]; fig. 13 (Family: none)	4-6

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G10L25/63(2013.01)i, G10L25/57(2013.01)i, G10L25/84(2013.01)i, H04R1/40(2006.01)i, H04R3/00(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. G10L25/63, G10L25/57, G10L25/84, H04R1/40, H04R3/00

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	1971-2017年
日本国実用新案登録公報	1996-2017年
日本国登録実用新案公報	1994-2017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 A	JP 2015-029241 A (パナソニック IPマネジメント株式会社) 2015.02.12, [0010]、[0049]、[図6] & US 2014/0376740 A1, [0011], [0076], FIG.6 & CN 104243779 A	1-2, 4-8 3
A	JP 2011-002704 A (日本電信電話株式会社) 2011.01.06, [003 2] - [0056] (ファミリーなし)	1-8
A	JP 2003-248837 A (株式会社メガチップス) 2003.09.05, [017 3] - [0226]、[図4] (ファミリーなし)	1-8

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

29.03.2017

国際調査報告の発送日

11.04.2017

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)

千本 潤介
 電話番号 03-3581-1101 内線 3591

5Z

3457

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	WO 2014/192457 A1 (ソニー株式会社) 2014.12.04, [0049] & US 2016/0142767 A1, [0071] & CN 105247879 A	1-8
A	JP 2010-169925 A (株式会社コナミデジタルエンタテインメント) 2010.08.05, [0020] & WO 2010/084830 A1 & TW 201040940 A	1-8
A	JP 2004-248145 A (株式会社メガチップスシステムソリューションズ) 2004.09.02, [0087] - [0103], [図13] (ファミリーなし)	4-6