

(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11月9日(09.11.2017)



(10) 国際公開番号  
**WO 2017/191711 A1**

- (51) 国際特許分類:  
G10L 15/28 (2013.01) G10L 15/00 (2013.01)  
G06F 3/16 (2006.01) G10L 15/10 (2006.01)  
G06F 17/28 (2006.01)
- (21) 国際出願番号: PCT/JP2017/010187
- (22) 国際出願日: 2017年3月14日(14.03.2017)
- (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 2016-092412 2016年5月2日(02.05.2016) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 河野 真一 (KAWANO, Shinichi); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 市川 美和 (ICHIKAWA, Miwa); 〒1080075 東京

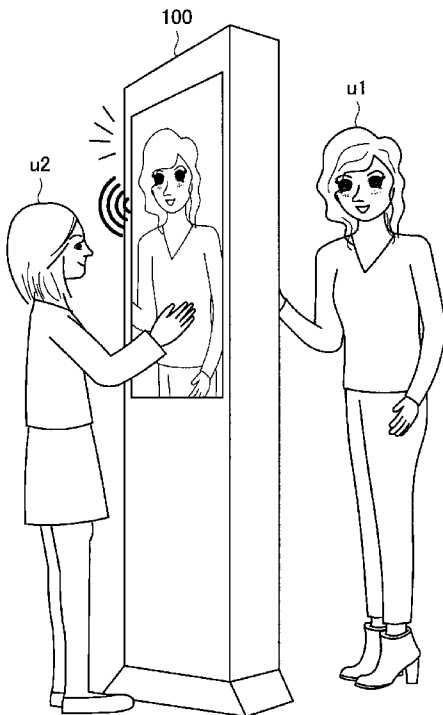
都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 亀谷 美明, 外 (KAMEYA, Yoshiaki et al.); 〒1600004 東京都新宿区四谷3-1-3 第一富澤ビル はづき国際特許事務所 四谷オフィス 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, 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.

(54) Title: CONTROL DEVICE, CONTROL METHOD, AND COMPUTER PROGRAM

(54) 発明の名称: 制御装置、制御方法およびコンピュータプログラム



(57) Abstract: [Problem] To provide a control device which is capable of causing translation processing to continue smoothly. [Solution] Provided is a control device provided with a control unit which, in cases when speech content from a speaker reaches a level at which it is determined whether translation processing of the speech content ought to continue, determines that the translation processing is to continue on the basis of the speech content and the speech state of the speaker.

(57) 要約: 【課題】 翻訳処理をスムーズに継続させることが可能な制御装置を提供する。【解決手段】 発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行う制御部を備える、制御装置が提供される。



WO 2017/191711 A1

(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 :

一 国際調査報告 (条約第21条(3))

## 明 細 書

**発明の名称**： 制御装置、制御方法およびコンピュータプログラム

### 技術分野

[0001] 本開示は、制御装置、制御方法およびコンピュータプログラムに関する。

### 背景技術

[0002] 文字や音声を認識し、その文字や音声を他の言語に翻訳する技術の開発が進んでいる。そして、そのような翻訳の技術を様々な分野に適用するための技術の開発も進んでいる。例えば、特許文献1には、サイネージ端末に対して翻訳技術を適用した技術の例が開示されている。

### 先行技術文献

#### 特許文献

[0003] 特許文献1：特開2015-219768号公報

### 発明の概要

#### 発明が解決しようとする課題

[0004] 発話を行う者（以下「発話者」とも称する）が言い淀んだり、考えたりしている場合に、発話内容に対する翻訳処理を止めてしまうと、円滑なコミュニケーションが出来なくなる。

[0005] そこで、本開示では、翻訳処理をスムーズに継続させることが可能な、新規かつ改良された制御装置、制御方法およびコンピュータプログラムを提案する。

#### 課題を解決するための手段

[0006] 本開示によれば、発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行う制御部を備える、制御装置が提供される。

[0007] また本開示によれば、発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発

話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行うことを含む、制御方法が提供される。

[0008] また本開示によれば、発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断をコンピュータに実行させる、コンピュータプログラムが提供される。

### 発明の効果

[0009] 以上説明したように本開示によれば、翻訳処理をスムーズに継続させることが可能な、新規かつ改良された制御装置、制御方法およびコンピュータプログラムを提供することができる。

[0010] なお、上記の効果は必ずしも限定的なものではなく、上記の効果とともに、または上記の効果に代えて、本明細書に示されたいずれかの効果、または本明細書から把握され得る他の効果が奏されてもよい。

### 図面の簡単な説明

[0011] [図1]本開示の実施の形態に係る、サイネージ端末100を用いたコミュニケーションシステムについて示す説明図である。

[図2]同実施の形態に係る、サイネージ端末100を用いたコミュニケーションシステム1の全体構成例を示す説明図である。

[図3]サイネージ端末100を上方向から見た場合の例を示す説明図である。

[図4]同実施の形態に係るサイネージ端末100の機能構成例を示す説明図である。

[図5]同実施の形態に係る情報処理装置200の機能構成例を示す説明図である。

[図6]同実施の形態に係る音声情報処理装置300の機能構成例を示す説明図である。

[図7]同実施の形態に係る翻訳処理装置400の機能構成例を示す説明図である。

[図8]同実施の形態に係るコミュニケーションシステム1の動作例を説明する

説明図である。

[図9]同実施の形態に係るコミュニケーションシステム1の動作例を説明する説明図である。

[図10]ユーザインターフェースの例を示す説明図である。

[図11]ユーザu1から見てサインージ端末100の反対側にユーザu2が立った状態の例を示す説明図である。

[図12]ディスプレイ103の表示例を示す説明図である。

[図13]ユーザインターフェースの例を示す説明図である。

[図14]ユーザu1が顔を高感度マイク位置に近付ける様子を示す説明図である。

[図15]ユーザインターフェースの例を示す説明図である。

[図16]ユーザインターフェースの例を示す説明図である。

[図17]ユーザインターフェースの例を示す説明図である。

[図18]ユーザインターフェースの例を示す説明図である。

[図19]コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。

[図20]コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。

[図21]コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。

[図22]ユーザインターフェースの例を示す説明図である。

[図23]コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。

[図24]コミュニケーションシステム1が聞き手側の処理あふれの判断を行うことを説明するための説明図である。

[図25]コミュニケーションシステム1が聞き手側の処理あふれの判断を行うことを説明するための説明図である。

[図26]コミュニケーションシステム1が聞き手側の処理あふれの判断を行う

ことを説明するための説明図である。

[図27]コミュニケーションシステム1が聞き手側の処理あふれの判断を行うことを説明するための説明図である。

[図28]ユーザインターフェースの例を示す説明図である。

[図29]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図30A]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図30B]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図31]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図32]口の位置情報、左目の位置情報、右目の位置情報、および鼻の最下部位置情報を示す説明図である。

[図33]左の耳位置および右の耳位置を示す説明図である。

[図34]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図35]高感度マイク位置の決定方法を説明するための説明図である。

[図36A]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図36B]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図37]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図38]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図39A]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図39B]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図40]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図41]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図42]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図43]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図44]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図45]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図46]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図47]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図48]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図49]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図50]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図51]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

[図52]同実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。

## 発明を実施するための形態

[0012] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0013] なお、説明は以下の順序で行うものとする。

### 1. 本開示の実施の形態

#### 1. 1. 概要

#### 1. 2. 構成例

#### 1. 3. 動作例

### 2. まとめ

[0014] <1. 本開示の実施の形態>

#### [1. 1. 概要]

本開示の実施の形態について詳細に説明する前に、本開示の実施の形態の概要を説明する。

[0015] 上述したように、文字や音声を認識し、その文字や音声を他の言語に翻訳する技術の開発が進んでいる。そして、そのような翻訳の技術を様々な分野に適用するための技術の開発も進んでいる。例えば、上記特許文献1には、透明なディスプレイを備えるサイネージ端末に対して翻訳技術を適用して、人間同士がコミュニケーションを取りやすくするための技術の例が開示されている。

[0016] 例えば、街中に置かれているサイネージ端末や、遠隔地同士を結ぶテレビ電話などを介して、人間同士が正対してコミュニケーション、特に、会話及び翻訳結果の提示を行う状況を考える。このようにサイネージ端末を介して人間同士が正対してコミュニケーションをとる場合、以下に示すような点を考慮する必要がある。

[0017] (1) 発話開始時

発話者の口と、サイネージ端末に設けられるマイクとの距離が遠いと、サ

イネージ端末は音声認識を失敗する可能性が高くなる。また、サイネージ端末は、音声認識処理を常時起動させておくと、音声認識の対象となる区間を特定しにくくなるため、音声認識を失敗する可能性が高くなる。さらに、ディスプレイサイズが大きいサイネージ端末において、マイクの位置を全ての人に向けて最適化することは現実的では無く、誰かによっては、発話時に体の移動が大きくなる位置にマイクを配置せざるを得ない。

[0018] つまり、発話者は口をマイクに近付ける必要があるが、発話者が体をマイクに近付れたり、口をマイクに近付れたりすることで発話者の姿勢が崩れ、対話相手は違和感を覚えやすくなる。また、音声認識を起動させるためのユーザインターフェースをディスプレイに表示させる場合、発話者がユーザインターフェースを探すために視線を移動させたり、ユーザインターフェースを操作したりする必要性が生じる。これにより、正対して会話をしたいのに発話者が対話者の顔とは違う方向を見ることとなり、そもそもこれらの行動は本来会話には不要な行動なので、対話として不自然な状況となる。

[0019] ディ스플레이の全面がマイクになる等の技術を開発することで、発話者が発話時に不自然な場所を見ることに起因する不自然な対話を回避することは出来るかもしれない。しかし、人によって口の位置が異なるために、特定のマイクの感度だけを高めて音声認識の成功率を高めるのは難しく、またマイクが多くなればなるほど周囲の雑音などが拾われやすくなるため、かえって音声認識を失敗する可能性が高くなってしまう。ビームフォーミング技術によって特定の方向のマイク感度だけを上げることも可能ではあるが、この場合でも周囲の雑音などの影響を受けやすくなり、音声認識を失敗する可能性が高くなってしまう。

[0020] (2) 発話者が発話中における発話側の考慮点

サイネージ端末を介して、人間同士が正対して、発話内容を翻訳しながらコミュニケーションする場合、発話における風習や慣習の違い、使用言語の違いなどにより、発話側の発話内容を理解できない聞き手側は、受け取った音声と、その音声の翻訳結果とが一致しているかどうかを理解することが非

常に困難である。発話者が単に独り言を呟いただけかもしれないし、考え込んで間が空いた状態に変な認識や翻訳が行われただけかもしれないが、聞き手側はそのことを完全に把握することは困難である。そして、発話者の発話が止まると聞き手側が話し始める場合もありうるが、その行動が発話者の意図に沿わないことも考えられ、お互いにとって期待されたコミュニケーションとならないおそれがある。

[0021] しかし、サインージ端末が翻訳処理中であることを発話側が聞き手側に伝え続けるのは手間が多い。また人間は不要な言葉を発話しやすいので、発話した言葉が翻訳すべきものなのかをサインージ端末が区別するのも非常ニコン案である。そもそも、サインージ端末が翻訳処理中であることを発話側が聞き手側に伝えられたとしても、そのこと自体はコミュニケーションには不要であり、そもそも会話としては不自然である。

[0022] 発話者は、発話する内容に困りつつも、考えながら話したり、質問したい内容を思い出しながら話したりする状況も多い。従って、発話していない場合でも、聞き手側から会話を遮って欲しくないことが多い。しかし、そこまでの意図を発話者が発話等で自然に聞き手側に伝えることは難しく、特に翻訳が発生する状況であれば尚更である。その結果、発話者の発話を遮って聞き手側が話し始めてしまう可能性がある。

[0023] すなわち、サインージ端末を介した、翻訳を要するコミュニケーションにおいて、余計な言葉を翻訳せず、必要な部分だけを翻訳し、サインージ端末が翻訳処理中であることを聞き手側に簡単に伝えることで、会話を自然な状態に保つことが強く求められる。

[0024] (3) 発話者が発話中における聞き手側の考慮点

サインージ端末を介した、翻訳を要するコミュニケーションにおいて、サインージ端末が次々に発話者の発話内容を翻訳し続け、翻訳結果を多量に提示すると、翻訳結果の受け手は処理ができなくなり、コミュニケーションが破綻する。

[0025] 一方、発話側は自分の発話量を把握し続けることは難しく、発話に夢中に

なると、気が付かない内に情報量が大きく膨れあがってしまう。聞き手側に提示される情報量は翻訳されているため、その翻訳後の情報量の大小を発話側が把握することは困難であるとともに、聞き手側が発話量の多さを感じていることを把握することも困難である。聞き手側としては、発話側に適切な量で発話を止めて欲しいので、聞き手側は、言語の壁によらずにそのことを発話側に伝えることが求められる。

[0026] また、発話側の発話量や、翻訳後の情報量が大きく膨れあがった際に、聞き手側は、発話者に対して発話を止めさせる必要がある。しかし、聞き手側が情報量の多さを感じていることを発話者に通知すると、コミュニケーションが不自然となり、そもそも通知自体をためらいがちになる。さらに発話者がマイクに口を近付けていると正対している状態が崩れるため、受け手側が積極的にマイクに近づきづらい。

[0027] そこで本件開示者は、上述した点に鑑みて、サインージ端末等を介して、人間同士が正対してコミュニケーション、特に、会話及び翻訳結果の提示を行う状況において、自然なコミュニケーションを可能にするための技術について鋭意検討を行った。その結果、本件開示者は、サインージ端末等を介して、人間同士が正対してコミュニケーションを行う際に、発話者や対話者の状況を検出することで自然なコミュニケーションを可能にするための技術を考案するに至った。

[0028] 以上、本開示の実施の形態の概要について説明した。続いて、本開示の実施の形態についてより詳細に説明する。

[0029] [1. 2. 構成例]

図1は、本開示の実施の形態に係る、サインージ端末100を用いたコミュニケーションシステムについて示す説明図である。

[0030] サインージ端末100は、街中、例えば駅、デパート、ショッピングモール、スタジアムなどの人が多く集まる場所に設置される、大型のディスプレイを備えた情報表示端末である。本実施形態に係るサインージ端末100は、正対して人間同士のコミュニケーションを行える機能を有しており、特に

、人間の発話内容を解析し、翻訳して、対話相手に伝える機能を有する。本実施形態では、2人の人間（ユーザu1、u2）が、サイネージ端末100を挟んで対話しているケースを前提として説明する。

[0031] （コミュニケーションシステム1）

図2は、本開示の実施の形態に係る、サイネージ端末100を用いたコミュニケーションシステム1の全体構成例を示す説明図である。以下、図2を用いて、本開示の実施の形態に係るコミュニケーションシステム1の全体構成例について説明する。

[0032] 図2に示したように、本開示の実施の形態に係るコミュニケーションシステム1は、サイネージ端末100と、情報処理装置200と、音声情報処理装置300と、翻訳処理装置400と、を含んで構成される。

[0033] サイネージ端末100は、人間同士のコミュニケーションを手助けする機能を有する情報表示端末である。本実施形態では、サイネージ端末100は、一方の面に、カメラ101、マイク102a～102h、ディスプレイ103を備える。カメラ101は、サイネージ端末100を利用する人間の姿を撮像する。マイク102a～102hは、サイネージ端末100を利用する人間の発話内容を集音する。本実施形態では、マイク102a～102dが一方の辺に設けられ、マイク102e～102hが反対の辺に設けられているが、マイクの数及び配置は係る例に限定されるものではない。以下の説明では、マイク102a～102hと単にマイク102と総称する場合がある。

[0034] なお、マイク102a～102hは、全てが集音可能状態となっている必要はなく、ユーザu1、u2の顔の位置に応じて少なくとも一つだけが集音可能な状態となってもよい。集音可能な状態とするマイク102の選択方法については後に詳述する。

[0035] ディスプレイ103は、種々の情報を表示する表示デバイスであり、液晶ディスプレイ、有機ELディスプレイなどで構成される。ディスプレイ103は透明であっても良く非透明であってもよい。本実施形態では、ディスプ

レイ 103 は、ユーザ u 1、u 2 の発話内容を翻訳したものを表示したり、ユーザ u 1、u 2 の発話状況を相手に伝えるための情報を表示したりする。

[0036] 図 3 は、サインージ端末 100 を上方向から見た場合の例を示す説明図である。このように、サインージ端末 100 は、一方の面に、カメラ 101、マイク 102 a ~ 102 h、ディスプレイ 103 を備え、その面の反対側の面に、カメラ 101'、マイク 102 a' ~ 102 h'、ディスプレイ 103' を備える。

[0037] 情報処理装置 200 は、主に、サインージ端末 100 に表示させる情報の表示に関する処理を実行する装置である。情報処理装置 200 は、サインージ端末 100 との間で何らかの通信路により接続されている。通信路は有線であっても良く、無線であっても良い。またサインージ端末 100 と情報処理装置 200 との間の通信における通信プロトコルは問わない。

[0038] 音声情報処理装置 300 は、主に、サインージ端末 100 のマイク 102 a ~ 102 h で集音された音声に対する処理を実行する装置である。音声情報処理装置 300 は、情報処理装置 200 との間で何らかの通信路により接続されている。通信路は有線であっても良く、無線であっても良い。また情報処理装置 200 と音声情報処理装置 300 との間の通信における通信プロトコルは問わない。

[0039] 翻訳処理装置 400 は、主に、翻訳に関する処理を実行する装置であり、特に、ユーザ u 1、u 2 によって発話され、音声情報処理装置 300 によって認識された内容に対する翻訳処理を実行する装置である。翻訳処理装置 400 は、音声情報処理装置 300 との間で何らかの通信路により接続されている。通信路は有線であっても良く、無線であっても良い。音声情報処理装置 300 と翻訳処理装置 400 との間の通信における通信プロトコルは問わない。

[0040] なお、図 2 に示した例では、サインージ端末 100 が情報処理装置 200、音声情報処理装置 300 および翻訳処理装置 400 とネットワークで接続されている構成を示したが、本開示に係る例に限定されるものではない。情

報処理装置 200、音声情報処理装置 300 および翻訳処理装置 400 に設けられる機能の少なくとも一部がサイネージ端末 100 に設けられていてもよく、逆にサイネージ端末 100 に設けられる機能の一部が情報処理装置 200、音声情報処理装置 300 または翻訳処理装置 400 に設けられても良い。また図 2 には、情報処理装置 200、音声情報処理装置 300 および翻訳処理装置 400 の 3 つの装置を示したが、本開示に係る例に限定されるものではなく、いずれかの装置に設けられる機能の少なくとも一部が別の装置に設けられても良い。

[0041] 以上、図 2 を用いて、本開示の実施の形態に係るコミュニケーションシステム 1 の全体構成例について説明した。続いて、本開示の実施の形態に係るサイネージ端末 100 の機能構成例について説明する。

[0042] (サイネージ端末 100)

図 4 は、本開示の実施の形態に係るサイネージ端末 100 の機能構成例を示す説明図である。以下、図 4 を用いて本開示の実施の形態に係るサイネージ端末 100 の機能構成例について説明する。

[0043] 図 4 に示したように、本開示の実施の形態に係るサイネージ端末 100 は、カメラ 101、マイク 102、ディスプレイ 103、通信部 104、及び制御部 105 を含んで構成される。

[0044] カメラ 101 は、上述したように、サイネージ端末 100 を利用する人間の姿を動画像で撮像する。カメラ 101 が撮像する動画像は制御部 105 に送られたり、また通信部 104 を介して情報処理装置 200 に送られたりする。

[0045] マイク 102 は、上述したように、サイネージ端末 100 を利用する人間が発話する声を集音する。マイク 102 が集音した音は、制御部 105 に送られたり、また通信部 104 を介して音声情報処理装置 300 に送られたりする。

[0046] ディスプレイ 103 は、上述したように、種々の情報を表示する表示デバイスであり、液晶ディスプレイ、有機 EL ディスプレイなどで構成される。

ディスプレイ103は透明であっても良く非透明であってもよい。本実施形態では、ディスプレイ103は、ユーザu1、u2の発話内容を翻訳したものを表示したり、ユーザu1、u2の発話状況を相手に伝えるための情報を表示したりする。ディスプレイ103への情報の表示は、例えば後述の制御部105や、情報処理装置200によって制御される。

[0047] 通信部104は、他の装置との間の情報を授受するための通信インターフェースである。通信部104は、他の装置との間で有線通信を行っても良く、無線通信であっても良い。また通信部104は、他の装置との間の通信における通信プロトコルに任意のものを適用することができる。

[0048] 制御部105は、例えばCPU (Central Processing Unit) 等のプロセッサや、ROM、RAMなどで構成され、サインエージ端末100の各部の動作を制御する。

[0049] 本実施形態では、制御部105は、表示制御部106と、集音制御部107と、検出部108と、を含んで構成される。

[0050] 表示制御部106は、ディスプレイ103への情報の表示を制御する。表示制御部106によってディスプレイ103へ表示される情報の例は後に詳述する。

[0051] 集音制御部107は、複数のマイク102に対して、集音機構のオン、オフを切り替える制御を行う。集音制御部107は、後述する高感度マイク位置取得フローによって高感度マイク位置として決定されたマイク102のみ、集音機構をオンさせる処理を行う。集音制御部107は、例えば、カメラ101で撮像された人間の顔の位置に基づいて、集音機構をオン、またはオフさせるマイク102を選択する。集音制御部107による、マイク102に対する集音機構のオン、オフの切り替えの具体例は後に詳述する。

[0052] 検出部108は、カメラ101が撮像した画像を用いて様々な検出処理を行う。本実施形態では、検出部108が行う検出処理として、人物の検出、顔の検出、顔のパーツの検出、耳の位置の推定、動き量の検出、集音制御部107が集音機構をオンさせたマイク102への口の近接の判断、ユーザの

視線検出、ユーザの相槌やうなずきの検出、ユーザの掌の検出、等がある。検出部108によって検出される情報は、後述の様々な処理、例えば、初期処理フロー、耳位置推定フロー、高感度マイク位置取得フロー、翻訳モード処理フロー、球オブジェクトアニメーション処理フロー、翻訳モード継続判断処理フロー、処理量あふれ判断処理フローなどに用いられうる。

[0053] 以上、図4を用いて本開示の実施の形態に係るサイネージ端末100の機能構成例について説明した。続いて、本開示の実施の形態に係る情報処理装置200の機能構成例について説明する。

[0054] (情報処理装置200)

図5は、本開示の実施の形態に係る情報処理装置200の機能構成例を示す説明図である。以下、図5を用いて本開示の実施の形態に係る情報処理装置200の機能構成例について説明する。

[0055] 図5に示したように、本開示の実施の形態に係る情報処理装置200は、通信部210と、制御部220と、を含んで構成される。

[0056] 通信部210は、他の装置との間の情報を授受するための通信インターフェースである。通信部210は、他の装置との間で有線通信を行っても良く、無線通信であっても良い。また通信部210は、他の装置との間の通信における通信プロトコルに任意のものを適用することができる。

[0057] 制御部220は、例えばCPU等のプロセッサや、ROM、RAMなどで構成され、情報処理装置200の各部の動作を制御する。制御部220は、後述するコミュニケーションシステム1の処理フロー、例えば、初期処理フロー、耳位置推定フロー、高感度マイク位置取得フロー、翻訳モード処理フロー、球オブジェクトアニメーション処理フロー、翻訳モード継続判断処理フロー、処理量あふれ判断処理フローなどを実行する。

[0058] 本実施形態では、制御部220は、表示制御部221と、画像生成部222と、判断部223と、を含んで構成される。

[0059] 表示制御部221は、サイネージ端末100のディスプレイ103への情報の表示を制御する。表示制御部221は、例えば、サイネージ端末100

を使用している人物が写っている画像に対する画像処理を行う。表示制御部 221 によって行われうる画像処理の内容については後に詳述する。

[0060] 画像生成部 222 は、例えば表示制御部 221 によって行われうる画像処理の結果に応じて、ディスプレイ 103 へ表示する画像を生成する。画像生成部 222 によって行われうる画像生成処理の内容については後に詳述する。

[0061] 判断部 223 は、コミュニケーションシステム 1 において行われるコミュニケーションに関する処理、具体的には、後述するコミュニケーションシステム 1 の処理フローにおける様々な判断処理を実行する。判断部 223 によって行われうる判断処理の内容については後に詳述するが、一例を挙げれば、例えば、発話者が発話を一時的に止めた場合に、翻訳処理を継続すべきか否かの判断や、発話内容または翻訳結果を受け取っている受け手が発話量の多さを感じているか否かの判断などを行いうる。

[0062] 判断部 223 は、例えば、発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、発話者の発話内容及び発話状況に基づいて、翻訳処理の継続判断を行う。判断部 223 が翻訳処理の継続判断を行った結果、翻訳処理を継続すべきであると判断した場合は、情報処理装置 200 は、翻訳処理装置 400 に翻訳処理を継続させ、翻訳処理を終了すべきであると判断した場合は、翻訳処理装置 400 に翻訳処理を終了させる。

[0063] 以上、図 5 を用いて本開示の実施の形態に係る情報処理装置 200 の機能構成例について説明した。続いて、本開示の実施の形態に係る音声情報処理装置 300 の機能構成例について説明する。

[0064] (音声情報処理装置 300)

図 6 は、本開示の実施の形態に係る音声情報処理装置 300 の機能構成例を示す説明図である。以下、図 6 を用いて本開示の実施の形態に係る音声情報処理装置 300 の機能構成例について説明する。

[0065] 図 6 に示したように、本開示の実施の形態に係る音声情報処理装置 300

は、通信部 310 と。制御部 320 と、を含んで構成される。

[0066] 通信部 310 は、他の装置との間の情報を授受するための通信インターフェースである。通信部 310 は、他の装置との間で有線通信を行っても良く、無線通信であっても良い。また通信部 310 は、他の装置との間の通信における通信プロトコルに任意のものを適用することができる。

[0067] 制御部 320 は、例えば CPU 等のプロセッサや、ROM、RAM など構成され、音声情報処理装置 300 の各部の動作を制御する。

[0068] 本実施形態では、制御部 320 は、検出部 321 と、解析部 322 と、生成部 323 と、を含んで構成される。

[0069] 検出部 321 は、サインエッジ端末 100 に向かって発せられた人間の声を検出する。検出部 321 が検出した声は解析部 322 において解析が行われる。検出部 321 は、人間の声を検出する際に、発話の内容を検出する他、言い淀みの検出、所定の言葉（フィラーワード）の検出、発話のトーンの検出、無音区間の検出などを行いうる。

[0070] 解析部 322 は、検出部 321 によって検出された人間の声に対して様々な解析処理を実行する。解析部 322 が実行する解析処理としては、例えば、音声情報の解析、言語の解析、発話内容に含まれる形態素や句の解析などが行われうる。

[0071] 生成部 323 は、解析部 322 による解析結果に基づいて、サインエッジ端末 100 に向かって発せられた内容をテキストにしたものを生成する。生成部 323 が生成したテキストは、翻訳処理装置 400 での翻訳処理に用いられる。

[0072] 以上、図 6 を用いて本開示の実施の形態に係る音声情報処理装置 300 の機能構成例について説明した。続いて、本開示の実施の形態に係る翻訳処理装置 400 の機能構成例について説明する。

[0073] （翻訳処理装置 400）

図 7 は、本開示の実施の形態に係る翻訳処理装置 400 の機能構成例を示す説明図である。以下、図 7 を用いて本開示の実施の形態に係る翻訳処理装

置 4 0 0 の機能構成例について説明する。

- [0074] 図 7 に示したように、本開示の実施の形態に係る翻訳処理装置 4 0 0 は、通信部 4 1 0 と。制御部 4 2 0 と、を含んで構成される。
- [0075] 通信部 4 1 0 は、他の装置との間の情報を授受するための通信インターフェースである。通信部 4 1 0 は、他の装置との間で有線通信を行っても良く、無線通信であっても良い。また通信部 4 1 0 は、他の装置との間の通信における通信プロトコルに任意のものを適用することができる。
- [0076] 制御部 4 2 0 は、例えば CPU 等のプロセッサや、ROM、RAM など構成され、翻訳処理装置 4 0 0 の各部の動作を制御する。
- [0077] 本実施形態では、制御部 4 2 0 は、翻訳処理部 4 2 1 と、スコア生成部 4 2 2 と、を含んで構成される。
- [0078] 翻訳処理部 4 2 1 は、音声情報処理装置 3 0 0 がサインエージ端末 1 0 0 に向かって発せられた人間の声から生成したテキスト、またはサインエージ端末 1 0 0 に向かって発せられた人間の声を直接用いた、別の言語への翻訳処理を行う。
- [0079] スコア生成部 4 2 2 は、翻訳処理部 4 2 1 における翻訳処理の際のスコアを生成する。スコア生成部 4 2 2 によるスコアの生成処理については後に詳述する。
- [0080] 以上、図 7 を用いて本開示の実施の形態に係る翻訳処理装置 4 0 0 の機能構成例について説明した。続いて、本開示の実施の形態に係るコミュニケーションシステム 1 の動作例を説明する。
- [0081] [ 1 . 3 . 動作例 ]  
( ユースケース例 )
- 本開示の実施の形態に係るコミュニケーションシステム 1 の動作例を説明するにあたり、以下で説明するようなユースケースを用いる。以下の説明では、ユーザ u 1 の母国語は英語、ユーザ u 2 の母国語は英語以外の言語、例えば日本語であるとする。
- [0082] 図 8 は、本開示の実施の形態に係るコミュニケーションシステム 1 の動作

例を説明する説明図である。例えば、外国から来たユーザu1が、目的地への行き方や、食べ物を買える場所などが分からずに困っている状態を考える。ユーザu1は、街中を彷徨っていると、翻訳をしながらコミュニケーションが可能である旨のメッセージが書かれたサイネージ端末100を見つけたので、誰かに教えて貰うことを期待してその前に立つ。このとき、ディスプレイ103'には、カメラ101が撮像するユーザu1の姿が映し出されている。

[0083] 図9は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を説明する説明図である。困った顔をしてサイネージ端末100の前に立っているユーザu1を見つけたユーザu2は、そのユーザu1を助けたいと思っているが、英語が苦手である。しかし、ユーザu2は、翻訳をしながらコミュニケーションが可能なサイネージ端末100を介してなら何とかなるかもしれないと思い、意を決してサイネージ端末100の前に立つ。

[0084] すると、ユーザu1は、サイネージ端末100の反対側にユーザu2が現れたので、サイネージ端末100に向かって話しかける。サイネージ端末100は、ユーザu1の発話内容を取得して、バックにあるサーバでの翻訳処理の内容を、ユーザu2の側にあるディスプレイに出力する。ユーザu2は、サイネージ端末100が出力する翻訳処理の内容を確認することで、ユーザu1の発話内容を知ることが出来る。

[0085] (初期処理)

まず初期処理について説明する。コミュニケーションシステム1は、ユーザがサイネージ端末100の前に立つと、初期処理として、会話モードへの移行を行う。コミュニケーションシステム1は、例えば、カメラ101が撮像した画像から人を検出し、顔がカメラ101に正対していて顔検出が可能な状態となり、さらにその人が所定時間以上その場にとどまったことを条件に会話モードへと移行する。

[0086] コミュニケーションシステム1は、会話モードへ移行すると、サイネージ端末100に正対しているユーザの画像情報から種々の処理を初期処理とし

て実行する。コミュニケーションシステム1は、初期処理として、例えばユーザの顔をシミュレートした画像の生成、ユーザの頭、顔、体の輪郭情報を用いたユーザインターフェースの生成、ユーザの目、口、鼻の位置の情報の取得及び口並びの鼻の位置の情報に基づく耳の位置の計算などを行う。これらの初期処理は、例えば情報処理装置200が実行しうる。

[0087] また、コミュニケーションシステム1は、初期処理として、マイク位置の情報をユーザインターフェースに反映させる処理を行う。具体的には、コミュニケーションシステム1は、サイネージ端末100に備えられるマイク102の位置の情報と、取得したユーザの口の位置の情報とから、ユーザに口を近付けて発話してもらうマイク102の位置を示すためのユーザインターフェースを生成する。

[0088] 図10は、サイネージ端末100のディスプレイ103に表示されるユーザインターフェースの例を示す説明図である。図10に示したのは、ユーザu1がサイネージ端末100の前に立った場合におけるユーザインターフェースの例である。コミュニケーションシステム1は、サイネージ端末100に正対しているユーザu1の画像情報から、ユーザu1の正対身体輪郭U111を生成し、ディスプレイ103に表示させる。また、コミュニケーションシステム1は、マイク102の位置の情報と、取得したユーザの口の位置の情報とから、ユーザに口を近付けて発話してもらうマイク102の位置（高感度マイク位置と称する）を決定し、その高感度マイク位置を示すマイクU112を生成し、ディスプレイ103に表示させる。

[0089] コミュニケーションシステム1は、正対身体輪郭U111における口に該当する位置から遠くなる位置にあるマイク102の位置を、ユーザに口を近付けて発話してもらう高感度マイク位置として決定しても良い。これは、発話時にはなるべく正対身体輪郭U111から外れて発話することをユーザに促すためである。

[0090] コミュニケーションシステム1は、マイクU112をディスプレイ103に表示させる際に、正対身体輪郭U111における口に該当する位置か

ら、ユーザに口を近付けて発話してもらう高感度マイク位置へ移動させるよう表示しても良い。またコミュニケーションシステム1は、マイクU1112をディスプレイ103に表示させる際に、点滅して表示させたり、上下に動くよう表示させたりしてもよい。

[0091] 図11は、ユーザu1から見てサインージ端末100の反対側にユーザu2が立った状態の例を示す説明図である。そして図12は、図10に示したユーザインターフェースがディスプレイ103に表示されている場合において、ユーザu1から見てサインージ端末100の反対側にユーザu2が立った際の、ディスプレイ103の表示例を示す説明図である。このようにディスプレイ103には、反対側に設けられたカメラ101'が撮像した画像が表示される。

[0092] 図13は、サインージ端末100のディスプレイ103'に表示されるユーザインターフェースの例を示す説明図である。図13に示したのは、ユーザu2がサインージ端末100の前に立った場合におけるユーザインターフェースの例である。コミュニケーションシステム1は、サインージ端末100に正対しているユーザu2の画像情報から、ユーザu2の正対身体輪廓U1111'を生成し、ディスプレイ103'に表示させる。また、コミュニケーションシステム1は、マイク102'の位置の情報と、取得したユーザの口の位置の情報とから、ユーザに口を近付けて発話してもらうマイク102'の位置（高感度マイク位置）を示すマイクU1112'を生成し、ディスプレイ103'に表示させる。

[0093] このように、サインージ端末100の前にユーザu1、u2が立つと、コミュニケーションシステム1は初期処理を実行する。初期処理が完了すると、続いてコミュニケーションシステム1は、翻訳モードへの移行、及び音声認識の起動処理を行う。

[0094] （翻訳モードへの移行、音声認識の起動）

コミュニケーションシステム1は、サインージ端末100の前に立つユーザの口の位置と、高感度マイク位置との距離が所定の閾値以下となると、翻

訳モードへ移行させる。図14は、ユーザu1が顔を高感度マイク位置に近付ける様子を示す説明図である。このようにユーザu1が顔を高感度マイク位置に近付けて、ユーザの口の位置と、高感度マイク位置との距離が所定の閾値以下となることが検出されると、コミュニケーションシステム1は、翻訳モードへ移行させる。

[0095] そして、コミュニケーションシステム1は、翻訳モードへと移行させると、音声認識を起動させる。コミュニケーションシステム1は、音声認識を起動させる際には、初期処理で生成していた、ユーザu1の顔をシミュレートした画像を、相手側のディスプレイ103'における、ユーザu1の正対身体輪郭U1111の顔の位置に表示させる処理を実行する。

[0096] 図15は、コミュニケーションシステム1が音声認識を起動させた際に、ディスプレイ103'に表示されるユーザインターフェースの例を示す説明図である。ユーザu1が顔を高感度マイク位置に近付けて、体の位置が正対身体輪郭から外れると、コミュニケーションシステム1は、ユーザu1の顔の位置に、ユーザu1の顔をシミュレートした画像114を表示させる。

[0097] このように、ユーザu1の顔の位置に、ユーザu1の顔をシミュレートした画像114を表示させることで、コミュニケーションシステム1は、ユーザu1が顔を高感度マイク位置に近付けている場合であっても、ユーザu1とユーザu2とが、サインエージ端末100を挟んで向かい合って対話しているように見せることが出来る。

[0098] コミュニケーションシステム1は、ユーザu1の顔をシミュレートした画像114を表示させる際、ユーザu2の方に見えるユーザu1の姿の一部、または全部を画像114に置き換えても良い。ユーザu1の顔をシミュレートした画像114は、顔全体であってもよく、口元や目元だけなど顔の一部であっても良い。またコミュニケーションシステム1は、ユーザu1の顔をシミュレートした画像114を表示させる際、その画像114に対して、話しているように口の部分を動かしたり、ランダムな間隔で瞬きをさせたりする画像処理を行っても良い。

- [0099] なお、コミュニケーションシステム1は、ユーザu1がサインージ端末100に正対した時点で、ユーザu2の方に見えているユーザu1の姿の一部、または全部を画像114に置き換えても良い。
- [0100] またコミュニケーションシステム1は、音声認識を起動させる際には、高感度マイク位置以外の位置にあるマイク102をオフさせる。またコミュニケーションシステム1は、音声認識を起動させると、音声認識を起動させている間、ユーザu1の視線情報を取得し続ける。またコミュニケーションシステム1は、音声認識を起動させると、音声認識を起動させている間、ユーザu1の発話の周波数情報を取得し続ける。
- [0101] またコミュニケーションシステム1は、音声認識を起動させる際には、音声認識を終了させるためのユーザインターフェースをディスプレイ103に表示させる。コミュニケーションシステム1は、音声認識を終了させるためのユーザインターフェースとして、例えば、正対身体輪郭U1111を点滅させたり、現在の身体輪郭の情報を身体輪郭ユーザインターフェースとして表示させたりしてもよい。
- [0102] (音声入力、翻訳処理)
- コミュニケーションシステム1は、ユーザu1が高感度マイク位置のマイク102に向かって発話を行うと、その発話の内容が相手のユーザu2に届いていることを示すユーザインターフェースを表示させても良い。
- [0103] 図16は、ユーザu1の側のディスプレイ103に表示されるユーザインターフェースの例を示す説明図である。ユーザu1が高感度マイク位置のマイク102に向かって発話を行うと、コミュニケーションシステム1は、その発話の認識結果を表示する球U1115をディスプレイ103に表示させる。この球U1115は、ユーザu1の口の高さから、ディスプレイ103に表示されているユーザu2の耳に向かって移動するように表示されるユーザインターフェースである。
- [0104] このような演出を行うことで、コミュニケーションシステム1は、ユーザu1の発話の内容が相手のユーザu2に届いていることを示すことができる

- 。
- [0105] またコミュニケーションシステム1は、ユーザu1が高感度マイク位置のマイク102に向かって発話を行うと、その発話の音声認識結果をディスプレイ103に表示させてもよい。図16には、ユーザu1の発話の音声認識結果116がディスプレイ103に表示されている例が示されている。図16に示した例では、ユーザu1が英語で発話していることを音声情報処理装置300が認識し、ユーザu1の発話内容が英語としてディスプレイ103にされている。ユーザがどの言語で発話したかどうかは、音声情報処理装置300が自動的に認識しても良いし、音声情報処理装置300が発話内容を解析することによって判定してもよい。また、ユーザにサインエージ端末100を操作させて、どの言語で発話するかをサインエージ端末100に対して指定させてもよい。
- [0106] コミュニケーションシステム1は、ユーザu1の発話の音声認識結果を用いて翻訳処理を実行する。コミュニケーションシステム1は、翻訳処理の実行の際には、翻訳処理のスコアを取得し続ける。スコアの取得方法は後に詳述する。またコミュニケーションシステム1は、翻訳処理の実行の際には、音声認識結果の全文を翻訳してもよく、音声認識結果を要約した上でその要約を翻訳しても良い。コミュニケーションシステム1は、音声認識結果を要約する際には、例えば、音声認識結果に対して形態素解析を行い、要点と思われる品詞を抽出してもよい。
- [0107] 一方、コミュニケーションシステム1は、ユーザu1が発話中であることをユーザu2に対して示すユーザインターフェースを表示させても良い。
- [0108] 図17は、ユーザu2の側のディスプレイ103'に表示されるユーザインターフェースの例を示す説明図である。ユーザu1が高感度マイク位置のマイク102に向かって発話を行うと、コミュニケーションシステム1は、ユーザu1の顔をシミュレートした画像114の口の位置から、発話の認識結果を表示する球U117をディスプレイ103'に表示させる。
- [0109] このような演出を行うことで、コミュニケーションシステム1は、ユーザ

u 1 が発話中であることを相手のユーザ u 2 に示すことができる。

[0110] またコミュニケーションシステム 1 は、ユーザ u 1 が高感度マイク位置のマイク 102 に向かって発話を行うと、その発話の音声認識結果に対する翻訳結果をディスプレイ 103' に表示させてもよい。図 17 には、ユーザ u 1 の発話の音声認識結果に対する翻訳結果 118 がディスプレイ 103' に表示されている例が示されている。

[0111] なお、図 17 には、翻訳結果 118 として「私の国の電車は定刻通りに到着しない」と示されているが、上述した要約の翻訳を行うと、例えば翻訳結果 118 として「電車 定刻 到着しない」のように表示されうる。

[0112] コミュニケーションシステム 1 は、翻訳結果 118 をサイネージ端末 100 に表示する際に、聞き手の正対状態が崩れないような位置に表示させることが望ましい。図 18 に示した例では、翻訳結果 118 が、発話者の画像 114 の近くに表示されている。翻訳結果 118 が、発話者の画像 114 の近くに表示されていることで、聞き手であるユーザ u 2 は、正対状態を崩さずに翻訳結果を確認することが出来る。

[0113] コミュニケーションシステム 1 は、音声認識結果を時間軸上に並べて表示させてもよい。図 18 は、ユーザ u 1 の側のディスプレイ 103 に表示されるユーザインターフェースの例を示す説明図である。コミュニケーションシステム 1 は、ユーザ u 1 の音声認識結果 116 を、図 18 に示すように時間軸上に並べて表示させてもよい。ユーザ u 1 の音声認識結果を時系列で表示させることで、コミュニケーションシステム 1 は、発話者に対して過去の発話内容を確認させることができる。

[0114] またコミュニケーションシステム 1 は、ユーザ u 1 の音声認識結果 116 を時間軸上に並べて表示させる際に、図 18 に示すように発話者の写真を併せて表示させても良い。コミュニケーションシステム 1 は、音声認識結果 116 に発話者の写真を併せて表示させることで、過去に誰がその内容を話したのかを発話者に確認させることができる。

[0115] (翻訳モードの継続判断 (1))

コミュニケーションシステム1は、ユーザu1の発話が途絶えたことを検出した場合に、例えば、ユーザu1の過去の発話内容及びユーザu1の発話状況を用いて、翻訳モードを継続すべきかどうかの判断を行う。まず、翻訳モードを継続すべきかどうかの最初の判断例を示す。

[0116] コミュニケーションシステム1は、音声入力が無くなった、すなわち、ユーザu1の発話が途絶えた場合でも、ユーザu1の発話内容に対する翻訳スコアが所定値より低く、かつ、発話者であるユーザu1の視線がユーザu2の方を向いていないときは、翻訳モードを継続する。これは、伝える情報が分からなくなると、人と目を合わせない、または視線が泳ぐようになるという人間の心理に基づくものである。

[0117] 図19は、コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。図19には、ユーザu1の側のディスプレイ103に、ユーザu2の姿が表示されている状態が例示されている。

[0118] 図19には、ユーザu2の正対身体輪郭領域に、所定の閾値 $d\_surrounding\_thresh$ を加えた領域119が示されている。この領域119は、ディスプレイ103に表示されていなくても良い。コミュニケーションシステム1は、ユーザu1の発話内容に対する翻訳スコアが所定値より低く、かつ、ユーザu1の視線が、領域119の外側に、所定の時間 $t\_surrounding\_thresh$ 以上存在している場合は、翻訳モードを継続すると判断する。

[0119] 一方、コミュニケーションシステム1は、ユーザu1の発話内容に対する翻訳スコアが所定値より低く、かつ、ユーザu1の視線が、領域119の内側に、所定の時間 $t\_surrounding\_thresh$ 以上存在している場合は、翻訳モードを解除すると判断する。

[0120] なお、図19に示した例では、ユーザu2の正対身体輪郭領域の全体に、一律に所定の閾値 $d\_surrounding\_thresh$ を加えた例が示されているが、この所定の閾値は、ユーザu2の体の部位に応じて変化させても良い。コミュニケーションシステム1は、例えば、首より上と首より下とで、この所定の閾

値を変化させても良い。

- [0121] またコミュニケーションシステム1は、ユーザu1の視線の先に応じて、翻訳モードを継続すべきかどうかの判断を変化させても良い。すなわち、コミュニケーションシステム1は、ユーザu1が、ユーザu2の顔を見ているか、体を見ているか、手を見ているか、等に応じて、翻訳モードを継続すべきかどうかの判断を変化させても良い。
- [0122] またコミュニケーションシステム1は、ユーザu2の背景の動きの有無に応じて、翻訳モードを継続すべきかどうかの判断を変化させても良い。これによりコミュニケーションシステム1は、ユーザu1が話す内容が分からなくなっていて目が泳いでいるのか、ユーザu2の背景に目を取られたのか、に応じて、翻訳モードを継続すべきかどうかの判断を変化できる。
- [0123] コミュニケーションシステム1は、コミュニケーションシステム1を利用するユーザが過去にもコミュニケーションシステム1を利用したことがあれば、そのユーザの癖を用いて翻訳モードを継続すべきかどうか判断しても良い。例えば、視線を動かしやすいユーザであれば、ユーザが対話相手から視線を逸らしても翻訳モードを継続すべきであると判断しても良い。
- [0124] コミュニケーションシステム1は、ユーザ同士がコミュニケーションを行っているシチュエーションによって、翻訳モードを継続すべきかどうかの判断に用いる閾値を変化させても良い。例えば、コミュニケーションシステム1は、コミュニケーションを行っているシチュエーションが、道案内であるのか、会議であるのか、等に応じて、翻訳モードを継続すべきかどうかの判断に用いる閾値を変化させても良い。また例えば、コミュニケーションシステム1は、コミュニケーションを行っているシチュエーションとして、ユーザが急いでいる状況なのか、時間に余裕がある状況なのか、等に応じて、翻訳モードを継続すべきかどうかの判断に用いる閾値を変化させても良い。
- [0125] コミュニケーションシステム1は、コミュニケーションを行っている相手との関係によって、翻訳モードを継続すべきかどうかの判断に用いる閾値を変化させても良い。例えば、コミュニケーションシステム1は、コミュニケ

ーションを行っている相手が、初対面の人物であるのか、異性であるのか、取引先の高い役職の人であるのか、等に応じて、翻訳モードを継続すべきかどうかの判断に用いる閾値を変化させても良い。

[0126] (翻訳モードの継続判断(2))

次に、翻訳モードを継続すべきかどうかの2つ目の判断例を示す。

[0127] コミュニケーションシステム1は、音声入力が無くなった、すなわち、ユーザu1の発話が途絶えた場合でも、ユーザu1が発話した文章が完結していなければ、例えば、ユーザu1の発話内容に言い淀みが検出されれば翻訳モードを継続する。またコミュニケーションシステム1は、無音になる前に所定のフィラーワードが含まれており、そのフィラーワードが発せられてから所定の時間t\_fillerwords\_threshが経過するまでは、翻訳モードを継続する。

[0128] 図20は、コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。図20には、コミュニケーションシステム1が翻訳モードを継続すべきかどうかの3つの判断例を示している。1つ目の例は、ユーザu1の発話が途絶えてから所定の時間が経過すると、コミュニケーションシステム1が翻訳モードを解除する例である。2つ目の例は、ユーザu1の発話が途絶えても、言い淀みを検出するとコミュニケーションシステム1が翻訳モードを継続する例である。3つ目の例は、ユーザu1が所定のフィラーワードを発したことを検出すると、コミュニケーションシステム1は、そのフィラーワードが発せられてから所定の時間t\_fillerwords\_threshが経過するまでは、翻訳モードを継続する例である。

[0129] (翻訳モードの継続判断(3))

次に、翻訳モードを継続すべきかどうかの3つ目の判断例を示す。

[0130] コミュニケーションシステム1は、音声入力が無くなった、すなわち、ユーザu1の発話が途絶えた場合でも、ユーザの声質に基づいて翻訳モードを継続するかどうか判断する。例えば、コミュニケーションシステム1は、音

声入力が無くなった、すなわち、ユーザ u 1 の発話が途絶えた場合でも、最後の発話の周波数情報が、それまで取得し続けた周波数の平均と比べてある閾値以下となっていた場合は、翻訳モードを継続する。

- [0131] 具体的には、コミュニケーションシステム 1 は、音声入力が無くなった、すなわち、ユーザ u 1 の発話が途絶えた際に、最後の発話の周波数情報 `info_last_voice_freq` を算出し、それまで行われた発話すべての平均の周波数情報 `f_all_voice` を算出する。そして、コミュニケーションシステム 1 は、`info_last_voice_freq` < `info_voice_freq_ave - f_thresh` を満たす場合は声のトーンが下がった、つまり、発話者が自信を無くしている状態であると判断し、翻訳モードを継続する。一方、コミュニケーションシステム 1 は、`info_last_voice_freq` >= `info_voice_freq_ave - f_thresh` を満たす場合は、翻訳モードを解除する。

- [0132] (翻訳モードの継続判断 (4))

次に、翻訳モードを継続すべきかどうかの 4 つ目の判断例を示す。

- [0133] コミュニケーションシステム 1 は、音声入力が無くなる前に音声認識を行った結果、所定のデータベース（ここでは「会話終了判断ワードデータベース」とする）に含まれる者と一致していれば、ユーザからの音声入力がないまま所定の時間 `t_convendwords_thresh` が経過するまでは翻訳モードを継続し、所定の時間 `t_convendwords_thresh` が経過すると翻訳モードを終了する。

- [0134] 図 2 1 は、コミュニケーションシステム 1 が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。図 2 1 には、ユーザが、発話を停止する直前に、会話終了判断ワードデータベースに登録されている” What do you think of that?” という語句を発話した場合の例が示されている。この語句の発話を検出したコミュニケーションシステム 1 は、ユーザからの音声入力がないまま所定の時間 `t_convendwords_thresh` が経過するまでは翻訳モードを継続する。そしてコミュニケーションシステム 1 は、ユーザからの音声入力がないまま所定の時間 `t_convendwords_thresh` が経過すると翻訳モードを終了する。

[0135] (翻訳モードの継続判断(5))

次に、翻訳モードを継続すべきかどうかの5つ目の判断例を示す。

[0136] コミュニケーションシステム1は、ユーザが意図的に翻訳モードを解除するための動作を行った場合は、翻訳モードを解除する。意図的に翻訳モードを解除するための動作には、例えばディスプレイ103に表示された終了ボタンを押す、体を正対状態に戻す、発話を終えて視線を対話相手に戻す、等があり得るが、これらの例に限定されるものではない。

[0137] 図22は、コミュニケーションシステム1がディスプレイ103に表示させるユーザインターフェースの例を示す説明図である。図22には、ディスプレイ103に、「STOP」と表示された、翻訳モードを解除するための終了ボタン120が表示されている例が示されている。コミュニケーションシステム1は、ユーザが終了ボタン120をタッチしたこと、または終了ボタン120に指を近接させたことを検出すると、翻訳モードを解除する。なお、終了ボタン120に表示される言葉は、ユーザu1が発話している言語に応じて変化させても良い。

[0138] ユーザが終了ボタン120をタッチしたこと、または終了ボタン120に指を近接させたことを検出すると、コミュニケーションシステム1は、例えば、ユーザに対して体を正対状態に戻すよう促すメッセージを、文字でディスプレイ103に表示したり、音声で出力したりしても良い。

[0139] 図23は、コミュニケーションシステム1が翻訳モードを継続すべきかどうかの判断を行うことを説明するための説明図である。図23には、ユーザが正対身体輪郭U1111に対応する位置に姿勢を戻した場合の例が示されている。コミュニケーションシステム1は、ユーザが正対身体輪郭U1111に対応する位置に姿勢を戻し、姿勢を戻してから所定時間経過したことを検出すると、翻訳モードを解除する。

[0140] コミュニケーションシステム1は、ユーザが正対身体輪郭U1111に対応する位置に完全に姿勢を戻していなくても、例えば体の5割以上が正対身体輪郭U1111に対応する範囲に含まれていれば、体が正対状態に戻った

と判断しても良い。

[0141] その他、コミュニケーションシステム1は、例えばサイネージ端末100の周囲の騒音や雑音の有無によって翻訳モードを継続の可否を判断しても良い。発話者が発話を停止した場合に、サイネージ端末100の周囲で騒音や雑音が発生していれば、発話者はもしかしてその騒音や雑音に気を取られたのかもしれない。従って、コミュニケーションシステム1は、発話者が発話を停止した場合に、サイネージ端末100の周囲で騒音や雑音が発生していれば、翻訳モードを継続すると判断しても良い。

[0142] その他、コミュニケーションシステム1は、例えば相手からの回答に対する内容に基づいて、翻訳モードを継続の可否を判断しても良い。例えば、相手からの回答が質問であれば、ユーザはその回答に対して発話しようとするので、ユーザがしばらく発話を行わなかったとしても、コミュニケーションシステム1は、翻訳モードを継続すると判断する。一方、相手からの回答が質問でなければ、ユーザがしばらく発話を行わなかったとすると、コミュニケーションシステム1は、翻訳モードを終了すると判断する。

[0143] (聞き手側の処理あふれ判断)

上述の「(3) 発話者が発話中における聞き手側の考慮点」において指摘したように、サイネージ端末を介した、翻訳を要するコミュニケーションにおいて、サイネージ端末が次々に発話者の発話内容を翻訳し続け、翻訳結果を多量に提示すると、翻訳結果の受け手は処理ができなくなり、コミュニケーションが破綻する。

[0144] そこでコミュニケーションシステム1は、聞き手側の状況を検出して、聞き手側の処理あふれの有無を判断する。そして聞き手側が処理あふれを起こしていると判断した場合は、コミュニケーションシステム1は、発話者側に、発話を一時的に止めるよう通知する。

[0145] (聞き手側の処理あふれ判断(1))

まず、聞き手側の処理あふれ判断の最初の例を示す。

[0146] コミュニケーションシステム1は、発話者であるユーザu1が発話中に、

聞き手側であるユーザ u 2 の視線がユーザ u 1 の方を向いておらず、ユーザ u 1 の正対身体輪郭情報から所定の閾値以上の距離の外側にあり、その状態が所定時間以上継続したときは、ユーザ u 2 が処理あふれを起こしていると判断する。

[0147] 図 2 4 は、コミュニケーションシステム 1 が聞き手側の処理あふれの判断を行うことを説明するための説明図である。図 2 4 には、ユーザ u 2 の側のディスプレイ 1 0 3' に、ユーザ u 1 の姿が表示されている状態が例示されている。

[0148] 図 2 4 には、ユーザ u 1 の正対身体輪郭領域に、所定の閾値  $d_{surrounding\_thresh}$  を加えた領域 1 1 9' が示されている。この領域 1 1 9' は、ディスプレイ 1 0 3' に表示されていなくても良い。コミュニケーションシステム 1 は、発話者であるユーザ u 1 が発話中に、ユーザ u 2 の視線が、領域 1 1 9' の外側に、所定の時間  $t_{surrounding\_thresh}$  以上存在している場合は、ユーザ u 2 が処理あふれを起こしていると判断する。一方、発話者であるユーザ u 1 が発話中に、ユーザ u 2 の視線が、領域 1 1 9' の内側に存在していれば、また領域 1 1 9' の外側に視線があっても、所定の時間  $t_{surrounding\_thresh}$  の経過前に領域 1 1 9' の内側に視線が戻れば、ユーザ u 2 が処理あふれを起こしていないと判断する。

[0149] (聞き手側の処理あふれ判断 (2))

次に、聞き手側の処理あふれ判断の 2 つ目の例を示す。

[0150] コミュニケーションシステム 1 は、発話者であるユーザ u 1 が発話中に、聞き手側であるユーザ u 2 が相槌を打たなくなってから所定時間以上経過したときは、ユーザ u 2 が処理あふれを起こしていると判断する。

[0151] 図 2 5 は、コミュニケーションシステム 1 が聞き手側の処理あふれの判断を行うことを説明するための説明図である。ユーザ u 2 がある時点で相槌を打たなくなり、その後、相槌を検出しなくなってから所定時間以上経過したときは、コミュニケーションシステム 1 は、ユーザ u 2 が処理あふれを起こしていると判断する。

[0152] (聞き手側の処理あふれ判断(3))

次に、聞き手側の処理あふれ判断の3つ目の例を示す。

[0153] コミュニケーションシステム1は、発話者であるユーザu1が発話中に、聞き手側であるユーザu2が掌をユーザu1に向け、その状態が所定時間以上継続したときは、ユーザu2が処理あふれを起こしていると判断する。

[0154] 図26は、コミュニケーションシステム1が聞き手側の処理あふれの判断を行うことを説明するための説明図である。ユーザu2がある時点で掌をユーザu1に向け、その後、その状態が所定時間以上継続したときは、コミュニケーションシステム1は、ユーザu2が処理あふれを起こしていると判断する。

[0155] (聞き手側の処理あふれ判断(4))

次に、聞き手側の処理あふれ判断の4つ目の例を示す。

[0156] コミュニケーションシステム1は、発話者であるユーザu1が発話中に、音声認識結果のテキスト、または音声認識結果を翻訳した際のテキストが、所定の条件を満たした場合に、ユーザu2が処理あふれを起こしていると判断する。所定の条件としては、例えば、句解析を行った結果、句の数がある閾値を超えた(分かち書き言語であれば単語の数がある閾値を超えた)、形態素(英語ならば、単語)のうち、主要品詞である名詞及び動詞の数がある閾値を超えた、文字数がある閾値を超えた、などがある。

[0157] 図27は、コミュニケーションシステム1が聞き手側の処理あふれの判断を行うことを説明するための説明図である。ユーザu1が発話中に、音声認識結果のテキスト、または音声認識結果を翻訳した際のテキストが、所定の条件を満たした場合は、コミュニケーションシステム1は、ユーザu2が処理あふれを起こしていると判断する。

[0158] コミュニケーションシステム1は、発話者が発話した内容、または発話内容の翻訳結果に、専門性の高い単語が含まれているなどして、内容が難しい場合は、上記の句の数を少なくしてもよい。また、コミュニケーションシステム1は、発話者が発話した内容が難しい場合は、専門性の高い単語に重み

を付けることで、聞き手側の処理あふれを起こしているかどうかの判断を行っても良い。

[0159] 以上、聞き手側が処理あふれを起こしているかどうかを判断する5つの例を示した。コミュニケーションシステム1は、上述したいずれかの方法により、また上述した方法を複数組み合わせることにより、聞き手側が処理あふれを起こしているかどうかを判断する。なお上述したものは聞き手側が処理あふれを起こしているかどうかを判断するための一例に過ぎないことはいうまでも無い。

[0160] また、聞き手の属性に応じて処理あふれを起こすまでの量は異なる。大人であれば処理あふれを起こすまでの量は多いが、子供であれば処理あふれを起こすまでの量は少ない。従って、コミュニケーションシステム1は、聞き手の属性に応じて処理あふれを起こすと判断するまでの閾値を変化させても良い。

[0161] またコミュニケーションシステム1は、聞き手側の生体情報に基づいて処理あふれを起こしているかどうかを判断してもよい。例えば、心拍数が上昇した、発汗量が増えた、等の情報を、聞き手側が所持しているデバイスに備えられるセンサで取得し、そのデバイスから情報を得ることで、コミュニケーションシステム1は、聞き手側が処理あふれを起こしているかどうかを判断してもよい。

[0162] (処理あふれを起こしたと判断した場合のユーザインターフェース)

コミュニケーションシステム1は、聞き手側が処理あふれを起こしていることを発話者側に伝えるためのユーザインターフェースを発話者側のディスプレイ103に提示する。

[0163] 上述した例では、コミュニケーションシステム1は、発話者が発話していると、球U1115が聞き手側の耳に入っていくユーザインターフェースを示した。このようなユーザインターフェースを表示している場合、聞き手側が処理あふれを起こしていることを判断すると、コミュニケーションシステム1は、球U1115が聞き手側の耳で跳ね返り、こぼれ落ちていくような

ユーザインターフェースを表示させてもよい。

[0164] 図28は、ユーザu1の側のディスプレイ103に表示されるユーザインターフェースの例を示す説明図である。コミュニケーションシステム1は、上述の判断によって聞き手側であるユーザu2が処理あふれを起こしていることを判断すると、球U1115がユーザu2の耳で跳ね返り、こぼれ落ちていくようなユーザインターフェースを表示させる。このようなユーザインターフェースを発話者側のディスプレイ103に表示させることで、コミュニケーションシステム1は、聞き手側が処理あふれを起こしていることを発話者側に伝えることが可能となる。

[0165] ユーザu1の側のディスプレイ103に表示されるユーザインターフェースとしては、この他にも、例えば発話内容の認識結果の表示を止める、球U1115自体の表示を止める、などがあり得る。また、コミュニケーションシステム1は、聞き手側が処理あふれを起こしていることを発話者側に伝えるために、発話者の発話音を打ち消すような音を出力してもよい。

[0166] 図28には、聞き手側が処理あふれを起こしていることを発話者側に伝えるためのユーザインターフェースを発話者側のディスプレイ103に提示する例を示したが、コミュニケーションシステム1は、発話者の発話を遮るためのユーザインターフェースを聞き手側に表示させても良い。例えば、コミュニケーションシステム1は、発話内容の翻訳結果を、あえて聞き手の正対状態が崩れる位置、例えばマイクU1の近くに表示させてもよい。コミュニケーションシステム1は、視線を外させたり、体の向きを変えさせて聞き手の正対状態を敢えて崩させたりすることで、聞き手側が処理あふれを起こしていることを聞き手自身に伝えることが可能となる。

[0167] コミュニケーションシステム1は、処理あふれを起こした条件に応じて、ディスプレイ103に表示させるユーザインターフェースを変化させてもよい。例えば、視線が発話者側を向かなくなったことで処理あふれを起こしたと判断すれば、コミュニケーションシステム1は、聞き手側の姿の周囲に汗が飛び出ているようなCGをディスプレイ103に表示させてもよい。また例

例えば、傾きが無くなったことで処理あふれを起こしたと判断すれば、コミュニケーションシステム1は、聞き手側の頭の上に、聞き手側が困っているような表現を表すCGをディスプレイ103に表示させてもよい。

[0168] 以上、本開示の実施の形態に係るコミュニケーションシステム1の動作例を、一つのユースケースを取り上げて説明した。続いて、本開示の実施の形態に係るコミュニケーションシステム1の動作例をより詳細に説明する。

[0169] (発話者側の全体フロー)

図29は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図29に示したのは、発話者側の全体フローを示したものである。

[0170] 発話者（上述の例では、ユーザu1）がサインージ端末100の前に立つと、コミュニケーションシステム1は初期処理を実行する（ステップS101）。この初期処理は、サインージ端末100で取得された情報に基づいて、情報処理装置200が実行しうる。初期処理の詳細については後に詳述する。

[0171] コミュニケーションシステム1は、上記ステップS101で初期処理を実行すると、続いて、相手側（上述の例では、ユーザu2）の初期フローが終了しているかどうか判断する（ステップS102）。この判断処理は例えば情報処理装置200（例えば、判断部223）が実行しうる。

[0172] 上記ステップS102の判断の結果、相手側の初期フローが終了していれば（ステップS102, Yes）、コミュニケーションシステム1は、翻訳モードの処理を実行する（ステップS103）。一方、上記ステップS102の判断の結果、相手側の初期フローが終了していなければ（ステップS102, No）、コミュニケーションシステム1は、相手側の初期フローが終了するまで待機する。なお、コミュニケーションシステム1は、相手が現れないなどの理由で、所定時間経過しても相手側の初期フローが終了していなければ、初期フローを終了させても良い。

[0173] (初期処理フロー)

続いて、図29のステップS101で示した初期処理の詳細な流れを説明する。図30A、30Bは、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図30A、30Bに示したのは、図29のステップS101で示した初期処理の詳細なフローを示したものである。なお、以下で示す初期処理は、サイネージ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、初期処理は、コミュニケーションシステム1を構成するどの装置で行われても良い。

- [0174] まずコミュニケーションシステム1は、ユーザがサイネージ端末100の前にとどまった時間カウンタ $t\_stay$ をリセット( $t\_stay=0$ )し、会話モードかどうかを示すフラグ $f\_conv\_mode$ をリセット( $f\_conv\_mode=false$ )し、マイク102の位置情報 $p\_mic\_i$ ( $i=0, \dots, n$ )を取得する(ステップS111)。
- [0175] 続いて、コミュニケーションシステム1は、カメラ101が撮像したカメラ画像を用いた顔検出処理によって顔が検出できたかどうか判断する(ステップS112)。
- [0176] 上記ステップS112の判断の結果、顔が検出できない場合は(ステップS112、No)、続いてコミュニケーションシステム1は、上記ステップS111に戻る。一方、上記ステップS112の判断の結果、顔が検出できた場合は(ステップS112、Yes)、続いてコミュニケーションシステム1は、時間カウンタ $t\_stay$ をインクリメントする(ステップS113)。
- [0177] 上記ステップS113で時間カウンタ $t\_stay$ をインクリメントすると、続いてコミュニケーションシステム1は、時間カウンタ $t\_stay$ が閾値 $t\_stay\_thresh$ より大きいかどうか判断する(ステップS114)。
- [0178] 上記ステップS114の判断の結果、時間カウンタ $t\_stay$ が閾値 $t\_stay\_thresh$ 未満の場合は(ステップS114、No)、コミュニケーションシステム1は上記ステップS112の処理に戻る。一方、上記ステップS114の判断の結果、時間カウンタ $t\_stay$ が閾値 $t\_stay\_thresh$ より大きい場合は(ステップS114、Yes)、コミュニケーションシステム1は、会話モードかどうかを示すフラグ $f\_conv\_mode$ をtrueにセットし、(ステップS115)

、正体時のユーザの顔情報info\_faceを取得する（ステップS 1 1 6）。

[0179] フラグf\_conv\_modeをtrueにセットし、正体時のユーザの顔情報info\_faceを取得すると、続いてコミュニケーションシステム1は、正対時の人の輪郭情報info\_outline\_confrontを生成して正対身体輪郭UIとして画面へ反映させる（ステップS 1 1 7）。

[0180] 正対時の人の輪郭情報info\_outline\_confrontを生成すると、続いてコミュニケーションシステム1は、正体時のユーザの顔情報info\_faceや正対時の人の輪郭情報info\_outline\_confrontから、目情報info\_eye、口情報info\_mouth及び鼻情報info\_noseを取得する（ステップS 1 1 8）。

[0181] 目情報info\_eye、口情報info\_mouth及び鼻情報info\_noseを取得すると、続いてコミュニケーションシステム1は、info\_faseから表情を様々なものに変えた顔情報info\_facial\_expを生成する（ステップS 1 1 9）。ここでの表情としては、例えば感情を出していない状態の表情、笑っている表情、不安な表情などがありうるが、これらの表情に限定されるものではない。

[0182] 上記ステップS 1 1 9で顔情報info\_facial\_expを生成すると、続いてコミュニケーションシステム1は、耳位置推定フローを実行する（ステップS 1 2 0）。耳位置推定フローの詳細は後に詳述する。

[0183] 上記ステップS 1 2 0で耳位置推定フローを実行すると、続いてコミュニケーションシステム1は、高感度マイク位置取得フローを実行する（ステップS 1 2 1）。高感度マイク位置取得フローの詳細は後に詳述する。

[0184] 上記ステップS 1 2 1で高感度マイク位置取得フローを実行すると、続いてコミュニケーションシステム1は、口の位置情報p\_mouthから高感度マイク位置p\_h\_sensitive\_micに向けて、マイクUIをアニメーションで移動させる（ステップS 1 2 2）。

[0185] 以上、図29のステップS 1 0 1で示した初期処理の詳細な流れを説明した。続いて、上記ステップS 1 2 0の耳位置推定フローの詳細な流れを説明する。

[0186] （耳位置推定フロー）

図31は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図31に示したのは、図30BのステップS120の耳位置推定フローの詳細な流れである。なお、以下で示す耳位置推定フローは、サインージ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、耳位置推定フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0187] コミュニケーションシステム1は、耳位置推定フローを実行する際に、まず口情報info\_mouthから口の位置情報p\_mouthを取得する(ステップS131)。さらにコミュニケーションシステム1は、目情報info\_eyeから、左目の位置情報p\_eye\_lと、右目の位置情報p\_eye\_rと、を取得する(ステップS132)。さらにコミュニケーションシステム1は、鼻情報info\_noseから鼻の最下部位置情報p\_nose\_dを取得する(ステップS133)。

[0188] 図32は、口の位置情報p\_mouth、左目の位置情報p\_eye\_l、右目の位置情報p\_eye\_r、および鼻の最下部位置情報p\_nose\_dを示す説明図である。コミュニケーションシステム1は、サインージ端末100の前に立つユーザの目情報info\_eye、口情報info\_mouth及び鼻情報info\_noseから、図32に示したように口の位置情報p\_mouth、左目の位置情報p\_eye\_l、右目の位置情報p\_eye\_r、および鼻の最下部位置情報p\_nose\_dを取得する。

[0189] 続いてコミュニケーションシステム1は、右目の位置情報p\_eye\_rから左目の位置情報p\_eye\_lへのベクトルvec\_eye\_rlを2倍したものと、右目の位置情報p\_eye\_rから鼻の最下部位置情報p\_nose\_dへのベクトルvec\_r\_noseを0.5倍したものととの和をとって、左の耳位置p\_ear\_lを取得する(ステップS134)。

[0190] 同様に、コミュニケーションシステム1は、左目の位置情報p\_eye\_lから右目の位置情報p\_eye\_rへのベクトルvec\_eye\_lrを2倍したものと、p\_eye\_lからp\_nose\_dへのベクトルvec\_l\_noseを0.5倍したものととの和をとって、右の耳位置p\_ear\_rを取得する(ステップS135)。

[0191] 図33は、口の位置情報p\_mouth、左目の位置情報p\_eye\_l、右目の位置情

報 $p_{eye\_r}$ 、および鼻の最下部位置情報 $p_{nose\_d}$ から算出される左の耳位置 $p_{ear\_l}$ 及び右の耳位置 $p_{ear\_r}$ を示す説明図である。このように、コミュニケーションシステム1はユーザのおおよその左の耳位置 $p_{ear\_l}$ 及び右の耳位置 $p_{ear\_r}$ を算出する。

[0192] 以上、図30BのステップS120で示した耳位置推定フローの詳細な流れを説明した。続いて、図30BのステップS121の、高感度マイク位置取得フローの詳細な流れを説明する。

[0193] (高感度マイク位置取得フロー)

図34は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図34に示したのは、図30BのステップS121の、高感度マイク位置取得フローの詳細な流れである。なお、以下で示す高感度マイク位置取得フローは、サインエッジ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、高感度マイク位置取得フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0194] コミュニケーションシステム1は、高感度マイク位置取得フローを実行する際に、まず口の位置情報 $p_{mouth}$ の高さより低い位置にあるマイク102の位置 $p_{mic\_lower\_i}$ ( $i=0, \dots, n$ )を取得する(ステップS141)。

[0195] マイク102の位置 $p_{mic\_lower\_i}$ ( $i=0, \dots, n$ )を取得すると、続いてコミュニケーションシステム1は、ユーザの口の位置情報 $p_{mouth}$ と、マイク102の位置 $p_{mic\_lower\_i}$ との距離が最も近い位置のマイク102の位置 $p_{mic\_nearest}$ を取得する(ステップS142)。

[0196] 続いてコミュニケーションシステム1は、ユーザの口の位置情報 $p_{mouth}$ と、マイク102の位置 $p_{mic\_lower\_i}$ との距離が2番目に近い位置のマイク102の位置 $p_{mic\_next}$ を取得する(ステップS143)。

[0197] そしてコミュニケーションシステム1は、ステップS143で取得したマイク102の位置 $p_{mic\_next}$ を、高感度マイク位置 $p_{h\_sensitive\_mic}$ に決定する(ステップS144)。

- [0198] 図35は、高感度マイク位置の決定方法を説明するための説明図である。図35に示したようにユーザの顔及びマイク102が配置されている場合、`p_mic_lower_i`は、`p_mic_2`、`p_mic_3`、`p_mic_6`、`p_mic_7`である。そして、`p_mic_nearest`は`p_mic_2`となり、`p_mic_next`は`p_mic_6`となる。従って、図35に示したようにユーザの顔及びマイク102が配置されている場合、コミュニケーションシステム1は、`p_mic_6`の位置にあるマイクを高感度マイク位置`p_h_sensitive_mic`に決定する。
- [0199] コミュニケーションシステム1は、このようにユーザの顔の位置に基づいて高感度マイク位置を決定することで、マイクの位置やユーザの背の高さなどに依存せず、ユーザの音声入力を成功させることが可能となる。
- [0200] 以上、図30BのステップS121で示した高感度マイク位置取得フローの詳細な流れを説明した。続いて、図29のステップS103の、翻訳モード処理フローの詳細な流れを説明する。
- [0201] (翻訳モード処理フロー)
- 図36A、36Bは、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図36A、36Bに示したのは、図29のステップS103の、翻訳モード処理フローの詳細な流れである。なお、以下で示す翻訳モード処理フローは、サインエージ端末100で取得された情報や、翻訳処理装置400で生成された情報に基づいて、情報処理装置200が実行するものとするが、翻訳モード処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。
- [0202] コミュニケーションシステム1は、翻訳モード処理フローを実行する際には、まず翻訳モードかどうかを示すフラグ`f_trans_mode`をリセット(`f_trans_mode=false`)する(ステップS151)。
- [0203] フラグ`f_trans_mode`をリセットすると、続いてコミュニケーションシステム1は翻訳モード起動処理を実行する(ステップS152)。翻訳モード起動処理については後に詳述する。
- [0204] 翻訳モード起動処理を実行すると、続いてコミュニケーションシステム1

は高感度マイク位置`p_h_sensitive_mic`と口の位置情報`p_mouth`との距離`d_mic_mouth`が、閾値`d_mic_mouth_thresh`よりも小さいかどうかを判断する（ステップS153）。

[0205] 上記ステップS153の判断の結果、高感度マイク位置`p_h_sensitive_mic`と口の位置情報`p_mouth`との距離`d_mic_mouth`が、閾値`d_mic_mouth_thresh`以上であれば（ステップS153、No）、続いてコミュニケーションシステム1は、その時点の身体輪郭情報`info_outline_current`と正対時の人の輪郭情報`info_outline_confront`とから、現在の身体輪郭情報が正対輪郭情報内に閾値`t_stay_confront_thresh`よりも多くとどまったかどうかを判断する（ステップS154）。

[0206] 上記ステップS154の判断の結果、現在の身体輪郭情報が正対輪郭情報内に閾値`t_stay_confront_thresh`よりも多くとどまっていれば（ステップS154、Yes）、コミュニケーションシステム1は翻訳モード終了処理を実行する（ステップS155）。翻訳モード終了処理については後に詳述する。一方、上記ステップS154の判断の結果、現在の身体輪郭情報が正対輪郭情報内に閾値`t_stay_confront_thresh`よりも多くとどまっていなければ（ステップS154、No）、コミュニケーションシステム1は上記ステップS153の判断処理に戻る。

[0207] 一方、上記ステップS153の判断の結果、高感度マイク位置`p_h_sensitive_mic`と口の位置情報`p_mouth`との距離`d_mic_mouth`が、閾値`d_mic_mouth_thresh`より小さければ（ステップS153、Yes）、続いてコミュニケーションシステム1は、正対身体輪郭UIに対し、フェードインとフェードアウトのアニメーションを繰り返す（ステップS156）。

[0208] 正対身体輪郭UIに対し、フェードインとフェードアウトのアニメーションを繰り返すと、続いてコミュニケーションシステム1は、その時点の身体輪郭情報`info_outline_current`を生成し、現在の身体輪郭UIとして画面へ反映させる（ステップS157）。

[0209] 身体輪郭情報`info_outline_current`を生成し、現在の身体輪郭UIとして

画面へ反映させると、続いてコミュニケーションシステム1は、マイク102が集音した音情報info\_voiceの音量が音声入力判断閾値recognition\_vol\_threshを超えているかどうか判断する（ステップS158）。

[0210] ステップS158の判断の結果、音情報info\_voiceの音量が音声入力判断閾値recognition\_vol\_threshを超えていれば（ステップS158、Yes）、コミュニケーションシステム1は、音声翻訳フィードバック処理を実行する（ステップS159）。音声翻訳フィードバック処理の詳細については後に詳述する。音声翻訳フィードバック処理を実行すると、コミュニケーションシステム1は、上記ステップS153の判断処理に戻る。

[0211] 一方、ステップS158の判断の結果、音情報info\_voiceの音量が音声入力判断閾値recognition\_vol\_threshを超えていなければ（ステップS158、No）、コミュニケーションシステム1は、翻訳モードの継続判断処理を実行する（ステップS160）。翻訳モードの継続判断処理は後に詳述する。

[0212] 上記ステップS160の翻訳モードの継続判断処理を実行すると、続いてコミュニケーションシステム1は、翻訳モードかどうかを示すフラグf\_trans\_modeの値がtrueかどうか判断する（ステップS161）。フラグf\_trans\_modeの値がtrueであれば（ステップS161、Yes）、コミュニケーションシステム1は、上記ステップS153の判断処理に戻る。一方、フラグf\_trans\_modeの値がtrueでない、すなわちfalseであれば（ステップS161、No）、コミュニケーションシステム1は、上記ステップS155の翻訳モード終了処理を実行する。

[0213] 以上、図36A、36Bを用いて、図29のステップS103の、翻訳モード処理フローの詳細な流れを説明した。続いて、図36AのステップS152の翻訳モード起動処理フローの詳細な流れを説明する。

[0214] （翻訳モード起動処理フロー）

図37は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図37に示したのは、図36AのステップS15

2の翻訳モード起動処理フローの詳細な流れである。なお、以下で示す翻訳モード起動処理フローは、サインージ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、翻訳モード起動処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

- [0215] コミュニケーションシステム1は、高感度マイク位置 $p\_h\_sensitive\_mic$ と、口の位置情報 $p\_mouth$ との距離 $d\_mic\_mouth$ が、閾値 $d\_mic\_mouth\_thresh$ よりも小さいかどうか判断する（ステップS171）。
- [0216] ステップS171の判断の結果、高感度マイク位置 $p\_h\_sensitive\_mic$ と、口の位置情報 $p\_mouth$ との距離 $d\_mic\_mouth$ が、閾値 $d\_mic\_mouth\_thresh$ より小さくなければ（ステップS171、No）、コミュニケーションシステム1は、ステップS171の処理を繰り返す。一方、ステップS171の判断の結果、高感度マイク位置 $p\_h\_sensitive\_mic$ と、口の位置情報 $p\_mouth$ との距離 $d\_mic\_mouth$ が、閾値 $d\_mic\_mouth\_thresh$ より小さければ（ステップS171、Yes）、続いてコミュニケーションシステム1は、ユーザの顔情報 $info\_face$ の画像の、相手側の画面への反映を開始する（ステップS172）。
- [0217] 続いてコミュニケーションシステム1は、高感度マイク位置 $p\_h\_sensitive\_mic$ にならなかったマイクのマイク処理、すなわち集音処理を停止する（ステップS173）。
- [0218] 高感度マイク位置 $p\_h\_sensitive\_mic$ にならなかったマイクのマイク処理を停止すると、続いてコミュニケーションシステム1は、カメラ101で撮像された画像を用いて、ユーザの視線情報 $info\_gaze$ の取得を開始する（ステップS173）。
- [0219] ユーザの視線情報 $info\_gaze$ の取得を開始すると、続いてコミュニケーションシステム1は、高感度マイク位置のマイク102で集音されるユーザの発話の周波数情報 $info\_voice\_freq$ の取得を開始する（ステップS174）。
- [0220] 周波数情報 $info\_voice\_freq$ の取得を開始すると、続いてコミュニケーション

ンシステム 1 は、高感度マイク位置 `p_h_sensitive_mic` に対する音情報 `info_voice` での音声認識を起動する（ステップ S 1 7 5）。

[0221] 音声認識を起動すると、続いてコミュニケーションシステム 1 は、翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `true` に設定する（ステップ S 1 7 6）。なお、翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `true` に設定するのはこのタイミングに限定されるものではない。また、図 3 7 に示した処理フローにおける各処理の開始や起動のタイミングは、図 3 7 に示した順序に限定されるものではない。

[0222] 以上、図 3 6 A のステップ S 1 5 2 の翻訳モード起動処理フローの詳細な流れを説明した。続いて、図 3 6 A のステップ S 1 5 5 の翻訳モード終了フローの詳細な流れを説明する。

[0223] （翻訳モード終了フロー）

図 3 8 は、本開示の実施の形態に係るコミュニケーションシステム 1 の動作例を示す流れ図である。図 3 8 に示したのは、図 3 6 A のステップ S 1 5 5 の翻訳モード終了フローの詳細な流れである。なお、以下で示す翻訳モード終了フローは、サインページ端末 1 0 0 で取得された情報に基づいて、情報処理装置 2 0 0 が実行するものとするが、翻訳モード終了フローは、コミュニケーションシステム 1 を構成するどの装置で行われても良い。

[0224] コミュニケーションシステム 1 は、翻訳モード終了フローを実行する場合には、まず翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `false` に設定する（ステップ S 1 8 1）。なお、翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `false` に設定するのはこのタイミングに限定されるものではない。

[0225] 翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `false` に設定すると、続いてコミュニケーションシステム 1 は、ユーザの身体輪郭情報 `info_outline_current` の、ディスプレイ 1 0 3 への提示を停止する（ステップ S 1 8 2）。

[0226] ユーザの身体輪郭情報 `info_outline_current` の、ディスプレイ 1 0 3 への

提示を停止すると、続いてコミュニケーションシステム1は、正対身体輪郭UIに対する、フェードインとフェードアウトのアニメーションの繰り返しを停止する（ステップS183）。

[0227] 正対身体輪郭UIに対する、フェードインとフェードアウトのアニメーションの繰り返しを停止すると、続いてコミュニケーションシステム1は、高感度マイク位置p\_h\_sensitive\_micに対する音情報info\_voiceでの音声認識を停止する（ステップS184）。

[0228] 高感度マイク位置p\_h\_sensitive\_micに対する音情報info\_voiceでの音声認識を停止すると、続いてコミュニケーションシステム1は、高感度マイク位置のマイク102で集音されるユーザの発話の周波数情報info\_voice\_freqの取得を停止する（ステップS185）。

[0229] ユーザの発話の周波数情報info\_voice\_freqの取得を停止すると、続いてコミュニケーションシステム1は、ユーザの視線情報info\_gazeの取得を停止する（ステップS186）。

[0230] ユーザの視線情報info\_gazeの取得を停止すると、続いてコミュニケーションシステム1は、高感度マイク位置p\_h\_sensitive\_micにならなかったマイクのマイク処理、すなわち集音処理の停止を解除する（ステップS187）。

[0231] 高感度マイク位置p\_h\_sensitive\_micにならなかったマイクのマイク処理の停止を解除すると、続いてコミュニケーションシステム1は、ユーザの顔情報info\_faceの画像の、相手側の画面への反映を停止する（ステップS188）。

[0232] なお、図38に示した処理フローにおける各処理の停止のタイミングは、図38に示した順序に限定されるものではない。

[0233] 以上、図36AのステップS155の翻訳モード終了フローの詳細な流れを説明した。続いて、図36BのステップS159の音声翻訳フィードバック処理の詳細な流れを説明する。

[0234] （音声翻訳フィードバック処理フロー）

図39A、39Bは、本開示の実施の形態に係るコミュニケーションシス

テム1の動作例を示す流れ図である。図39A、39Bに示したのは、図36BのステップS159の音声翻訳フィードバック処理の詳細な流れである。なお、以下で示す音声翻訳フィードバック処理フローは、サイネージ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、音声翻訳フィードバック処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0235] コミュニケーションシステム1は、高感度マイク位置p\_h\_sensitive\_micのマイク102に入力される音情報info\_voiceの音量が閾値voice\_vol\_threshを超えたかどうか判断する（ステップS201）。

[0236] ステップS201の判断の結果、音情報info\_voiceの音量が閾値voice\_vol\_threshを超えていれば（ステップS201、Yes）、続いてコミュニケーションシステム1は、音量の大きさに応じた半径を持つ球オブジェクトobj\_ballを生成する（ステップS202）。球オブジェクトobj\_ballを生成すると、続いてコミュニケーションシステム1は、球オブジェクトアニメーション処理を実行する（ステップS203）。球オブジェクトアニメーション処理の詳細については後に詳述する。

[0237] 上記ステップS203の球オブジェクトアニメーション処理を実行すると、または、ステップS201の判断の結果、音情報info\_voiceの音量が閾値voice\_vol\_threshを超えていなければ（ステップS201、No）、続いてコミュニケーションシステム1は、ユーザの過去の発話の周波数平均情報としてinfo\_voice\_freq\_aveを計算する（ステップS204）。

[0238] 周波数平均情報info\_voice\_freq\_aveを計算すると、続いてコミュニケーションシステム1は、音声情報処理装置300で音声認識結果voice\_resultを得られたかどうか判断する（ステップS205）。

[0239] 上記ステップS205の判断の結果、音声情報処理装置300で音声認識結果voice\_resultを得られた場合は（ステップS205、Yes）、続いてコミュニケーションシステム1は、音声認識結果表示処理を実行する（ステップS206）。音声認識結果表示処理の詳細については後に詳述する。

- [0240] 音声認識結果表示処理を実行すると、続いてコミュニケーションシステム 1 は、音声認識結果voice\_resultに対する翻訳処理装置 400 での翻訳処理を起動する（ステップ S 207）。
- [0241] 音声認識結果voice\_resultに対する翻訳処理装置 400 での翻訳処理を起動すると、続いてコミュニケーションシステム 1 は、音声認識結果voice\_resultに対する翻訳処理装置 400 での翻訳結果trans\_resultを得られたかどうか判断する（ステップ S 208）。
- [0242] 上記ステップ S 208 での判断の結果、音声認識結果voice\_resultに対する翻訳処理装置 400 での翻訳結果trans\_resultを得られた場合は（ステップ S 208、Yes）、続いてコミュニケーションシステム 1 は、翻訳処理装置 400 での翻訳結果trans\_resultをサインージ端末 100 に表示する（ステップ S 209）。
- [0243] 翻訳処理装置 400 での翻訳結果trans\_resultをサインージ端末 100 に表示すると、続いてコミュニケーションシステム 1 は、翻訳結果trans\_resultを得たときのスコアtrans\_result\_scoreを記憶する（ステップ S 210）。翻訳スコアは、翻訳処理装置 400 のスコア生成部 422 が、翻訳結果trans\_resultの確からしさに応じて生成する。
- [0244] 翻訳結果trans\_resultを得たときのスコアtrans\_result\_scoreを記憶すると、続いてコミュニケーションシステム 1 は、翻訳結果表示処理を実行する（ステップ S 211）。翻訳結果表示処理の詳細については後に詳述する。
- [0245] 翻訳結果表示処理を実行すると、続いてコミュニケーションシステム 1 は、処理量あふれ判断処理を実行する（ステップ S 212）。処理量あふれ判断処理の詳細については後に詳述する。
- [0246] 処理量あふれ判断処理を実行すると、続いてコミュニケーションシステム 1 は、処理量あふれ提示処理を実行する（ステップ S 213）。処理量あふれ提示処理の詳細については後に詳述する。処理量あふれ提示処理を実行すると、コミュニケーションシステム 1 は、音声翻訳フィードバック処理を終了する。

[0247] 上記ステップS 205の判断の結果、音声情報処理装置300で音声認識結果voice\_resultを得られなかった場合（ステップS 205、No）、または上記ステップS 208での判断の結果、音声認識結果voice\_resultに対する翻訳処理装置400での翻訳結果trans\_resultを得られなかった場合は（ステップS 208、No）、コミュニケーションシステム1は、音声翻訳フィードバック処理を終了する。

[0248] 以上、図36BのステップS 159の音声翻訳フィードバック処理の詳細な流れを説明した。続いて、図39AのステップS 203の球オブジェクトアニメーション処理の詳細な流れを説明する。

[0249] （球オブジェクトアニメーション処理フロー）

図40は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図40に示したのは、図39AのステップS 203の球オブジェクトアニメーション処理の詳細な流れである。なお、以下で示す球オブジェクトアニメーション処理フローは、サイネージ端末100で取得された情報に基づいて、情報処理装置200が実行するものとするが、球オブジェクトアニメーション処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0250] 球オブジェクトアニメーション処理を実行する場合、コミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する（ステップS 211）。

[0251] 上記ステップS 211の判断の結果、ユーザが発話中である場合は（ステップS 211、Yes）、続いてコミュニケーションシステム1は、球オブジェクトobj\_ballを相手の左の耳位置p\_ear\_l及び右の耳位置p\_ear\_rの位置へ向かってアニメーションしながら移動させる（ステップS 212）。

[0252] 一方、上記ステップS 211の判断の結果、ユーザが発話中でない場合は（ステップS 211、No）、続いてコミュニケーションシステム1は、球オブジェクトobj\_ballを相手の口の位置情報p\_mouthから、左の耳位置p\_ear\_l及び右の耳位置p\_ear\_rの位置へ向かってアニメーションしながら移動させ

る（ステップS 2 1 3）。

[0253] 以上、図39AのステップS 2 0 3の球オブジェクトアニメーション処理の詳細な流れを説明した。続いて、図39BのステップS 2 1 2の処理量あふれ判断処理の詳細な流れを説明する。

[0254] （処理量あふれ判断処理フロー（1））

図41は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図41に示したのは、図39BのステップS 2 1 2の処理量あふれ判断処理の詳細な流れである。なお、以下で示す処理量あふれ判断処理フローは、サインージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、処理量あふれ判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0255] 処理量あふれ判断処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する（ステップS 2 2 1）。

[0256] 上記ステップS 2 2 1の判断の結果、ユーザが発話中であった場合は（ステップS 2 2 1、Y e s）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS 2 2 2）。

[0257] 一方、上記ステップS 2 2 1の判断の結果、ユーザが発話中でなかった場合は（ステップS 2 2 1、N o）、続いてコミュニケーションシステム1は、相手の正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に、ユーザの視線情報info\_gazeが入っているかどうか判断する（ステップS 2 2 3）。

[0258] 上記ステップS 2 2 3の判断の結果、相手の正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に、ユーザの視線情報info\_gazeが入っていなければ（ステップS 2 2 3、N o）、続いてコミュニケーションシステム1は、ユーザの視線情報info\_gazeが、正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に入っている時間が、閾値t\_s

urrounding\_threshよりも大きいかどうか判断する（ステップS 2 2 4）。

[0259] 上記ステップS 2 2 4の判断の結果、ユーザの視線情報info\_gazeが、正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に入っている時間が、閾値t\_surrounding\_threshよりも大きくなければ（ステップS 2 2 4、N o）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS 2 2 2）。

[0260] 一方、上記ステップS 2 2 3の判断の結果、相手の正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に、ユーザの視線情報info\_gazeが入っていれば（ステップS 2 2 3、Y e s）、または、上記ステップS 2 2 4の判断の結果、ユーザの視線情報info\_gazeが、正対身体輪郭情報領域に閾値d\_surrounding\_threshの距離を加えた領域内に入っている時間が、閾値t\_surrounding\_threshよりも大きければ（ステップS 2 2 4、Y e s）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をtrueに設定する（ステップS 2 2 5）。

[0261] （処理量あふれ判断処理フロー（2））

処理量あふれ判断処理フローの2つ目の例を説明する。図4 2は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図4 2に示したのは、図3 9 BのステップS 2 1 2の処理量あふれ判断処理の詳細な流れである。なお、以下で示す処理量あふれ判断処理フローは、サインページ端末1 0 0で取得された情報などに基づいて、情報処理装置2 0 0が実行するものとするが、処理量あふれ判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0262] 処理量あふれ判断処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する（ステップS 2 3 1）。

[0263] 上記ステップS 2 3 1の判断の結果、ユーザが発話中であった場合は（ステップS 2 3 1、Y e s）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステッ

プS 2 3 2)。

[0264] 一方、上記ステップS 2 3 1の判断の結果、ユーザが発話中でなかった場合は(ステップS 2 3 1、No)、続いてコミュニケーションシステム1は、ユーザの相槌を検出できたかどうか判断する(ステップS 2 3 3)。

[0265] 上記ステップS 2 3 3の判断の結果、ユーザの相槌を検出できなければ(ステップS 2 3 3、No)、続いてコミュニケーションシステム1は、最後に相槌を検出した時間から現時点までの時間が閾値t\_nod\_threshよりも大きいかどうか判断する(ステップS 2 3 4)。

[0266] 上記ステップS 2 3 4の判断の結果、最後に相槌を検出した時間から現時点までの時間が閾値t\_nod\_threshよりも大きくなければ(ステップS 2 3 4、No)、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する(ステップS 2 3 2)。

[0267] また、上記ステップS 2 3 3の判断の結果、ユーザの相槌を検出できれば(ステップS 2 3 3、Yes)、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する(ステップS 2 3 2)。

[0268] 一方、上記ステップS 2 3 4の判断の結果、最後に相槌を検出した時間から現時点までの時間が閾値t\_nod\_threshよりも大きければ(ステップS 2 3 4、Yes)、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をtrueに設定する(ステップS 2 3 5)。

[0269] (処理量あふれ判断処理フロー(3))

処理量あふれ判断処理フローの3つ目の例を説明する。図43は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図43に示したのは、図39BのステップS 2 1 2の処理量あふれ判断処理の詳細な流れである。なお、以下で示す処理量あふれ判断処理フローは、サイネージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、処理量あふれ判断処理フローは、コミュニ

ケーションシステム1を構成するどの装置で行われても良い。

- [0270] 処理量あふれ判断処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する（ステップS241）。
- [0271] 上記ステップS241の判断の結果、ユーザが発話中であった場合は（ステップS241、Yes）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS242）。
- [0272] 一方、上記ステップS241の判断の結果、ユーザが発話中でなかった場合は（ステップS241、No）、続いてコミュニケーションシステム1は、ユーザの掌を検出できたかどうか判断する（ステップS243）。
- [0273] 上記ステップS243の判断の結果、ユーザの掌を検出できれば（ステップS243、Yes）、続いてコミュニケーションシステム1は、ユーザの掌を検出した時間の累積時間t\_palm\_accumを更新する（ステップS244）。ユーザの掌を検出した時間の累積時間t\_palm\_accumを更新すると、続いてコミュニケーションシステム1は、ユーザの掌を検出した時間の累積時間t\_palm\_accumが閾値t\_palm\_accum\_threshよりも大きいかどうか判断する（ステップS245）。
- [0274] 上記ステップS245の判断の結果、ユーザの掌を検出した時間の累積時間t\_palm\_accumが閾値t\_palm\_accum\_threshよりも大きくなければ（ステップS245、No）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS242）。
- [0275] また、上記ステップS243の判断の結果、ユーザの掌を検出できなければ（ステップS243、No）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS242）。
- [0276] 一方、上記ステップS245の判断の結果、ユーザの掌を検出した時間の累積時間t\_palm\_accumが閾値t\_palm\_accum\_threshよりも大きければ（ステッ

プS 2 4 5、 Y e s )、続いてコミュニケーションシステム 1 は、処理量あふれを判断するフラグf\_task\_over\_modeの値をtrueに設定する (ステップS 2 4 6)。

[0277] (処理量あふれ判断処理フロー (4))

処理量あふれ判断処理フローの4つ目の例を説明する。図4 4 は、本開示の実施の形態に係るコミュニケーションシステム 1 の動作例を示す流れ図である。図4 4 に示したのは、図3 9 BのステップS 2 1 2の処理量あふれ判断処理の詳細な流れである。なお、以下で示す処理量あふれ判断処理フローは、サインページ端末1 0 0で取得された情報などに基づいて、情報処理装置2 0 0が実行するものとするが、処理量あふれ判断処理フローは、コミュニケーションシステム 1 を構成するどの装置で行われても良い。

[0278] 処理量あふれ判断処理を実行する場合、まずコミュニケーションシステム 1 は、ユーザが発話中であるかどうかを判断する (ステップS 2 5 1)。

[0279] 上記ステップS 2 5 1の判断の結果、ユーザが発話中であった場合は (ステップS 2 5 1、 Y e s )、続いてコミュニケーションシステム 1 は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する (ステップS 2 5 2)。

[0280] 一方、上記ステップS 2 5 1の判断の結果、ユーザが発話中でなかった場合は (ステップS 2 5 1、 N o )、続いてコミュニケーションシステム 1 は、音声認識結果voice\_resultを、自然言語解析処理によって句単位に分割する (ステップS 2 5 3)。

[0281] 音声認識結果voice\_resultを、自然言語解析処理によって句単位に分割すると、続いてコミュニケーションシステム 1 は、その音声認識結果voice\_resultの句の数num\_phraseを、それまでの累積数num\_accum\_phraseへ加算する (ステップS 2 5 4)。

[0282] 音声認識結果voice\_resultの句の数num\_phraseを、それまでの累積数num\_accum\_phraseへ加算すると、続いてコミュニケーションシステム 1 は、加算後の累積数num\_phrase\_accumが、閾値num\_phrase\_accum\_threshを超えたかどうか

か判断する（ステップS 2 5 5）。

[0283] 上記ステップS 2 5 5の判断の結果、加算後の累積数num\_phrase\_accumが、閾値num\_phrase\_accum\_threshを超えていなければ（ステップS 2 5 5、N o）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をfalseに設定する（ステップS 2 5 2）。

[0284] 一方、上記ステップS 2 5 5の判断の結果、加算後の累積数num\_phrase\_accumが、閾値num\_phrase\_accum\_threshを超えていれば（ステップS 2 5 5、Y e s）、続いてコミュニケーションシステム1は、処理量あふれを判断するフラグf\_task\_over\_modeの値をtrueに設定する（ステップS 2 5 6）。

[0285] 以上、処理量あふれ判断処理の詳細なフローを説明した。なお、ここでは5つの処理量あふれ判断処理フローの例を示したが、コミュニケーションシステム1は、この5つのフローの中から1つを選択して処理量あふれ判断処理を実行しても良く、複数のフローを選択して処理量あふれ判断処理フローを実行しても良い。複数のフローを選択する場合は、フローの結果が異なる場合があり得る。従って、複数のフローを選択する場合、コミュニケーションシステム1は、奇数個のフローを選択して、フラグf\_task\_over\_modeの値の設定の結果が多い方を採用してもよい。続いて、図39BのステップS 2 1 3の処理量あふれ提示処理の詳細な流れを説明する。

[0286] （処理量あふれ提示処理フロー）

図45は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図45に示したのは、図39BのステップS 2 1 3の処理量あふれ提示処理の詳細な流れである。なお、以下で示す処理量あふれ提示処理フローは、サインページ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、処理量あふれ提示処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0287] 処理量あふれ提示処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する（ステップS 2 6 1）。

[0288] 上記ステップS 2 6 1の判断の結果、ユーザが発話中であれば（ステップS 2 6 1、Y e s）、コミュニケーションシステム1は、ユーザの発話の音量の大きさに応じた半径を持つ球オブジェクトobj\_ballを、相手の右の耳位置p\_ear\_rと左の耳位置p\_ear\_lの位置へ向かってアニメーションながら移動し、その後、その位置から垂直下方向へアニメーションしながら移動させるよう、サイネージ端末100に表示させる（ステップS 2 6 2）。

[0289] 一方、上記ステップS 2 6 1の判断の結果、ユーザが発話中でなければ（ステップS 2 6 1、N o）、コミュニケーションシステム1は、そのまま処理量あふれ提示処理を終了する。

[0290] 以上、処理量あふれ提示処理の詳細な流れを説明した。続いて、音声認識結果表示処理の詳細な流れを説明する。

[0291] （音声認識結果表示処理フロー）

図46は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図46に示したのは、図39AのステップS 2 0 6の音声認識結果表示処理の詳細な流れである。なお、以下で示す音声認識結果表示処理フローは、サイネージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、音声認識結果表示処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0292] 音声認識結果表示処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であったかどうかを判断する（ステップS 2 7 1）。

[0293] 上記ステップS 2 7 1の判断の結果、ユーザが発話中であった場合は（ステップS 2 7 1、Y e s）、コミュニケーションシステム1は、音声認識結果を、サイネージ端末100における、相手の右の耳位置p\_ear\_rまたは左の耳位置p\_ear\_lの付近に表示させる（ステップS 2 7 2）。

[0294] 一方、上記ステップS 2 7 1の判断の結果、ユーザが発話中でなかった場合は（ステップS 2 7 1、N o）、コミュニケーションシステム1は、そのまま音声認識結果表示処理を終了する。

[0295] 以上、音声認識結果表示処理の詳細な流れを説明した。続いて、翻訳結果表示処理の詳細な流れを説明する。

[0296] (翻訳結果表示処理フロー)

図47は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図47に示したのは、図39BのステップS211の翻訳結果表示処理の詳細な流れである。なお、以下で示す翻訳結果表示処理フローは、サイネージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、翻訳結果表示処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0297] 翻訳結果表示処理を実行する場合、まずコミュニケーションシステム1は、ユーザが発話中であるかどうかを判断する(ステップS281)。

[0298] 上記ステップS281の判断の結果、ユーザが発話中でない場合は(ステップS281、No)、コミュニケーションシステム1は、翻訳結果を、サイネージ端末100における、相手の口の位置情報p\_mouthの付近に表示させる(ステップS282)。

[0299] 一方、上記ステップS281の判断の結果、ユーザが発話中である場合は(ステップS281、Yes)、コミュニケーションシステム1は、そのまま翻訳結果表示処理を終了する。

[0300] 以上、翻訳結果表示処理の詳細な流れを説明した。続いて、図36BのステップS160の翻訳モード継続判断処理の詳細な流れを説明する。

[0301] (翻訳モード継続判断処理フロー(1))

図48は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図48に示したのは、図36BのステップS160の翻訳モード継続判断処理の詳細な流れである。なお、以下で示す翻訳モード継続判断処理フローは、サイネージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、翻訳モード継続判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

- [0302] 翻訳モード継続判断処理フローを実行する場合、まずコミュニケーションシステム1は、翻訳スコア`trans_score`が閾値`trans_score_thresh`より低いかどうか判断する（ステップS291）。
- [0303] 上記ステップS291の判断の結果、翻訳スコア`trans_score`が閾値`trans_score_thresh`より低ければ（ステップS291、Yes）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグ`f_trans_mode`の値を`false`に設定して、翻訳モード継続判断処理フローを終了する（ステップS292）。
- [0304] 一方、上記ステップS291の判断の結果、翻訳スコア`trans_score`が閾値`trans_score_thresh`より低くなければ（ステップS291、No）、続いてコミュニケーションシステム1は、ユーザの視線情報`info_gaze`が相手の正対身体輪郭情報領域内に入っているかどうか判断する（ステップS293）。
- [0305] 上記ステップS293の判断の結果、ユーザの視線情報`info_gaze`が相手の正対身体輪郭情報領域内に入っていなければ（ステップS293、No）、続いてコミュニケーションシステム1は、ユーザの視線情報`info_gaze`が、相手の正対身体輪郭情報領域に閾値`d_surrounding_thresh`の距離を加えた領域内に入っているかどうか判断する（ステップS294）。
- [0306] 上記ステップS294の判断の結果、ユーザの視線情報`info_gaze`が、相手の正対身体輪郭情報領域に閾値`d_surrounding_thresh`の距離を加えた領域内に入っていなければ（ステップS294、No）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグ`f_trans_mode`の値を`false`に設定して、翻訳モード継続判断処理フローを終了する（ステップS292）。
- [0307] 一方、上記ステップS293の判断の結果、ユーザの視線情報`info_gaze`が相手の正対身体輪郭情報領域内に入っていれば（ステップS293、Yes）、または、上記ステップS294の判断の結果、ユーザの視線情報`info_gaze`が、相手の正対身体輪郭情報領域に閾値`d_surrounding_thresh`の距離を加えた領域内に入っていれば（ステップS294、Yes）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグ`f_trans_mode`の値を`true`

eに設定して、翻訳モード継続判断処理フローを終了する（ステップS 2 9 5）。

[0308] （翻訳モード継続判断処理フロー（2））

翻訳モード継続判断処理フローの2つ目の例を説明する。図49は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図49に示したのは、図36BのステップS 1 6 0の翻訳モード継続判断処理の詳細な流れである。なお、以下で示す翻訳モード継続判断処理フローは、サインエージ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、翻訳モード継続判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0309] 翻訳モード継続判断処理フローを実行する場合、まずコミュニケーションシステム1は、ユーザの発話によって得られる音情報info\_voiceにユーザの言いよどみが検出できるかどうか判断する（ステップS 3 0 1）。

[0310] 上記ステップS 3 0 1の判断の結果、音情報info\_voiceにユーザの言いよどみが検出できれば（ステップS 3 0 1、Y e s）、続いてコミュニケーションシステム1は、最後に取得できたユーザの音声認識結果voice\_resultが、フィルターワードデータベース内に入っているかどうか判断する（ステップS 3 0 2）。

[0311] 表1は、フィルターワードデータベースの内容の例を示す表である。コミュニケーションシステム1は、ユーザの音声認識結果voice\_resultが、フィルターワードデータベース内に入っているフィルターワード群と一致するかどうかだけでなく、類似音フィルターワード群と一致するかどうかを判断しても良い。

[0312]

[表1]

フィルターワード群	類似音フィルターワード群
あの	あのう、あの一、アノ、
えっと	ええと、えと、エッと、干支、
um	Uh, huh,
You know	Know, you non

(表1 : フィルターワードデータベースの例)

- [0313] 上記ステップS302の判断の結果、ユーザの音声認識結果voice\_resultが、フィルターワードデータベース内に入っていれば（ステップS302、Yes）、続いてコミュニケーションシステム1は、最後に発話したフィルターワードからの時間が閾値t\_fillerwords\_threshを超えていないかどうか判断する（ステップS303）。
- [0314] 上記ステップS301の判断の結果、音情報info\_voiceにユーザの言いよどみが検出できれば（ステップS301、No）、または、上記ステップS302の判断の結果、ユーザの音声認識結果voice\_resultが、フィルターワードデータベース内に入っていなければ（ステップS302、No）、または、上記ステップS303の判断の結果、最後に発話したフィルターワードからの時間が閾値t\_fillerwords\_threshを超えていれば（ステップS303、No）、コミュニケーションシステム1は、音情報info\_voiceとして無音区間と判断する条件を満たすかどうか判断する（ステップS304）。無音区間と判断する条件としては、例えば音量が極めて小さい区間が継続しているかどうか、等があり得る。
- [0315] 上記ステップS304の判断の結果、音情報info\_voiceとして無音区間と判断する条件を満たしていれば（ステップS304、Yes）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグf\_trans\_modeの値をfalseに設定して、翻訳モード継続判断処理フローを終了する（ステップS305）。
- [0316] 一方、上記ステップS303の判断の結果、最後に発話したフィルターワードからの時間が閾値t\_fillerwords\_threshを超えていなければ（ステップS

303、Yes)、または、上記ステップS304の判断の結果、音情報info\_voiceとして無音区間と判断する条件を満たしていなければ(ステップS304、No)、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグf\_trans\_modeの値をtrueに設定して、翻訳モード継続判断処理フローを終了する(ステップS306)。

[0317] (翻訳モード継続判断処理フロー(3))

翻訳モード継続判断処理フローの3つ目の例を説明する。図50は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図50に示したのは、図36BのステップS160の翻訳モード継続判断処理の詳細な流れである。なお、以下で示す翻訳モード継続判断処理フローは、サインページ端末100で取得された情報などに基づいて、情報処理装置200が実行するものとするが、翻訳モード継続判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0318] 翻訳モード継続判断処理フローを実行する場合、まずコミュニケーションシステム1は、最後に取得した発話の周波数情報info\_last\_voice\_freqが過去の平均情報であるinfo\_voice\_freq\_aveから閾値f\_threshを引いたものよりも大きいかどうか判断する(ステップS311)。

[0319] 上記ステップS311の判断の結果、最後に取得した発話の周波数情報info\_last\_voice\_freqが過去の平均情報であるinfo\_voice\_freq\_aveから閾値f\_threshを引いたものよりも大きければ(ステップS311、Yes)、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグf\_trans\_modeの値をfalseに設定して、翻訳モード継続判断処理フローを終了する(ステップS312)。

[0320] 一方、上記ステップS311の判断の結果、最後に取得した発話の周波数情報info\_last\_voice\_freqが過去の平均情報であるinfo\_voice\_freq\_aveから閾値f\_threshを引いたものよりも大きくなければ(ステップS311、No)、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグf\_trans\_modeの値をtrueに設定して、翻訳モード継続判断処理フローを終了する

(ステップS 3 1 3)。

[0321] (翻訳モード継続判断処理フロー (4))

翻訳モード継続判断処理フローの4つ目の例を説明する。図5 1は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図5 1に示したのは、図3 6 BのステップS 1 6 0の翻訳モード継続判断処理の詳細な流れである。なお、以下で示す翻訳モード継続判断処理フローは、サインエージ端末1 0 0で取得された情報などに基づいて、情報処理装置2 0 0が実行するものとするが、翻訳モード継続判断処理フローは、コミュニケーションシステム1を構成するどの装置で行われても良い。

[0322] 翻訳モード継続判断処理フローを実行する場合、まずコミュニケーションシステム1は、最後に取得できたユーザの音声認識結果voice\_resultが、会話終了判断ワードデータベース内に入っているかどうか判断する(ステップS 3 2 1)。

[0323] 表2は、会話終了判断ワードデータベースの内容の例を示す表である。コミュニケーションシステム1は、ユーザの音声認識結果voice\_resultが、会話終了判断ワードデータベース内に入っている会話終了判断ワード群と一致するかどうかだけではなく、類似音会話終了判断ワード群と一致するかどうかを判断しても良い。

[0324] [表2]

会話終了判断ワード群	類似音会話終了判断ワード群
どうですか	どうです、どう、そうですか
これでよろしいですか	これでよい、これでよろしい、
What do you think of that	What do you think, how do you think
Don' t you think so	Do you think, do you think so

(表2 : 会話終了判断ワードデータベースの例)

[0325] 上記ステップS 3 2 1の判断の結果、最後に取得できたユーザの音声認識結果voice\_resultが、会話終了判断ワードデータベース内に入っていれば(ステップS 3 2 1、Y e s)、続いてコミュニケーションシステム1は、最

後に発話した会話終了判断ワードからの時間が閾値`t_convendwords_thresh`を超えていないかどうか判断する（ステップS 3 2 2）。

[0326] 上記ステップS 3 2 1の判断の結果、最後に取得できたユーザの音声認識結果`voice_result`が、会話終了判断ワードデータベース内に入っていないならば（ステップS 3 2 1、No）、または、上記ステップS 3 2 2の判断の結果、最後に発話した会話終了判断ワードからの時間が閾値`t_convendwords_thresh`を超えていれば（ステップS 3 2 2、No）、コミュニケーションシステム1は、音情報`info_voice`として無音区間と判断する条件を満たすかどうか判断する（ステップS 3 2 3）。無音区間と判断する条件としては、例えば音量が極めて小さい区間が継続しているかどうか、等があり得る。

[0327] 上記ステップS 3 2 3の判断の結果、音情報`info_voice`として無音区間と判断する条件を満たしていれば（ステップS 3 2 3、Yes）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグ`f_trans_mode`の値を`false`に設定して、翻訳モード継続判断処理フローを終了する（ステップS 3 2 4）。

[0328] 一方、上記ステップS 3 2 2の判断の結果、最後に発話した会話終了判断ワードからの時間が閾値`t_convendwords_thresh`を超えていなければ（ステップS 3 2 2、Yes）、または、上記ステップS 3 2 3の判断の結果、音情報`info_voice`として無音区間と判断する条件を満たしていなければ（ステップS 3 2 3、No）、コミュニケーションシステム1は、翻訳モードかどうかを示すフラグ`f_trans_mode`の値を`true`に設定して、翻訳モード継続判断処理フローを終了する（ステップS 3 2 4）。

[0329] （翻訳モード継続判断処理フロー（5））

翻訳モード継続判断処理フローの5つ目の例を説明する。図5 2は、本開示の実施の形態に係るコミュニケーションシステム1の動作例を示す流れ図である。図5 2に示したのは、図3 6 BのステップS 1 6 0の翻訳モード継続判断処理の詳細な流れである。なお、以下で示す翻訳モード継続判断処理フローは、サインページ端末1 0 0で取得された情報などに基づいて、情報処

理装置 200 が実行するものとするが、翻訳モード継続判断処理フローは、コミュニケーションシステム 1 を構成するどの装置で行われても良い。

[0330] 翻訳モード継続判断処理フローを実行する場合、まずコミュニケーションシステム 1 は、画面上に出ている翻訳モード解除ボタンが押されたかどうか判断する（ステップ S 331）。

[0331] 上記ステップ S 331 の判断の結果、翻訳モード解除ボタンが押されていれば（ステップ S 331、Yes）、コミュニケーションシステム 1 は、翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `false` に設定して、翻訳モード継続判断処理フローを終了する（ステップ S 332）。

[0332] 一方、上記ステップ S 331 の判断の結果、翻訳モード解除ボタンが押されていないならば（ステップ S 331、No）、コミュニケーションシステム 1 は、翻訳モードかどうかを示すフラグ `f_trans_mode` の値を `true` に設定して、翻訳モード継続判断処理フローを終了する（ステップ S 333）。

[0333] 以上、翻訳モード継続判断処理の詳細な流れを説明した。なお、ここでは 5 つの翻訳モード継続判断処理フローの例を示したが、コミュニケーションシステム 1 は、この 5 つのフローの中から 1 つを選択して翻訳モード継続判断処理を実行しても良く、複数のフローを選択して翻訳モード継続判断処理フローを実行しても良い。複数のフローを選択する場合は、フローの結果が異なる場合があり得る。従って、複数のフローを選択する場合、コミュニケーションシステム 1 は、奇数個のフローを選択して、フラグ `f_trans_mode` の値の設定の結果が多い方を採用してもよい。

[0334] 本開示の実施の形態に係るコミュニケーションシステム 1 は、上述した動作を実行することで、サイネージ端末等を介して、人間同士が正対してコミュニケーションを行う際に、発話者や対話者の状況に基づいて自然なコミュニケーションを行わせることが可能となる。

[0335] <2. まとめ>

以上説明したように本開示の実施の形態によれば、サイネージ端末 100 を介して、翻訳をしながら人間同士が会話を行えるコミュニケーションシス

テム 1 において、仮に発話が途絶えたとしても、発話が途絶えたことのみを理由として翻訳を打ち切らないようにするコミュニケーションシステム 1 が提供される。

[0336] 本開示の実施の形態に係るコミュニケーションシステム 1 は、発話が途絶えたことのみを理由として翻訳を打ち切らないようにすることで、発話者側が発話内容を思い浮かべることが出来ない場合にも、翻訳したい対象だけを翻訳相手に伝えられるとともに、相手に発話を遮られないようにできる。

[0337] また本開示の実施の形態によれば、サイネージ端末 100 を介して、翻訳をしながら人間同士が会話を行えるコミュニケーションシステム 1 において、話し手側が発話している最中に、聞き手側の処理量があふれていることを、話し手側にサイネージ端末 100 を用いて通知するコミュニケーションシステム 1 が提供される。

[0338] 本開示の実施の形態に係るコミュニケーションシステム 1 は、話し手側が発話している最中に、聞き手側の処理量があふれていることを、話し手側にサイネージ端末 100 を用いて通知することで、聞き手側の情報の処理量のイメージを、話し手の発話を遮ること無く、話し手側に伝えることが可能となる。

[0339] すなわち、本開示の実施の形態に係るコミュニケーションシステム 1 は、自分の発話を相手に遮られることを回避したり、聞き手側の処理量があふれていることを、話し手側に伝えたりすることで、話し手側及び聞き手側にとって円滑なコミュニケーションを可能にさせる。

[0340] なお上記実施形態では、サイネージ端末 100 を介したコミュニケーションシステム 1 の例を示したが、本開示は係る例に限定されるものではない。本開示は、ディスプレイ、スピーカー、マイクを介してその場に無い人間同士がコミュニケーションするシステム、例えばテレビ電話システム、テレビ会議システム、AR（拡張現実）を利用したバーチャル会話支援システムなどにも適用可能である。

[0341] テレビ会議システムは、会議の参加者が会議中にカメラではなくマイクの

方を向いてしまうと不自然な状況になりやすい。そのため、本開示を適用するテレビ会議システムは、会議の参加者がマイクの方を向いた場合には、上述したように予め作成した正対時の画像に置き換えて相手側のディスプレイに提示する。このように正対時の画像に置き換えて相手側のディスプレイに提示することで、本開示を適用するテレビ会議システムは、不自然な状況となることを回避出来る。

[0342] また、本開示をテレビ会議システムに適用する場合、例えば、発話者の発話内容をログとして記録しておき、そのログを翻訳に活用しても良い。会議では頻繁に「あれ」や「さっきの」等といった句が発話されることが多いが、本開示を適用するテレビ会議システムは、その「あれ」や「さっきの」等の句を具体的な名詞等に置き換えて翻訳し、相手側に提示しても良い。

[0343] 本明細書の各装置が実行する処理における各ステップは、必ずしもシーケンス図またはフローチャートとして記載された順序に沿って時系列に処理する必要はない。例えば、各装置が実行する処理における各ステップは、フローチャートとして記載した順序と異なる順序で処理されても、並列的に処理されてもよい。

[0344] また、各装置に内蔵されるCPU、ROMおよびRAMなどのハードウェアを、上述した各装置の構成と同等の機能を発揮させるための、単体のコンピュータまたは複数のコンピュータが一連の処理を行うコンピュータシステムで実行されるプログラムも作成可能である。また、該コンピュータプログラムを記憶させた記憶媒体も提供されることが可能である。また、機能ブロック図で示したそれぞれの機能ブロックをハードウェアまたはハードウェア回路で構成することで、一連の処理をハードウェアまたはハードウェア回路で実現することもできる。

[0345] また上述の説明で用いた機能ブロック図で示したそれぞれの機能ブロックの一部又は全部は、たとえばインターネット等のネットワークを介して接続されるサーバ装置で実現されてもよい。また上述の説明で用いた機能ブロック図で示したそれぞれの機能ブロックの構成は、単独の装置で実現されても

よく、複数の装置が連携するシステムで実現されても良い。複数の装置が連携するシステムには、例えば複数のサーバ装置の組み合わせ、サーバ装置と端末装置との組み合わせ等が含まれ得る。

[0346] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0347] また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

[0348] なお、以下のような構成も本開示の技術的範囲に属する。

(1)

発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行う制御部を備える、制御装置。

(2)

前記制御部は、前記過去の発話内容及び前記発話状況に応じて、前記翻訳処理を終了させる、前記(1)に記載の制御装置。

(3)

前記制御部は、前記発話状況として前記発話者の視線に基づいて、前記翻訳処理を終了させる、前記(2)に記載の制御装置。

(4)

前記制御部は、前記発話状況として前記発話者の視線が該発話者の対話者に向いており、かつ前記発話内容として前記発話者の発話を所定時間検出しなくなると、前記翻訳処理を終了させる、前記(3)に記載の制御装置。

(5)

前記制御部は、前記発話内容として所定のキーワードが含まれていれば、前記翻訳処理を終了させる、前記(2)に記載の制御装置。

(6)

前記制御部は、前記発話状況として前記発話者の動作に基づいて、前記翻訳処理の継続判断を行う、前記(1)に記載の制御装置。

(7)

前記制御部は、前記発話状況として前記発話者の視線に基づいて、前記翻訳処理の継続判断を行う、前記(6)に記載の制御装置。

(8)

前記制御部は、前記発話状況として前記発話者の動作及び周囲の音に基づいて、前記翻訳処理の継続判断を行う、前記(6)に記載の制御装置。

(9)

前記制御部は、前記発話状況として前記発話者が置かれているシチュエーションに基づいて、前記翻訳処理の継続判断を行う、前記(1)に記載の制御装置。

(10)

前記制御部は、前記発話状況として前記発話者の対話相手との関係に基づいて、前記翻訳処理の継続判断を行う、前記(1)に記載の制御装置。

(11)

前記制御部は、前記発話内容として前記発話者の対話相手からの発話内容に対する回答の内容に基づいて、前記翻訳処理の継続判断を行う、前記(1)に記載の制御装置。

(12)

前記制御部は、前記発話内容として文章の成立の有無に基づいて前記翻訳処理の継続判断を行う、前記(1)に記載の制御装置。

(13)

前記制御部は、前記発話内容として前記発話者の声質に基づいての継続判

断を行う、前記（１）に記載の制御装置。

（１４）

前記制御部は、前記発話者の声質を周波数に基づいて判断する、前記（１３）に記載の制御装置。

（１５）

前記制御部は、前記レベルとして、前記発話内容及び前記発話状況に基づいて算出したスコアを用いる、前記（１）～（１４）に記載の制御装置。

（１６）

前記制御部は、算出した前記スコアが所定値以下の場合に、前記発話者の視線が前記発話者の対話者に向いているときは、前記翻訳処理を継続する、前記（１５）に記載の制御装置。

（１７）

前記制御部は、前記スコアを、前記発話者の視線に基づいて算出する、前記（１５）に記載の制御装置。

（１８）

前記制御部は、前記スコアを、前記発話者が発話した内容の翻訳結果の確からしさに基づいて算出する、前記（１５）に記載の制御装置。

（１９）

発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行うことを含む、制御方法。

（２０）

発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断をコンピュータに実行させる、コンピュータプログラム。

## 符号の説明

[0349] 1 コミュニケーションシステム

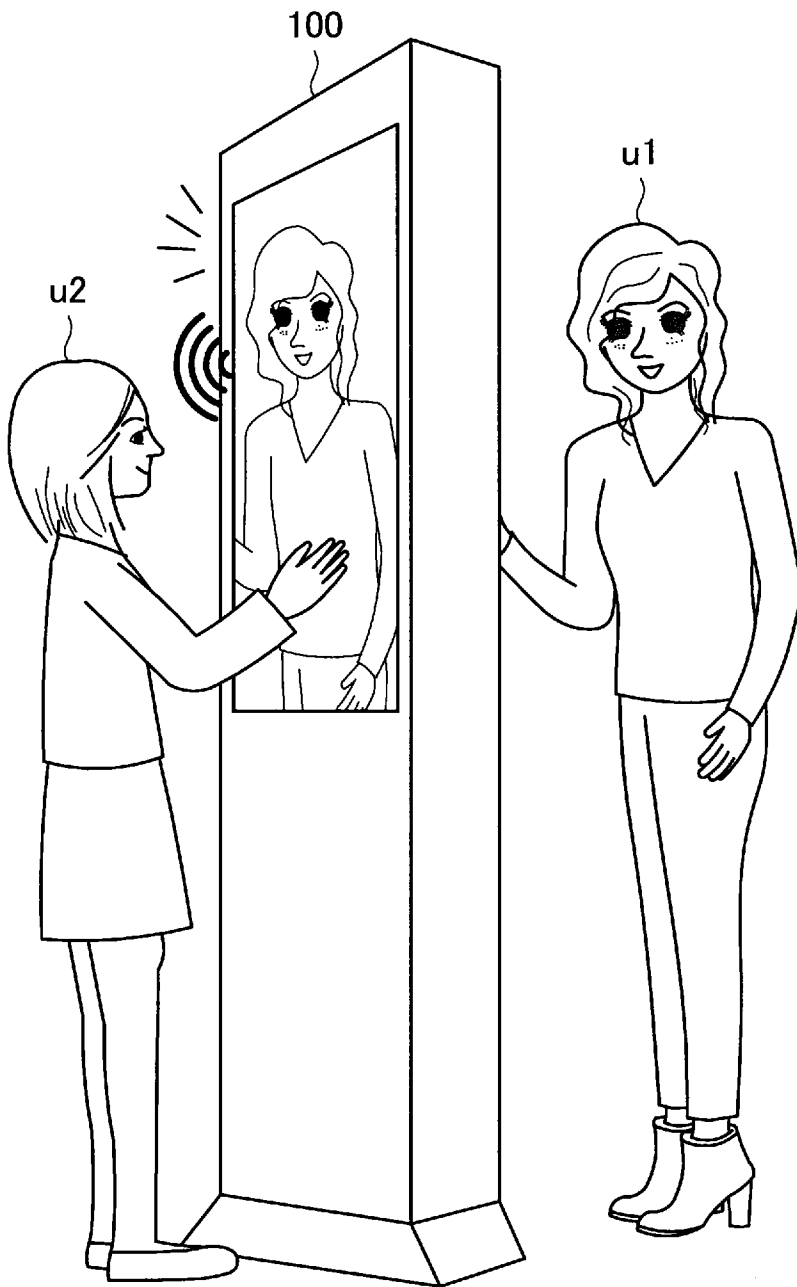
100 サイネージ端末

## 請求の範囲

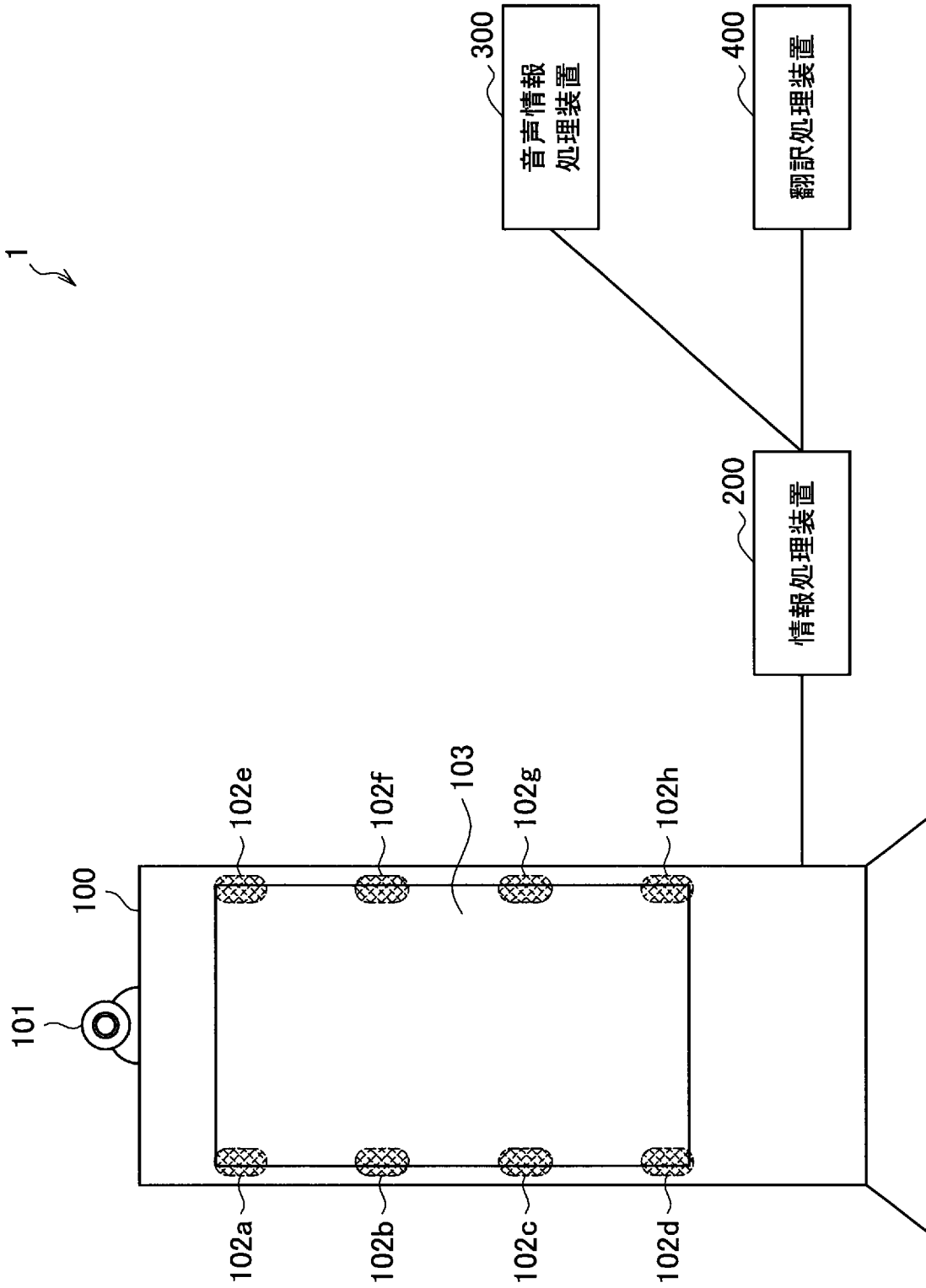
- [請求項1] 発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行う制御部を備える、制御装置。
- [請求項2] 前記制御部は、前記過去の発話内容及び前記発話状況に応じて、前記翻訳処理を終了させる、請求項1に記載の制御装置。
- [請求項3] 前記制御部は、前記発話状況として前記発話者の視線に基づいて、前記翻訳処理を終了させる、請求項2に記載の制御装置。
- [請求項4] 前記制御部は、前記発話状況として前記発話者の視線が該発話者の対話者に向いており、かつ前記発話内容として前記発話者の発話を所定時間検出しなくなると、前記翻訳処理を終了させる、請求項3に記載の制御装置。
- [請求項5] 前記制御部は、前記発話内容として所定のキーワードが含まれていれば、前記翻訳処理を終了させる、請求項2に記載の制御装置。
- [請求項6] 前記制御部は、前記発話状況として前記発話者の動作に基づいて、前記翻訳処理の継続判断を行う、請求項1に記載の制御装置。
- [請求項7] 前記制御部は、前記発話状況として前記発話者の視線に基づいて、前記翻訳処理の継続判断を行う、請求項6に記載の制御装置。
- [請求項8] 前記制御部は、前記発話状況として前記発話者の動作及び周囲の音に基づいて、前記翻訳処理の継続判断を行う、請求項6に記載の制御装置。
- [請求項9] 前記制御部は、前記発話状況として前記発話者が置かれているシチュエーションに基づいて、前記翻訳処理の継続判断を行う、請求項1に記載の制御装置。
- [請求項10] 前記制御部は、前記発話状況として前記発話者の対話相手との関係に基づいて、前記翻訳処理の継続判断を行う、請求項1に記載の制御装置。

- [請求項11] 前記制御部は、前記発話内容として前記発話者の対話相手からの発話内容に対する回答の内容に基づいて、前記翻訳処理の継続判断を行う、請求項1に記載の制御装置。
- [請求項12] 前記制御部は、前記発話内容として文章の成立の有無に基づいて前記翻訳処理の継続判断を行う、請求項1に記載の制御装置。
- [請求項13] 前記制御部は、前記発話内容として前記発話者の声質に基づいての継続判断を行う、請求項1に記載の制御装置。
- [請求項14] 前記制御部は、前記発話者の声質を周波数に基づいて判断する、請求項13に記載の制御装置。
- [請求項15] 前記制御部は、前記レベルとして、前記発話内容及び前記発話状況に基づいて算出したスコアを用いる、請求項1に記載の制御装置。
- [請求項16] 前記制御部は、算出した前記スコアが所定値以下の場合に、前記発話者の視線が前記発話者の対話者に向いているときは、前記翻訳処理を継続する、請求項15に記載の制御装置。
- [請求項17] 前記制御部は、前記スコアを、前記発話者の視線に基づいて算出する、請求項15に記載の制御装置。
- [請求項18] 前記制御部は、前記スコアを、前記発話者が発話した内容の翻訳結果の確からしさに基づいて算出する、請求項15に記載の制御装置。
- [請求項19] 発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断を行うことを含む、制御方法。
- [請求項20] 発話者による発話内容が、発話内容に対する翻訳処理を継続すべきかどうかを判断するレベルとなった場合に、前記発話者の発話内容及び発話状況に基づいて、前記翻訳処理の継続判断をコンピュータに実行させる、コンピュータプログラム。

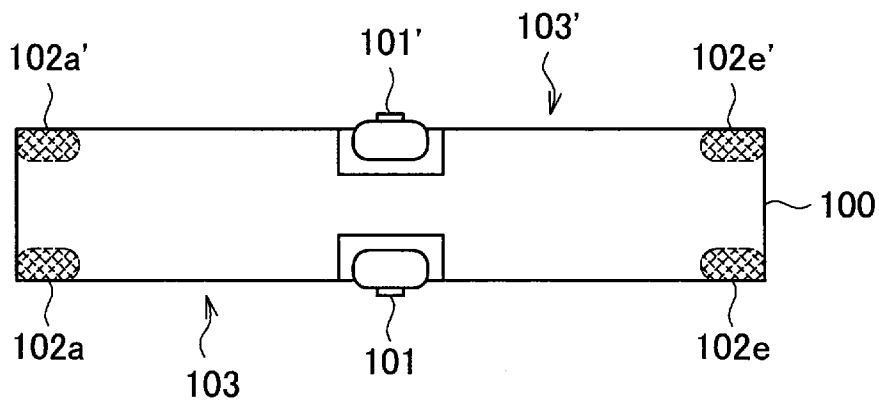
[図1]



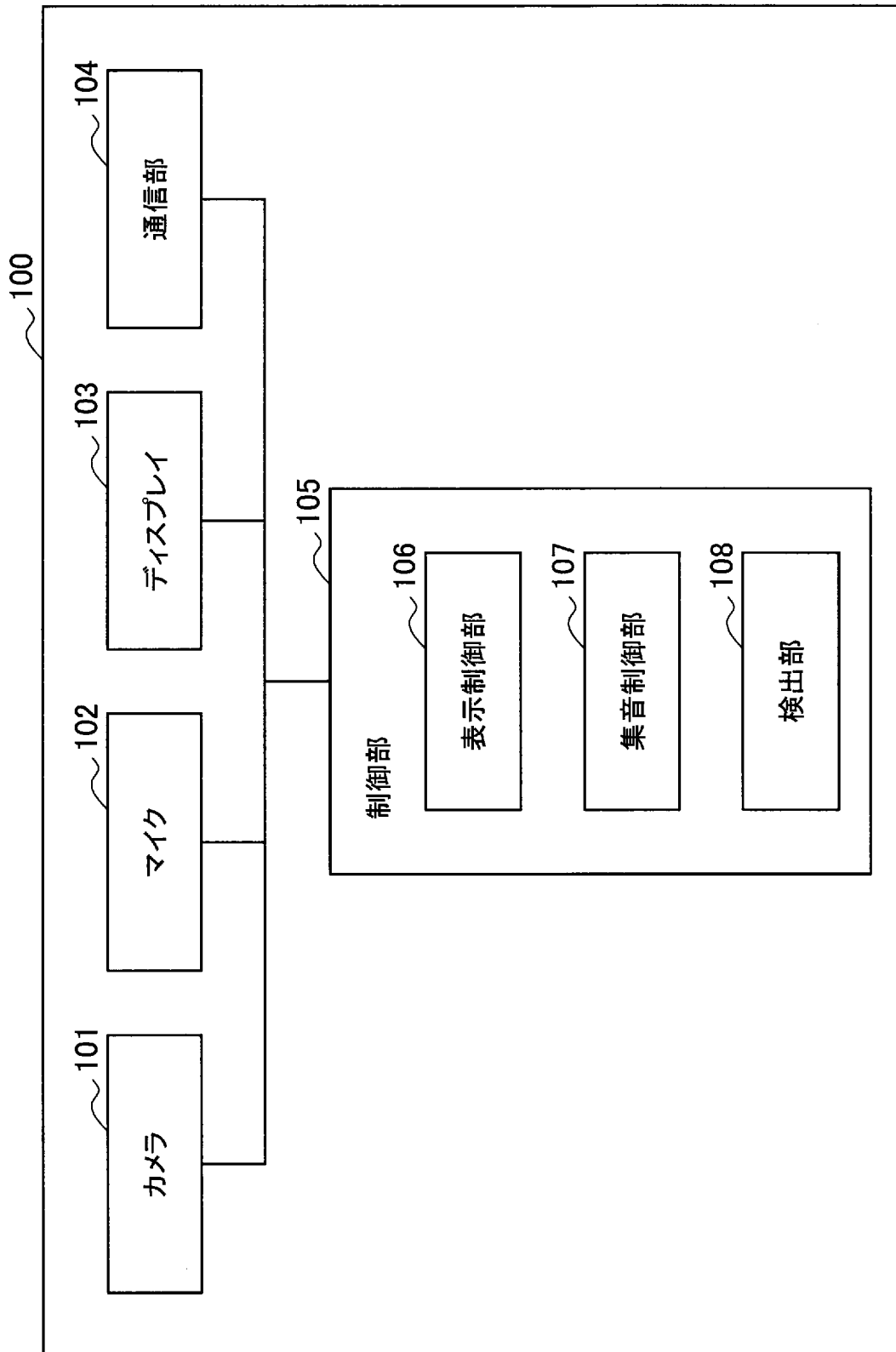
[図2]



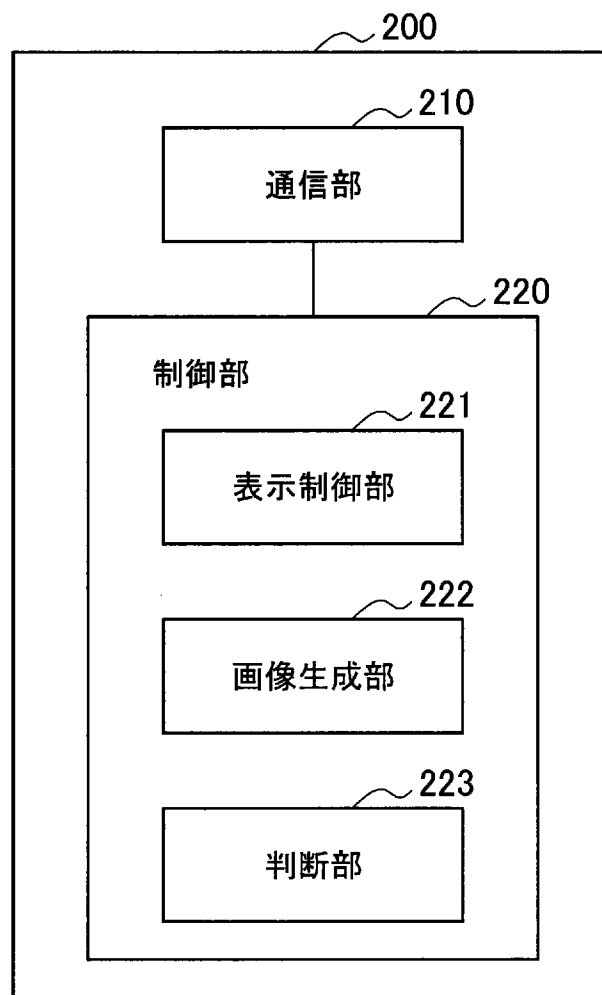
[図3]



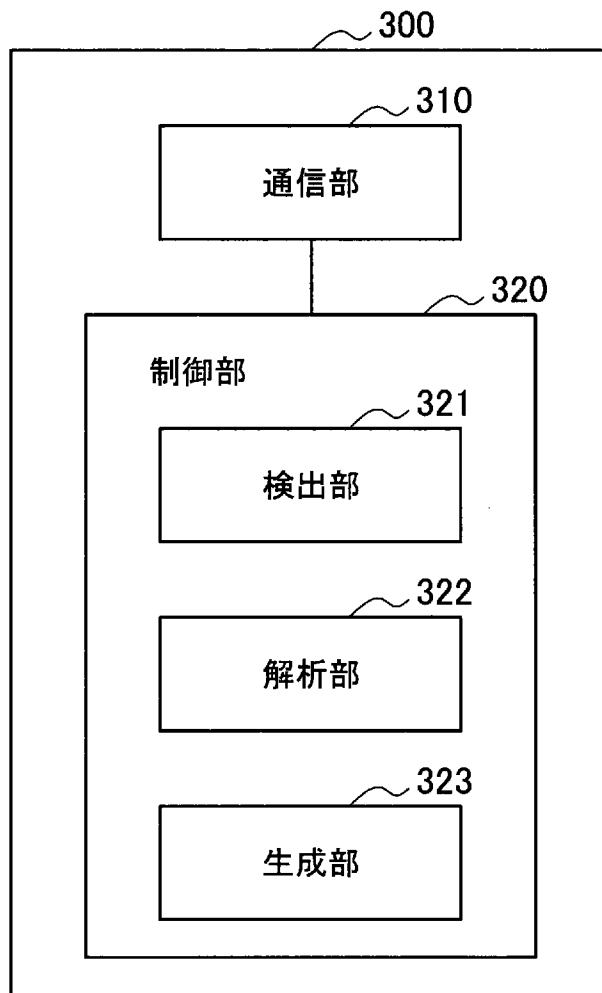
[図4]



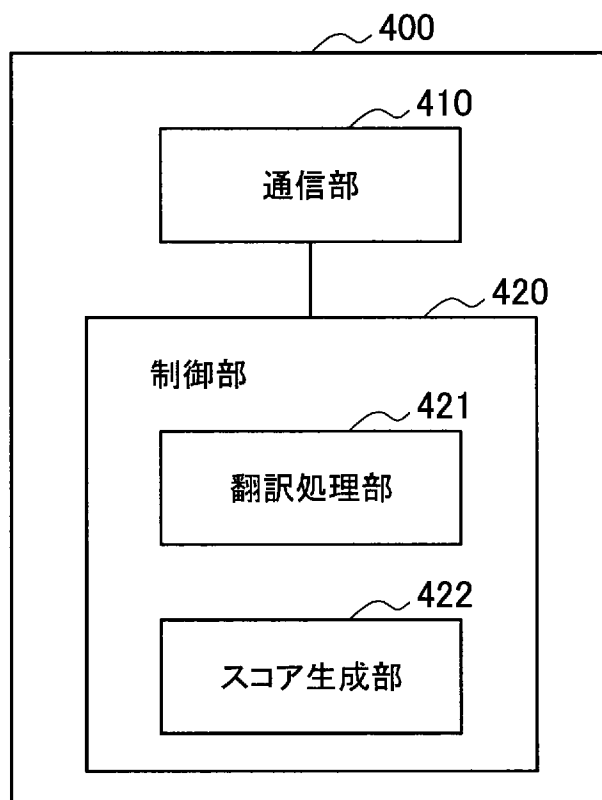
[図5]



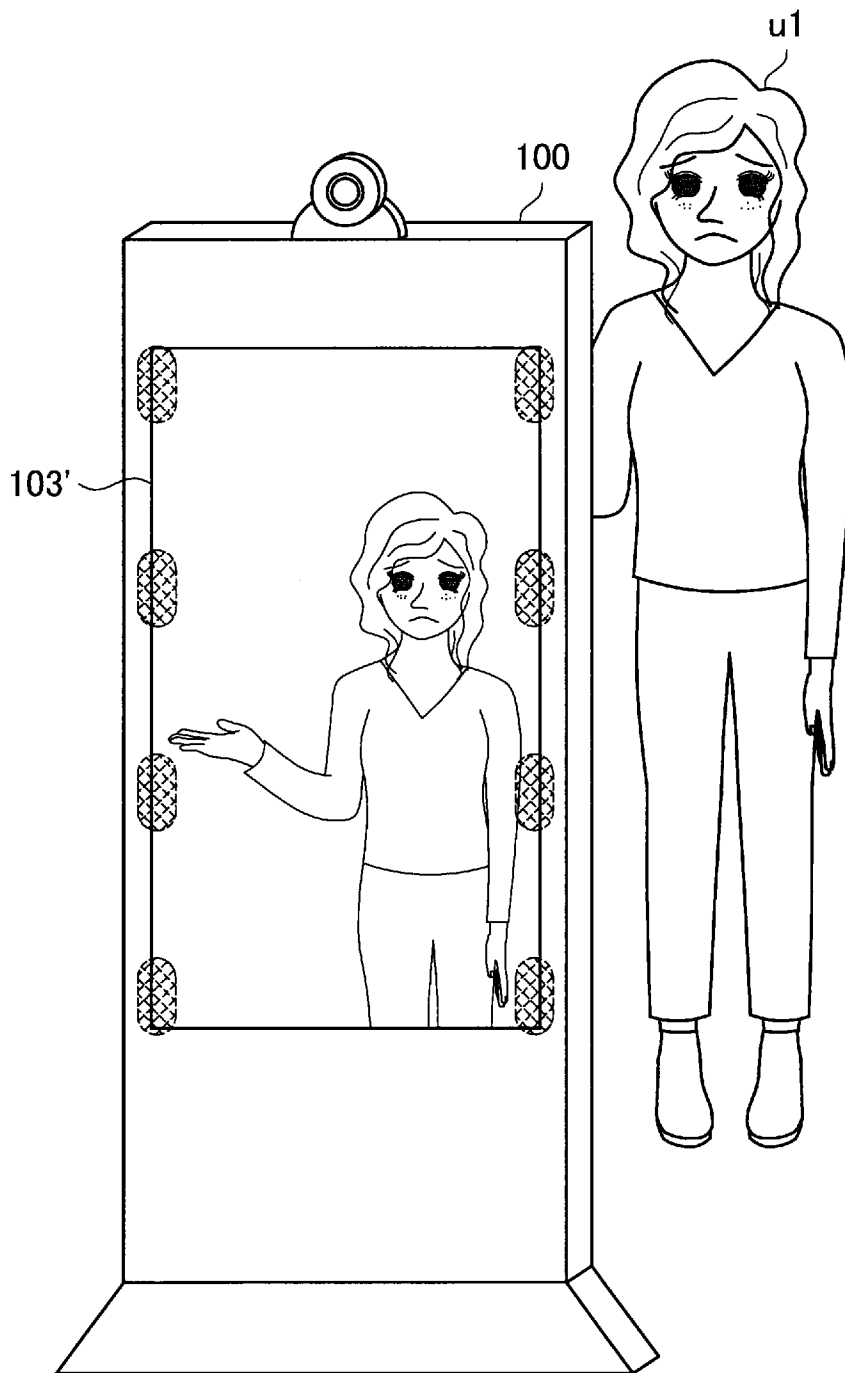
[図6]



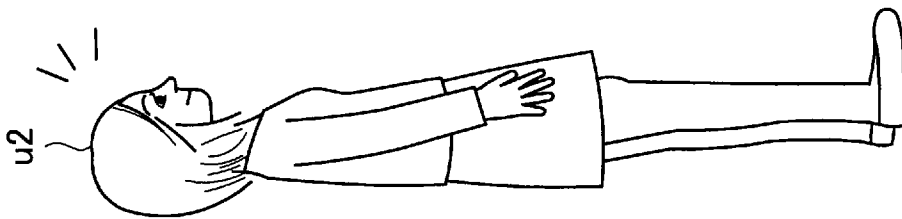
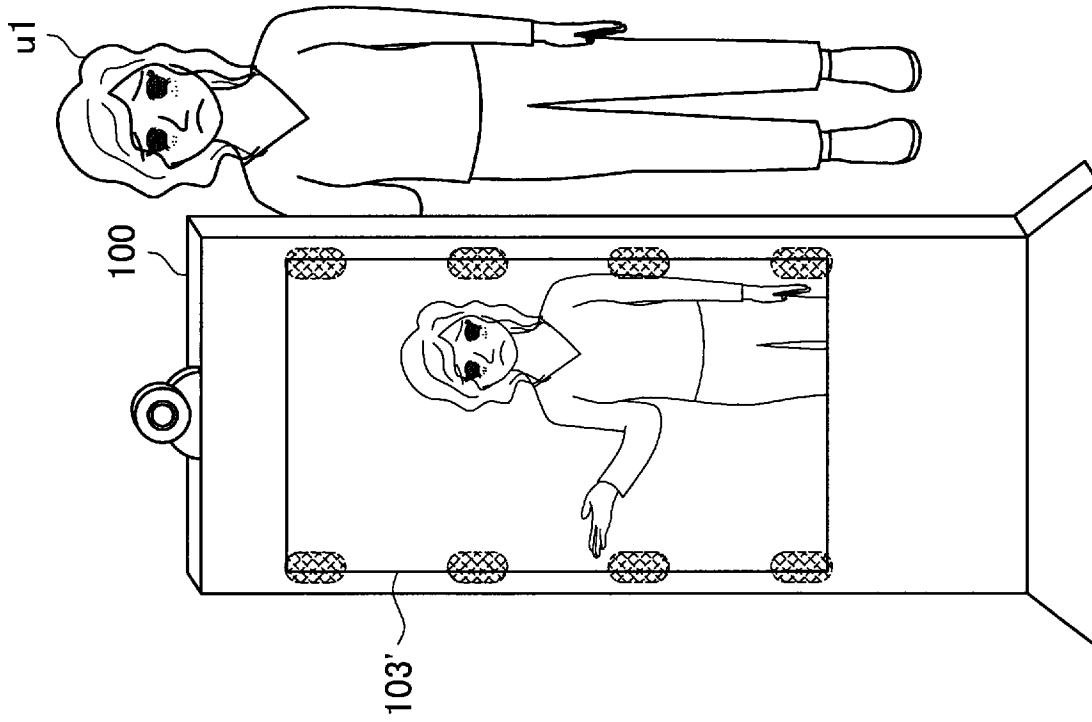
[図7]



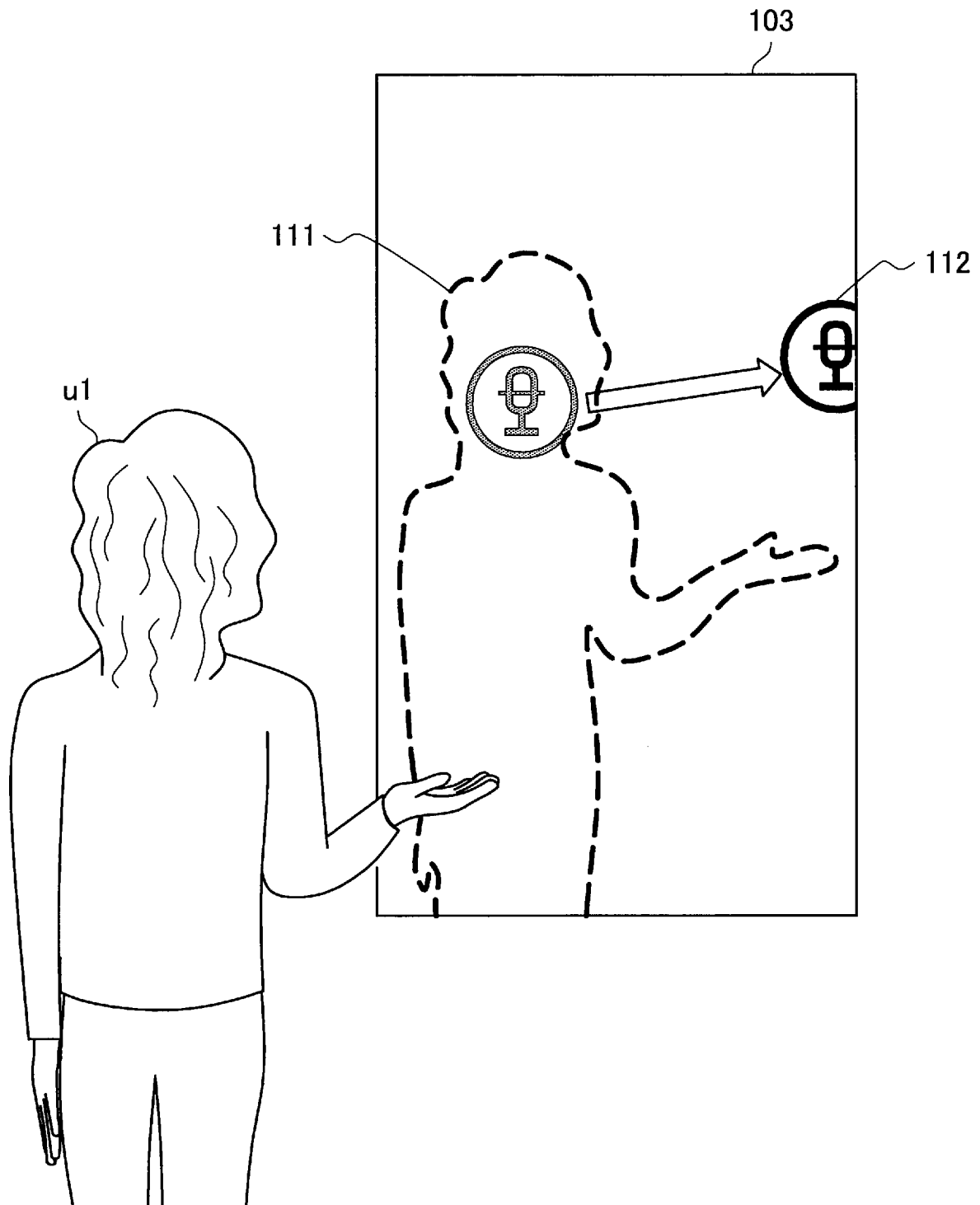
[図8]



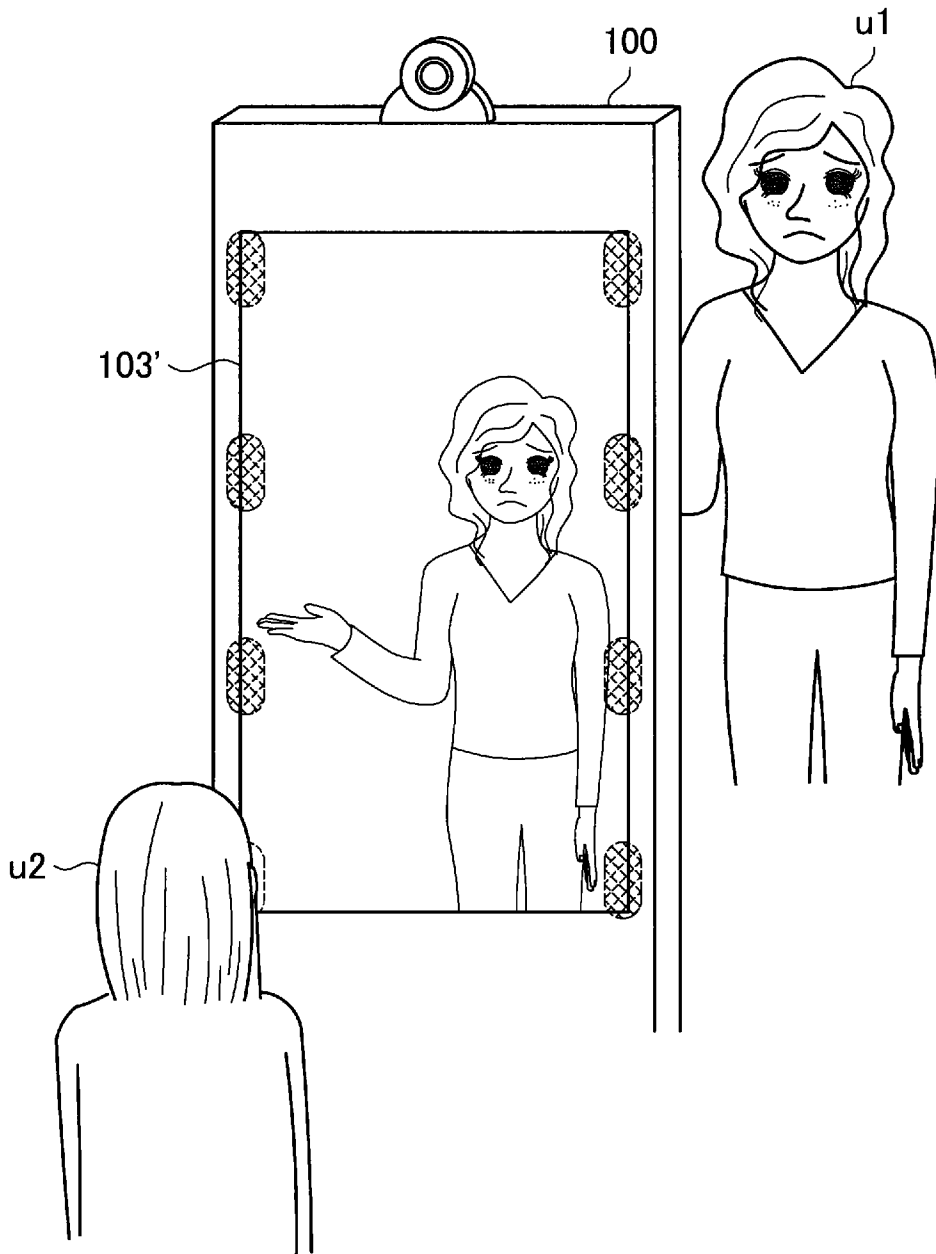
[図9]



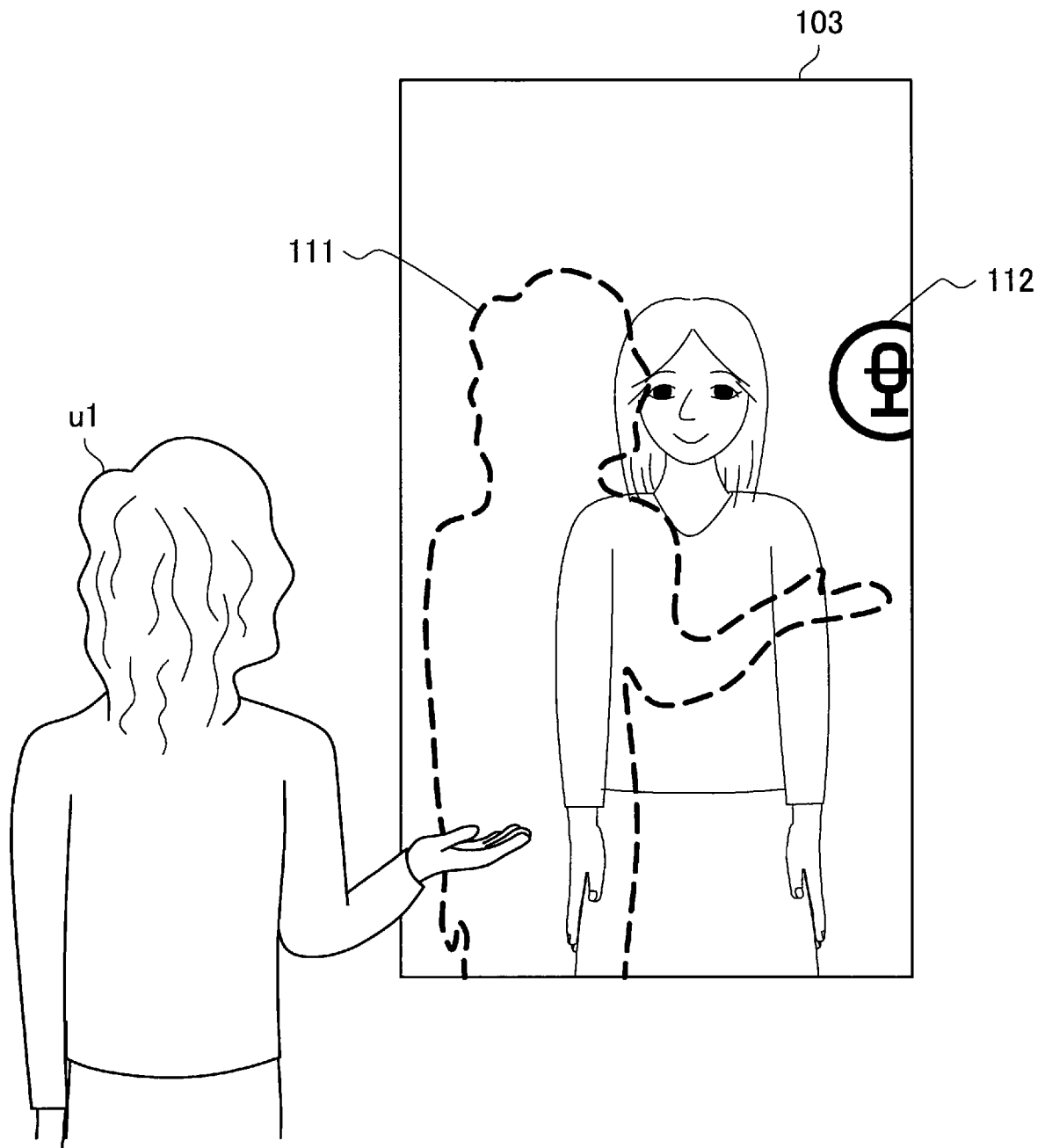
[図10]



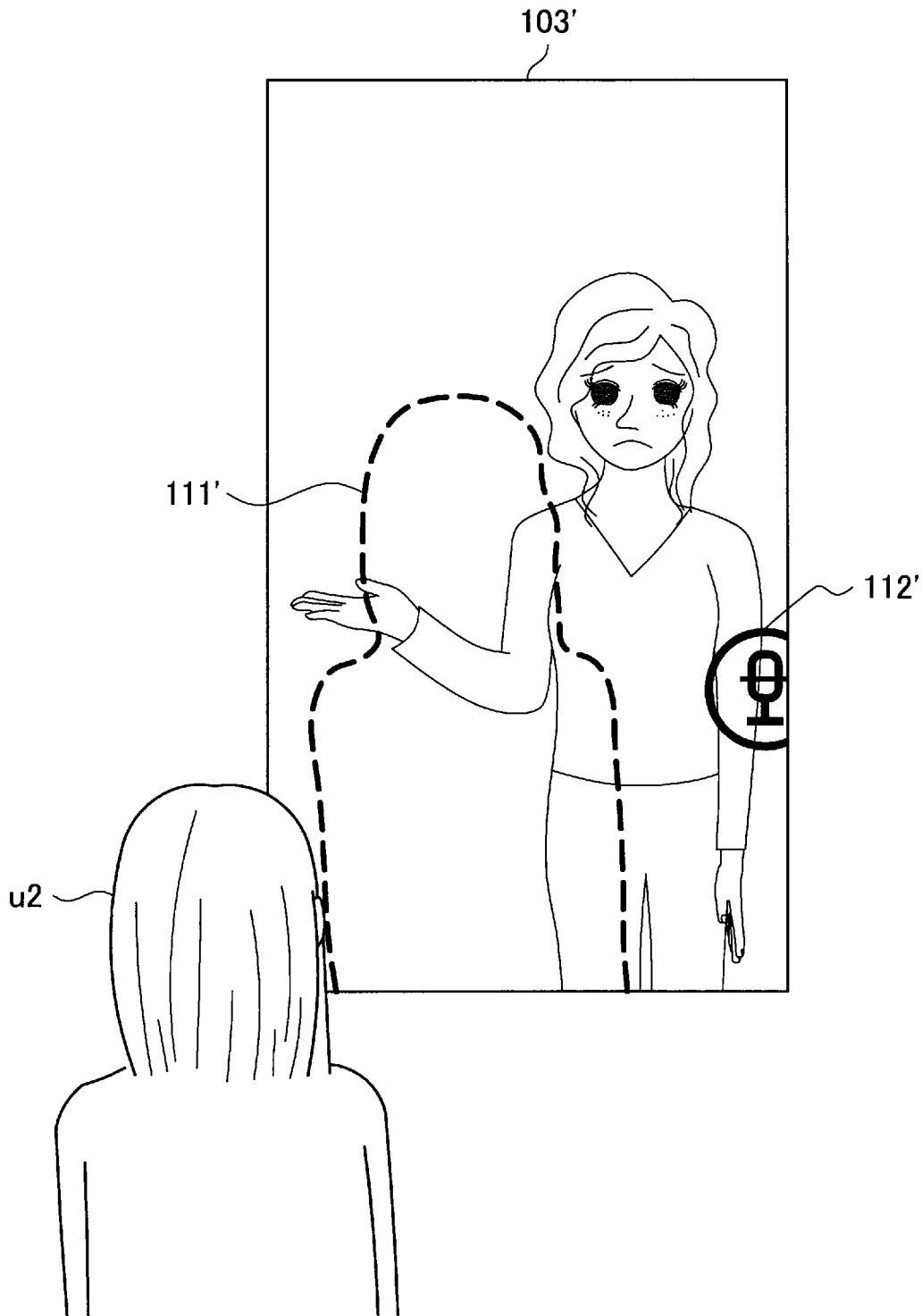
[図11]



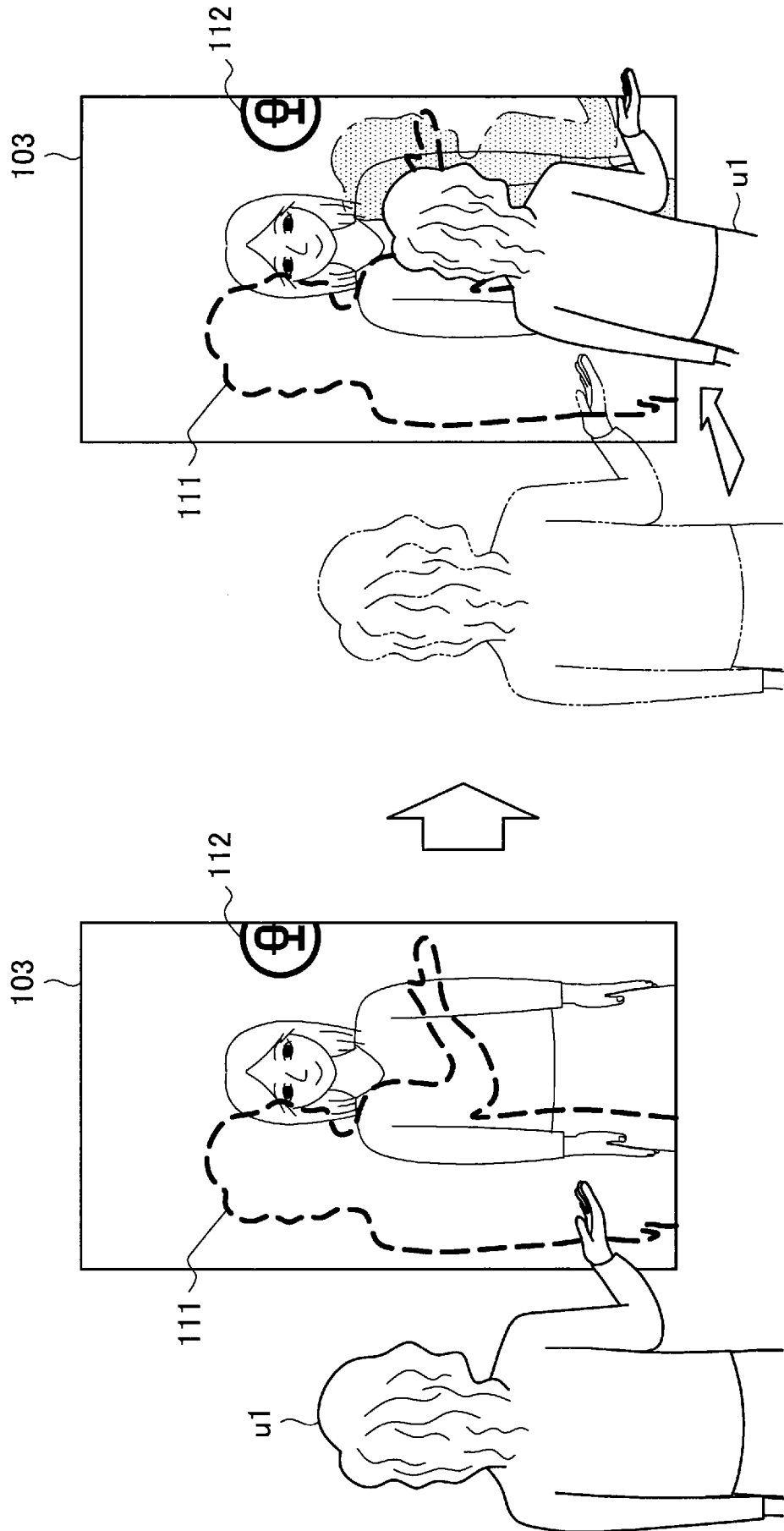
[図12]



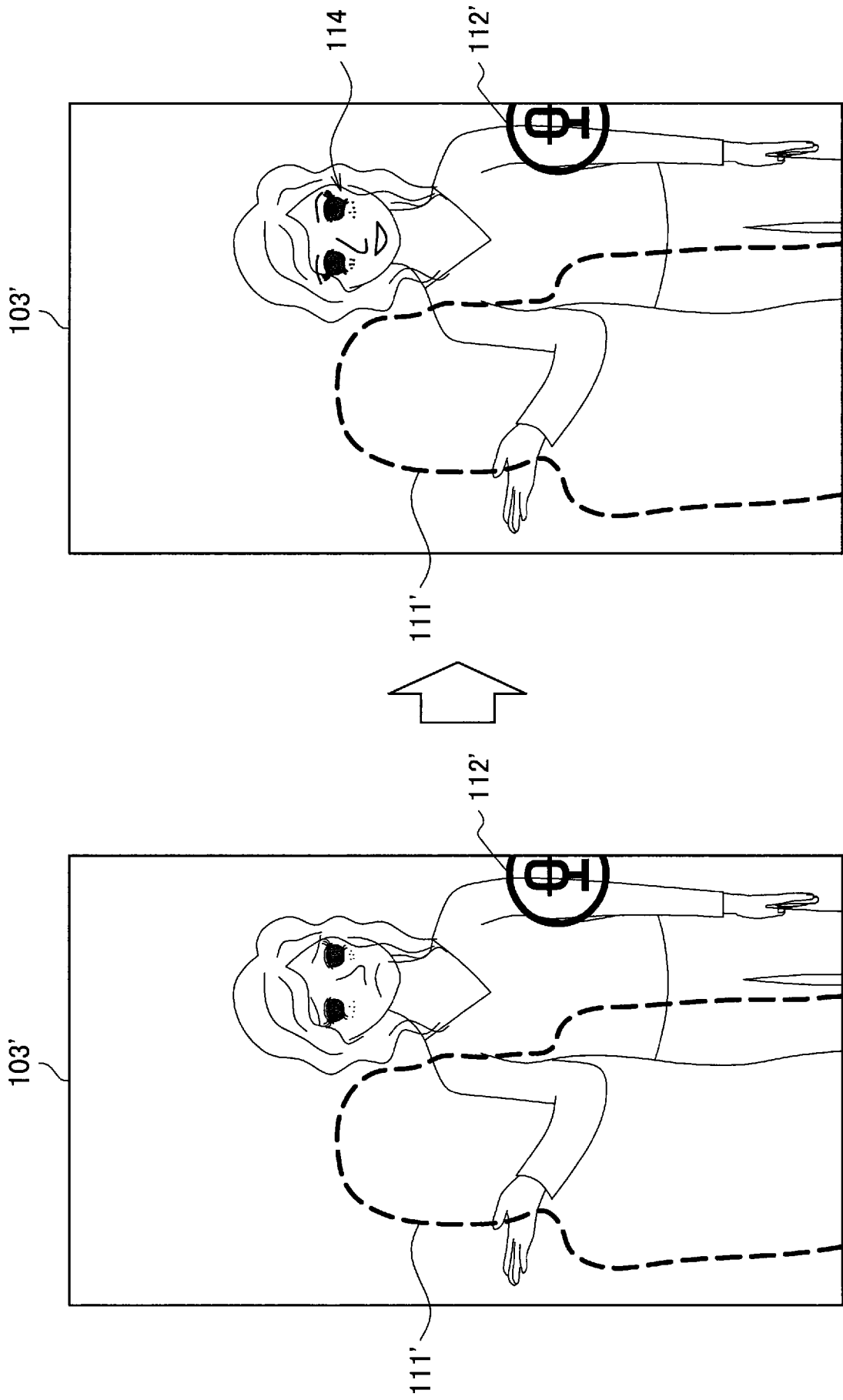
[図13]



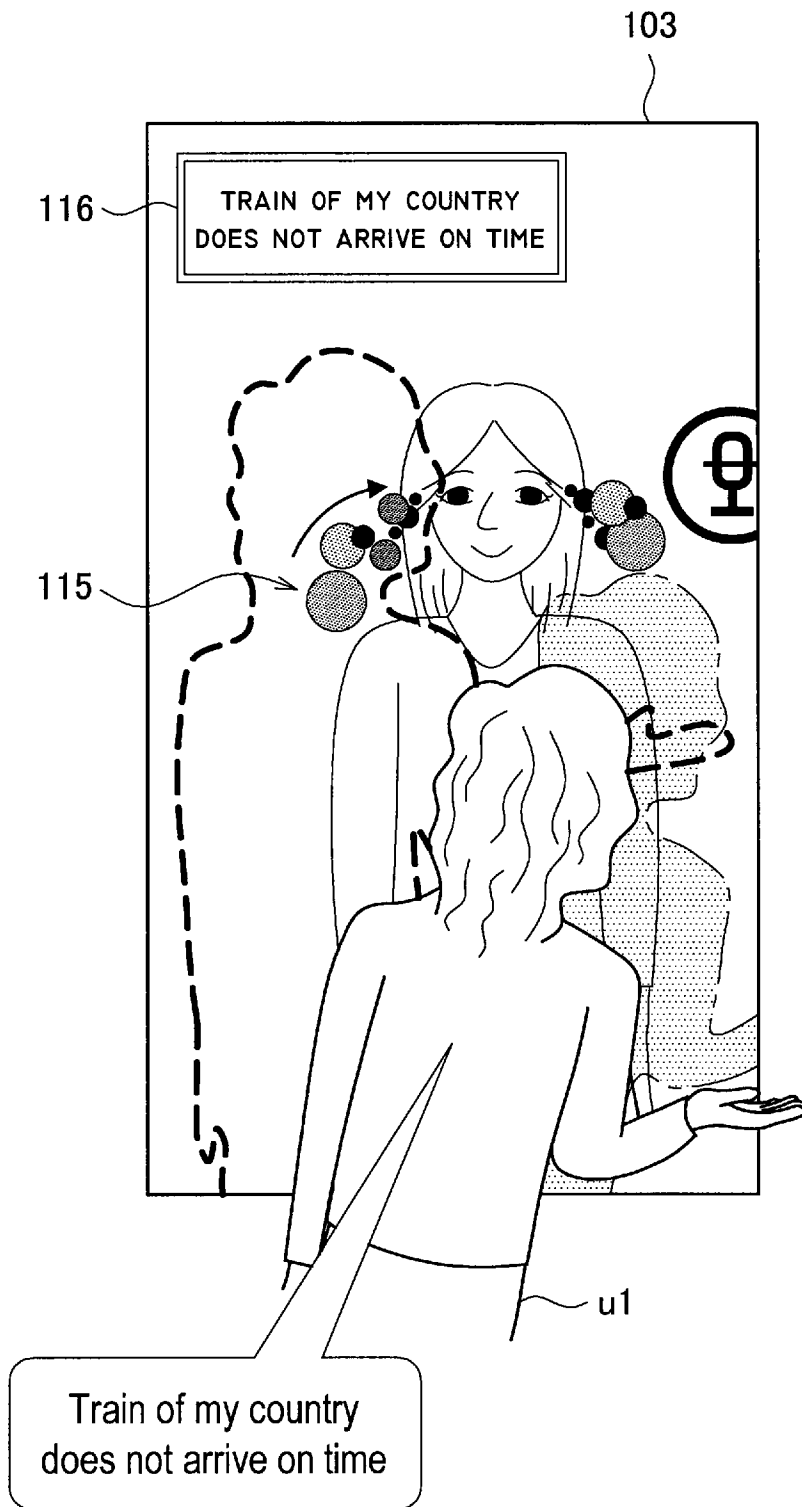
[図14]



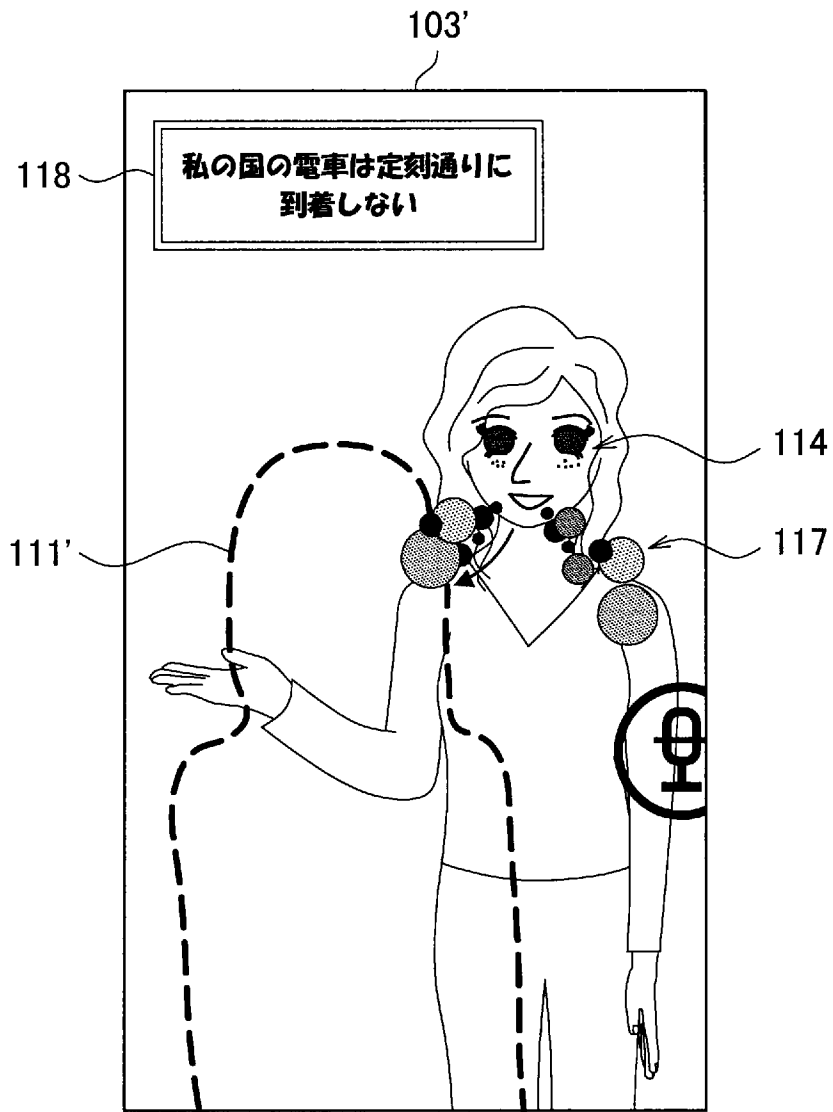
[図15]



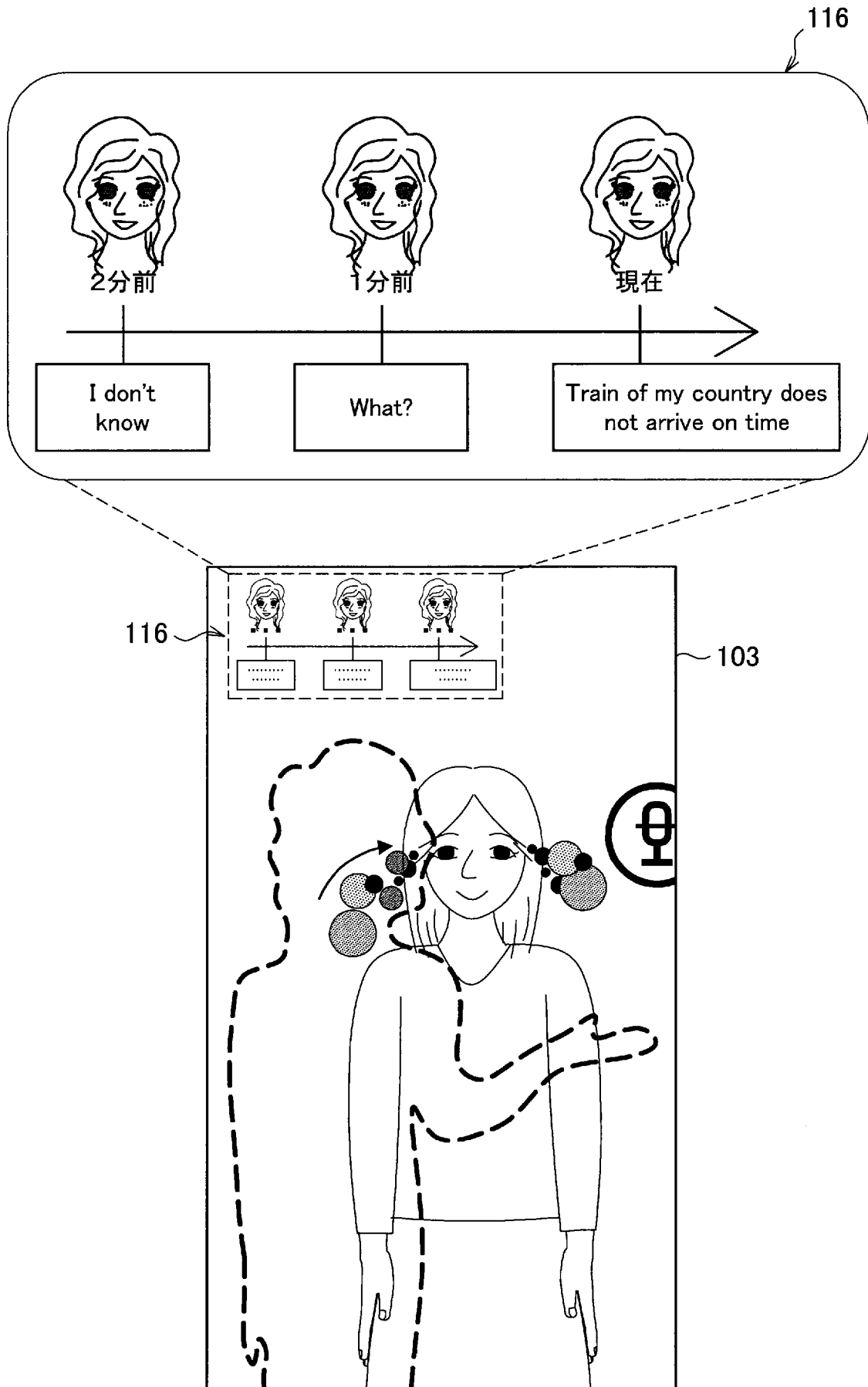
[図16]



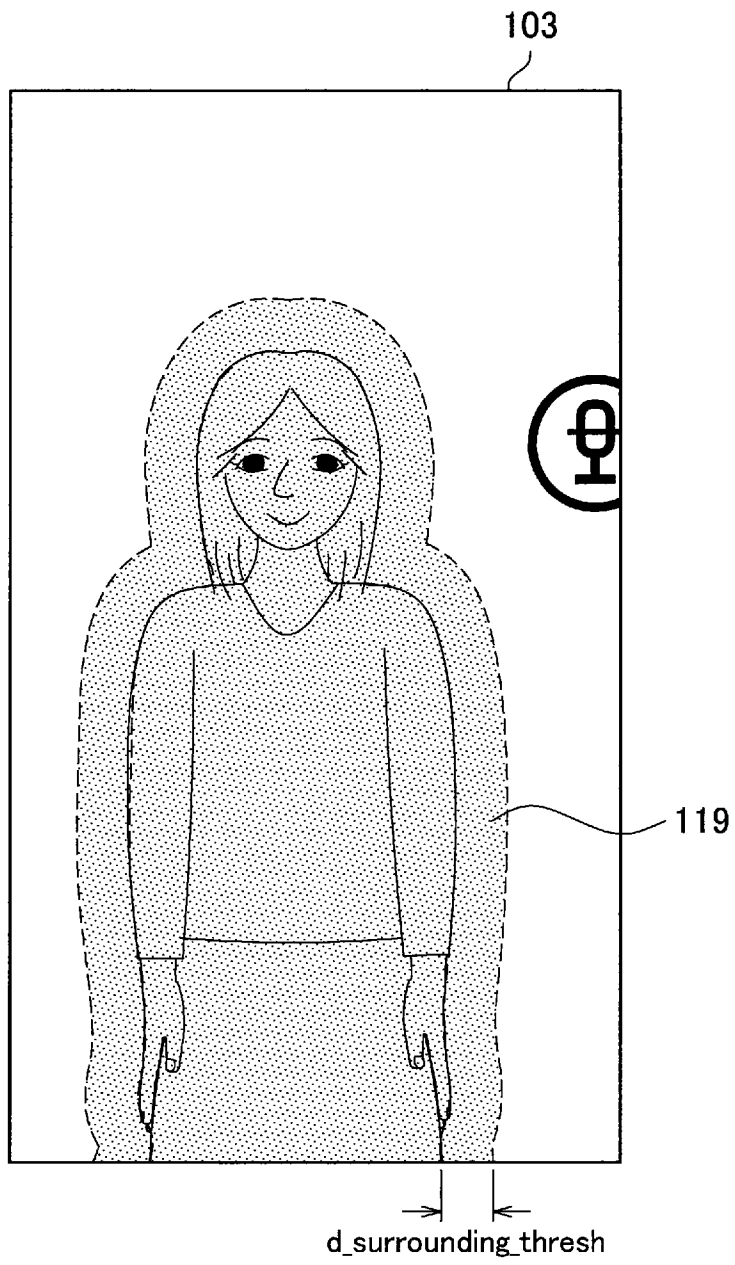
[図17]



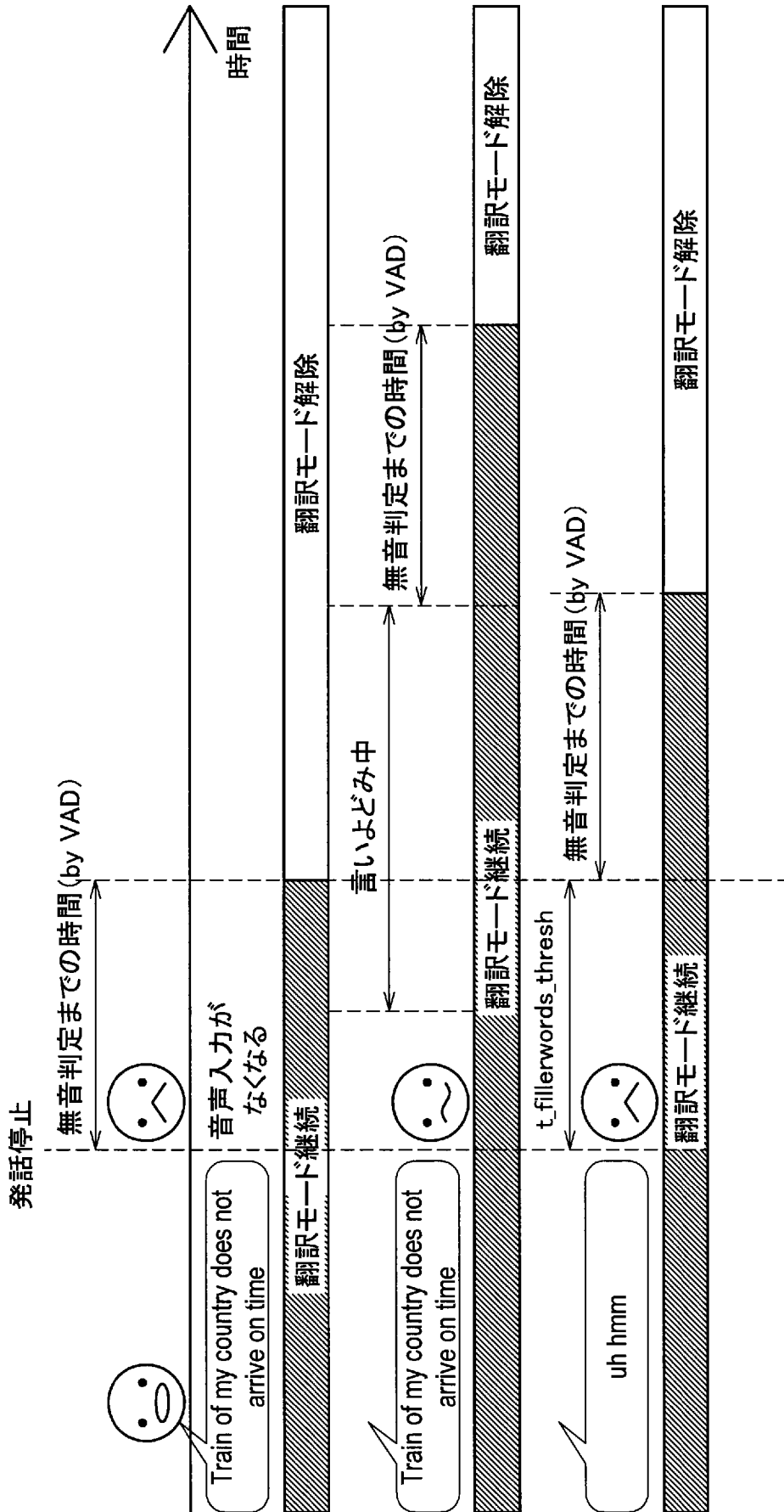
[図18]



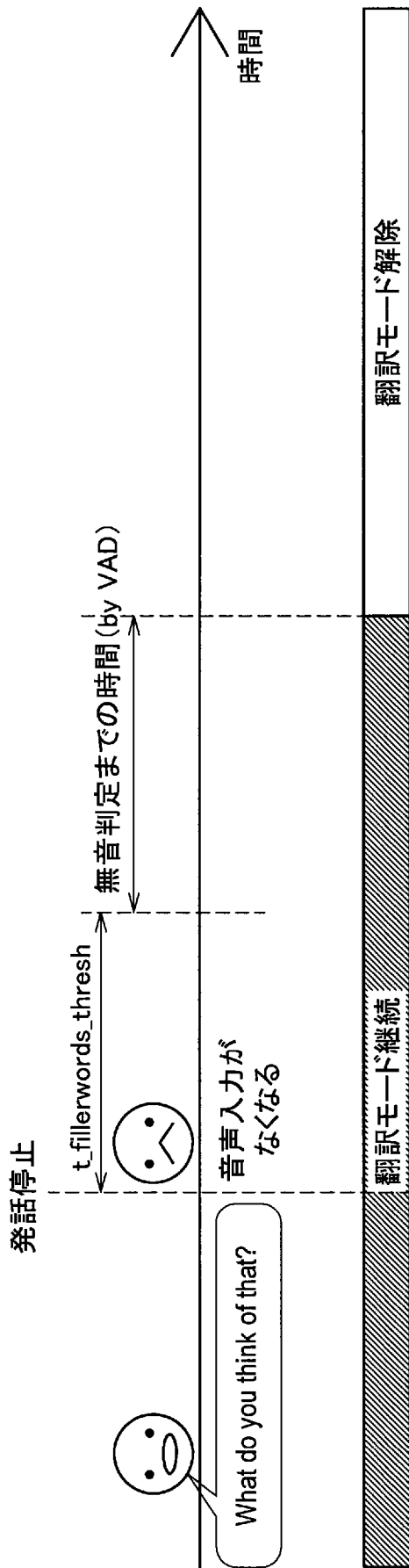
[図19]



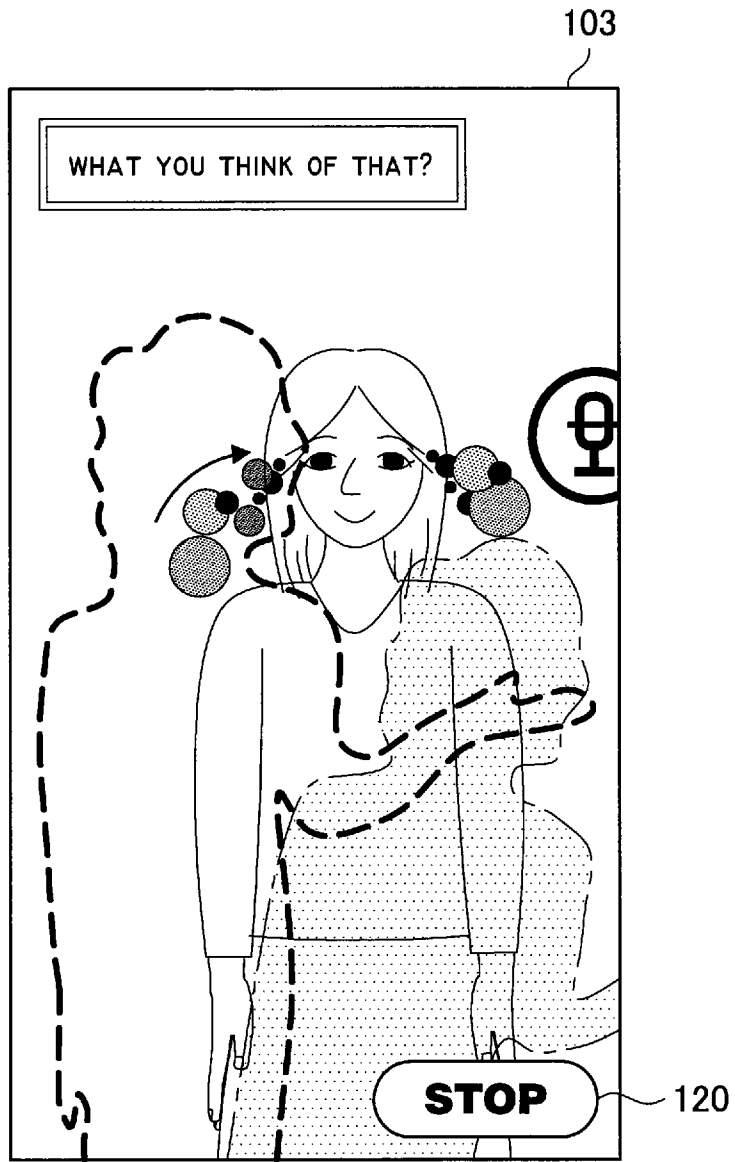
[図20]



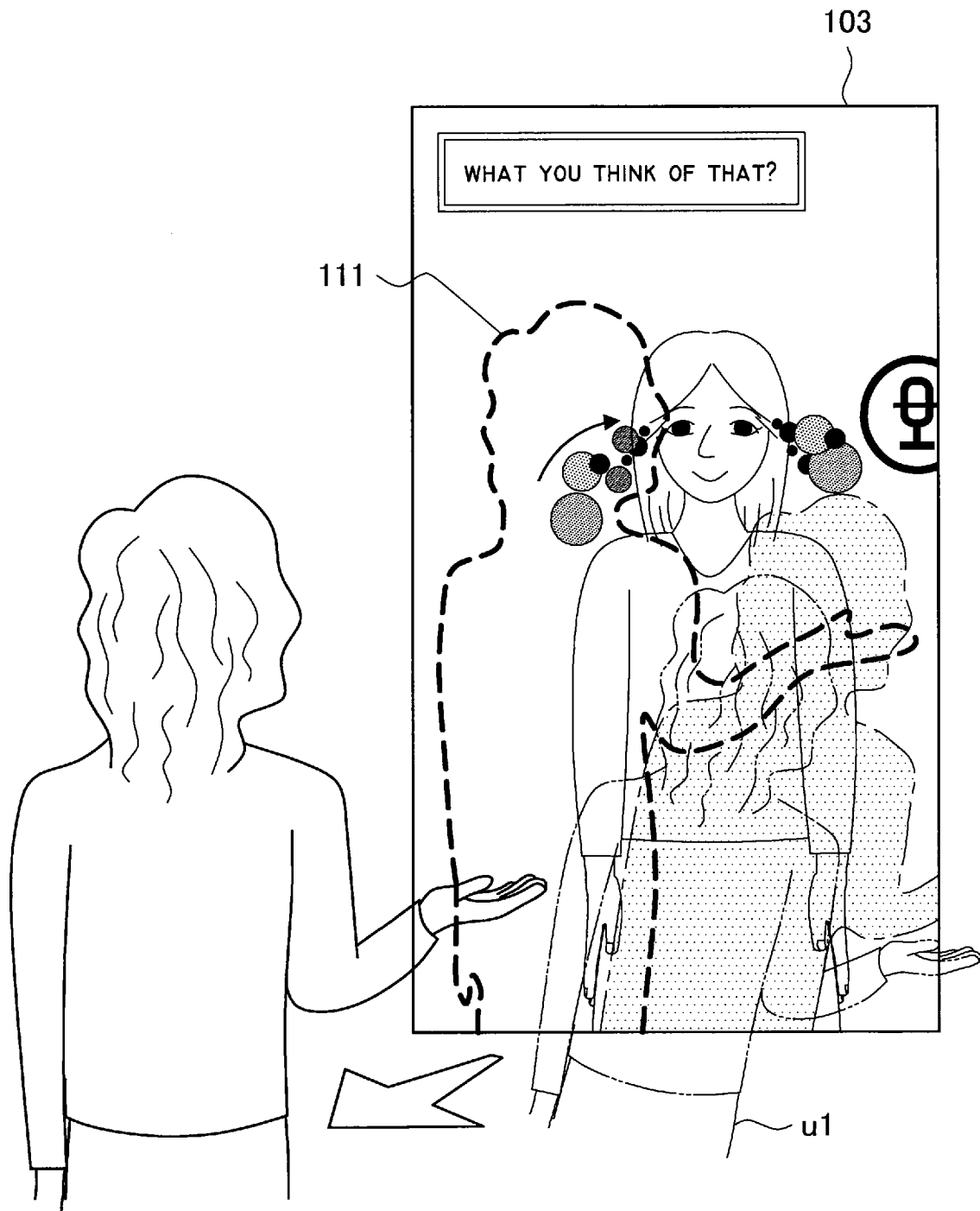
[図21]



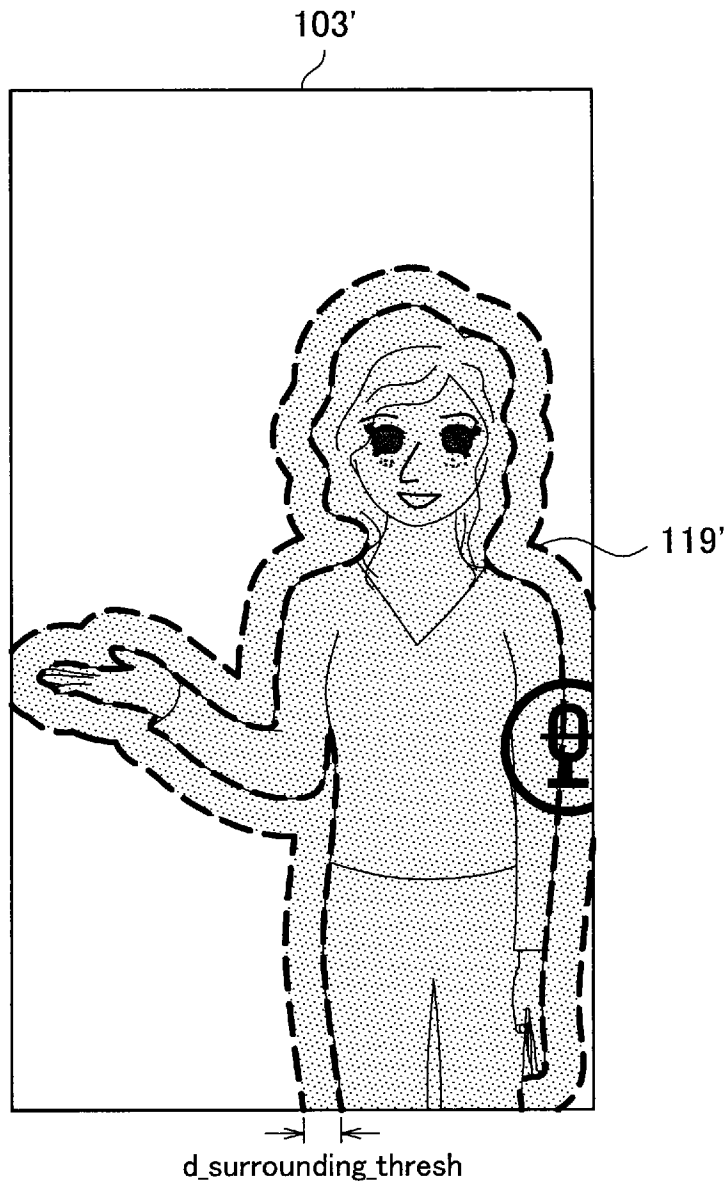
[図22]



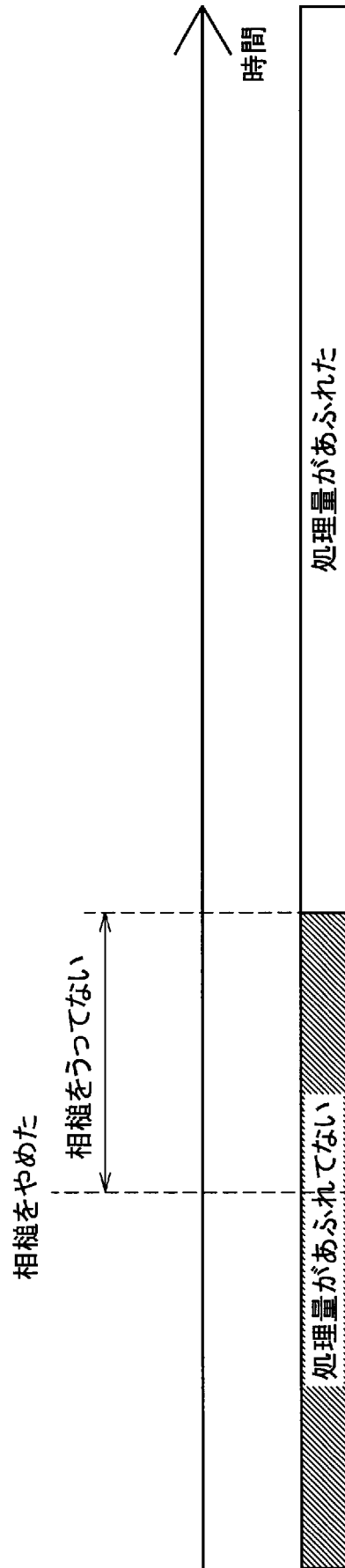
[図23]



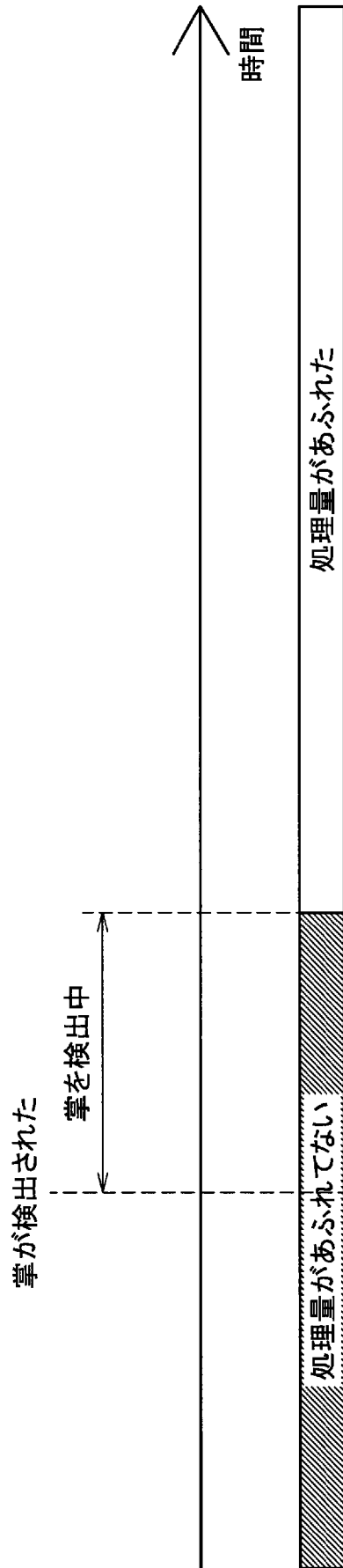
[図24]



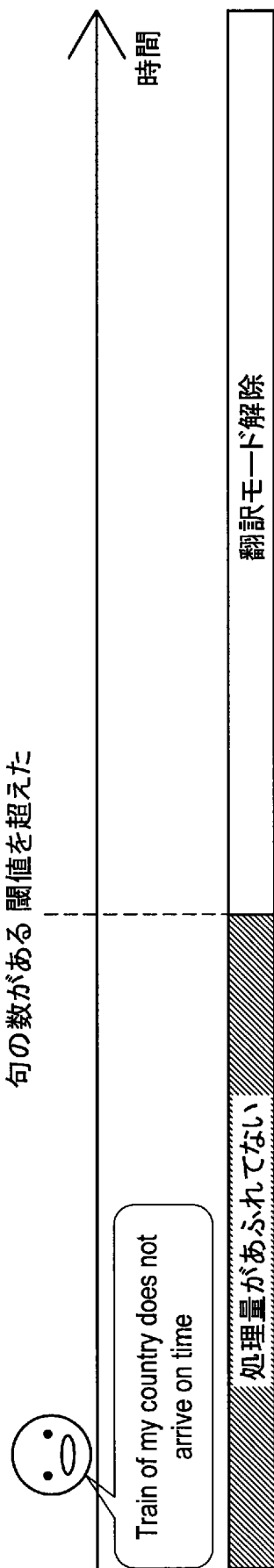
[図25]



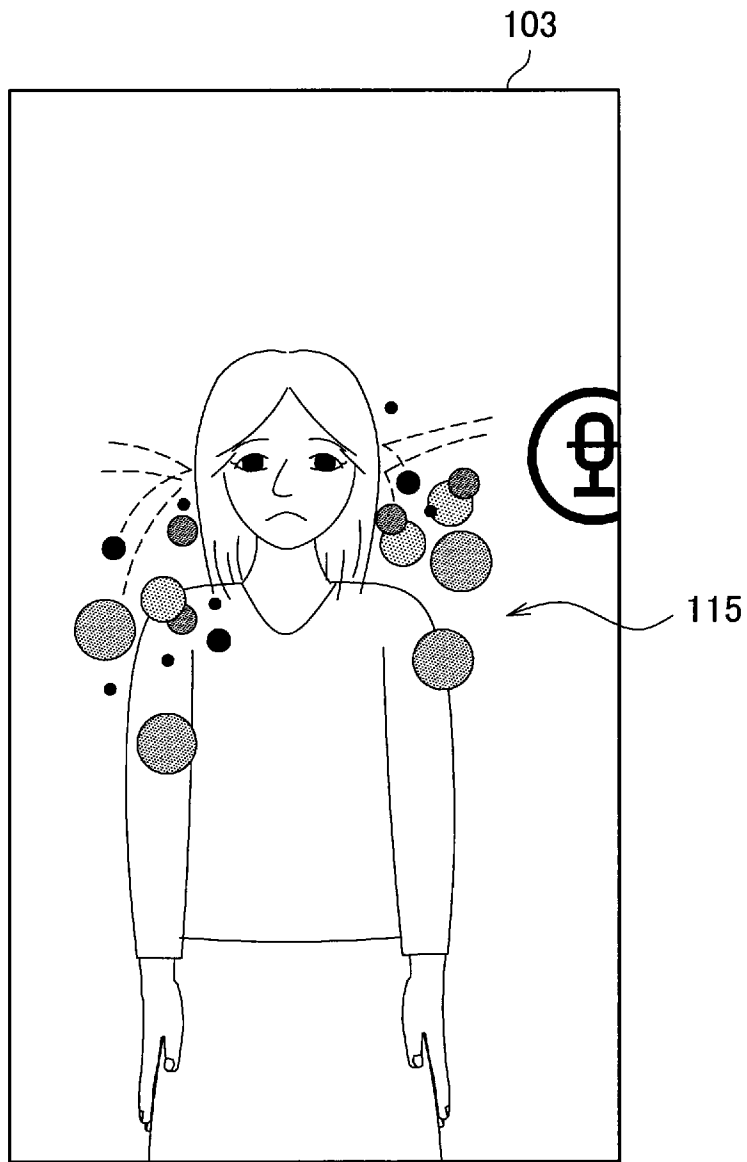
[図26]



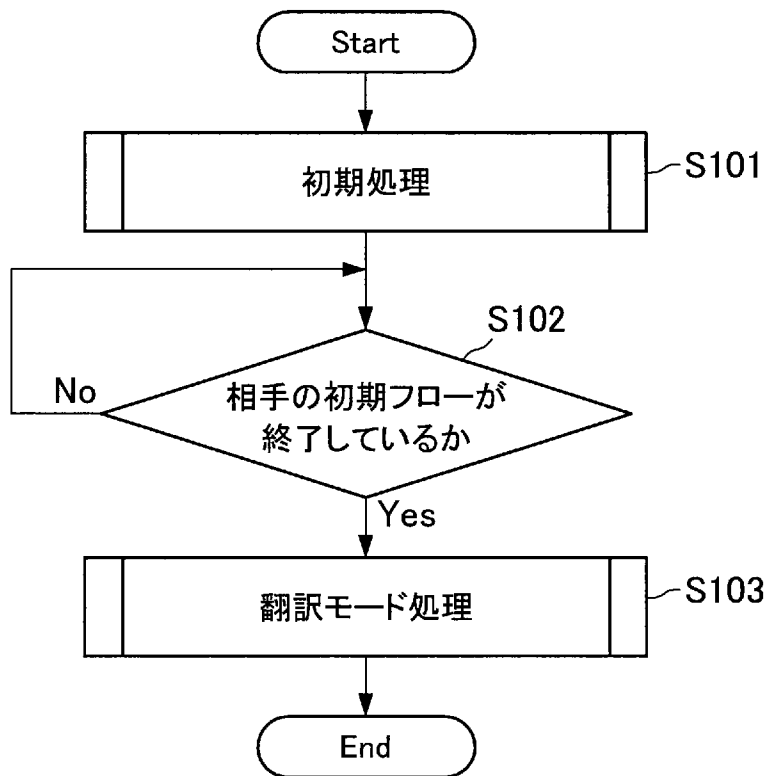
[図27]



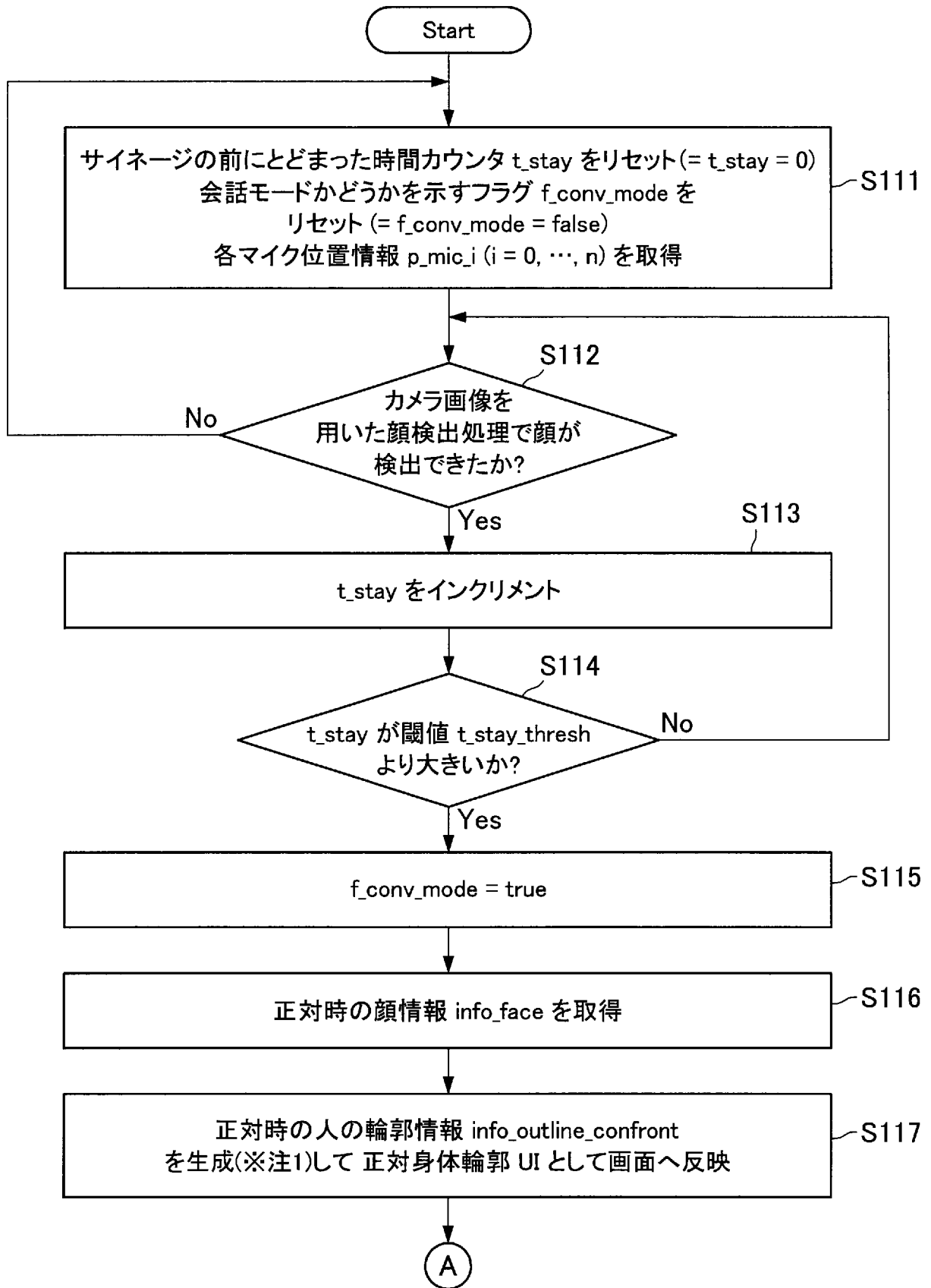
[図28]



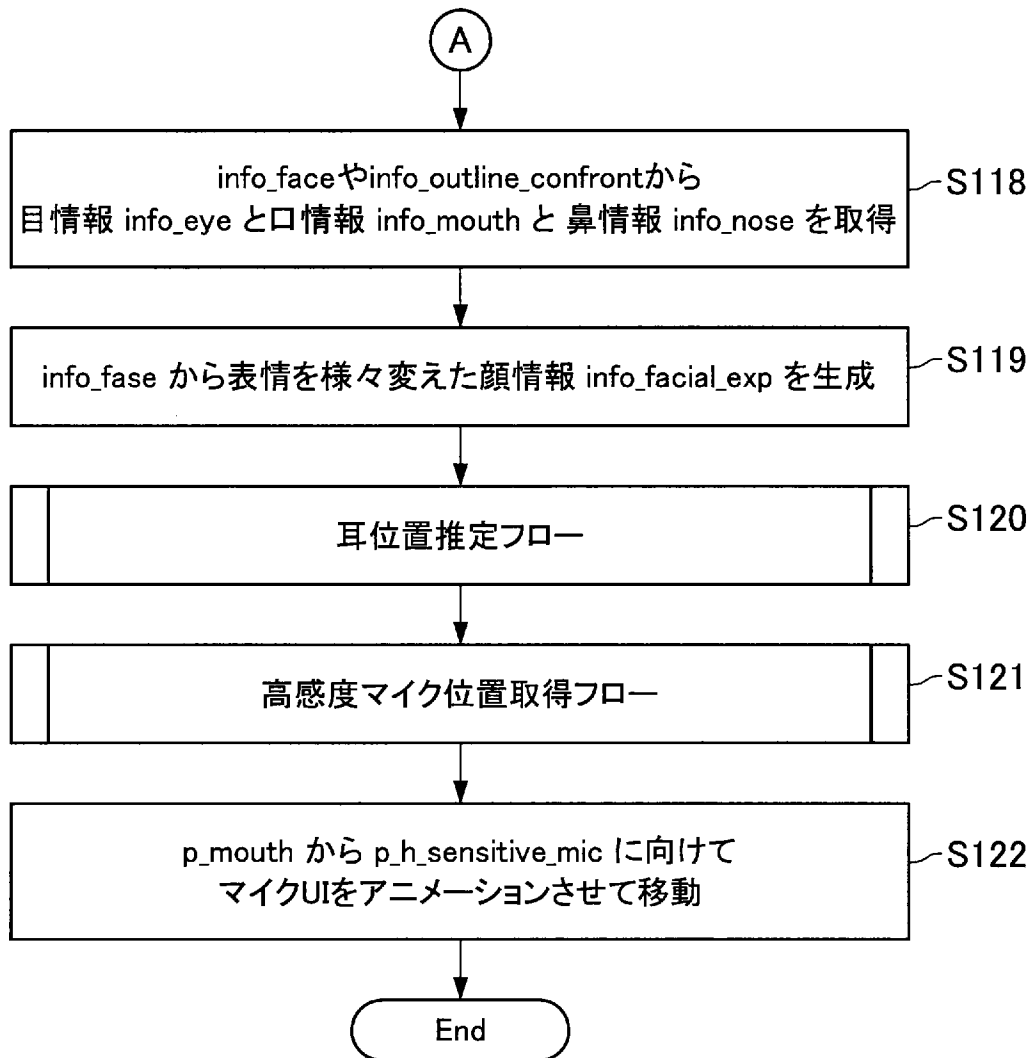
[図29]



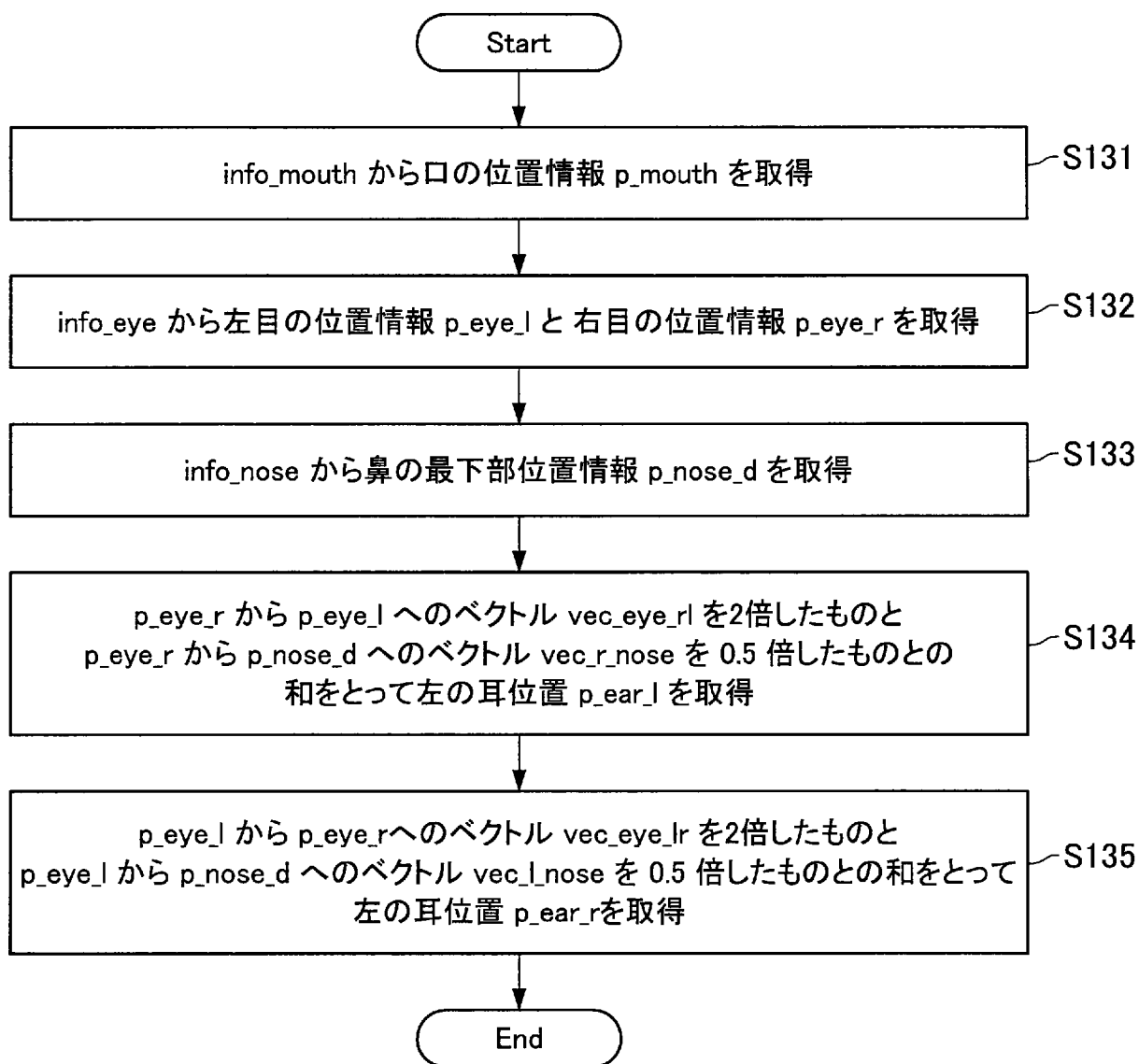
[図30A]



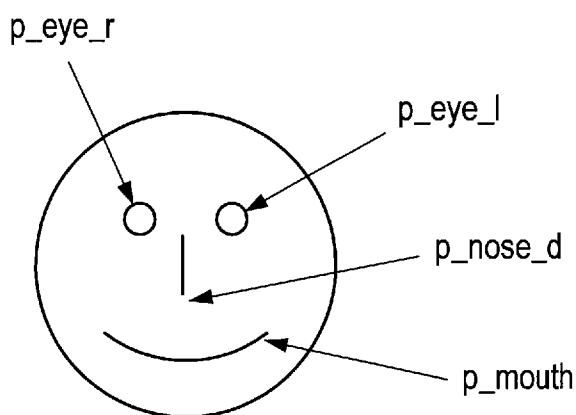
[図30B]



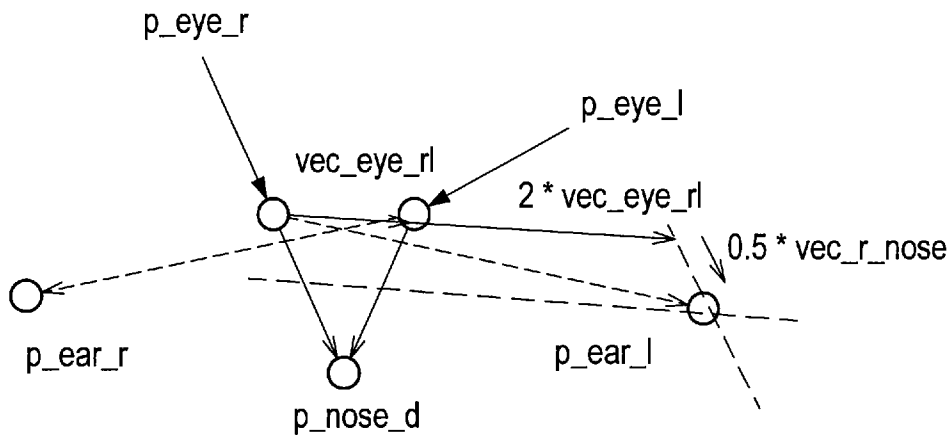
[図31]



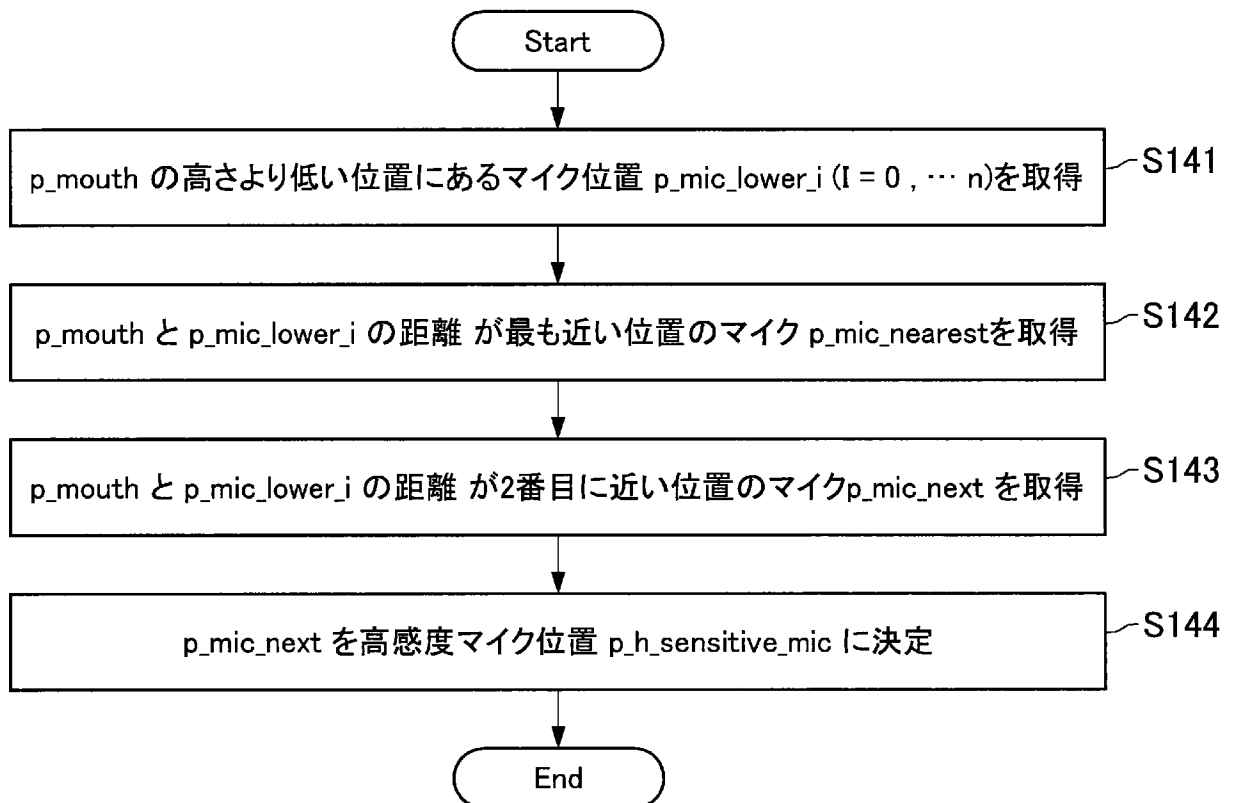
[図32]



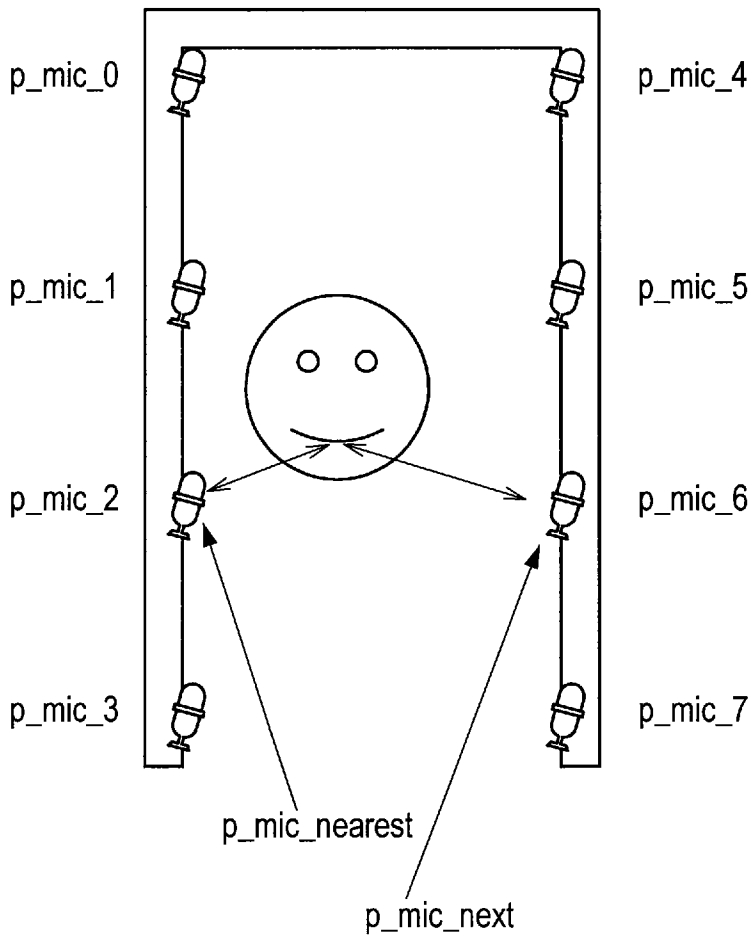
[図33]



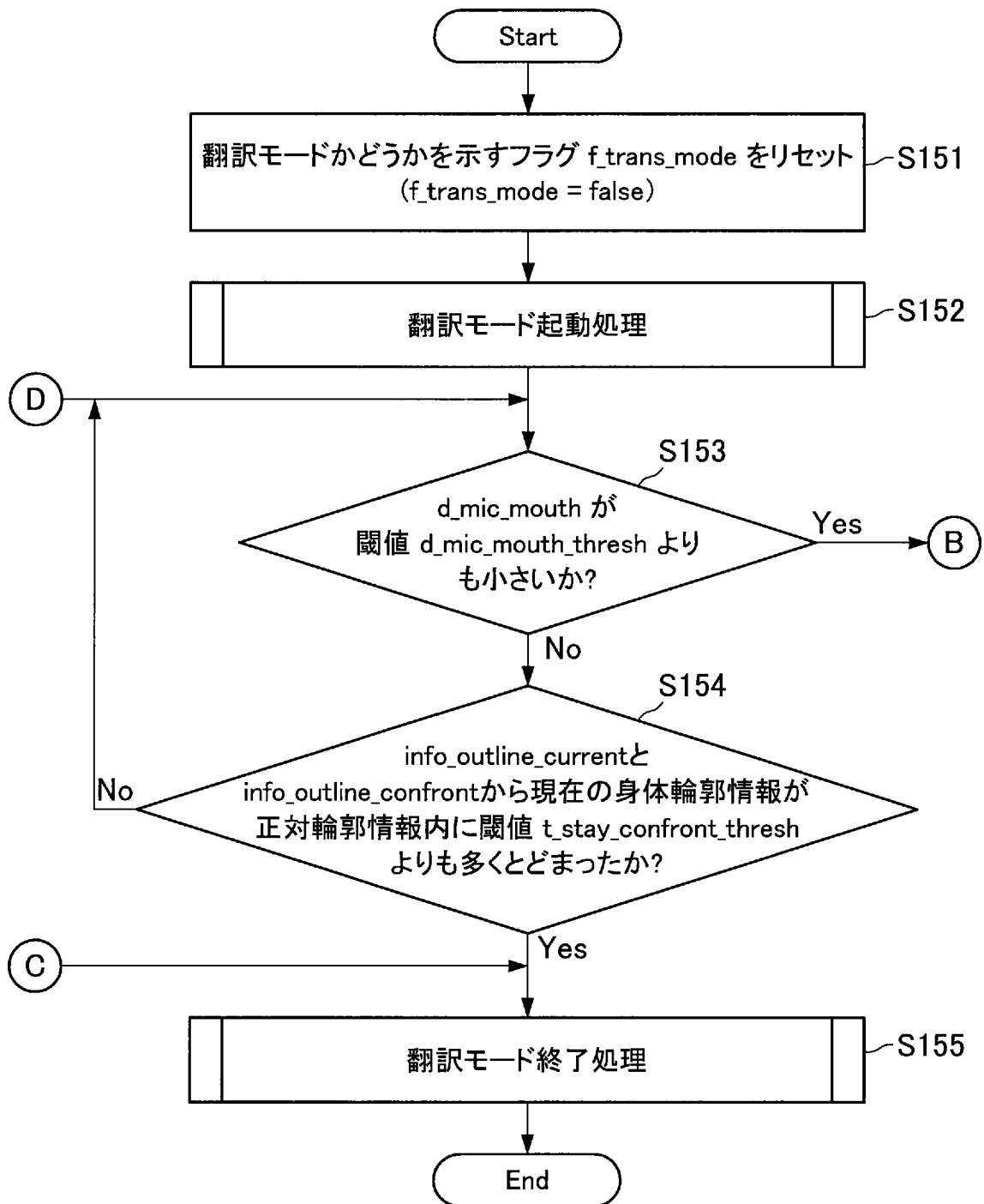
[図34]



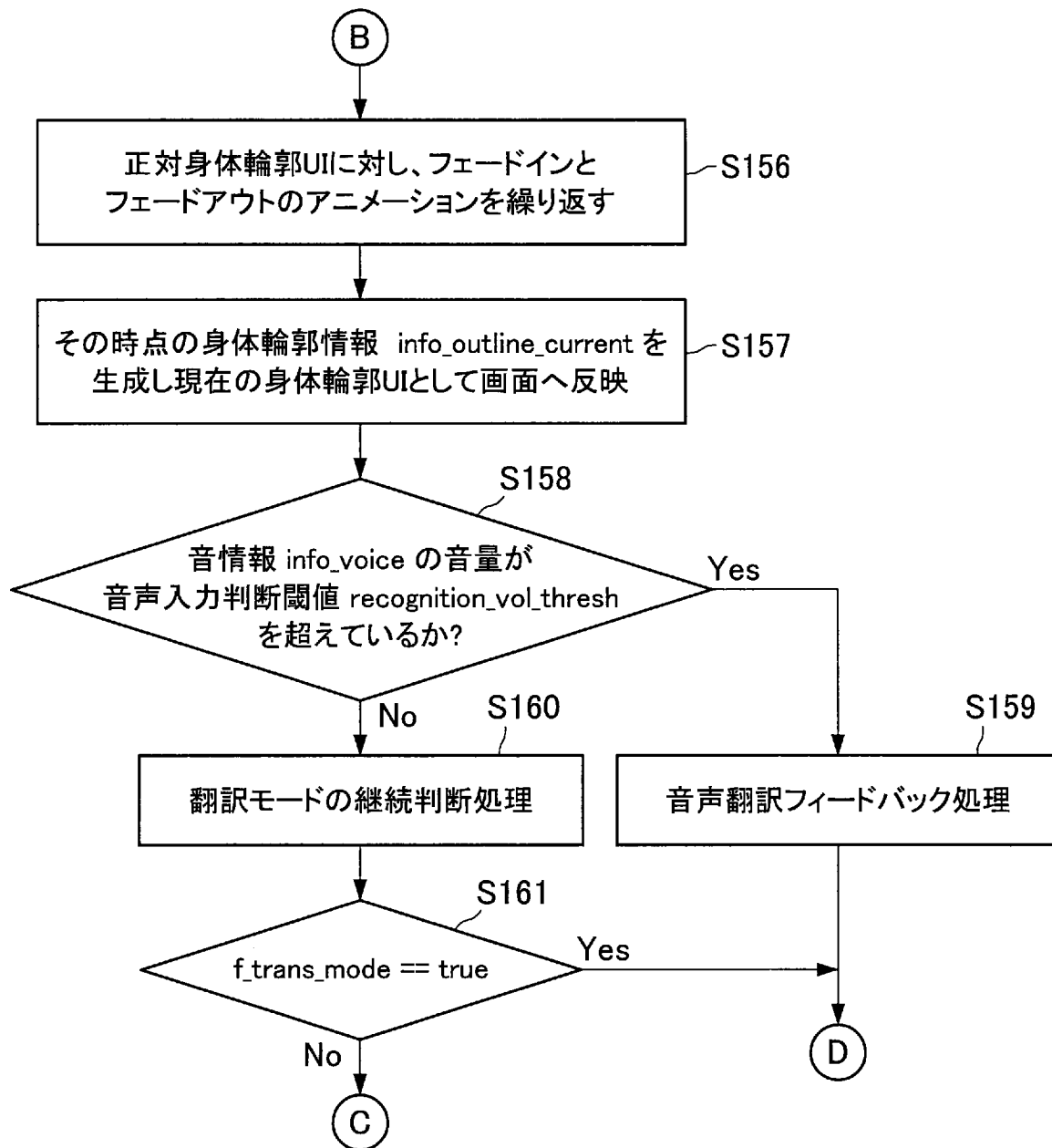
[図35]



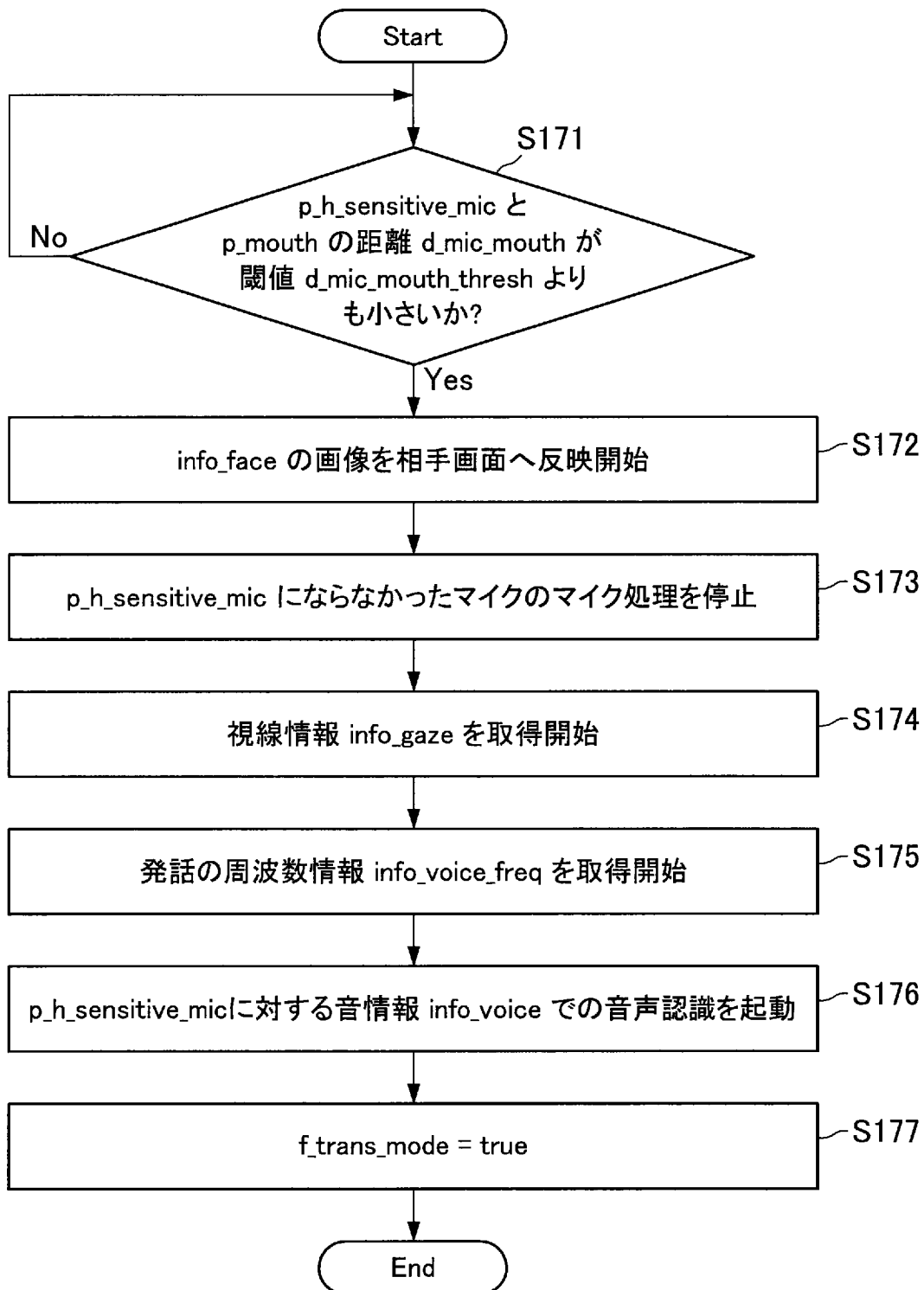
[図36A]



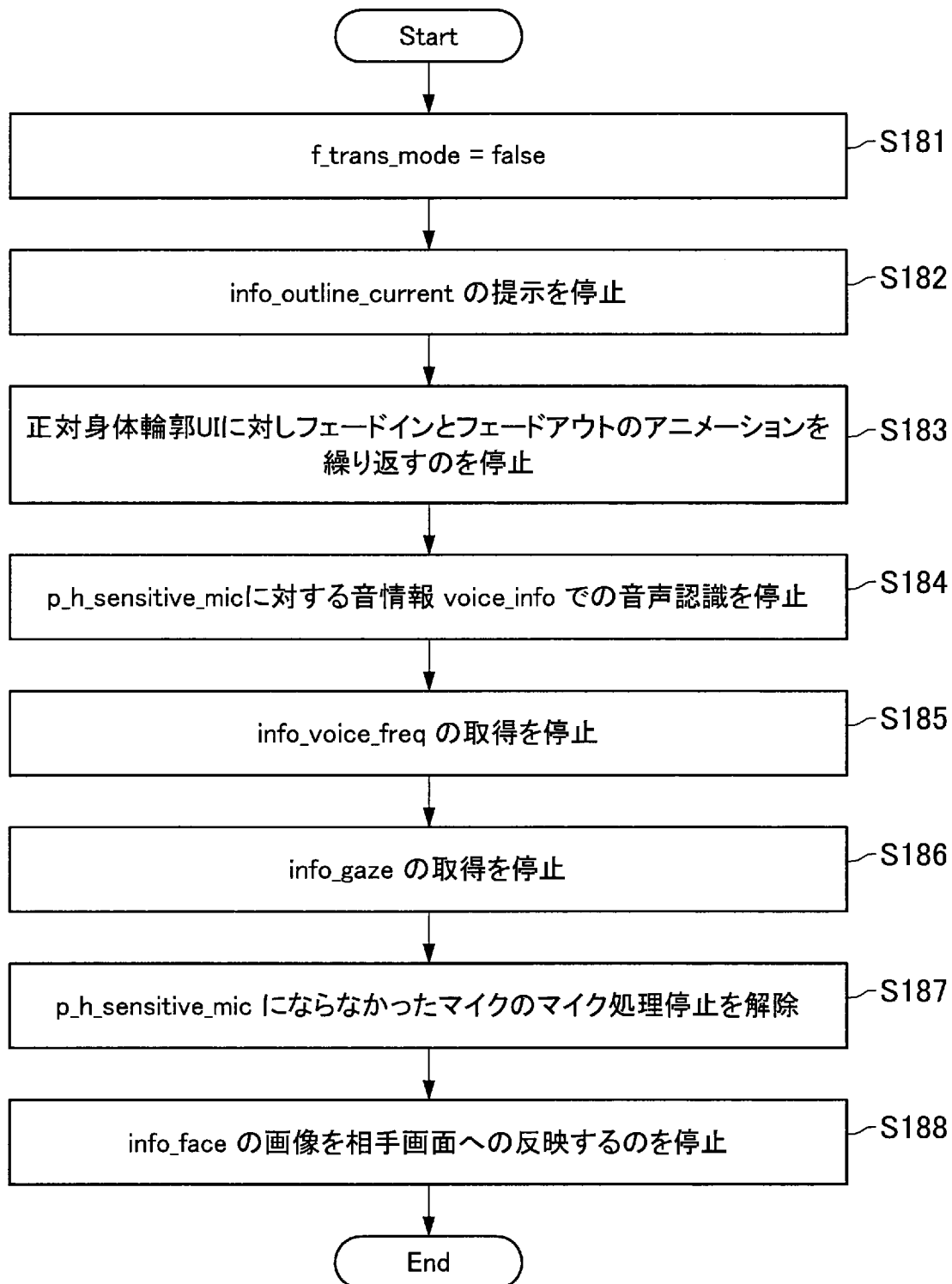
[図36B]



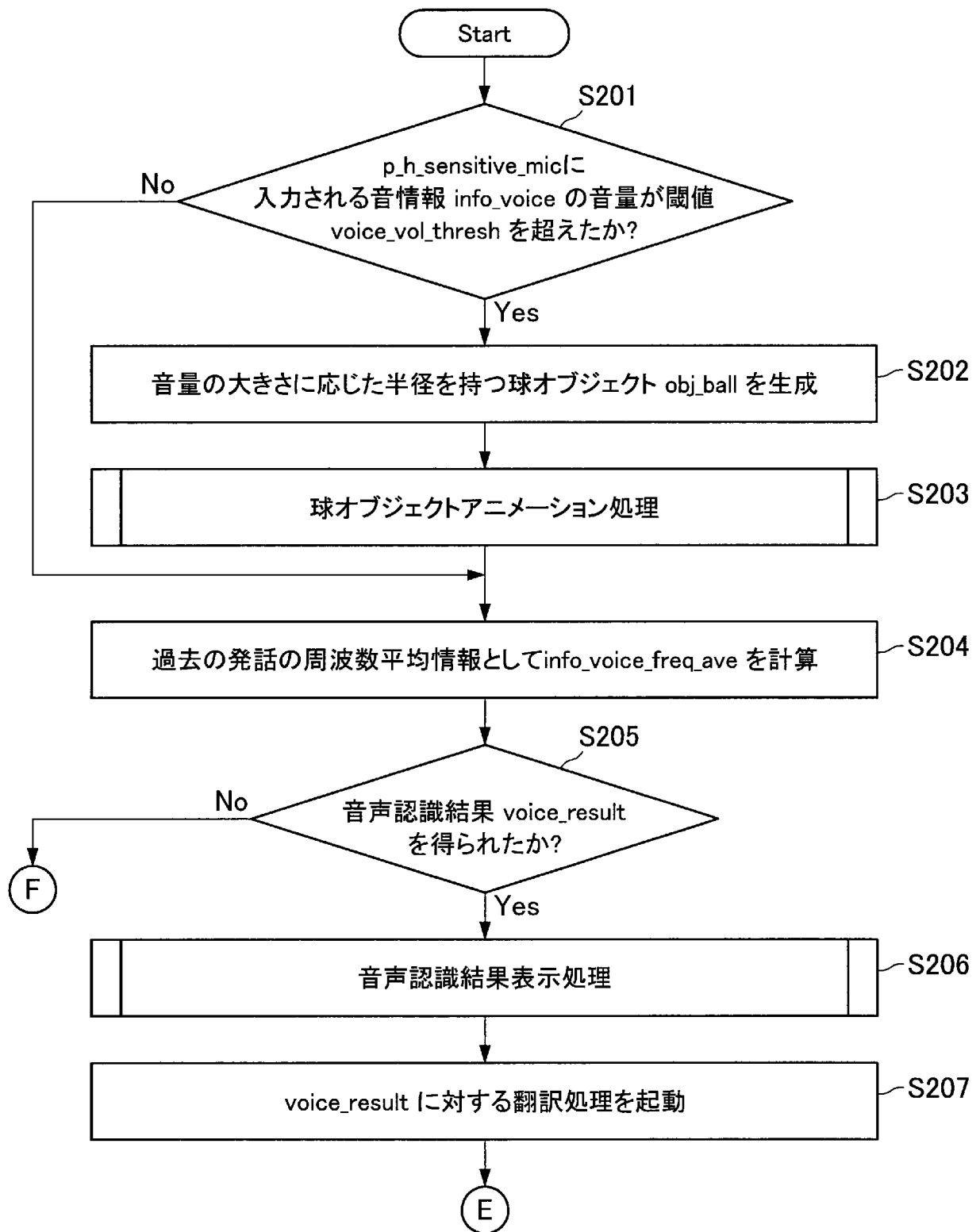
[図37]



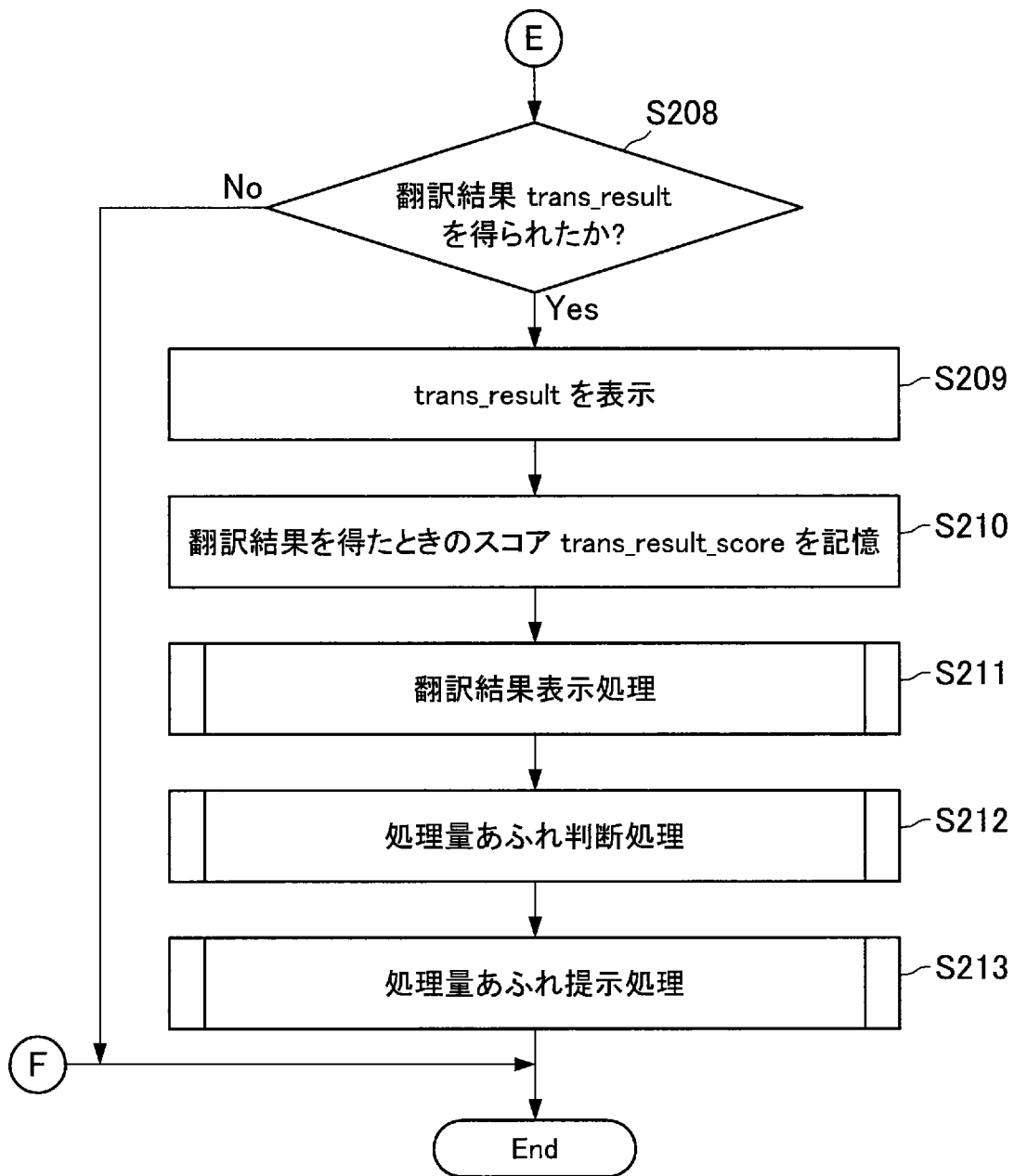
[図38]



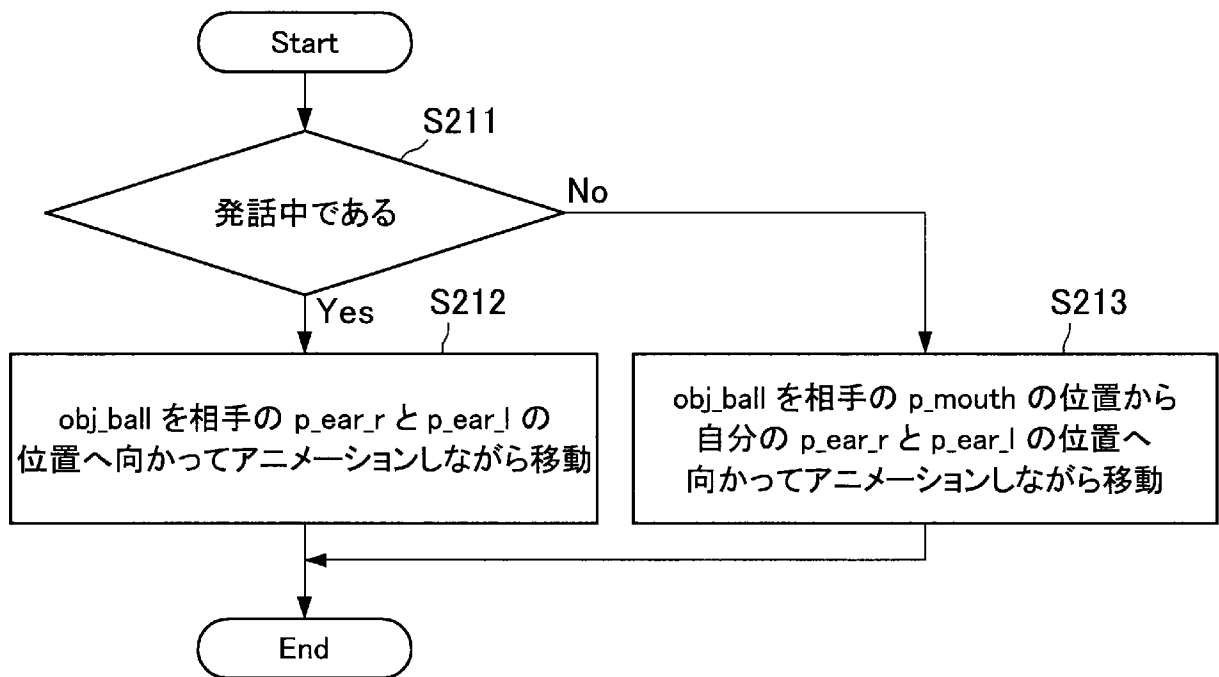
[図39A]



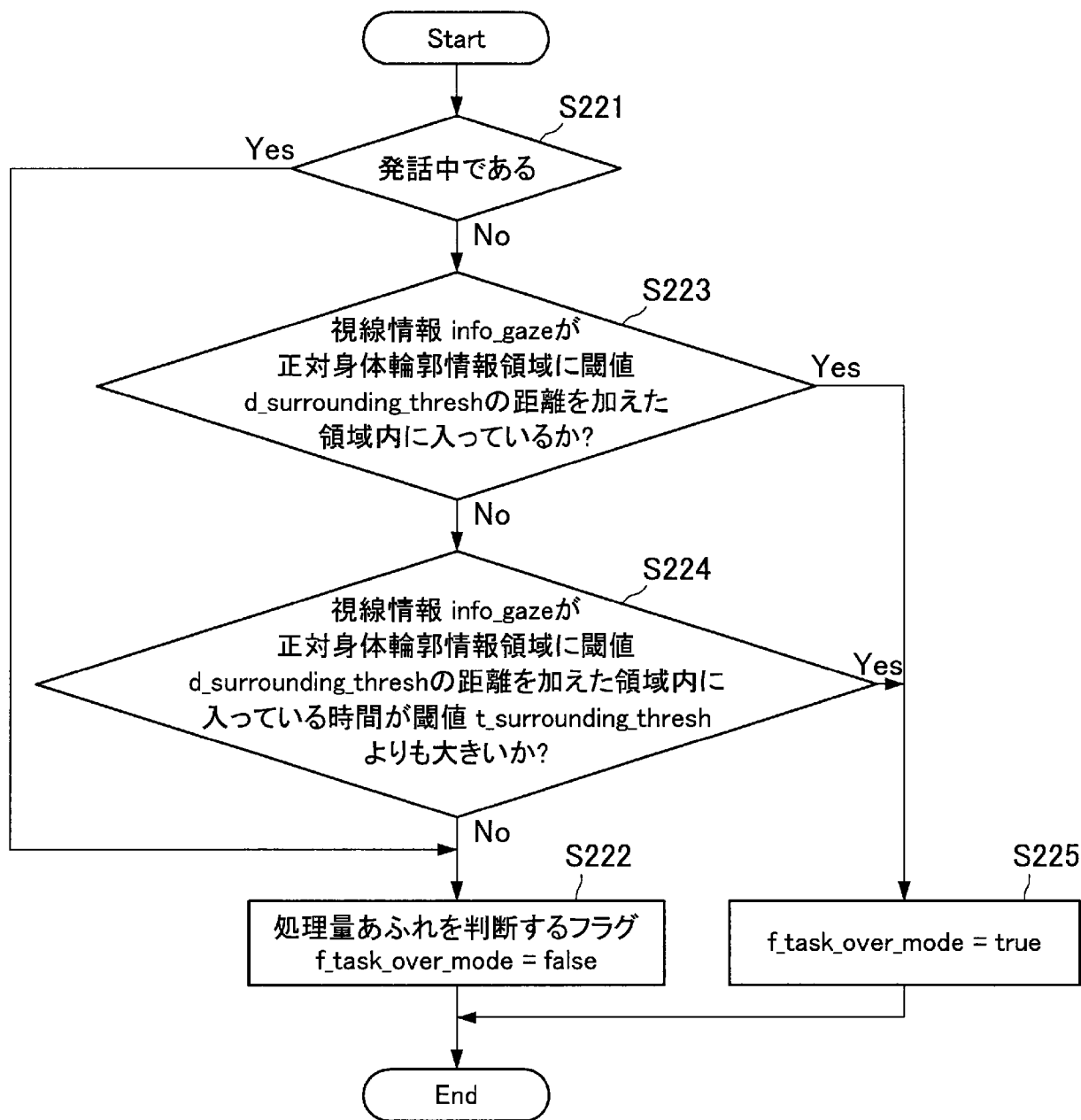
[図39B]



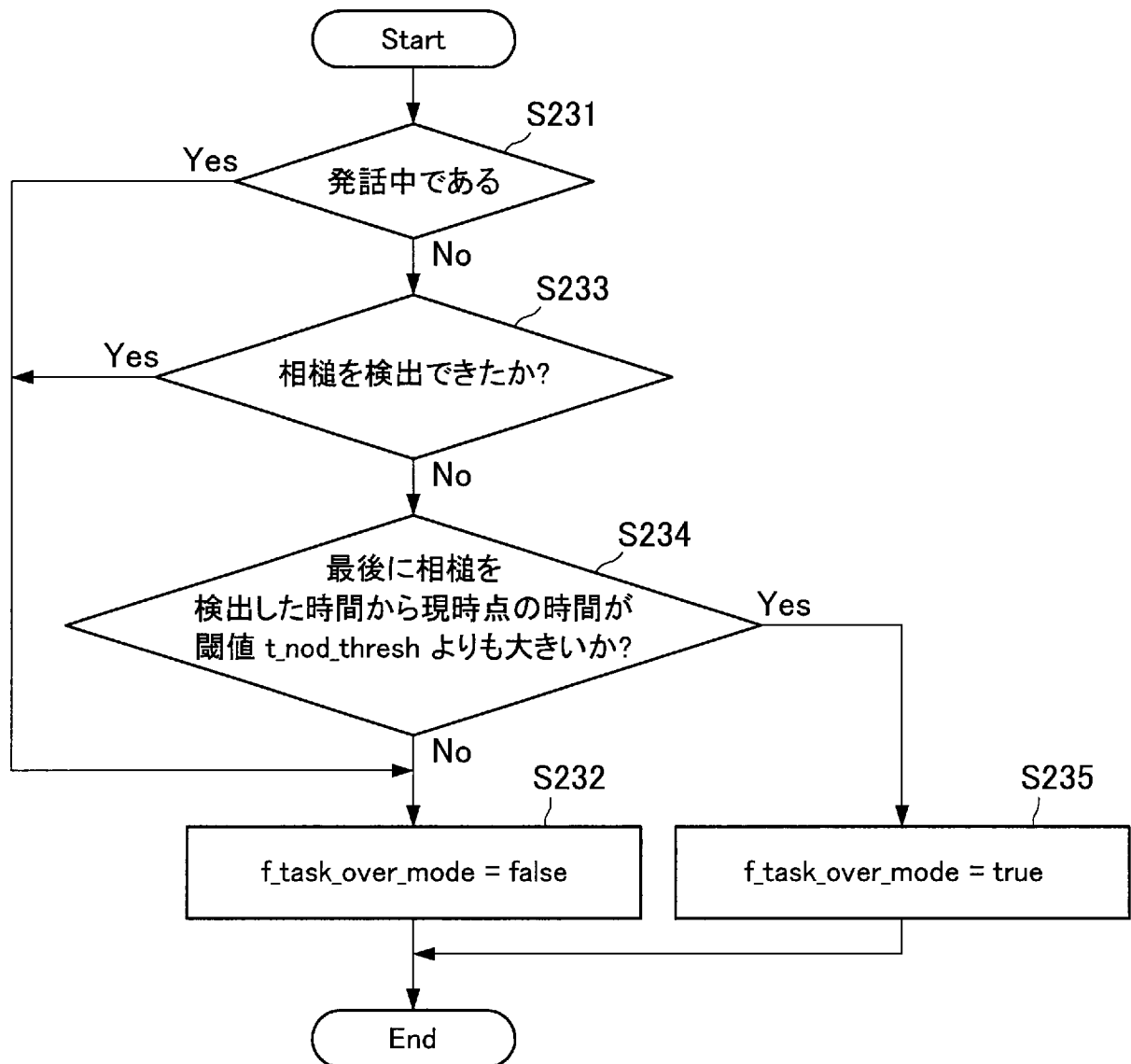
[図40]



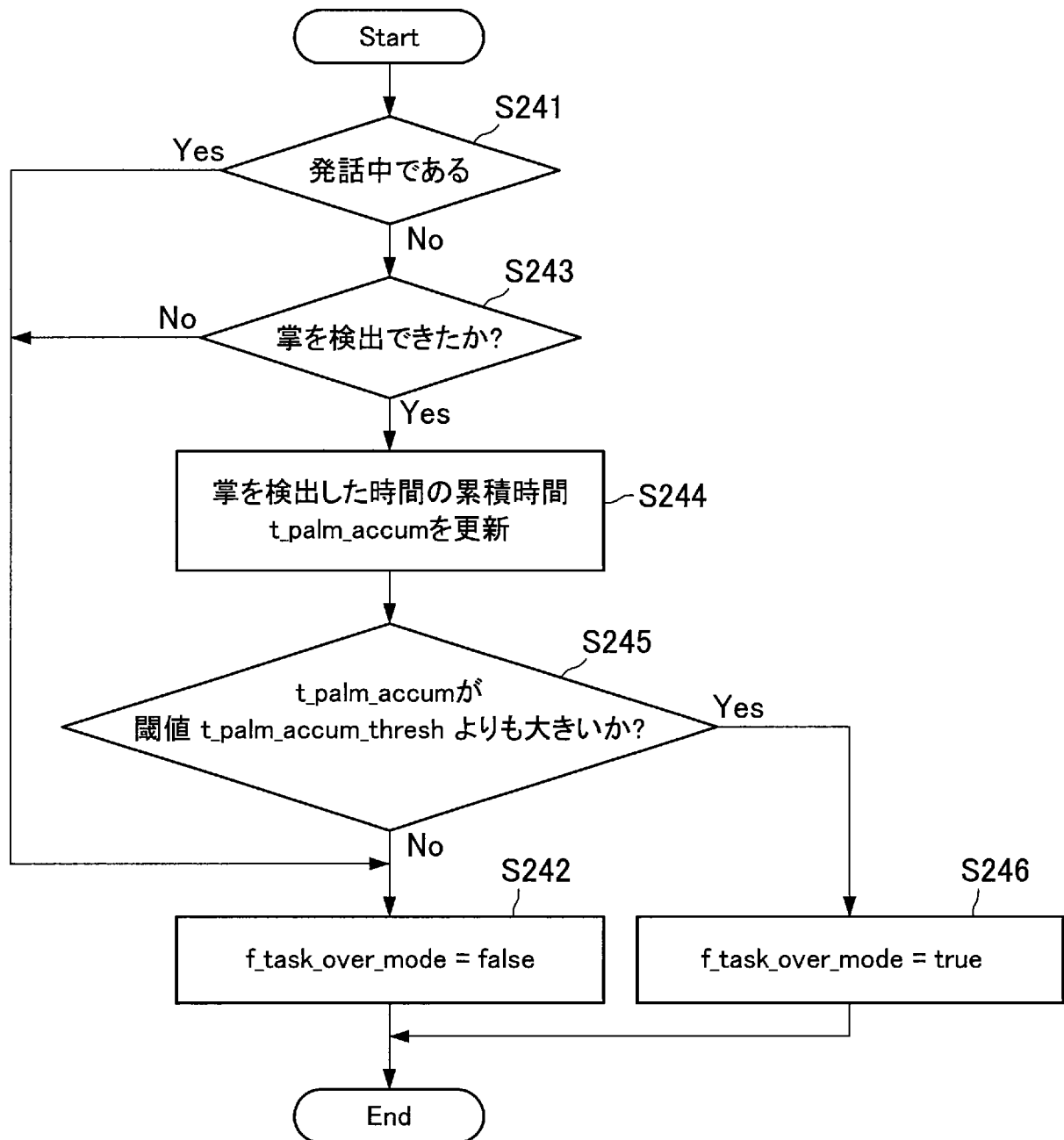
[図41]



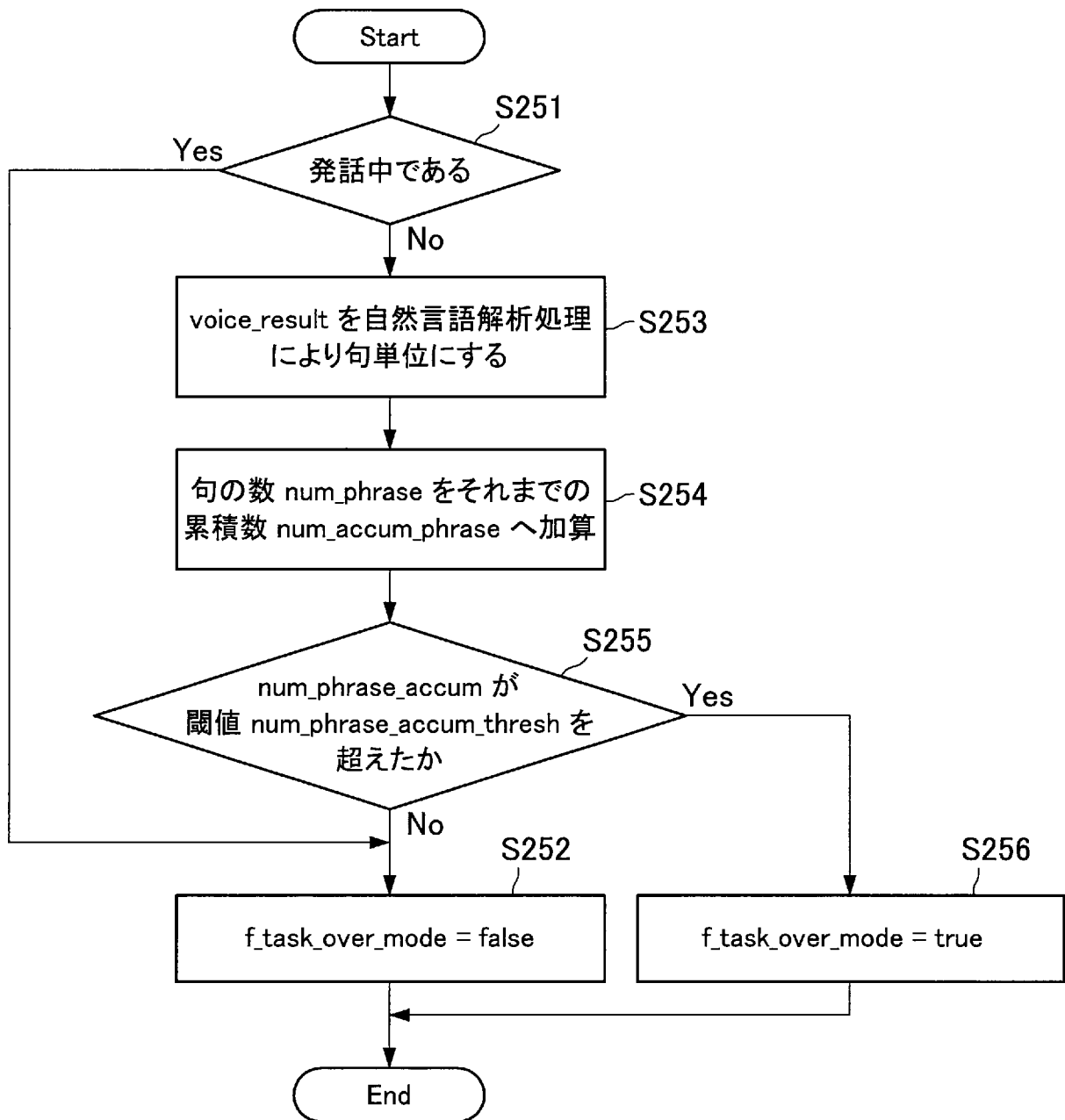
[図42]



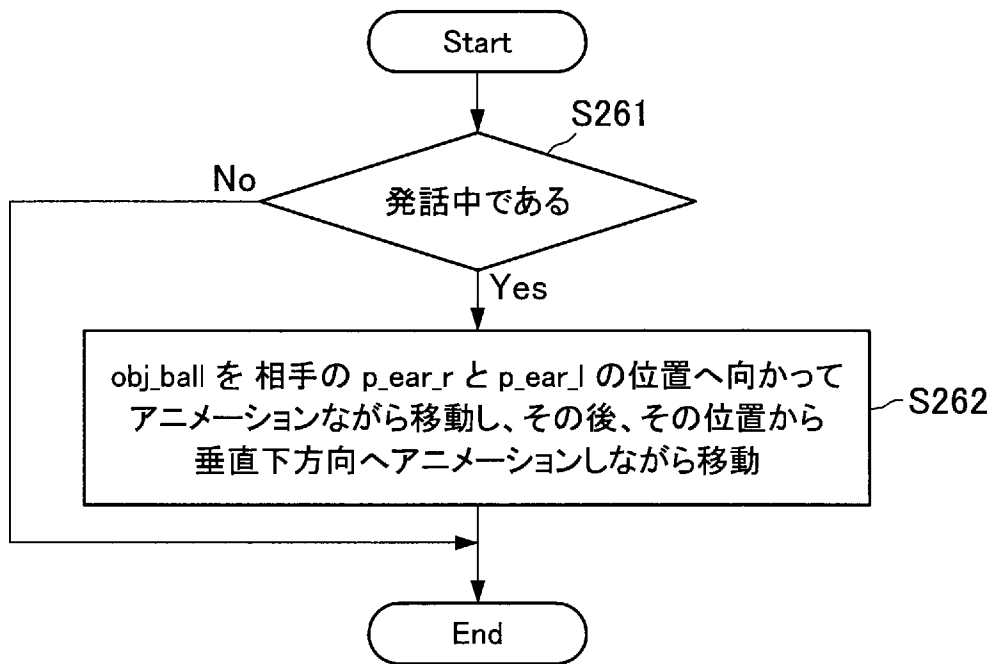
[図43]



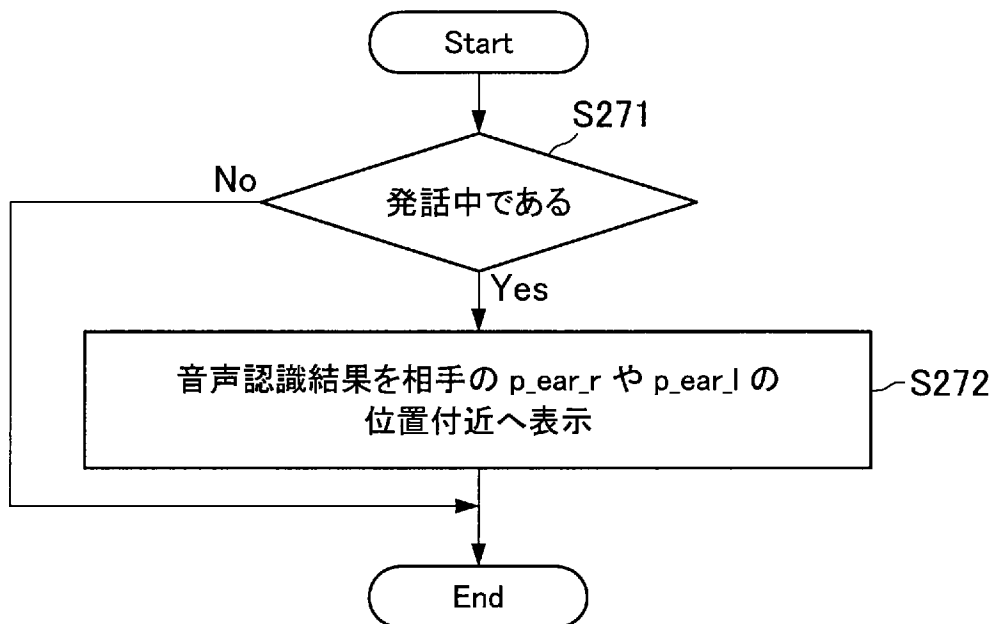
[図44]



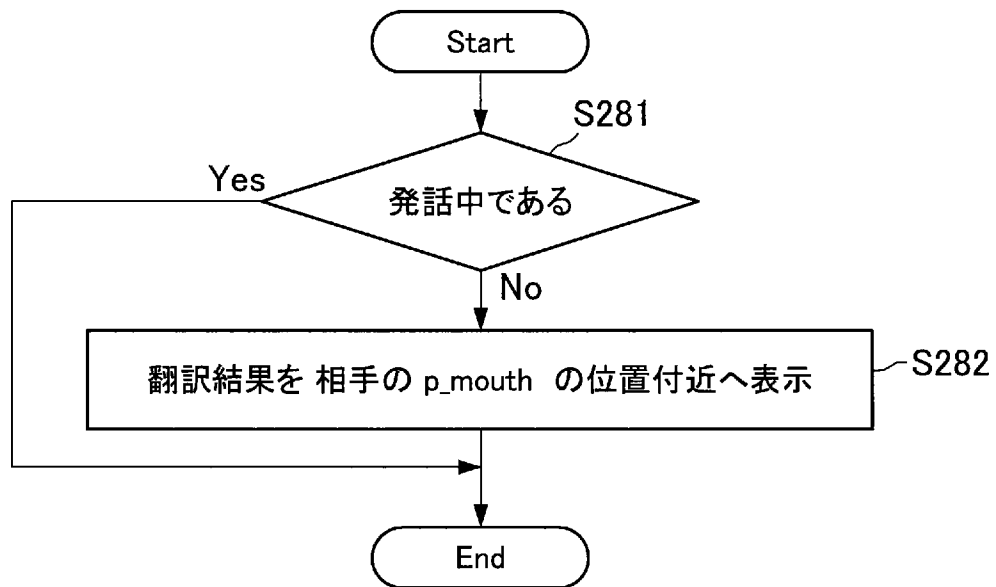
[図45]



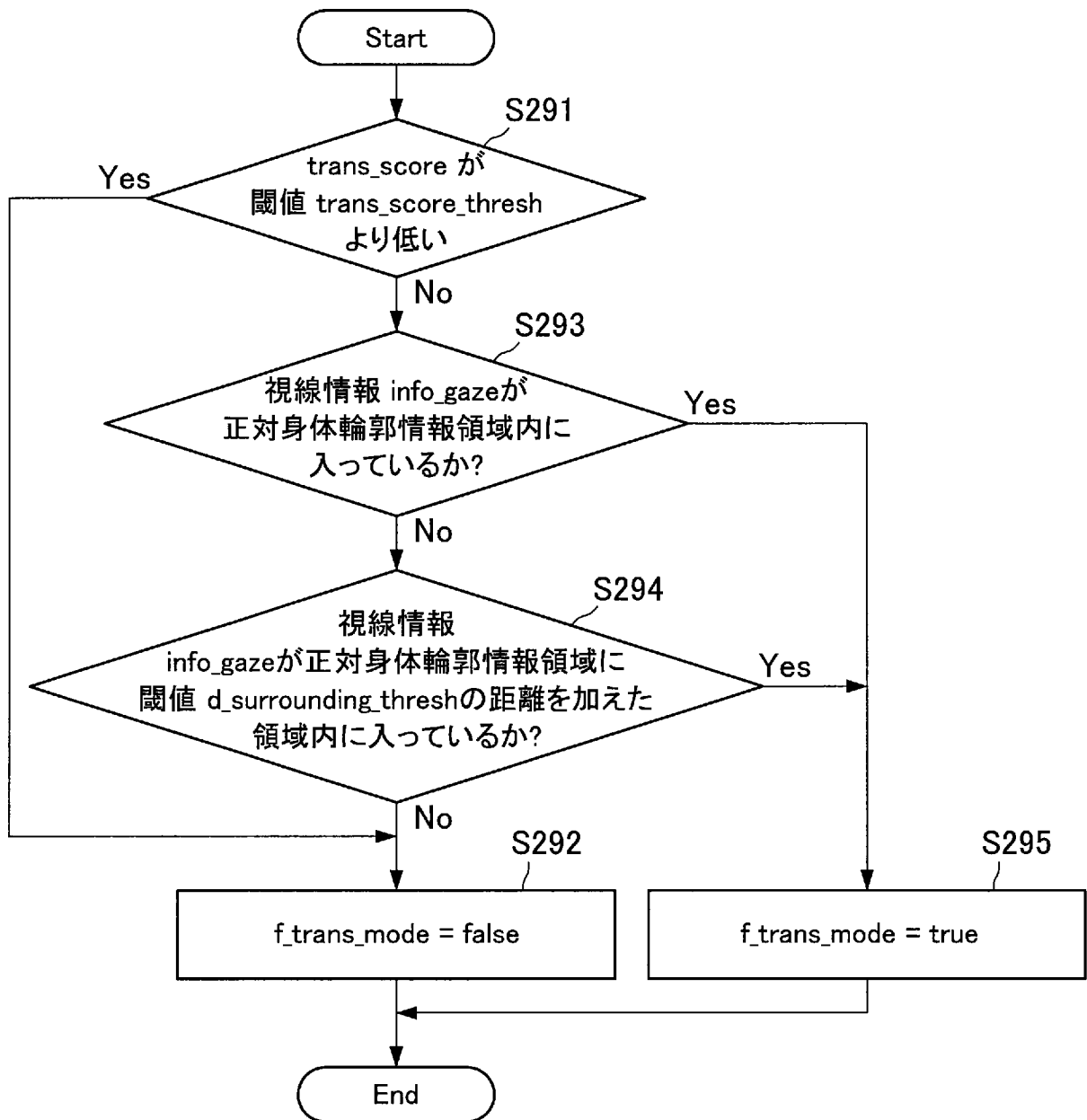
[図46]



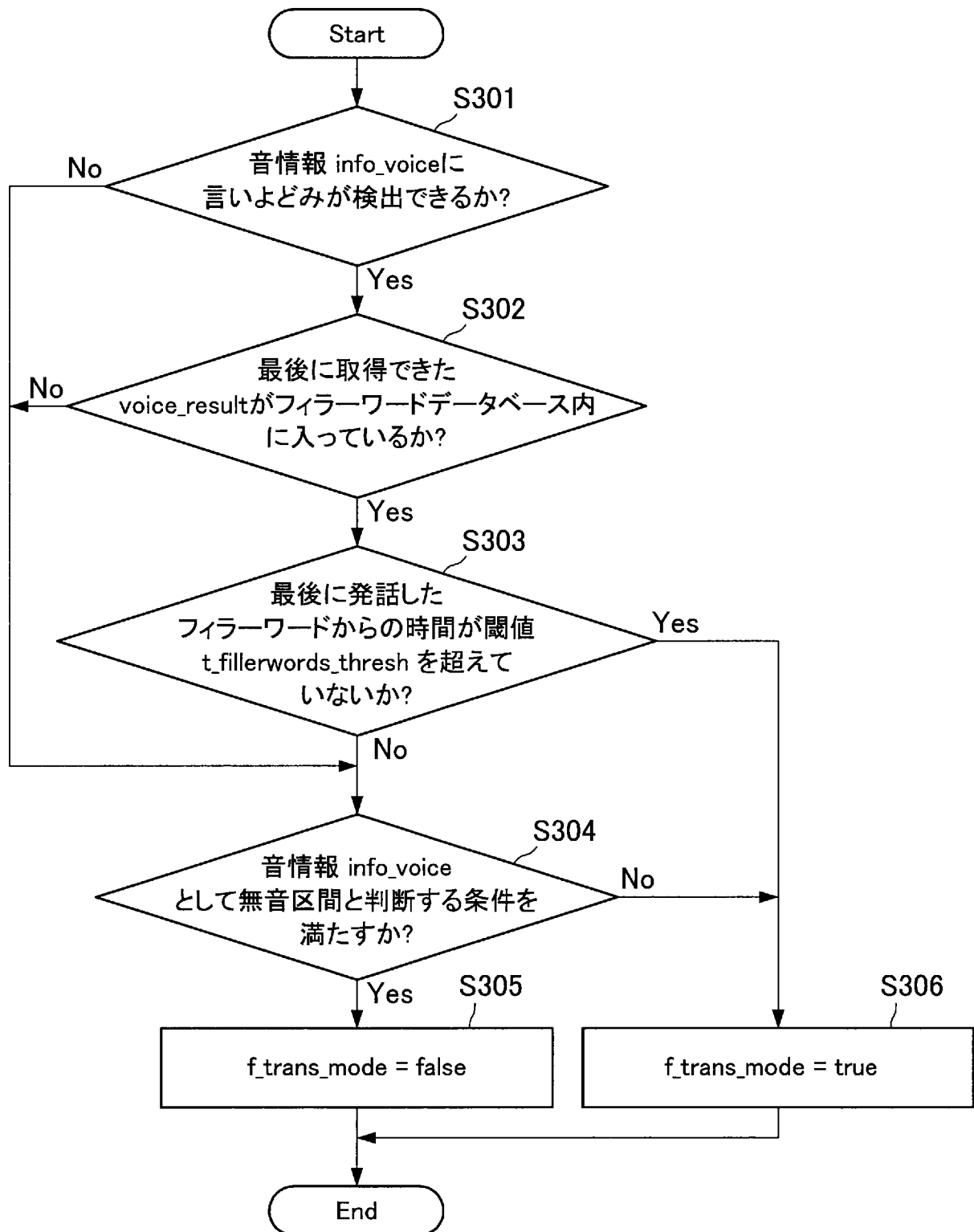
[図47]



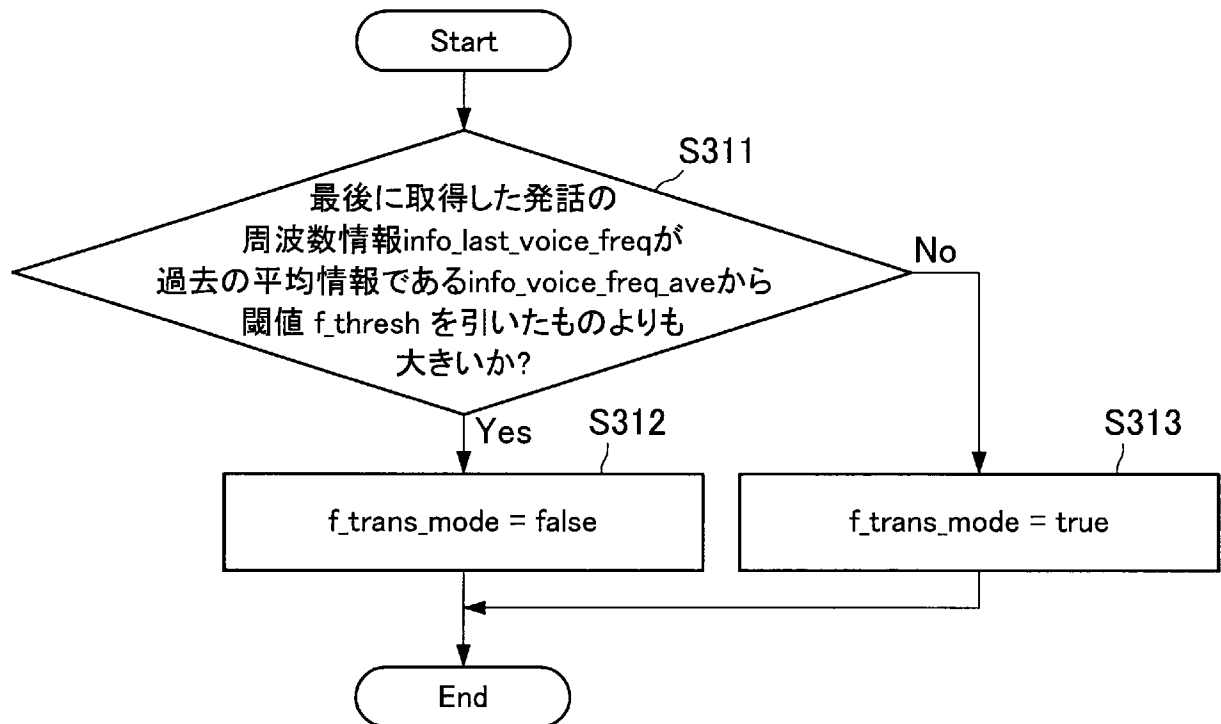
[図48]



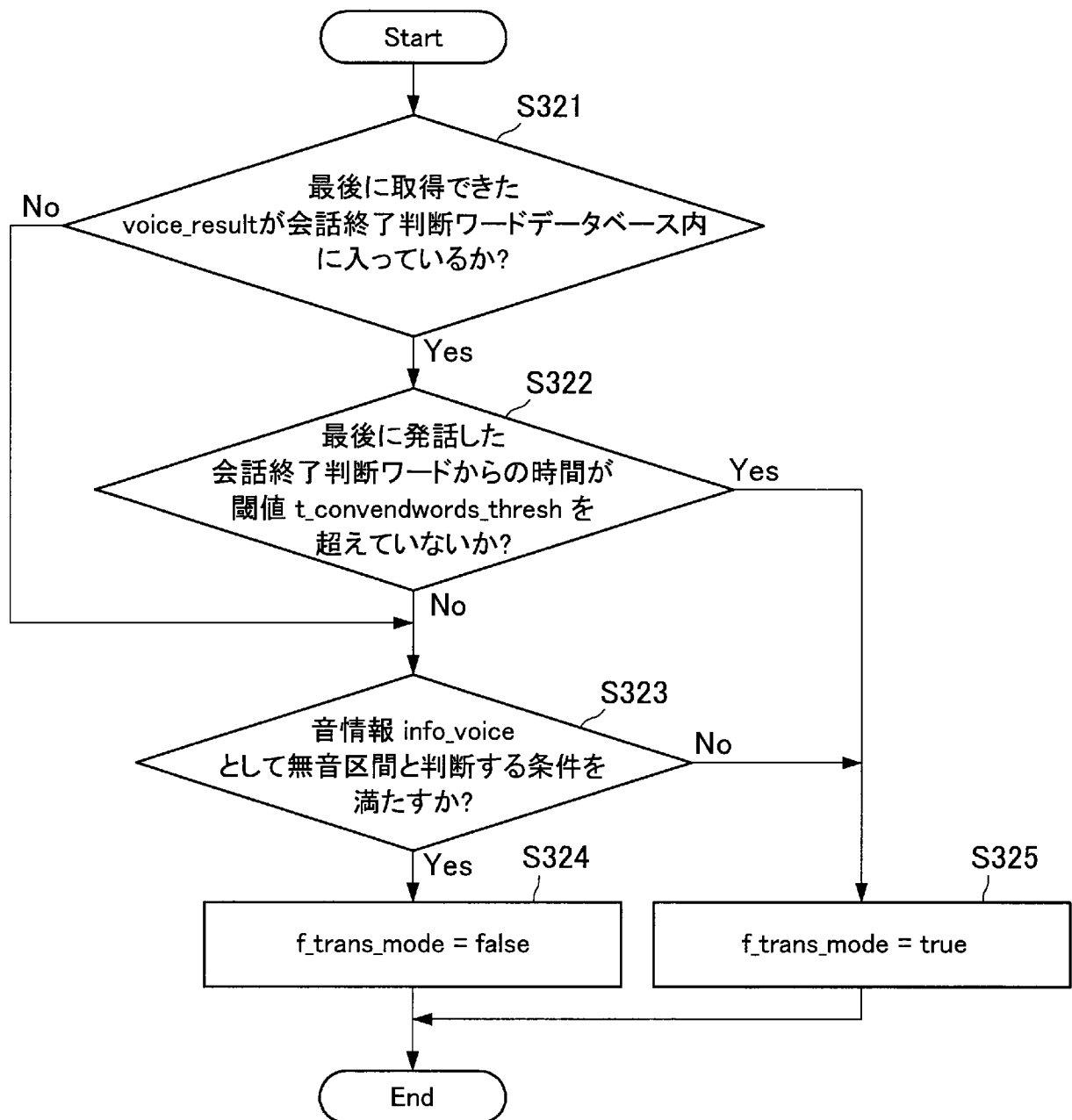
[図49]



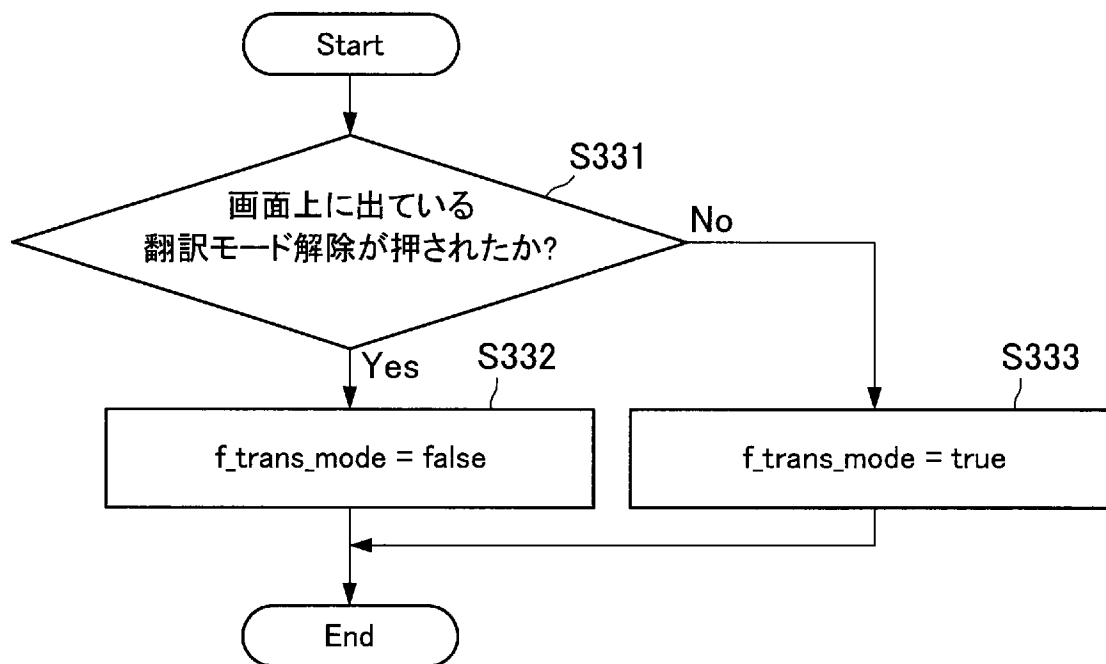
[図50]



[図51]



[図52]



**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/JP2017/010187

**A. CLASSIFICATION OF SUBJECT MATTER**  
G10L15/28(2013.01)i, G06F3/16(2006.01)i, G06F17/28(2006.01)i, G10L15/00(2013.01)i, G10L15/10(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)  
G10L15/28, G06F3/16, G06F17/28, G10L15/00, G10L15/10

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 Y A	JP 2000-222406 A (Sony Corp.), 11 August 2000 (11.08.2000), paragraphs [0034] to [0043]; fig. 7 (Family: none)	1-2, 15, 18-20 1-2, 5, 10-12, 19-20 3-4, 6-9, 13-14, 16-17
Y	JP 2000-293193 A (Canon Inc.), 20 October 2000 (20.10.2000), paragraph [0020]; fig. 3 (Family: none)	1-2, 5, 10-12, 19-20
Y	JP 2003-280688 A (Nissan Diesel Motor Co., Ltd.), 02 October 2003 (02.10.2003), paragraph [0037]; fig. 5 (Family: none)	2, 5, 12

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 12 May 2017 (12.05.17)	Date of mailing of the international search report 23 May 2017 (23.05.17)
---	--

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/JP2017/010187

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2000-242642 A (Sony Corp.), 08 September 2000 (08.09.2000), paragraphs [0050] to [0053]; fig. 6 to 7 & WO 2000/049523 A1	10-11

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G10L15/28(2013.01)i, G06F3/16(2006.01)i, G06F17/28(2006.01)i, G10L15/00(2013.01)i, G10L15/10(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. G10L15/28, G06F3/16, G06F17/28, G10L15/00, G10L15/10

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 Y A	JP 2000-222406 A (ソニー株式会社) 2000.08.11, [0034] - [0043]、第7図 (ファミリーなし)	1-2, 15, 18-20 1-2, 5, 10-12, 19-20 3-4, 6-9, 13-1 4, 16-17
Y	JP 2000-293193 A (キヤノン株式会社) 2000.10.20, [0020]、第3図 (ファミリーなし)	1-2, 5, 10-12, 19-20

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

12.05.2017

国際調査報告の発送日

23.05.2017

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)

大野 弘

5Z

9175

電話番号 03-3581-1101 内線 3591

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
Y	JP 2003-280688 A (日産ディーゼル工業株式会社) 2003. 10. 02, [0037]、第5図 (ファミリーなし)	2, 5, 12
Y	JP 2000-242642 A (ソニー株式会社) 2000. 09. 08, [0050] - [0053]、第6-7図 & WO 2000/049523 A1	10-11