

(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
2019年6月27日(27.06.2019)

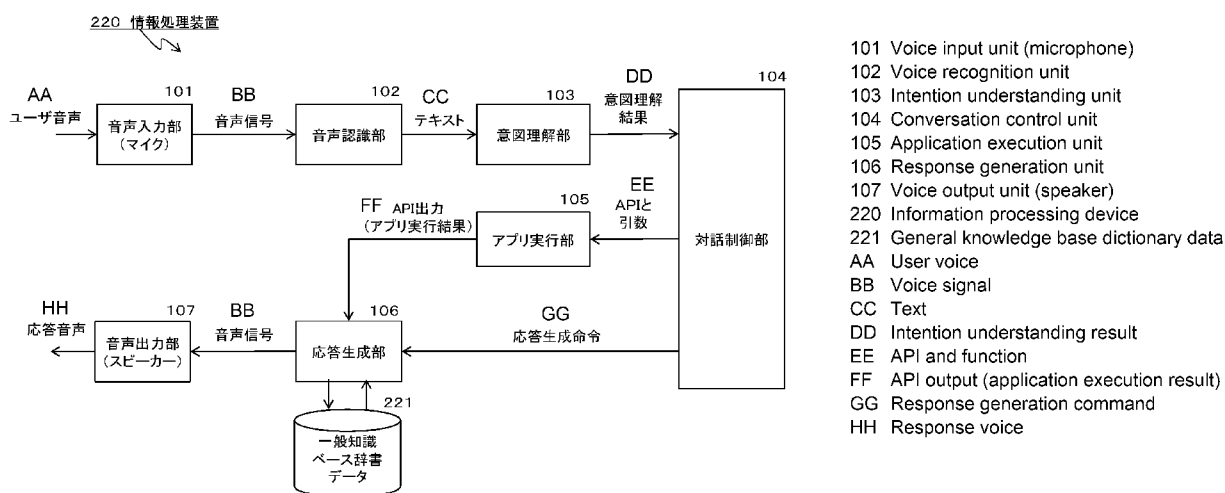


(10) 国際公開番号
WO 2019/123775 A1

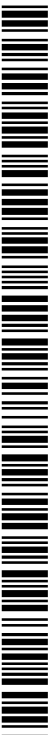
- (51) 国際特許分類:
G10L 15/22 (2006.01) G10L 13/08 (2013.01)
G06F 3/16 (2006.01) G10L 15/10 (2006.01)
G10L 13/00 (2006.01) G10L 15/30 (2013.01)
- (21) 国際出願番号: PCT/JP2018/037347
- (22) 国際出願日: 2018年10月5日(05.10.2018)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2017-245693 2017年12月22日(22.12.2017) 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: 東山 恵祐 (TOUYAMA, Keisuke); 〒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: 宮田 正昭, 外(MIYATA, Masaaki et al.); 〒1040032 東京都中央区八丁堀三丁目25番9号 Daiwa八丁堀駅前ビル西館8階 特許業務法人大同特許事務所 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, JO, 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,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING SYSTEM, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理システム、および情報処理方法、並びにプログラム



(57) Abstract: When generating a confirmation response to a user utterance, additional information acquired from dictionary data or an application execution result is added, thereby generating a confirmation response in which the possibility of a misunderstanding is reduced. The present invention has a voice input unit into which a user utterance is inputted, a confirmation response generation unit that generates a confirmation response seeking confirmation of the user utterance by the uttering user, and a voice output unit that outputs the confirmation response generated by the confirmation response generation unit, the confirmation response generation unit generating a confirmation response in which there is used the additional information obtained from the dictionary data and/or the application execution result. A confirmation response is generated in which there is used, for example, general knowledge base dictionary data or the execution result of an application, or information obtained from a captured image, sensor, etc.



WO 2019/123775 A1

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 :

- 一 国際調査報告 (条約第21条(3))

(57) 要約 : ユーザ発話に対する確認応答を生成する際に、辞書データやアプリケーション実行結果から取得される付加情報を追加することで誤解が生じる可能性を低減した確認応答を生成する。ユーザ発話を入力する音声入力部と、ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、確認応答生成部の生成した確認応答を出力する音声出力部を有し、確認応答生成部は、辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する。例えば、一般知識ベース辞書データや、アプリケーションの実行結果、あるいは、撮影画像やセンサー等から得られる情報を用いた確認応答を生成する。

明 細 書

発明の名称：

情報処理装置、情報処理システム、および情報処理方法、並びにプログラム

技術分野

[0001] 本開示は、情報処理装置、情報処理システム、および情報処理方法、並びにプログラムに関する。さらに詳細には、ユーザ発話の音声認識結果に基づいて応答を行う情報処理装置、情報処理システム、および情報処理方法、並びにプログラムに関する。

背景技術

[0002] 昨今、ユーザ発話の音声認識を行い、認識結果に基づいて応答を行う対話システムの利用が増大している。

この対話システムにおいては、例えばシステムが、マイクを介して入力するユーザ発話を認識理解して、それに応じた処理を行う。

例えばユーザが、「明日の天気を教えて」と発話した場合、天気情報提供サーバから天気情報を取得して、取得情報に基づく応答を生成してスピーカから出力する処理を行う。あるいは、ユーザが、「テレビを消して」と発話した場合、テレビをオフにする動作を行う。

[0003] しかし、このような対話システムにおいて、ユーザ発話の音声認識結果あるいは意図推定結果に対する信頼度が低い場合や、認識した結果の解釈が複数通りある場合等、システムが行動を一意に決定できない場合がある。

このような場合、システムは、ユーザに対してユーザ意図を確認するための確認応答を生成して出力することがある。

また、例えばシステムの記憶部に保存してあるデータの登録や削除など、失敗した場合のリスクが高い処理を行う場合には、システムはユーザに対して処理を行ってよいかどうかを再確認するための確認応答を出力することがある。

[0004] 例えば、特許文献1（特開2011-54088号公報）は、以下の構成を開示している。

聞き間違いをしやすいユーザ発話、例えば、

（例1）「get a cup」と、「put a cut」、

（例2）「銀閣寺について教えて」と、「金閣寺について教えて」、

このようなユーザ発話を入力した場合に、ユーザの意図を確認するための応答を出力する構成を開示している。

[0005] 特許文献1には、さらに、行動の対象物が複数あって決定できない場合の例として以下の例が記載されている。

（例3）「コップを取って」

上記（例3）のユーザ発話は認識できるが、コップが複数ある場合には、どのコップなのかがわからない。このような場合にもユーザの意図を確認するための応答を出力する。

特許文献1には、このような対話型の確認構成が開示されている。

[0006] 特許文献1に示す上記（例2）の場合にシステムが生成して出力する確認応答は、以下の応答である。

ユーザ：「銀閣寺（発音不明瞭）について教えて」

システム：「銀閣寺でよろしいですか？」

[0007] この先行技術の例では、ユーザ発話に対して「銀閣寺」という単語のみを利用したおうむ返しの応答文章を出力している。

ここで、システムの応答が音声でユーザに伝えられる場合を考えると、システムの応答が何らかの理由で聞き取りにくい場合には、システムが「銀閣寺」と発話したにも関わらず、ユーザは「金閣寺」と聞き間違いをすることがある。また、仮に音声だけではなく何らかの画面表示等も交えた応答出力を行う際でも、ユーザ自身が金閣寺と銀閣寺を勘違い等により逆に理解している場合には、本来は銀閣寺についての情報が欲しい場合でも、金閣寺の情報が欲しいと意志表明をすることになるため、最終的にユーザの意図とは異なる情報しか得られないことになる。

先行技術文献

特許文献

[0008] 特許文献1：特開2011-54088号公報

発明の概要

発明が解決しようとする課題

[0009] 上述したように、既存の対話システムにおいては、システムのユーザ発話に対する音声認識結果あるいは意図推定結果に対する信頼度が低い場合や、認識した結果の解釈が複数通りある場合等には、ユーザに対して確認応答を出力することがある。しかし、その際に、ユーザの発話をそのままおうむ返しする手法だと、ユーザがさらに聞き間違いをする可能性や、ユーザ自身が勘違いをしている場合にそれを正すことができず、本来ユーザが欲しい情報とは異なる情報を提供してしまう、といった問題が発生する。

[0010] 本開示は、例えば、上記問題点に鑑みてなされたものであり、ユーザ発話の認識結果があいまいな場合、あるいは複数の候補がある場合等、システム側が処理を決定できない場合等において、システムがユーザに対して曖昧性解消のための確認応答を提示する際に情報を付加した応答とすることにより、確認精度の向上を実現する情報処理装置、情報処理システム、および情報処理方法、並びにプログラムを提供することを目的とする。

課題を解決するための手段

[0011] 本開示の第1の側面は、
ユーザ発話を入力する音声入力部と、
前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、
前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、
前記確認応答生成部は、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理装置にある。

[0012] さらに、本開示の第2の側面は、
ユーザ端末と、データ処理サーバを有する情報処理システムであり、
前記ユーザ端末は、
ユーザ発話を入力する音声入力部と、
前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、
前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、
前記データ処理サーバは、
前記音声入力部の入力音声の認識処理を実行する音声認識部と、
前記音声認識部の認識結果に基づく意図理解処理を実行する意図理解部と、
、
前記意図理解部の生成した意図理解結果に基づく対話制御処理を実行するとともに、意図理解結果に基づいて前記ユーザ発話に対する確認応答が必要か否か判定し、確認応答が必要と判定した場合、前記ユーザ端末の確認応答生成部に対して確認応答生成命令を出力する対話制御部を有し、
前記ユーザ端末の前記確認応答生成部は、
前記データ処理サーバからの確認応答生成命令入力に応じて、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理システムにある。

[0013] さらに、本開示の第3の側面は、
情報処理装置において実行する情報処理方法であり、
音声入力部が、ユーザ発話を入力する音声入力ステップと、
確認応答生成部が、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成ステップと、
音声出力部が、前記確認応答生成ステップで生成した確認応答を出力する音声出力ステップを実行し、
前記確認応答生成ステップにおいて、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理方法にある。

[0014] さらに、本開示の第4の側面は、

ユーザ端末と、データ処理サーバを有する情報処理システムにおいて実行する情報処理方法であり、

前記ユーザ端末が、

ユーザ発話を入力する音声入力処理を実行し、

前記データ処理サーバが、

前記音声入力処理において入力された入力音声の音声認識処理と、

前記音声認識処理の生成した音声認識結果に基づく意図理解処理と、

前記意図理解処理の意図理解結果に基づいて、前記ユーザ発話に対する確認応答が必要か否か判定し、確認応答が必要と判定した場合、前記ユーザ端末に対して確認応答生成命令を出力し、

前記ユーザ端末が、

前記データ処理サーバからの確認応答生成命令入力に応じて、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成処理を実行し、

前記確認応答生成処理において、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成処理を実行する情報処理方法にある。

[0015] さらに、本開示の第5の側面は、

情報処理装置において情報処理を実行させるプログラムであり、

音声入力部に、ユーザ発話を入力させる音声入力ステップと、

確認応答生成部に、前記ユーザ発話に対する発話ユーザの確認を求めるための確認応答を生成させる確認応答生成ステップと、

音声出力部に、前記確認応答を出力させる音声出力ステップを実行させ、

前記確認応答生成ステップにおいては、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行させるプログラムにある。

[0016] なお、本開示のプログラムは、例えば、様々なプログラム・コードを実行可能な情報処理装置やコンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体によって提供可能なプログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、情報処理装置やコンピュータ・システム上でプログラムに応じた処理が実現される。

[0017] 本開示のさらに他の目的、特徴や利点は、後述する本開示の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

発明の効果

[0018] 本開示の一実施例の構成によれば、ユーザ発話に対する確認応答を生成する際に、辞書データやアプリケーション実行結果から取得される付加情報を追加することで誤解が生じる可能性を低減した確認応答を生成する構成が実現される。

具体的には、例えば、ユーザ発話を入力する音声入力部と、ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、確認応答生成部の生成した確認応答を出力する音声出力部を有し、確認応答生成部は、辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する。例えば、一般知識ベース辞書データや、アプリケーションの実行結果、あるいは、撮影画像やセンサー等から得られる情報を用いた確認応答を生成する。

本構成により、ユーザ発話に対する確認応答を生成する際に、辞書データやアプリケーション実行結果から取得される付加情報を追加することで誤解が生じる可能性を低減した確認応答を生成する構成が実現される。

なお、本明細書に記載された効果はあくまで例示であって限定されるもの

ではなく、また付加的な効果があってもよい。

図面の簡単な説明

[0019] [図1]ユーザ発話を認識して応答を行う情報処理装置の一般的な構成例について説明する図である。

[図2]実施例1の情報処理装置220の構成例を示す図である。

[図3]実施例1の情報処理装置220を適用した確認応答生成例について説明する図である。

[図4]実施例1の情報処理装置220を適用した確認応答生成例について説明する図である。

[図5]実施例2の情報処理装置240の構成例を示す図である。

[図6]実施例2の情報処理装置240を適用した確認応答生成例について説明する図である。

[図7]実施例2の情報処理装置240を適用した確認応答生成例について説明する図である。

[図8]実施例3の情報処理装置260の構成例を示す図である。

[図9]実施例3の情報処理装置260を適用した確認応答生成例について説明する図である。

[図10]実施例3の変形例としての情報処理装置280の構成例を示す図である。

[図11]実施例3の変形例としての情報処理装置280を適用した確認応答生成例について説明する図である。

[図12]実施例3の変形例としての情報処理装置280を適用した確認応答生成例について説明する図である。

[図13]実施例3の変形例としての情報処理装置280を適用した確認応答生成例について説明する図である。

[図14]実施例3の変形例としての情報処理装置290の構成例を示す図である。

[図15]実施例3の変形例としての情報処理装置280を適用した確認応答生

成例について説明する図である。

[図16]実施例4の情報処理装置310の構成例を示す図である。

[図17]実施例4の情報処理装置310を適用した確認応答生成例について説明する図である。

[図18]実施例4の変形例としての情報処理装置320の構成例を示す図である。

[図19]実施例4の変形例としての情報処理装置320を適用した確認応答生成例について説明する図である。

[図20]実施例5の情報処理装置を適用した確認応答生成例について説明する図である。

[図21]情報処理システムの構成例について説明する図である。

[図22]情報処理装置のハードウェア構成例について説明する図である。

発明を実施するための形態

[0020] 以下、図面を参照しながら本開示の情報処理装置、情報処理システム、および情報処理方法、並びにプログラムの詳細について説明する。なお、説明は以下の項目に従って行なう。

1. ユーザ発話を認識して応答や処理を行う情報処理装置の一般的な構成例について

2. (実施例1) 応答生成部が一般知識ベース辞書データを利用して確認応答を生成する実施例について

3. (実施例2) 応答生成部がアプリ(アプリケーション)実行部を利用して確認応答を生成する実施例について

4. (実施例3) 応答生成部が情報入力部の情報を利用して確認応答を生成する実施例について

5. (実施例4) 音声出力に併せて、画像や制御信号の出力を実行する実施例について

6. (実施例5) 対話履歴に基づいて、確認応答を生成する実施例について

7. 情報処理装置、および情報処理システムの構成例について
8. 情報処理装置のハードウェア構成例について
9. 本開示の構成のまとめ

[0021] [1. ユーザ発話を認識して応答や処理を行う情報処理装置の一般的な構成例について]

まず、図1を参照して、ユーザ発話を認識して応答や処理を行う情報処理装置の一般的な構成例について説明する。

[0022] 図1は、ユーザ発話を認識して応答を行う情報処理装置100の一般的な構成例を示す図である。

図1に示すように、情報処理装置100は、音声入力部（マイク）101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、応答生成部106、音声出力部（スピーカー）107を有する。

[0023] 音声入力部（マイク）101から入力されたユーザ発話音声は音声認識部102において音声データからテキストデータに変換される。

音声認識部102は、例えばASR（Automatic Speech Recognition）機能を有し、音声データを複数の単語から構成されるテキストデータに変換する。

[0024] 音声認識部102において生成されたテキストデータは、意図理解部103に入力される。

意図理解部103は、テキストに含まれるユーザの意図候補を選択して出力する。

意図理解部103は、例えば、NLU（Natural Language Understanding）等の自然言語認識機能を有し、テキストデータから、ユーザ発話の意図（Intent）の候補を選択する。

意図理解部103の生成した意図理解結果は、対話制御部104に入力される。

[0025] 対話制御部104は、意図理解部103の生成した意図理解結果を入力して、例えば、過去の対話履歴等から現在の対話状態を推定し、応答生成に必

要となる処理を実行させるためにコールすべきAPIとその引数を決定する。

[0026] これは、具体的には、例えば、意図理解部103の生成した意図理解結果として、

ユーザ発話が、「東京の天気を問い合わせる発話である」との意図理解結果が得られた場合、天気情報を取得可能なアプリケーションを起動させて、東京（引数）をアプリケーションに渡して、応答に必要な情報（東京の天気）を取得させるといった処理である。

[0027] 対話制御部104によるAPIコールと引数情報がアプリ実行部105に出力されて、APIコール対応のアプリケーションが実行され、引数に応じたAPI出力、すなわち、応答処理に必要なデータが応答生成部106に出力される。

[0028] 応答生成部106は、アプリ実行部105から入力したAPI出力、すなわち、応答処理に必要なデータを入力して、応答として出力すべき音声データを生成する。

応答生成部106は、TTS (Text to Speech) 機能に基づいて、応答データを構成するテキストデータを音声データに変換する。

応答生成部106の生成した音声データは、音声出力部（スピーカー）107を介してユーザに向けて出力される。

[0029] 上述した処理シーケンスは、ユーザ発話に基づく音声認識部102におけるテキスト変換、および意図理解部103における意図理解が問題なく行われた場合の処理であり、ユーザの意図に応じた正しい応答が出力された例である。

[0030] しかし、ユーザ発話に基づく音声認識部102におけるテキスト変換、および意図理解部103における意図理解が困難な場合がある。

このような場合には、対話制御部104は、応答生成部106に対して、ユーザ発話に対する発話ユーザの確認を求める確認応答の生成命令、すなわち「確認応答生成命令」を出力して、応答生成部106が、ユーザに再度、

確認のための発話を要求する。

[0031] すなわち、ユーザ発話の認識や理解の確信度が低い場合には、その内容をユーザに問い返して確認する応答文章を生成して提示することがある。

なお、音声認識部102におけるテキスト変換データや、意図理解部103における意図理解結果データには、信頼度を示す属性情報が付加され、対話制御部104に入力され、対話制御部104は、この属性情報に基づいて信頼度を判定することができる。

[0032] また、例えば、ユーザの発話が、スケジュール記録アプリケーションを起動して、予定を登録する処理や削除する処理を要求している場合等、スケジュール記録データベースに変更を加える必要がある場合等、重要な処理を行う場合は、その処理を行ってよいかどうかの確認を行うための応答文章を生成して出力する場合がある。

[0033] 以下、情報処理装置100が、ユーザの意図を確認するための確認応答を行う場合の具体的な処理例について説明する。

例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「8日（よ一か）に買い物の予定を入れてくれる？」

このようなユーザ発話があると、

情報処理装置100は、「よ一か」が、本当に「8日（よ一か）」を意図しているのか、あるいはユーザが「4日（よっか）」を意図して発話してしまったかを判別できない。

このため、情報処理装置100は、ユーザに対して確認を求める以下の確認応答を出力する。

装置の確認応答：「8日（よ一か）ですね？」

[0034] この例は、ユーザが発話した日時情報（8日）について、装置側の認識理解確信度が低い場合の例である。

[0035] このような確認応答は、従来の装置においても多く行われる。

上記の確認応答では、ユーザの発話情報（「8日（よ一か）」）をそのまま応答に用いている。このような応答を行う場合、以下に挙げるような問題

が発生する可能性がある。

(a) 装置が応答を、スピーカーを介して音声出力した場合、ユーザが「8日」を「4日」と聞き間違える可能性がある。

(b) ユーザの真の意図は8日ではないが、何らかの勘違い（例えば予定を入れたのは水曜日だが、本来水曜日は7日なのに8日であると思い違いしている等）により8日と発話してしまった場合、その勘違いに気付くことができない

[0036] このように、従来の装置は、ユーザ発話の確認を求める確認応答を出力する際、ユーザ発話をそのまま「おうむ返し」する構成としたものが多い。

この手法は、ユーザがさらに聞き間違いをする可能性や、ユーザ自身が勘違いをしている場合にそれをたどることができず、本来ユーザが欲しい情報とは異なる情報を提供してしまう、といった問題が発生する。

[0037] 本開示の情報処理装置は、このような問題を解決する構成を有する。

すなわち、ユーザ発話の認識結果があいまいな場合、あるいは複数の候補がある場合において、システムがユーザに対して曖昧性解消のための確認応答を提示する際に、情報を付加することにより確認精度を上げることを可能としたものである。

具体的には、例えば、辞書データまたはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する。

以下、本開示の情報処理装置の構成と処理の具体例について説明する。

[0038] [2. (実施例1) 応答生成部が一般知識ベース辞書データを利用して確認応答を生成する実施例について]

まず、実施例1として、応答生成部が一般知識ベース辞書データを利用して確認応答を生成する実施例について説明する。

[0039] 図2は、本実施例1の情報処理装置220の構成例を示す図である。

図2に示すように、情報処理装置220は、音声入力部（マイク）101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、応答生成部106、音声出力部（スピーカー）107を有し、さら

に、応答生成部106が応答生成に利用する一般知識ベース辞書データ221を有する。

[0040] 音声入力部（マイク）101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、応答生成部106、音声出力部（スピーカー）107、これらの構成部は、基本的には、先に図1を参照して説明した処理を行う。

[0041] すなわち、音声入力部（マイク）101から入力されたユーザ発話音声は音声認識部102において音声データからテキストデータに変換される。

音声認識部102は、例えばASR（Automatic Speech Recognition）機能を有し、音声データを複数の単語から構成されるテキストデータに変換する。

[0042] 音声認識部102において生成されたテキストデータは、意図理解部103に入力される。

意図理解部103は、テキストに含まれるユーザの意図候補を選択して出力する。

意図理解部103は、例えば、NLU（Natural Language Understanding）等の自然言語認識機能を有し、テキストデータから、ユーザ発話の意図（Intent）の候補を選択する。

意図理解部103の生成した意図理解結果は、対話制御部104に入力される。

[0043] 対話制御部104は、意図理解部103の生成した意図理解結果を入力して、例えば、過去の対話履歴等から現在の対話状態を推定し、応答生成に必要な処理を実行させるためにコールすべきAPIとその引数を決定する。

[0044] 対話制御部104によるAPIコールと引数情報がアプリ実行部105に出力されて、APIコール対応のアプリケーションが実行され、引数に応じたAPI出力、すなわち、応答処理に必要なデータが応答生成部106に出力される。

[0045] 応答生成部106は、アプリ実行部105から入力したAPI出力、すなわち、応答処理に必要なデータを入力して、応答として出力すべき音声データを生成する。

応答生成部106は、TTS (Text to Speech) 機能に基づいて、応答データを構成するテキストデータを音声データに変換する。

応答生成部106の生成した音声データは、音声出力部（スピーカー）107を介してユーザに向けて出力される。

[0046] 上述した処理シーケンスは、先に図1を参照して説明した処理シーケンスと同様の処理である。

すなわち、ユーザ発話に基づく音声認識部102におけるテキスト変換、および意図理解部103における意図理解が問題なく行われた場合の処理であり、ユーザの意図に応じた正しい応答が出力された例である。

[0047] 対話制御部104は、ユーザ発話に基づく音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い場合、あるいは、ユーザの発話がスケジュール設定等の重要処理である場合に、応答生成部106に対して確認応答の生成、出力を要求する。

本実施例1では、この応答生成部106の実行する確認応答生成処理が、先に図1を参照して説明した処理とは異なる。

[0048] 先に、図1を参照して説明した確認応答は、

ユーザ発話：「8日（よ一か）に買い物の予定を入れてくれる？」

このようなユーザ発話に対して、ユーザの意図が、「8日（よ一か）」であるか、「4日（よっか）」であるかを確認するため、装置の確認応答として、

「8日（よ一か）ですね？」

このようなおうむ返しの確認応答を行っていた。

[0049] この場合、前述したように、

(a) 装置が応答を、スピーカーを介して音声出力した場合、ユーザが「8日」を「4日」と聞き間違える可能性がある。

(b) ユーザの真の意図は8日ではないが、何らかの勘違い（例えば予定を入れたのは水曜日だが、本来水曜日は7日なのに8日であると思い違いしている等）により8日と発話してしまった場合、その勘違いに気付くことができない

このような問題が発生する。

[0050] 図2に示す本開示の情報処理装置220の応答生成部106は、このような問題の発生を低減させるため、一般知識ベース辞書データ221を利用して確認応答、すなわち、ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する。

なお、一般知識ベース辞書データ221は、様々な事実や、常識、習慣、経験等の一般的な知識をコンピュータ可読形式のデータとしてデータベース化したデータである。ナレッジベース辞書データとも呼ばれる。

[0051] なお、図2以下において説明する本開示の情報処理装置220の応答生成部106は、対話制御部104から応答生成命令を入力した場合に確認応答を生成して出力するのみならず、さらに、対話制御部104から確認応答生成命令ではない通常の応答生成命令を入力した場合にも、ユーザ発話内容を再確認するための確認応答を生成して出力する場合がある。

すなわち、応答生成部106が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107を介して出力する。

[0052] 本開示の情報処理装置220の応答生成部106の生成する確認応答の具体例について説明する。

まず、図3に示す3つの確認応答生成例101～103について説明する。

[0053] (確認応答生成例101)

図3(101)に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「8日（よーか）に買い物の予定を入れてくれる？」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。入力データには、音声認識部102におけるテキスト変換データや、意図理解部103における意図理解結果データには、信頼度を示す属性情報が付加されている。

[0054] 対話制御部104は、属性情報に基づいて、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い否かを判定する。信頼度が予め規定した基準値（しきい値）より低い場合、応答生成部106に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部106に対して確認応答の生成、出力を要求する。

[0055] 本処理例においては、例えば以下のような処理が行われることになる。

情報処理装置220の意図理解部103が、上記のユーザ発話に対する意図理解処理において、「よーか」が、本当に「8日（よーか）」を意図しているのか、あるいはユーザが「4日（よっか）」を意図して発話してしまったかを判別できず、ユーザの意図が曖昧である（信頼度が低い）という意図理解結果を対話制御部104に出力する。

[0056] 対話制御部104は、この意図理解部103からの入力、すなわち、ユーザの意図が曖昧であるという意図理解結果に基づいて、応答生成部106に確認応答生成命令を出力する。

[0057] なお、意図理解部103が、意図理解処理において、「よーか」が、「8日（よーか）」を意図しているとする高信頼度情報を属性情報として付加した結果を対話制御部104に出力した場合であっても、対話制御部104は、スケジュール設定の重要な処理であるとの判断に基づいて、日にちの再確認が必要であるとの判断を行う場合がある。この判断が行われた場合も、応答生成部106に確認応答生成命令を出力する。

[0058] また、対話制御部104から応答生成部106に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部

106が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107を介して出力する。

[0059] 本実施例において、応答生成部106は、一般知識ベース辞書データ221のデータを利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「8日（はちにち）に買い物の予定を登録しますね？」

[0060] 上記の例は、8日には「よーか」と「はちにち」の二通りの読み方がある、という知識を用いて、両方の読み方を伝える確認応答の例である。

この知識は、応答生成部106が、一般知識ベース辞書データ221のデータを利用して取得する。

応答生成部106は、一般知識ベース辞書データ221のデータを利用して得られる知識により、以下の確認応答、すなわち、

「8日（はちにち）に買い物の予定を登録しますね？」

このような確認応答を生成して出力する。

[0061] この確認応答は、音声出力部（スピーカ）107を介してユーザに対して出力される。

ユーザは、「8日」の2つの読み方（よーか、はちにち）の確認を求める確認応答を聞き、情報処理装置220が伝えたいのは4日ではなく8日であると確信することができ、聞き間違いを防ぐことができる。

[0062] （確認応答生成例102）

図2に示す情報処理装置220の応答生成部106が生成する確認応答の第2の具体例について説明する。

図3（102）に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「8日に買い物の予定を入れてくれる？」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話

制御部 104 に入力される。

[0063] 対話制御部 104 は、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解の信頼度が低い場合、あるいは、ユーザの発話がスケジュール設定等の重要処理である場合、応答生成部 106 に対して確認応答の生成、出力を要求する。

[0064] 本例では、意図理解部 103 における意図理解の結果として、ユーザ発話が「8日（よーか）」を意図しているか、「4日（よっか）」を意図しているかが曖昧である（信頼度が低い）との意図理解結果を対話制御部 104 が入力する。

さらに、対話制御部 104 は、ユーザ発話がスケジュール設定を要求する重要な処理であるとの判断から、日にちの再確認が必要であるとの判断を行い、この判断に基づいて、応答生成部 106 に確認応答生成命令を出力する。

[0065] なお、対話制御部 104 から応答生成部 106 に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部 106 が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107 を介して出力する。

[0066] 応答生成部 106 は、一般知識ベース辞書データ 221 のデータを利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「来週水曜日の8日ですね？」

[0067] 上記の例は、ユーザの意図が「8日（よーか）」であるか、「4日（よっか）」であるかを確認するために装置が実行する確認応答である。

情報処理装置 220 の応答生成部 106 は、一般知識ベース辞書データ 221 のデータを利用して、8日は来週水曜日であるという知識を取得して、取得した知識に基づく情報（8日は来週水曜日である）を付加した確認応答を生成する。

[0068] この確認応答は、音声出力部（スピーカ）107 を介してユーザに対し

て出力される。

これにより、ユーザは、情報処理装置 220 が、4 日ではなく 8 日と認識していることを確信することができる。

また、例えば、ユーザ自身の真の意図が 9 日の木曜日に買い物の予定を登録したいと思っていた場合に、ユーザ自身の勘違いにより 8 日と発話してしまった場合であっても、上記の確認応答を装置から聞くことで、ユーザは 8 日を木曜日だと勘違いしていたことを気づくことになる。

すなわち、8 日が正しくは水曜日であることに気付かせることができる。その結果として、ユーザが、本来、予定を登録したかった木曜日は 9 日であると認識でき、「9 日に予定を入れておいて」のように、真の意図に基づく発話をさせることが可能となる。

[0069] (確認応答生成例 103)

図 2 に示す情報処理装置 220 の応答生成部 106 の生成する確認応答の第 3 の具体例について説明する。

図 3 (103) に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「3 日後に買い物の予定を入れていて。」

このようなユーザ発話があると、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解が行われ、意図理解結果が対話制御部 104 に入力される。

[0070] 対話制御部 104 は、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部 106 に対して確認応答の生成、出力を要求する。あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部 106 に対して確認応答の生成、出力を要求する。

[0071] 本例では、対話制御部 104 は、スケジュール設定の重要な処理であるとの判断から、日にちの再確認が必要であるとの判断を行い、この判断に基づいて、応答生成部 106 に確認応答生成命令を出力する。

[0072] なお、対話制御部104から応答生成部106に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部106が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107を介して出力する。

[0073] 応答生成部106は、一般知識ベース辞書データ221のデータを利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「11日木曜日ですね？」

[0074] 上記の例は、ユーザ発話に含まれる「3日後」を、実際の日にと曜日（11日木曜日）に置き換えて確認を求める形式の確認応答の例である。

情報処理装置220の応答生成部106は、一般知識ベース辞書データ221のデータを利用して、「3日後」は「11日」であり、さらに「木曜日」であるという知識を取得する。

[0075] 応答生成部106は、この知識を用いて、取得した知識を付加した（ここでは置き換えた）応答を生成する。すなわち、

「11日木曜日ですね？」

この確認応答を生成して、音声出力部（スピーカ）107を介してユーザに対して出力する。

これにより、ユーザは、自分自身が発話した「3日後」に加え、「11日」「木曜日」という情報を取得する。

この付加的な情報によって、予定を登録する日時に対する確信度を上げることができる。

[0076] さらに、ユーザの真の意図は金曜日（4日後）への予定登録だった場合、すなわち、ユーザが「3日後が金曜日である」と誤解していた場合には、上記の情報処理装置220からの確認応答、すなわち、

「11日木曜日ですね？」

この確認応答によって、「木曜日」という情報がユーザに提供される。

この結果、ユーザは、「3日後は金曜日だと思っていたけど木曜日だった

のか。じゃあ、4日後の12日にしよう」と考えなおして、新たな発話、
「4日後に予定を入れておいて」
という訂正発話を行うことが可能となる。

[0077] このように、1つの情報に対して他の表現を加えて応答することにより、ユーザは複数の観点から確認することができるようになる。この例の場合、ユーザが自身の勘違いに気付く元ともなっており、結果としてユーザが求めているゴールに早く到達できることになる。また、システムがユーザの発話を正しく理解しているかどうかの確認も同時に行うことができる。

[0078] 以上、説明した3つの具体例（確認応答生成例101～103）は、日付に関する例であるが、日付以外にも、図2に示す一般知識ベース辞書データ221から得られる知識を用いて、ユーザの意図を正確に取得可能とした確認応答を生成することができる。

以下、日付の確認と異なる確認応答の生成例について説明する。

図4に示す2つの確認応答生成例104～105について説明する。

[0079] （確認応答生成例104）

図2に示す情報処理装置220の応答生成部106の生成する確認応答の第4の具体例について説明する。

図4（104）に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「金閣寺への行き方を教えて」

このようなユーザ発話があると、

まず、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。

[0080] 本例では、例えば、情報処理装置220の意図理解部103は、意図理解処理において、「金閣寺」が、本当に「金閣寺（キンカクジ）」を意図しているのか、あるいは「銀閣寺（ギンカクジ）」を意図して誤って発話してしまったかを判別できず、ユーザの意図が曖昧であるという意図理解結果を対

話制御部 104 に出力する。

[0081] 対話制御部 104 は、この意図理解結果に基づいて、応答生成部 106 に確認応答生成命令を出力する。

情報処理装置 220 の応答生成部 106 は、対話制御部 104 からの確認応答生成命令の入力に応じて確認応答を生成する。

[0082] なお、対話制御部 104 から応答生成部 106 に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部 106 が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107 を介して出力する。

[0083] 応答生成部 106 は、一般知識ベース辞書データ 221 のデータを利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「足利義満が作った金閣寺ですか？」

[0084] 上記の例は、ユーザの発話意図が「金閣寺（キンカクジ）」を意図しているのか、あるいは「銀閣寺（ギンカクジ）」を意図しているかが、曖昧だったので、確認をしたい場合の確認応答である。

[0085] この確認応答において、例えば、そのまま「金閣寺ですか？」と応答を出した場合、その応答が音声のみの場合には、ユーザがまた聞き間違える（金閣寺を銀閣寺と聞き間違える）可能性がある。

例えば情報処理装置が表示部を有する場合には、表示部に音声出力した文字を表示することにより、ユーザの聞き違いを防ぐこともできるが、ユーザが表示部を見るときは限らない。

従って、音声だけでも誤解が生じにくい応答にすることが望ましい。

[0086] よって、上記の例、すなわち、

装置の確認応答：「足利義満が作った金閣寺ですか？」

この確認応答のように、「足利義満が作った」という情報を追加することにより、ユーザにとっての確信度を上げることが可能となる。

[0087] なお、応答生成部 106 は、一般知識ベース辞書データ 221 から得られ

る知識に基づいて、様々な情報を付加した確認応答を生成することが可能である。

例えば、以下のような様々な確認応答を生成することが可能である。

「黄金色の金閣寺ですか？」

「三島由紀夫の小説になった金閣寺ですか？」

「正式名称は鹿苑寺と呼ばれる金閣寺ですか？」

例えば、これらの金閣寺に関する付加的な情報を、一般知識ベース辞書データ 221 から取得して、付加的情報を追加した確認応答を生成して出力することが可能である。

[0088] (確認応答生成例 105)

図 2 に示す情報処理装置 220 の応答生成部 106 の生成する確認応答の第 5 の具体例について説明する。

図 4 (105) に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「1チャンネル(いっちゃんねる)を5時から録画して」

このようなユーザ発話があると、

まず、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解が行われ、意図理解結果が対話制御部 104 に入力される。

[0089] 本例は、情報処理装置 220 の意図理解部 103 は、意図理解処理において、録画すべきチャンネルを指定していると理解したが、それが 1ch(いっちゃんねる)か 10ch(じゅっちゃんねる)かが曖昧なため、ユーザの意図が曖昧であるという意図理解結果を対話制御部 104 に出力する。

[0090] 対話制御部 104 は、この意図理解結果に基づいて、応答生成部 106 に確認応答生成命令を出力する。

情報処理装置 220 の応答生成部 106 は、対話制御部 104 からの確認応答生成命令の入力に応じて確認応答を生成する。

[0091] なお、対話制御部 104 から応答生成部 106 に対して、確認応答生成命

令ではない、通常の応答生成命令が入力された場合であっても、応答生成部 106 が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107 を介して出力する。

[0092] 応答生成部 106 は、一般知識ベース辞書データ 221 のデータを利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「XYZテレビでよろしいですか？」

[0093] 上記の例は、ユーザの発話意図は録画すべきチャンネルを指定していると理解したが、それが1chか10chかが曖昧なため、ユーザへの確認を行った例である。

応答生成部 106 は、一般知識ベース辞書データ 221 のデータを利用して、1チャンネルが「XYZテレビ」のチャンネルであるという知識を取得し、聞き間違いを起こしやすい数字+チャンネルを提示するのではなく、放送局名を含む確認応答を生成する。

この処理により、ユーザは確認応答の意図を正しく理解することができるようになる。

[0094] [3.（実施例2）応答生成部がアプリ（アプリケーション）実行部を利用して確認応答を生成する実施例について]

次に、実施例2として、応答生成部がアプリ（アプリケーション）実行部を利用して確認応答を生成する実施例について説明する。

[0095] 図5は、本実施例2の情報処理装置240の構成例を示す図である。

図5に示すように、情報処理装置240は、音声入力部（マイク）101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、応答生成部106、音声出力部（スピーカ）107を有し、さらに、応答生成部106が応答生成に利用するアプリ実行部241を有する。

[0096] 音声入力部（マイク）101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、音声出力部（スピーカ）107については、図1、図2を参照して説明したと同様の処理を行う構成であるの

で説明を省略する。

[0097] 図5に示す本実施例2の情報処理装置240の応答生成部106は、アプリ実行部241を利用して確認応答を生成する。

なお、アプリ実行部241は、例えば、情報処理装置240内のアプリケーションに限らず、ネットワークによって接続された様々な装置において実行されるアプリケーションを利用することを可能な構成である。

例えば、音楽再生アプリ、天気情報取得アプリ、レストラン予約アプリ等、様々なアプリを利用することができる。

[0098] 本実施例の情報処理装置240の応答生成部106の生成する確認応答の具体例について説明する。

まず、図6に示す2つの確認応答生成例201~202について説明する。

[0099] (確認応答生成例201)

図6(201)に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「10日に買い物の予定を入れてくれる？」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。

[0100] 対話制御部104は、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部106に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部106に対して確認応答の生成、出力を要求する。

[0101] 本例では、ユーザ発話がスケジュール設定等の重要処理であると判断し、日にちの再確認が必要であるとの判断を行い、この判断に基づいて、応答生成部106に確認応答生成命令を出力する。

[0102] なお、対話制御部104から応答生成部106に対して、確認応答生成命

令ではない、通常の応答生成命令が入力された場合であっても、応答生成部 106 が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107 を介して出力する。

[0103] 応答生成部 106 は、アプリ実行部 241 を利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「10日ですね。ライブに行く日ですね。」

上記の例は、ユーザ発話に含まれる買い物の日付について、再確認を行う確認応答である。

応答生成部 106 は、アプリ実行部 241 を利用して、ユーザのスケジュールが登録されているスケジュールアプリを起動して、ユーザ発話に含まれていた日付（10日）の予定を検索する。

その結果、ユーザの10日の予定として、「ライブに行く予定」が設定されていたことを確認する。

[0104] 応答生成部 106 は、アプリ実行部 241 を利用して取得したこの確認結果を利用して、ユーザへ提示する情報量を増やした確認応答を生成する。すなわち、以下の確認応答を生成する。

装置の確認応答：「10日ですね。ライブに行く日ですね。」

[0105] 音声出力部 107 を介して上記の確認応答を聞いたユーザは、ライブに行くのは9日だと思っていたけど、10日だったのか。じゃあ買い物は11日にしようと考え、

その結果、新たなユーザ発話、

「いや、11日にして」

を行い、予定をスムーズに修正することが可能となる。

[0106] （確認応答生成例 202）

図5に示す情報処理装置 240 の応答生成部 106 の生成する確認応答の第2の具体例について説明する。

図6（202）に示すように、例えば以下のような、「ユーザの発話」が

なされたとする。

ユーザ発話：「銀閣寺への行き方を教えて」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。

[0107] 対話制御部104は、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部106に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部106に対して確認応答の生成、出力を要求する。

[0108] 本例では、対話制御部104は、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い、すなわち、ユーザ発話が「金閣寺」なのか「銀閣寺」なのかについての確信度が低いとの判定に基づいて、どちらなのかを確認するため応答生成部106に対して確認応答の生成、出力を要求する。

[0109] なお、対話制御部104から応答生成部106に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部106が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107を介して出力する。

[0110] 応答生成部106は、アプリ実行部241を利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「昨年8月に行った金閣寺ですね？」

[0111] 応答生成部106は、アプリ実行部241において起動されたスケジュールアプリ（スケジュール登録アプリケーション）を利用して「金閣寺」に関する予定が過去になかったかどうかを検索し、その結果を合わせて、ユーザへ提示する情報量を増やした確認応答を生成する。すなわち、以下の確認応答を生成する。

装置の確認応答：「昨年 8 月に行った金閣寺ですね？」

[0112] 音声出力部 107 を介して上記の確認応答を聞いたユーザは、銀閣寺には行ったことがない、ということは、「装置側で「銀閣寺」を「金閣寺」と聞き間違えている」のだと気付くことができる。その結果、新たなユーザ発話、

「いや、金閣寺ではなくて銀閣寺」

を行い、正確なコミュニケーションを行うことが可能となる。

[0113] (確認応答生成例 203)

図 5 に示す情報処理装置 240 の応答生成部 106 の生成する確認応答の第 3 の具体例について説明する。

図 7 (203) に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「佐藤さん家族とバーベキューって入れて」

このようなユーザ発話があると、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解が行われ、意図理解結果が対話制御部 104 に入力される。

[0114] 対話制御部 104 は、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部 106 に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部 106 に対して確認応答の生成、出力を要求する。

[0115] 本例では、対話制御部 104 は、「佐藤さん」なのか「加藤さん」なのかについての確信度が低いとの判定に基づいて、どちらなのかを確認するため応答生成部 106 に対して確認応答の生成、出力を要求する。

[0116] なお、対話制御部 104 から応答生成部 106 に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部 106 が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）10

7を介して出力する。

[0117] 応答生成部106は、アプリ実行部241を利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「お向かいの佐藤さんでよろしいですか？」

[0118] 応答生成部106は、アプリ実行部241において起動された住所登録アプリや知人登録アプリに登録されているご近所さん情報（名前と場所）を参照して、その結果を合わせて、ユーザへ提示する情報量を増やした確認応答を生成する。すなわち、以下の確認応答を生成する。

装置の確認応答：「お向かいの佐藤さんでよろしいですか？」

[0119] 音声出力部107を介して上記の確認応答を聞いたユーザは、（お向かいの佐藤さんで合っている、システムは自分の発話を正しく推定している）と考えることができ、新たなユーザ発話、

「そうだよ」

のように回答することができる。

[0120] 以上、説明した確認応答生成例201～203は、音声認識や意図理解結果の信頼度が低い場合の例であるが、音声認識や意図理解結果の信頼度が高いものの、その内容が複数の候補を持つ場合についても、同様の手段により、問題を解決することができる。

以下、このような確認応答生成例について説明する。

[0121] （確認応答生成例204）

図5に示す情報処理装置240の応答生成部106の生成する確認応答の第4の具体例について説明する。

図7（204）に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「佐藤さん家族とバーベキューって入れて」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。

[0122] 対話制御部104は、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部106に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部106に対して確認応答の生成、出力を要求する。

[0123] 本例では、ユーザの発話がスケジュール設定等の重要処理であり、応答生成部106に対して確認応答の生成、出力を要求する。

なお、対話制御部104から応答生成部106に対して、確認応答生成命令ではない、通常の応答生成命令が入力された場合であっても、応答生成部106が、確認応答を実行すべきか否かを自ら判断して、確認応答を実行すべきと判断した場合には、確認応答を生成して音声出力部（スピーカ）107を介して出力する。

[0124] 応答生成部106は、アプリ実行部241を利用して、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「ご近所の佐藤さんでよろしいですか？」

[0125] 応答生成部106は、アプリ実行部241において起動された住所登録アプリや知人登録アプリに登録されている「佐藤さん」に該当する人データが2人ある。登録住所が近所の佐藤さんと、登録住所が遠方がかつ会社の同僚である佐藤さんが登録されており、どちらの佐藤さんなのかについて確認するための確認応答を生成する。すなわち、以下の確認応答を生成する。

装置の確認応答：「ご近所の佐藤さんでよろしいですか？」

[0126] 音声出力部107を介して上記の確認応答を聞いたユーザは、（そっか、佐藤さんは複数人いるのか、でもご近所の佐藤さんじゃない方で登録したい）と考えることができ、新たなユーザ発話、例えば

「いや、同僚の佐藤さん」

上記のように回答することができる。

[0127] なお、応答生成部106は、上述した例の他にも、例えば、住所登録アプリに登録されている佐藤さんの下の名前を用いて

「佐藤太郎さんでよろしいですか？」

このような応答を生成してもよい。

[0128] [4. (実施例3) 応答生成部が情報入力部の情報を利用して確認応答を生成する実施例について]

次に、実施例3として、応答生成部が、例えば画像入力部等、音声以外の情報入力部の情報を利用して確認応答を生成する実施例について説明する。

[0129] 図8は、本実施例3の情報処理装置260の構成例を示す図である。

図5に示すように、情報処理装置260は、音声入力部(マイク)101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、応答生成部106、音声出力部(スピーカー)107を有し、さらに、画像入力部(カメラ)261、画像認識部262、オブジェクト(画像)データベース263、オブジェクト関連知識ベース辞書データ264を有する。

[0130] 音声入力部(マイク)101、音声認識部102、意図理解部103、対話制御部104、アプリ実行部105、音声出力部(スピーカー)107については、図1、図2を参照して説明したと同様の処理を行う構成であるので説明を省略する。

[0131] 図8に示す本実施例3の情報処理装置260は、発話を実行するユーザの画像を含むユーザ周囲の画像を入力する画像入力部(カメラ)261を有する。

画像入力部(カメラ)261の撮影画像は、画像認識部262に入力され、画像認識部262は、予め様々なオブジェクト画像が登録されたデータベースであるオブジェクト(画像)データベース263の登録情報と対比して、撮影画像に含まれる様々なオブジェクト(人や物品)を識別する。

[0132] 応答生成部106は、画像認識部262が識別した発話ユーザの周囲環境のオブジェクト情報を入力するとともに、オブジェクト関連知識ベース辞書データ264を利用して確認応答を生成する。

[0133] 本実施例の情報処理装置260の応答生成部106の生成する確認応答の

具体例について説明する。

図9に示す確認応答生成例301について説明する。

[0134] (確認応答生成例301)

図9(301)に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「スピーカーから音楽を再生して」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力される。

[0135] 対話制御部104は、音声認識部102におけるテキスト変換、および意図理解部103における意図理解の信頼度が低い否かを判定し、信頼度が低い場合、応答生成部106に対して確認応答の生成、出力を要求する。

あるいは、ユーザの発話がスケジュール設定等の重要処理である場合にも、応答生成部106に対して確認応答の生成、出力を要求する。

[0136] 本例は、対話制御部104は、ユーザ発話の意図は、「スピーカーから音楽を再生して」という意図であることを理解できており、対話制御部104は、応答生成部106に対して確認応答の生成、出力を要求しない可能性がある。

[0137] しかし、応答生成部106は、画像認識部262が識別した発話ユーザの周囲環境のオブジェクト情報を入力し、オブジェクト関連知識ベース辞書データ264の辞書データを参照することにより、発話ユーザの周囲に複数のスピーカーが存在していることを把握することができる。

このような場合、応答生成部106自身で確認応答を実行すべきと判断し、確認応答を生成して音声出力部(スピーカ)107を介して出力する。

具体的には、ユーザに対して確認を求める以下の確認応答を生成して出力する。

装置の確認応答：「あなたの近くにあるスピーカーでよろしいですか？」

[0138] この例は、応答生成部106が、発話ユーザの近くにスピーカーがあるこ

とを画像認識部 262 からの入力によって知り、さらに、オブジェクト関連知識ベース辞書データ 264 に基づいて制御可能なスピーカーが複数あることを知ったことによって生成された確認応答である。

ユーザからは「スピーカー」とのみ指定されているため、どのスピーカーなのかについて確認をしている。

[0139] 画像認識部 262 が解析した画像情報より、発話ユーザの近くにスピーカーがあることが確認できたため、その結果を合わせた応答文章を出力し、ユーザへ提示する情報量を増やして、以下の確認応答を生成して出力する。

装置の確認応答：「あなたの近くにあるスピーカーでよろしいですか？」

[0140] 音声出力部 107 を介して上記の確認応答を聞いたユーザは（そうそう、このスピーカー）と考えることができ、その結果、新たなユーザ発話、

「うん、そうだよ」

のように回答することができる。

[0141] 次に、本実施例 3 の変形例として、情報入力部からの入力情報に基づいて話者識別を行う構成と処理について図 10 以下を参照して説明する。

図 10 に示す情報処理装置 280 は、

図 8 を参照して説明した構成と類似する構成であるが、以下の点が異なる。

。

図 8 の画像認識部 262 を、図 10 のユーザ識別部 282 に変更、

図 8 のオブジェクト（画像）データベース 263 を、図 10 のユーザ画像データベース 283 に変更、

図 8 のオブジェクト関連知識ベース辞書データ 264 を、図 10 のユーザ情報関連知識ベース辞書データ 284 に変更、

これらの点が異なる。

[0142] 図 10 に示す情報処理装置 280 は、発話を実行するユーザの周囲の人の顔画像を画像入力部（カメラ） 261 からユーザ識別部 282 に入力する。

ユーザ識別部 282 は、予め様々なユーザ画像が登録されたデータベースであるユーザ画像データベース 283 の登録情報と対比して、撮影画像に含ま

れる人を識別する。

[0143] 応答生成部106は、ユーザ識別部282が識別したユーザ情報を入力するとともに、ユーザ情報関連知識ベース辞書データ284を利用して確認応答を生成する。

[0144] 本実施例の情報処理装置280の応答生成部106の生成する確認応答の具体例について説明する。

図11に示す確認応答生成例302について説明する。

[0145] (確認応答生成例302)

ここでは、ユーザ発話の前段階で、ユーザと情報処理装置280による対話が行われ、今晚の食事場所を決定する処理が行われたとする。その後、図11(302)に示すように、以下の「ユーザの発話」がなされたとする。

ユーザ発話：「じゃあユウキにメールしておいて」

このようなユーザ発話があると、音声認識部102におけるテキスト変換、および意図理解部103における意図理解が行われ、意図理解結果が対話制御部104に入力され、対話制御部104は、応答生成部106に対して応答生成、出力を要求する。

[0146] しかし、応答生成部106は、ユーザ識別部282が識別したユーザの周囲にいる人物情報を入力し、さらに、ユーザ情報関連知識ベース辞書データ284の辞書データを参照することにより、発話ユーザの周囲に「ユウキ」がいないことを把握する。

[0147] さらに、ユーザ識別部282の識別情報と、ユーザ情報関連知識ベース辞書データ284の辞書データに基づいて、発話ユーザがパパであることを判別し、ママと娘のユウコも一緒にいることを認識する。

[0148] 応答生成部106は、発話ユーザであるパパのユーザ発話：「じゃあユウキにメールしておいて」に含まれる人名「ユウキ」が、「ユウコ」さんである可能性もあると判定し、メール先が「ユウキ」さんであるか、「ユウコ」さんであるかを確認させるための確認応答を生成する。すなわち以下の確認応答である。

装置の確認応答：「ユウキくんに送りますね？ユウキさんは今ここにはいないですよね？」

[0149] 上記の確認応答の例は、応答生成部106が、ユーザ識別部282の識別情報と、ユーザ情報関連知識ベース辞書データ284の辞書データに基づいて、対話しているのはパパであり、またママと娘のユウコも一緒に画面を見ていることを認識し、発話者の意図が「ユウコ」さんなのか「ユウキ」さんなのかについての確信度が低いと判定した結果に基づく確認応答の例である。

[0150] ユーザ情報関連知識ベース辞書データ284に登録されている家族は、パパ・ママ・ユウキ（息子）・ユウコ（娘）の4人であるが、ユーザ識別部282の識別情報に基づいて、ユウコさんが今ここにいることを認識しているため、その結果を合わせて応答文章とすることにより、ユーザへ提示する情報量を増やしている。その結果、ユーザは、（そう、ユウキに送って欲しい）と考えることができ、その結果、新たなユーザ発話、

「お願いね」

のように回答することができる。

[0151] さらに、図10に示す情報処理装置280における応答生成部106は、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用することで、以下のような処理を行うこともできる。

[0152] 応答生成部106は、応答文言を作成する際に発話ユーザの属性（年齢等）によって確認応答の構成を変えることができる。

具体例を図12に示す。

[0153] （確認応答生成例303）

確認応答例303は、応答生成部106が、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用して、発話ユーザの年齢が低いと判断した場合の例である。

図12（303）に示すように、以下の「ユーザの発話」がなされたとす

る。

ユーザ発話：「金閣寺への行き方を教えて」

[0154] 応答生成部106は、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用して、発話ユーザの年齢が低いと判断する。

この場合、応答生成部106は、以下の確認応答を生成して出力する。

装置の確認応答：「黄金色でピカピカしている金閣寺ですか？」

ユーザの年齢が低い場合、金閣寺や銀閣寺についての詳細な知識を持っていない可能性が高い。よって、例えば「足利義満が作った」というような知識を応答文に付加しても、ユーザがそれに対する知識を持ち合わせていなければ、かえって混乱の元になってしまう。そこで、そのようなユーザに向けては誰にでもわかりやすいような知識を付加することにより、曖昧性をより確実に解消することができる。

[0155] (確認応答生成例304)

確認応答例304は、応答生成部106が、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用して、発話ユーザの年齢が高いと判断した場合の例である。

図12(304)に示すように、以下の「ユーザの発話」がなされたとする。

ユーザ発話：「10日に外出の予定を登録」

[0156] 応答生成部106は、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用して、発話ユーザの年齢が高いと判断する。

この場合、応答生成部106は、以下の確認応答を生成して出力する。

装置の確認応答：「10日の旗日でよろしいですか？」

[0157] 「10日」が国民の祝日の場合、「祝日」「休みの日」「休日」などいろいろな言葉で表現することができるが、最近使われなくなった表現として「旗日」がある。ユーザの年代(や環境)によっては「旗日」の方が簡単に理

解できる場合もあり、そのような場合にはその年代向けの表現を用いて応答文を生成して出力する。

[0158] (確認応答生成例 305)

確認応答例 305 は、応答生成部 106 が、ユーザ識別部 282 のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ 284 の情報を利用して、発話ユーザがある中学校の学生であると判断した場合の例である。

図 12 (305) に示すように、以下の「ユーザの発話」がなされたとする。

ユーザ発話：「学校に電話かけて」

[0159] この発話からは、多数存在する学校のうち、どの学校なのかを特定することができない。また、ユーザの家族に学生が二人以上いる場合（例えば小学生と中学生等）にも、どちらの学校を指しているかを特定することができない。

応答生成部 106 は、ユーザ識別部 282 のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ 284 の情報を利用して、発話したユーザを特定し、かつ、そのユーザが、品川中学の学生である知識を取得する。この取得知識に基づいて、応答生成部 106 は、以下の確認応答を生成して出力する。

装置の確認応答：「品川中学校でよろしいですね」

[0160] このように、学校名を用いた確認応答を発話ユーザに提示することで、発話ユーザは、装置側が、ユーザ発話を正しく認識していることを容易に確認することが可能となる。

[0161] さらに、応答生成部 106 は、ユーザの状況や履歴に応じて、応答内容を切り替えることもできる。

具体例を図 13 に示す。

[0162] (確認応答生成例 306, 307)

図 13 に示す確認応答例 306 は、発話ユーザ（家族中の属性：父親）が一人で対話している場合の応答例であり、

確認応答例307は、発話ユーザ（家族中の属性：父親）が家族とともに対話している場合の応答例である。

[0163] 図13（306）に示すように、以下の「ユーザの発話」がなされたとする。

ユーザ発話：「おばあさんの家までの道路状況を教えて」

[0164] 応答生成部106は、ユーザ識別部282のユーザ識別情報や、ユーザ情報関連知識ベース辞書データ284の情報を利用して、発話したユーザが父親であること、さらに、確認応答例306の場合は、発話ユーザ（家族中の属性：父親）が一人で対話していることを識別したものとする。

この場合、応答生成部106は、以下の確認応答を生成して出力する。

装置の確認応答：「△△さん（ユーザの祖母）の家ですね？」

[0165] 一方、図14に示す（307）確認応答例307は、発話ユーザ（家族中の属性：父親）が家族とともに対話していることを識別した場合の例である。

この場合、応答生成部106は、以下の確認応答を生成して出力する。

装置の確認応答：「〇〇さん（ユーザの母親、ユーザの子供たちから見た祖母）の家ですね？」

[0166] これらの例は、ユーザ発話が「おばさん」「おばあさん」のどちらを発話したかの確信度が低い場合である。ここで、「おばあさん」という言葉が指す相手は一意ではなく、誰にとっての祖母であるかによりその対象相手が異なる。ここで、ユーザが発話している状況がユーザ一人だけの場合は、ユーザ自身の祖母であり、ユーザがその子供たちと一緒に発話している状況であれば、子供たちにとっての祖母を指す可能性が高いという判断の下に、「おばあさん」にあたると思われる人の名前を応答文に入れている。このように、ユーザの状況に応じて確認応答を切り替えることにより、その場にいる対話参加者がより容易に内容を確認することが可能となる。

[0167] 次に、本実施例3の変形例として、情報入力部を例えば、温度、照度、ユーザの脈拍等を検出するセンサーを用い、センサーからの入力情報に基づい

て確認応答を生成する構成と処理について図14以下を参照して説明する。

図14に示す情報処理装置290は、

図8を参照して説明した構成と類似する構成であるが、以下の点異なる。

。

図8の画像入力部（カメラ）261を、図14のセンサー291に変更、

図8の画像認識部262を、図14のセンサー検出信号解析部292に変更、

図8のオブジェクト（画像）データベース263を、図14では省略、

図8のオブジェクト関連知識ベース辞書データ264を、図14のセンサー検出情報関連知識ベース辞書データ293に変更、

これらの点異なる。

[0168] 図14に示す情報処理装置290は、発話を実行するユーザやユーザの周囲の様々なオブジェクトについての状況（コンテキスト）を検出するセンサー291の検出情報をセンサー検出信号解析部292に入力する。

センサー291は、例えば、発話ユーザの部屋の温度や湿度、照度等を検出するセンサー、あるいは発話ユーザの周辺にある様々な機器の動作状況を検出するセンサー、例えばテレビやエアコンの動作状況、さらに、ユーザの体温、心拍数、脈拍等の検出センサーである。

[0169] センサー検出信号解析部292は、これらのセンサーの検出信号を解析して、解析結果を応答生成部106に入力する。

応答生成部106は、センサー検出信号解析部292が解析したセンサー検出情報を入力するとともに、センサー検出情報関連知識ベース辞書データ293を利用して確認応答を生成する。

[0170] 本実施例の情報処理装置290の応答生成部106の生成する確認応答の具体例について説明する。

図15に示す確認応答生成例308について説明する。

[0171] （確認応答生成例308）

以下に説明する例は、センサー291として、発話ユーザの部屋の温度を

計測する温度センサーを利用した場合の例である。

センサー検出信号解析部 292 は、発話一ザの部屋の温度が 10℃（寒い）であることを検出して、応答生成部 106 にこの温度情報（10℃）を入力する。

その後、図 15（308）に示すように、以下の「ユーザの発話」がなされたとする。

ユーザ発話：「飲み物をください」

このようなユーザ発話があると、音声認識部 102 におけるテキスト変換、および意図理解部 103 における意図理解が行われ、意図理解結果が対話制御部 104 に入力され、対話制御部 104 は、応答生成部 106 に対して応答生成、出力を要求する。

[0172] ここで、応答生成部 106 は、センサー検出信号解析部 292 から入力した情報に基づいて、発話一ザの部屋の温度が 10℃（寒い）であると判断して、以下のような確認応答を生成して出力する。

装置の確認応答：「ホットコーヒーでよろしいですか？それともお茶がよろしいでしょうか？」

[0173] 上記の確認応答の例は、ユーザに提供可能な飲み物としてコーヒー、お茶等の温かい飲み物と、牛乳、ジュースなどの冷たい飲み物がある状況において、応答生成部 106 が、センサー検出信号解析部 292 から入力した情報に基づいて、気温が 10℃と低い状態であることを認識して、発話ユーザが冷たい飲み物よりも温かい飲み物を欲しがっている可能性が高いと判断する。

[0174] なお、寒い場合は、コーヒー、お茶等の温かい飲み物をを欲しがっている可能性が高く、暑い場合は牛乳、ジュースなどの冷たい飲み物を欲しがっている可能性が高いとの知識は、センサー検出情報関連知識ベース辞書データ 293 から取得する。

[0175] [5.（実施例 4）音声出力に併せて、画像や制御信号の出力を実行する実施例について]

次に、実施例4として、音声出力に併せて、画像や制御信号の出力を実行する実施例について説明する。

[0176] 以下に説明する実施例は、ユーザに対する音声による確認応答の出力に併せて画像や制御信号の出力を実行する実施例である。

ユーザが応答の内容を容易に理解できるように、音声以外の情報を用いた構成である。

[0177] 図16を参照して本実施例の情報処理装置310の構成例について説明する。

図16に示す情報処理装置310は、先に実施例1として説明した図2に示す情報処理装置220の構成に、画像出力部（モニタ）311を追加した構成である。

この情報処理装置310は、応答生成部106は、音声出力部（スピーカ）107を介して、確認応答の音声を出力するとともに、画像出力部（モニタ）311を介して、確認応答の音声に応じて説明画像を出力する。

[0178] 本実施例の情報処理装置310の応答生成部106の生成する確認応答の具体例について説明する。

図17に示す確認応答生成例401について説明する。

[0179] （確認応答生成例401）

図17に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「あしたという曲をかけて」

[0180] ここで、「あした」という曲が、複数の歌手によって歌われているとする。

すなわち、情報処理装置310は、ユーザがどの歌手の「あした」をかけてほしいかを確認する必要がある。

応答生成部106は、一般知識ベース辞書データ221のデータに基づいて、

「あした」という曲が、複数の歌手によって歌われていることを確認し、

その歌手名や、CDジャケット画像等を取得する。

[0181] 応答生成部106は、一般知識ベース辞書データ221のデータを利用して得られた「あした」を歌っている歌手の歌手名や、CDジャケット画像等を画像出力部311に出力しながら、音声出力部107に以下の確認応答を出力する。

装置の確認応答：「どちらの『あした』を再生しましょうか？」

[0182] この例は、「あした」という曲名検索に対して、複数の曲が見つかった例である。複数の例を音声だけでユーザに伝える場合、いくつかの内容を同時に伝える必要がある。例えば、表記がひらがなのか、あるいは漢字が使われているのか、また、誰の曲なのか、を伝えなくてはならないが、音声だけで理解するのは困難である。

ここで、応答を伝えるモジュールとして画像が使える場合には、曲のタイトルやアーティスト名、ジャケット写真等を同時に表示することにより、ユーザは容易にその内容を理解することが可能となる。ただし、この場合はユーザがその画面を見ていることが前提となるため、別途センサーを用いた認識などにより、ユーザが画面を見ているか、あるいは画面を見ることができている位置にいるかどうかを判定した上で、画面を用いるか音声のみを用いるかを決定する構成とすることが好ましい。

[0183] さらに、本実施例の変形例として、応答生成部106が、確認応答の出力に併せて外部機器に対する制御信号を出力する例について説明する。

図18を参照して本実施例の情報処理装置320の構成例について説明する。

図18に示す情報処理装置320は、先に実施例3として説明した図8に示す情報処理装置260の構成をベースとしており、応答生成部106が音声出力部（スピーカー）107を介した音声出力を実行するのみならず、外部機器321に対する制御信号を出力する構成としたものである。

[0184] この情報処理装置310は、応答生成部106は、音声出力部（スピーカー）107を介して、確認応答の音声を出力するとともに、外部機器321

に対する制御信号を出力して、外部機器 3 2 1 を制御する。

[0185] 本実施例の情報処理装置 3 2 0 の応答生成部 1 0 6 の生成する確認応答の具体例について説明する。

図 1 9 に示す確認応答生成例 4 0 2 について説明する。

[0186] (確認応答生成例 4 0 2)

図 1 9 (4 0 2) に示すように、例えば以下のような、「ユーザの発話」がなされたとする。

ユーザ発話：「スピーカーから音楽を再生して」

[0187] ここで、発話ユーザの周囲には、複数のスピーカーがあるとする。

すなわち、情報処理装置 3 2 0 は、ユーザ周囲のどのスピーカーを利用して音楽再生を行うかを確認する必要がある。

応答生成部 1 0 6 は、画像認識部 2 6 2 から入力する複数のスピーカーの位置情報と、オブジェクト関連知識ベース辞書データから、各スピーカーに対する制御信号の出力態様情報を取得する。

[0188] さらに、応答生成部 1 0 6 は、取得した情報を利用して、例えば、まず、ユーザに一番近いスピーカー（外部機器）の LED を点滅させる制御信号をそのスピーカー（外部機器）に出力して、LED を点滅させながら、音声出力部 1 0 7 を介して以下の確認応答を出力する。

装置の確認応答：「こちらのスピーカーでよろしいですか？」

[0189] この例は、ユーザ周囲には複数のスピーカーがあるが、ユーザからは「スピーカー」とのみ指定されているため、どのスピーカーなのかについて確認をするための確認応答である。

[0190] 応答生成部 1 0 6 は、画像情報に基づいて、発話ユーザの近くにスピーカーがあることが確認できたため、そのスピーカーに対して LED 点滅制御信号を出力し、確認応答に含まれる「こちらのスピーカー」がどのスピーカーであるかをユーザに確実に認識させることが可能となる。

[0191] その結果、ユーザは（そうそう、このスピーカー）と考えることができ、「うん、そうだよ」

のように回答することができる。

[0192] ただし、この場合はユーザがそのスピーカークの近くにいることが前提となるため、別途センサーを用いた認識などにより、ユーザがスピーカークの近くにいるか、あるいはスピーカークを見ることが出来る位置にいるかどうかを判定した上で、最適な処理を行うように制御することが好ましい。

また、応答生成部106は、さらに例えばセンサーによる認識・識別結果を入力して、応答や制御信号の出力態様をセンサーによる認識・識別結果に応じて変更する構成としてもよい。

[0193] [6. (実施例5) 対話履歴に基づいて、確認応答を生成する実施例について]

次に、対話履歴に基づいて、確認応答を生成する実施例について説明する。

次に説明する実施例は、上述した実施例1～4のいずれにおいても利用可能な実施例であり、対話履歴に基づいて、確認応答を生成する実施例である。

[0194] 図20を参照して本実施例の応答生成部106の生成する確認応答の具体例について説明する。

図20には、最終的な装置の確認応答の前に、ユーザと装置間で以下の対話が行われた例を示している。

ユーザ発話：「横浜で遊べるところを教えて」

装置応答：「横浜みなとみらいに横浜コスモワールドという遊園地がありますよ」

ユーザ発話：「横浜みなとみらいのレストランを探して」

装置応答：「例えば、ランドマークタワーの中にいくつかレストランがありますよ。詳細をお知らせしましょうか？」

ユーザ発話：「港南区の地図を見せて」

[0195] ここで、「こうなん」という発話に対して、愛知県江南市、高知県香南市、横浜市港南区、新潟市江南区といった候補があり、一意に決められるだけ

の確信度がないものとする。

しかし、応答生成部106は、この発話以前に、横浜に関する対話を続けていたことにより、その履歴から横浜市港南区である可能性が高いと判断する。

そこで、応答生成部106は、『横浜の』を付与した確認応答、すなわち、以下の確認応答を生成して出力する。

装置確認応答：「横浜の港南区でよろしいですか？」

[0196] このような対話履歴に基づく付加情報（横浜の）を追加して確認応答を行うことで、ユーザは装置が正しく場所を認識したことを確認できる。

[0197] なお、図2～図20を参照して、本開示の情報処理装置の構成と処理の複数の例について説明したが、これらの各実施例は、任意の複数の実施例を組み合わせた構成とすることが可能である。さらに、以下において説明するように、図2他に示す情報処理装置の機能や処理は、複数の装置に分散させた構成とすることも可能である。

[0198] [7. 情報処理装置、および情報処理システムの構成例について]

複数の実施例について説明したが、これらの実施例において説明した例えば、図2他に示す情報処理装置の各処理機能は、すべて一つの装置、例えばユーザの所有するスマホやPC等の装置内に構成することも可能であるが、その一部をサーバ等において実行する構成とすることも可能である。

[0199] 図21にシステム構成例を示す。

図21（1）情報処理システム構成例1は、図2他に示す情報処理装置のほぼ全ての機能を一つの装置、例えばユーザの所有するスマホやPC、あるいは音声の入出力と画像の入出力機能を持つエージェント機器等のユーザ端末である情報処理装置410内に構成した例である。

ユーザ端末に相当する情報処理装置410は、例えば応答文生成時に外部アプリを利用する場合にのみ、アプリ実行サーバ420と通信を実行する。

[0200] 一方、図21（2）情報処理システム構成例2は、図2他に示す情報処理装置の機能の一部をユーザの所有するスマホやPC、エージェント機器等の

ユーザ端末である情報処理装置410内に構成し、一部を情報処理装置と通信可能なデータ処理サーバ460において実行する構成としたシステム例である。

例えば、図2に示す装置中の音声入力部101、音声出力部102のみ、あるいはこれらに画像入出力部や応答生成部106を含む機能のみをユーザ端末側の情報処理装置410側に設け、その他の機能をすべてサーバ側で実行するといった構成等が可能である。

[0201] [8. 情報処理装置のハードウェア構成例について]

次に、図22を参照して、情報処理装置のハードウェア構成例について説明する。

図22を参照して説明するハードウェアは、先に図2他を参照して説明した情報処理装置のハードウェア構成例であり、また、図21を参照して説明したデータ処理サーバ460を構成する情報処理装置のハードウェア構成の一例である。

[0202] CPU (Central Processing Unit) 501は、ROM (Read Only Memory) 502、または記憶部508に記憶されているプログラムに従って各種の処理を実行する制御部やデータ処理部として機能する。例えば、上述した実施例において説明したシーケンスに従った処理を実行する。RAM (Random Access Memory) 503には、CPU 501が実行するプログラムやデータなどが記憶される。これらのCPU 501、ROM 502、およびRAM 503は、バス504により相互に接続されている。

[0203] CPU 501はバス504を介して入出力インタフェース505に接続され、入出力インタフェース505には、各種スイッチ、キーボード、マウス、マイクロホン、センサーなどよりなる入力部506、ディスプレイ、スピーカーなどよりなる出力部507が接続されている。CPU 501は、入力部506から入力される指令に対応して各種の処理を実行し、処理結果を例えば出力部507に出力する。

[0204] 入出力インタフェース505に接続されている記憶部508は、例えばハードディスク等からなり、CPU501が実行するプログラムや各種のデータを記憶する。通信部509は、Wi-Fi通信、ブルートゥース（登録商標）（BT）通信、その他インターネットやローカルエリアネットワークなどのネットワークを介したデータ通信の送受信部として機能し、外部の装置と通信する。

[0205] 入出力インタフェース505に接続されているドライブ510は、磁気ディスク、光ディスク、光磁気ディスク、あるいはメモリカード等の半導体メモリなどのリムーバブルメディア511を駆動し、データの記録あるいは読み取りを実行する。

[0206] [9. 本開示の構成のまとめ]

以上、特定の実施例を参照しながら、本開示の実施例について詳解してきた。しかしながら、本開示の要旨を逸脱しない範囲で当業者が実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本開示の要旨を判断するためには、特許請求の範囲の欄を参酌すべきである。

[0207] なお、本明細書において開示した技術は、以下のような構成をとることができる。

(1) ユーザ発話を入力する音声入力部と、

前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、

前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、

前記確認応答生成部は、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理装置。

[0208] (2) 前記確認応答生成部は、

一般知識ベース辞書データから取得した一般知識によって構成される付加情報を用いた確認応答の生成を実行する(1)に記載の情報処理装置。

- [0209] (3) 前記一般知識は日付の読み方情報、またはカレンダー情報の少なくともいずれかである(2)に記載の情報処理装置。
- [0210] (4) 前記一般知識は歴史情報、または組織情報の少なくともいずれかである(2)に記載の情報処理装置。
- [0211] (5) 前記確認応答生成部は、
アプリケーション実行部において実行されるアプリケーションから得られる情報を用いた確認応答の生成を実行する(1)～(4)いずれかに記載の情報処理装置。
- [0212] (6) 前記アプリケーション実行部において実行されるアプリケーションはスケジュール登録アプリケーションであり、
前記確認応答生成部は、
前記スケジュール登録アプリケーションから得られる発話ユーザのスケジュール情報を用いた確認応答の生成を実行する(5)に記載の情報処理装置。
- [0213] (7) 前記アプリケーション実行部において実行されるアプリケーションは知人登録アプリケーションであり、
前記確認応答生成部は、
前記知人登録アプリケーションから得られる発話ユーザの知人情報を用いた確認応答の生成を実行する(5)に記載の情報処理装置。
- [0214] (8) 前記情報処理装置は、さらに、
前記発話ユーザの周辺画像を撮影する画像入力部、または前記発話ユーザの周辺情報を入力するセンサーを有し、
前記確認応答生成部は、
前記画像入力部からの入力画像、またはセンサー検出情報を解析し、解析結果から得られる情報を用いた確認応答を生成する(1)～(7)いずれかに記載の情報処理装置。
- [0215] (9) 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人または物を識別して、識別結

果を用いた確認応答を生成する（８）に記載の情報処理装置。

- [0216] （１０） 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の属性を識別して、識別属性に応じた確認応答の生成を実行する（８）または（９）に記載の情報処理装置。
- [0217] （１１） 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の年齢を識別して、識別年齢に応じた確認応答の生成を実行する（８）～（１０）いずれかに記載の情報処理装置。
- [0218] （１２） 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の構成を識別して、識別した人の構成に応じた確認応答の生成を実行する（８）～（１１）いずれかに記載の情報処理装置。
- [0219] （１３） 前記確認応答生成部は、
確認応答の音声出力に併せて、確認応答のための画像を生成して画像出力部へ出力する（１）～（１２）いずれかに記載の情報処理装置。
- [0220] （１４） 前記確認応答生成部は、
確認応答の音声出力に併せて、確認応答のための機器制御信号を生成して、制御対象となる外部機器へ出力する（１）～（１３）いずれかに記載の情報処理装置。
- [0221] （１５） 前記確認応答生成部は、
対話履歴情報に基づく確認応答生成処理を実行する（１）～（１４）いずれかに記載の情報処理装置。
- [0222] （１６） 前記情報処理装置は、
前記音声入力部の入力音声の認識処理を実行する音声認識部と、
前記音声認識部の認識結果に基づく意図理解処理を実行する意図理解部と、
、
前記意図理解部の生成した意図理解結果に基づく対話制御処理を実行する

対話制御部を有し、

前記対話制御部は、

前記意図理解部の生成した意図理解結果に基づいて、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を行うか否かを判定し、

前記確認応答生成部は、

前記対話制御部において確認応答を行うとの判定がなされた場合に、確認応答の生成処理を実行する（１）～（１５）いずれかに記載の情報処理装置

。

[0223] （１７） ユーザ端末と、データ処理サーバを有する情報処理システムであり、

前記ユーザ端末は、

ユーザ発話を入力する音声入力部と、

前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、

前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、

前記データ処理サーバは、

前記音声入力部の入力音声の認識処理を実行する音声認識部と、

前記音声認識部の認識結果に基づく意図理解処理を実行する意図理解部と

、

前記意図理解部の生成した意図理解結果に基づく対話制御処理を実行するとともに、意図理解結果に基づいて前記ユーザ発話に対する確認応答が必要か否かを判定し、確認応答が必要と判定した場合、前記ユーザ端末の確認応答生成部に対して確認応答生成命令を出力する対話制御部を有し、

前記ユーザ端末の前記確認応答生成部は、

前記データ処理サーバからの確認応答生成命令入力に応じて、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理システム。

[0224] （１８） 情報処理装置において実行する情報処理方法であり、

音声入力部が、ユーザ発話を入力する音声入力ステップと、
確認応答生成部が、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成ステップと、
音声出力部が、前記確認応答生成ステップで生成した確認応答を出力する音声出力ステップを実行し、
前記確認応答生成ステップにおいて、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理方法。

[0225] (19) ユーザ端末と、データ処理サーバを有する情報処理システムにおいて実行する情報処理方法であり、

前記ユーザ端末が、
ユーザ発話を入力する音声入力処理を実行し、
前記データ処理サーバが、
前記音声入力処理において入力された入力音声の音声認識処理と、
前記音声認識処理の生成した音声認識結果に基づく意図理解処理と、
前記意図理解処理の意図理解結果に基づいて、前記ユーザ発話に対する確認応答が必要か否か判定し、確認応答が必要と判定した場合、前記ユーザ端末に対して確認応答生成命令を出力し、

前記ユーザ端末が、
前記データ処理サーバからの確認応答生成命令入力に応じて、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成処理を実行し、

前記確認応答生成処理において、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成処理を実行する情報処理方法。

[0226] (20) 情報処理装置において情報処理を実行させるプログラムであり、

、
音声入力部に、ユーザ発話を入力させる音声入力ステップと、

確認応答生成部に、前記ユーザ発話に対する発話ユーザの確認を求めるための確認応答を生成させる確認応答生成ステップと、

音声出力部に、前記確認応答を出力させる音声出力ステップを実行させ、前記確認応答生成ステップにおいては、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行させるプログラム。

[0227] また、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。例えば、プログラムは記録媒体に予め記録しておくことができる。記録媒体からコンピュータにインストールする他、LAN (Local Area Network)、インターネットといったネットワークを介してプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

[0228] なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

産業上の利用可能性

[0229] 以上、説明したように、本開示の一実施例の構成によれば、ユーザ発話に対する確認応答を生成する際に、辞書データやアプリケーション実行結果から取得される付加情報を追加することで誤解が生じる可能性を低減した確認応答を生成する構成が実現される。

具体的には、例えば、ユーザ発話を入力する音声入力部と、ユーザ発話に

対する発話ユーザの確認を求め確認応答を生成する確認応答生成部と、確認応答生成部の生成した確認応答を出力する音声出力部を有し、確認応答生成部は、辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する。例えば、一般知識ベース辞書データや、アプリケーションの実行結果、あるいは、撮影画像やセンサー等から得られる情報を用いた確認応答を生成する。

本構成により、ユーザ発話に対する確認応答を生成する際に、辞書データやアプリケーション実行結果から取得される付加情報を追加することで誤解が生じる可能性を低減した確認応答を生成する構成が実現される。

符号の説明

- [0230] 100 情報処理装置
- 101 音声入力部（マイク）
- 102 音声認識部
- 103 意図理解部
- 104 対話制御部
- 105 アプリ実行部
- 106 応答生成部
- 107 音声出力部（スピーカ）
- 220, 240, 260, 280, 290, 310, 320 情報処理装置
- 221 一般知識ベース辞書データ
- 241 アプリ実行部
- 261 画像入力部（カメラ）
- 262 画像認識部
- 263 オブジェクト（画像）データベース
- 264 オブジェクト関連知識ベース辞書データ
- 281 画像入力部（カメラ）
- 282 ユーザ識別部

- 2 8 3 ユーザ画像データベース
- 2 8 4 ユーザ情報関連知識ベース辞書データ
- 2 9 1 センサー
- 2 9 2 センサー検出信号解析部
- 2 9 3 センサー検出情報関連知識ベース辞書データ
- 3 1 1 画像出力部（モニタ）
- 3 2 1 外部機器
- 4 1 0 情報処理装置
- 4 2 0 アプリ実行サーサーバ
- 4 6 0 データ処理サーバ
- 5 0 1 C P U
- 5 0 2 R O M
- 5 0 3 R A M
- 5 0 4 バス
- 5 0 5 入出力インタフェース
- 5 0 6 入力部
- 5 0 7 出力部
- 5 0 8 記憶部
- 5 0 9 通信部
- 5 1 0 ドライブ
- 5 1 1 リムーバブルメディア

請求の範囲

- [請求項1] ユーザ発話を入力する音声入力部と、
前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成部と、
前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、
前記確認応答生成部は、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理装置。
- [請求項2] 前記確認応答生成部は、
一般知識ベース辞書データから取得した一般知識によって構成される付加情報を用いた確認応答の生成を実行する請求項1に記載の情報処理装置。
- [請求項3] 前記一般知識は日付の読み方情報、またはカレンダー情報の少なくともいずれかである請求項2に記載の情報処理装置。
- [請求項4] 前記一般知識は歴史情報、または組織情報の少なくともいずれかである請求項2に記載の情報処理装置。
- [請求項5] 前記確認応答生成部は、
アプリケーション実行部において実行されるアプリケーションから得られる情報を用いた確認応答の生成を実行する請求項1に記載の情報処理装置。
- [請求項6] 前記アプリケーション実行部において実行されるアプリケーションはスケジュール登録アプリケーションであり、
前記確認応答生成部は、
前記スケジュール登録アプリケーションから得られる発話ユーザのスケジュール情報を用いた確認応答の生成を実行する請求項5に記載の情報処理装置。

- [請求項7] 前記アプリケーション実行部において実行されるアプリケーションは知人登録アプリケーションであり、
前記確認応答生成部は、
前記知人登録アプリケーションから得られる発話ユーザの知人情報を用いた確認応答の生成を実行する請求項5に記載の情報処理装置。
- [請求項8] 前記情報処理装置は、さらに、
前記発話ユーザの周辺画像を撮影する画像入力部、または前記発話ユーザの周辺情報を入力するセンサーを有し、
前記確認応答生成部は、
前記画像入力部からの入力画像、またはセンサー検出情報を解析し、解析結果から得られる情報を用いた確認応答を生成する請求項1に記載の情報処理装置。
- [請求項9] 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人または物を識別して、識別結果を用いた確認応答を生成する請求項8に記載の情報処理装置。
- [請求項10] 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の属性を識別して、識別属性に応じた確認応答の生成を実行する請求項8に記載の情報処理装置。
- [請求項11] 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の年齢を識別して、識別年齢に応じた確認応答の生成を実行する請求項8に記載の情報処理装置。
- [請求項12] 前記確認応答生成部は、
前記画像入力部からの入力画像に含まれる人の構成を識別して、識別した人の構成に応じた確認応答の生成を実行する請求項8に記載の情報処理装置。

- [請求項13] 前記確認応答生成部は、
確認応答の音声出力に併せて、確認応答のための画像を生成して画像出力部に出力する請求項1に記載の情報処理装置。
- [請求項14] 前記確認応答生成部は、
確認応答の音声出力に併せて、確認応答のための機器制御信号を生成して、制御対象となる外部機器に出力する請求項1に記載の情報処理装置。
- [請求項15] 前記確認応答生成部は、
対話履歴情報に基づく確認応答生成処理を実行する請求項1に記載の情報処理装置。
- [請求項16] 前記情報処理装置は、
前記音声入力部の入力音声の認識処理を実行する音声認識部と、
前記音声認識部の認識結果に基づく意図理解処理を実行する意図理解部と、
前記意図理解部の生成した意図理解結果に基づく対話制御処理を実行する対話制御部を有し、
前記対話制御部は、
前記意図理解部の生成した意図理解結果に基づいて、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を行うか否かを判定し、
前記確認応答生成部は、
前記対話制御部において確認応答を行うとの判定がなされた場合に、確認応答の生成処理を実行する請求項1に記載の情報処理装置。
- [請求項17] ユーザ端末と、データ処理サーバを有する情報処理システムであり、
前記ユーザ端末は、
ユーザ発話を入力する音声入力部と、
前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成

する確認応答生成部と、

前記確認応答生成部の生成した確認応答を出力する音声出力部を有し、

前記データ処理サーバは、

前記音声入力部の入力音声の認識処理を実行する音声認識部と、

前記音声認識部の認識結果に基づく意図理解処理を実行する意図理解部と、

前記意図理解部の生成した意図理解結果に基づく対話制御処理を実行するとともに、意図理解結果に基づいて前記ユーザ発話に対する確認応答が必要か否か判定し、確認応答が必要と判定した場合、前記ユーザ端末の確認応答生成部に対して確認応答生成命令を出力する対話制御部を有し、

前記ユーザ端末の前記確認応答生成部は、

前記データ処理サーバからの確認応答生成命令入力に応じて、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理システム。

[請求項18]

情報処理装置において実行する情報処理方法であり、

音声入力部が、ユーザ発話を入力する音声入力ステップと、

確認応答生成部が、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成ステップと、

音声出力部が、前記確認応答生成ステップで生成した確認応答を出力する音声出力ステップを実行し、

前記確認応答生成ステップにおいて、

辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行する情報処理方法。

[請求項19]

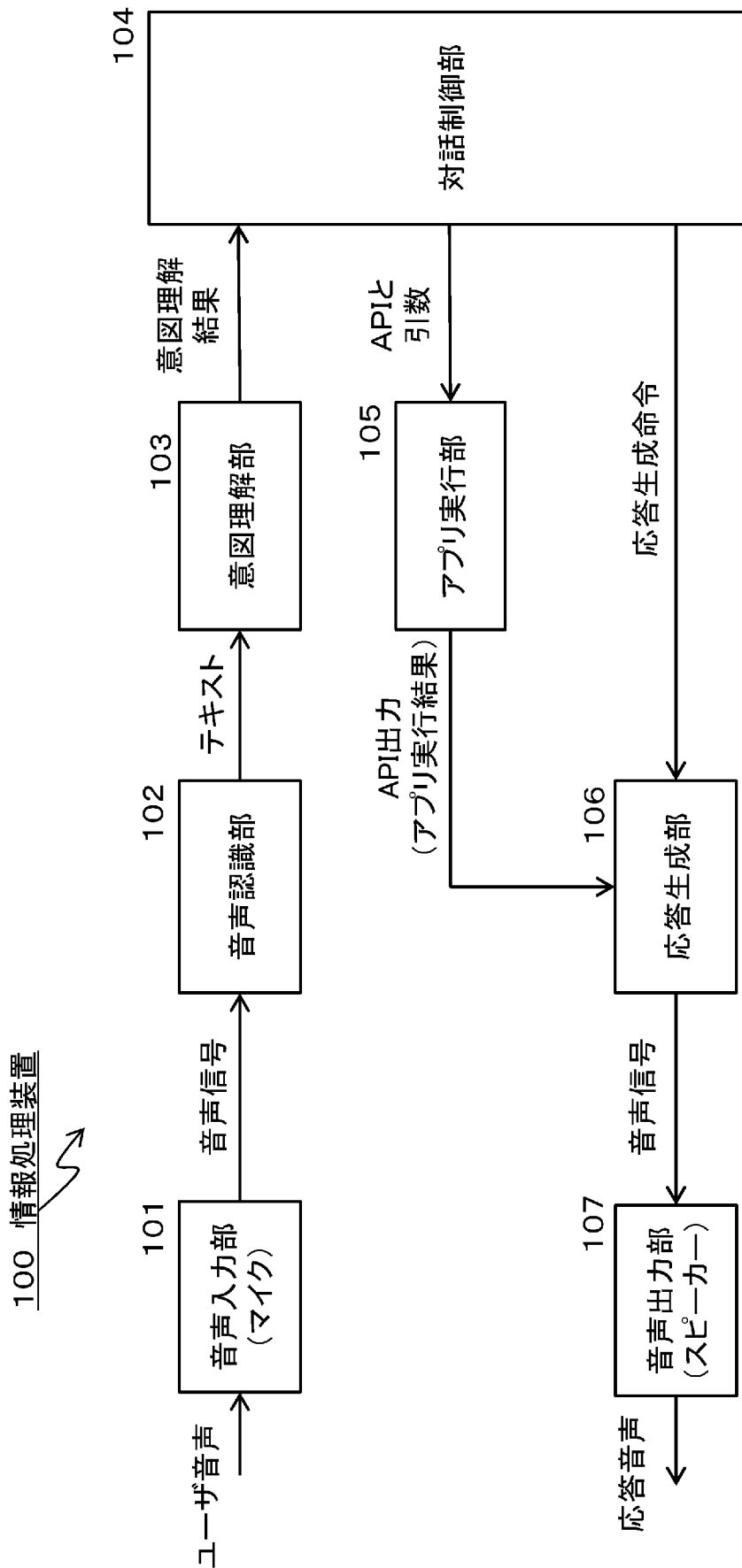
ユーザ端末と、データ処理サーバを有する情報処理システムにおい

て実行する情報処理方法であり、
前記ユーザ端末が、
ユーザ発話を入力する音声入力処理を実行し、
前記データ処理サーバが、
前記音声入力処理において入力された入力音声の音声認識処理と、
前記音声認識処理の生成した音声認識結果に基づく意図理解処理と
、
前記意図理解処理の意図理解結果に基づいて、前記ユーザ発話に対する確認応答が必要か否か判定し、確認応答が必要と判定した場合、
前記ユーザ端末に対して確認応答生成命令を出力し、
前記ユーザ端末が、
前記データ処理サーバからの確認応答生成命令入力に応じて、前記ユーザ発話に対する発話ユーザの確認を求める確認応答を生成する確認応答生成処理を実行し、
前記確認応答生成処理において、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成処理を実行する情報処理方法。

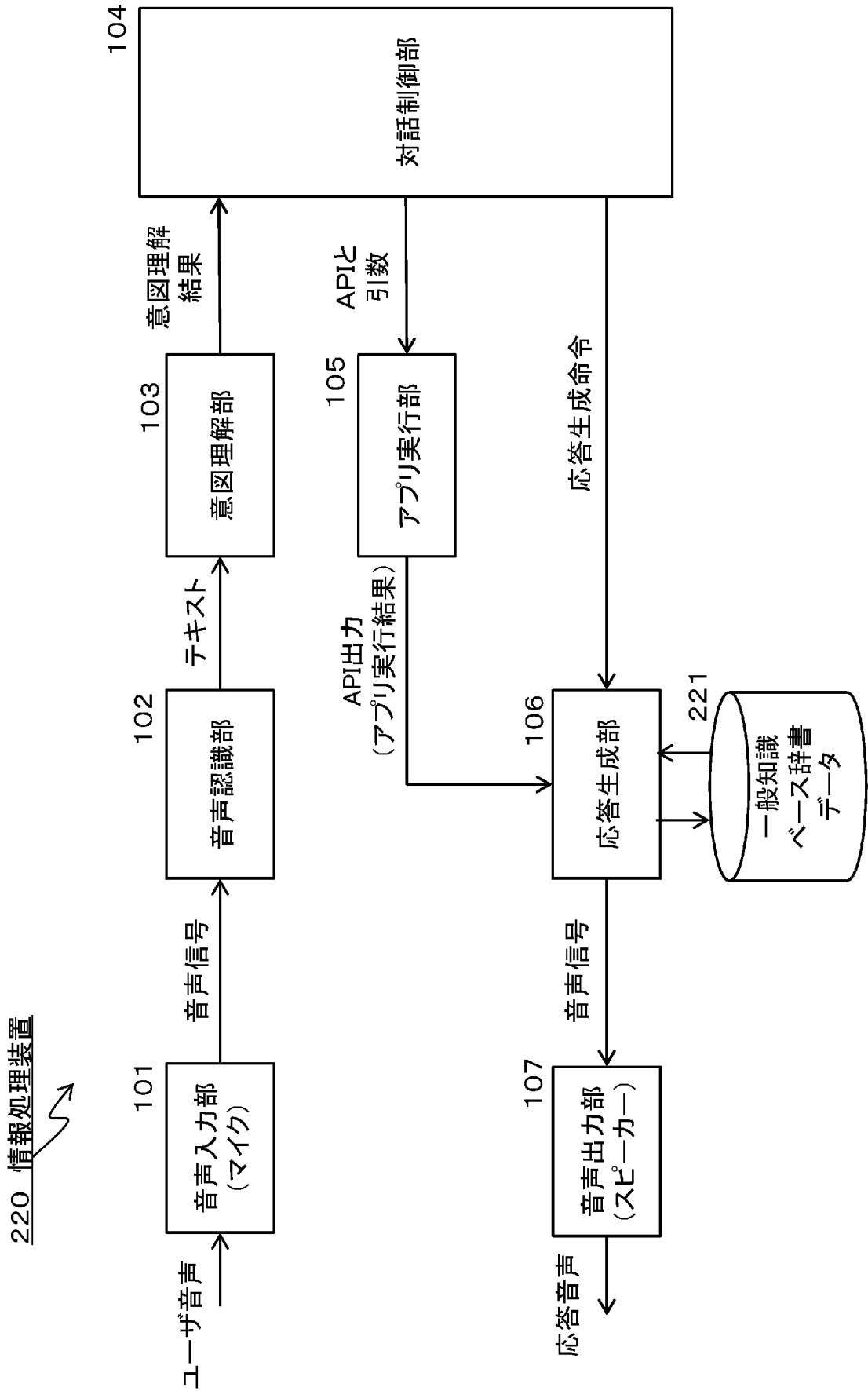
[請求項20]

情報処理装置において情報処理を実行させるプログラムであり、
音声入力部に、ユーザ発話を入力させる音声入力ステップと、
確認応答生成部に、前記ユーザ発話に対する発話ユーザの確認を求めるための確認応答を生成させる確認応答生成ステップと、
音声出力部に、前記確認応答を出力させる音声出力ステップを実行させ、
前記確認応答生成ステップにおいては、
辞書データ、またはアプリケーション実行結果の少なくともいずれかから得られる付加情報を用いた確認応答の生成を実行させるプログラム。

[図1]



[図2]



[図3]

(101) 確認応答生成例101

ユーザ発話	「8日(よ一か)に買い物の予定を入れてくれる？」
装置の確認応答	「8日(はちにち)に買い物の予定を登録しますね？」

(102) 確認応答生成例102

ユーザ発話	「8日(よ一か)に買い物の予定を入れてくれる？」
装置の確認応答	「来週水曜日の8日(よ一か)ですね？」

(103) 確認応答生成例103

ユーザ発話	「3日後に買い物の予定を入れておいて」
装置の確認応答	「11日木曜日ですね？」

[図4]

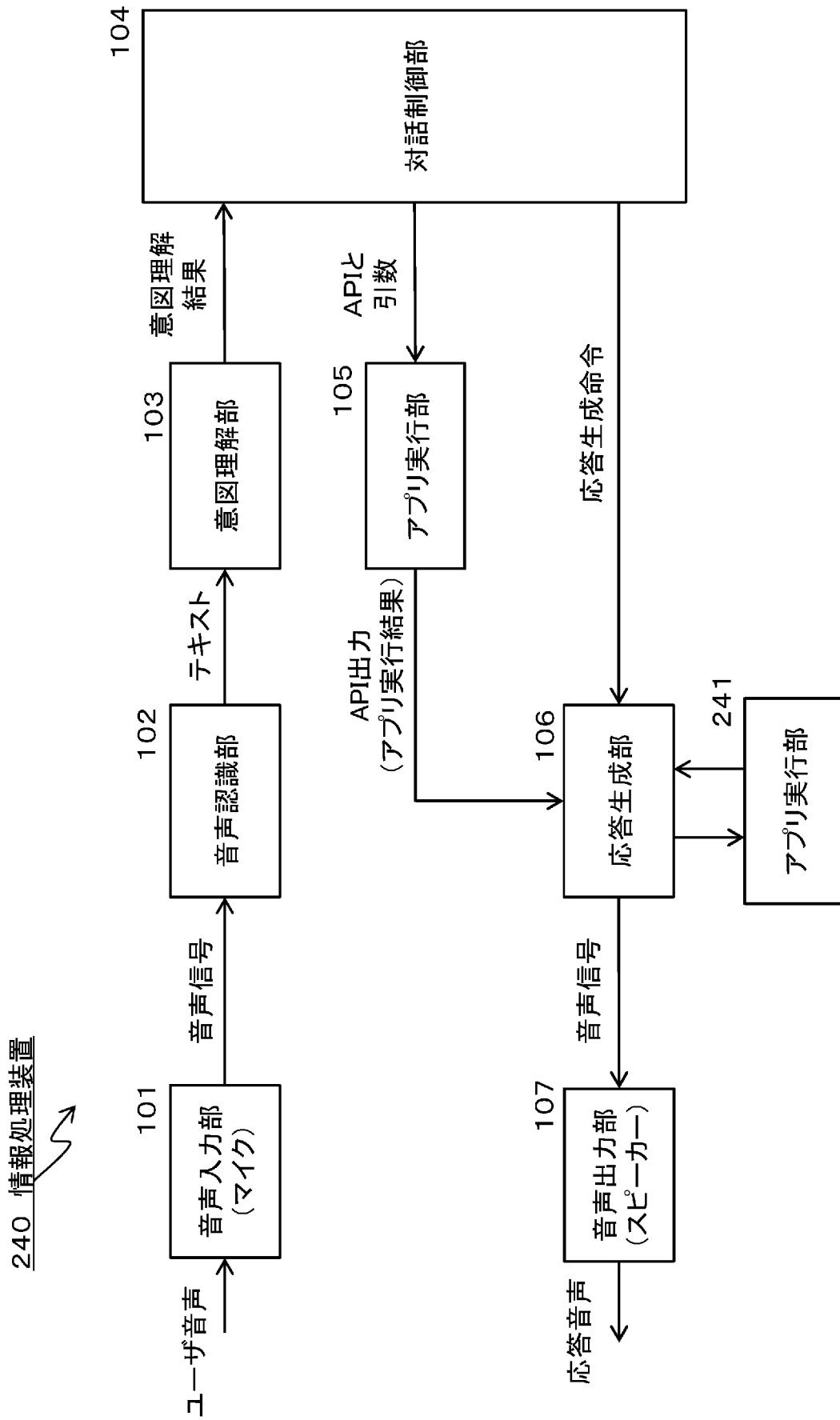
(104) 確認応答生成例104

ユーザ発話	「金閣寺への行き方を教えて？」
装置の確認応答	「足利義満が作った金閣寺ですか？」

(105) 確認応答生成例105

ユーザ発話	「1チャンネル(いっちゃんねる)を5時から録画して？」
装置の確認応答	「XYZテレビでよろしいですか？」

[図5]



[図6]

(201) 確認応答生成例201

ユーザ発話	「10日に買い物の予定を入れてくれる？」
装置の確認応答	「10日ですね、ライブに行く日ですね？」

(202) 確認応答生成例202

ユーザ発話	「銀閣寺への行き方を教えて？」
装置の確認応答	「昨年8月に行った金閣寺ですね？」

[図7]

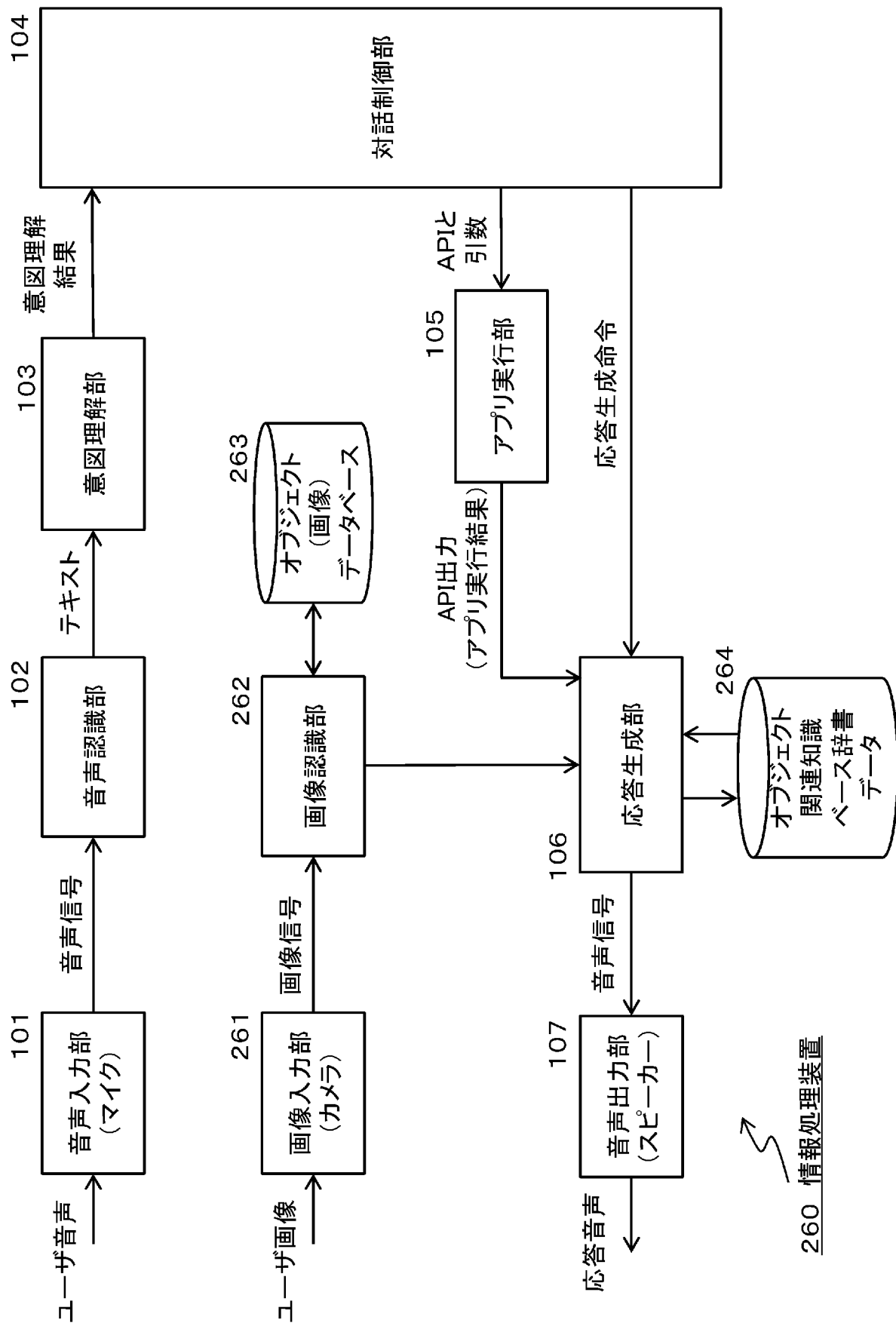
(203) 確認応答生成例 203

ユーザ発話	「佐藤さん家族とバーベキューって入れて」
装置の確認応答	「お向かいの佐藤さんでよろしいですか？」

(204) 確認応答生成例 204

ユーザ発話	「佐藤さん家族とバーベキューって入れて」
装置の確認応答	「ご近所の佐藤さんでよろしいですか？」

[図8]

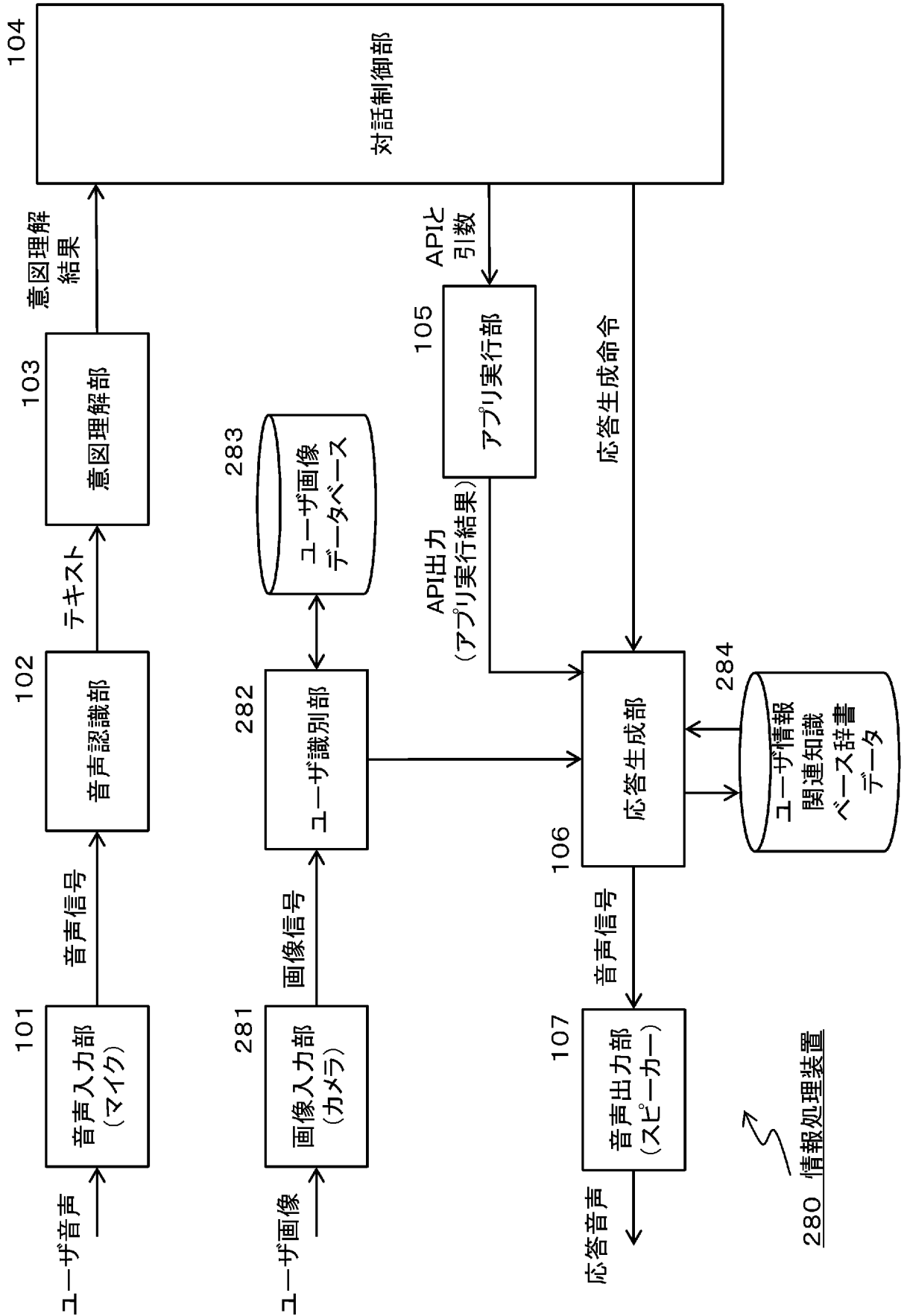


[図9]

(301) 確認応答生成例301

ユーザ発話	「スピーカーから音楽を再生して」
装置の確認応答	「あなたの近くにあるスピーカーカーでよろしいですか？」

[図10]



[図11]

(302) 確認応答生成例302

ユーザ発話	「じゃあユウキにメールしておいて」
装置の確認応答	「ユウキくんに送りますね、 ユウキさんは今ここにはいないですよね？」

[図12]

(303) 確認応答生成例303

ユーザ発話	「金閣寺への行き方を教えて」
装置の確認応答	「黄色でピカピカしている金閣寺ですか？」

(304) 確認応答生成例304

ユーザ発話	「10日に外出の予定を登録」
装置の確認応答	「10日の旗日でよろしいですか？」

(305) 確認応答生成例305

ユーザ発話	「学校に電話かけて」
装置の確認応答	「品川中学校でよろしいですね」

[図13]

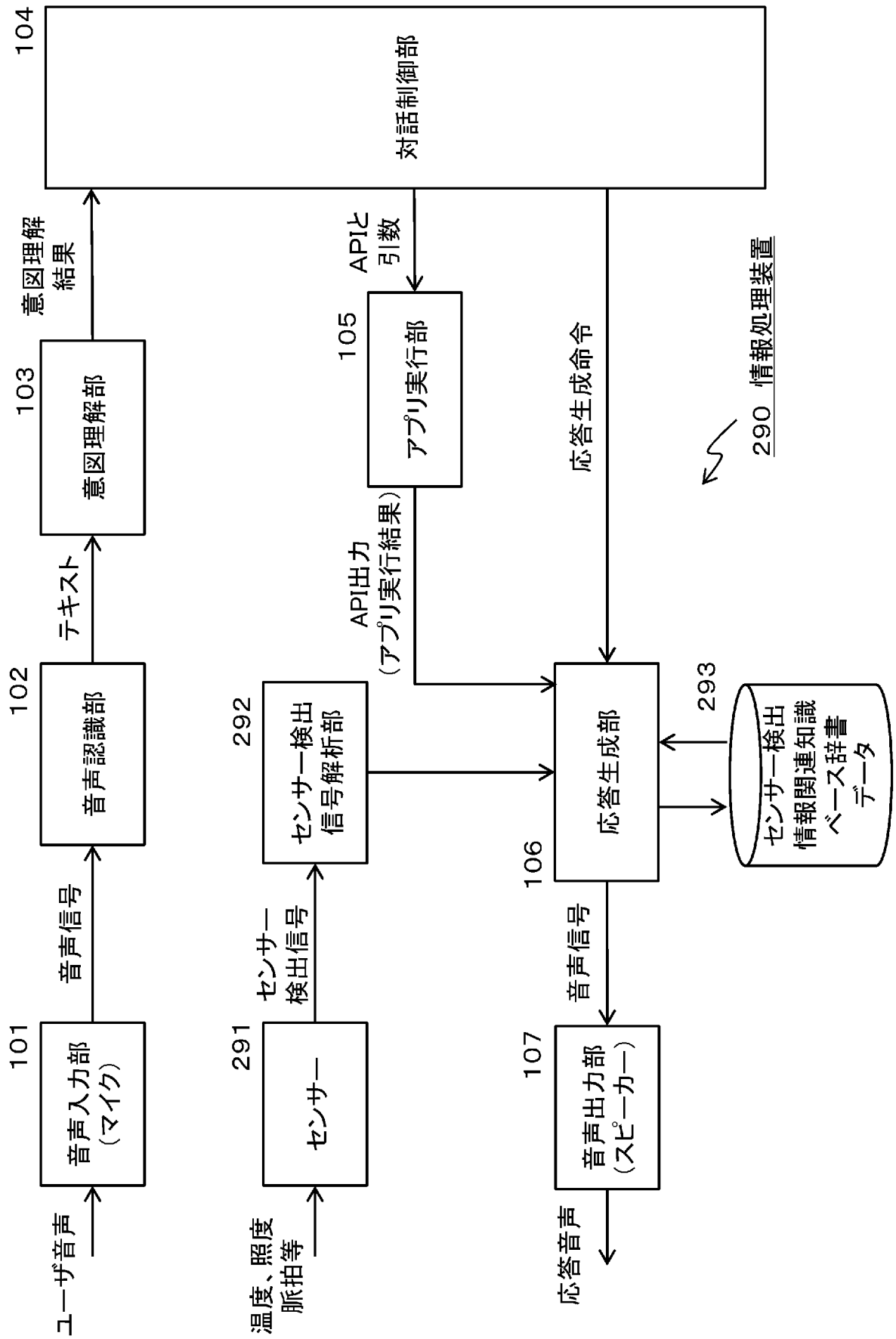
(306) 確認応答生成例306

ユーザ発話	「おばあさんの家までの道路状況を教えて」
装置の確認応答	「△△さん(ユーザの祖母)の家ですね？」

(307) 確認応答生成例307

ユーザ発話	「おばあさんの家までの道路状況を教えて」
装置の確認応答	「〇〇さん(ユーザの母親。ユーザの子供たちから見た祖母)の家ですね？」

[図14]

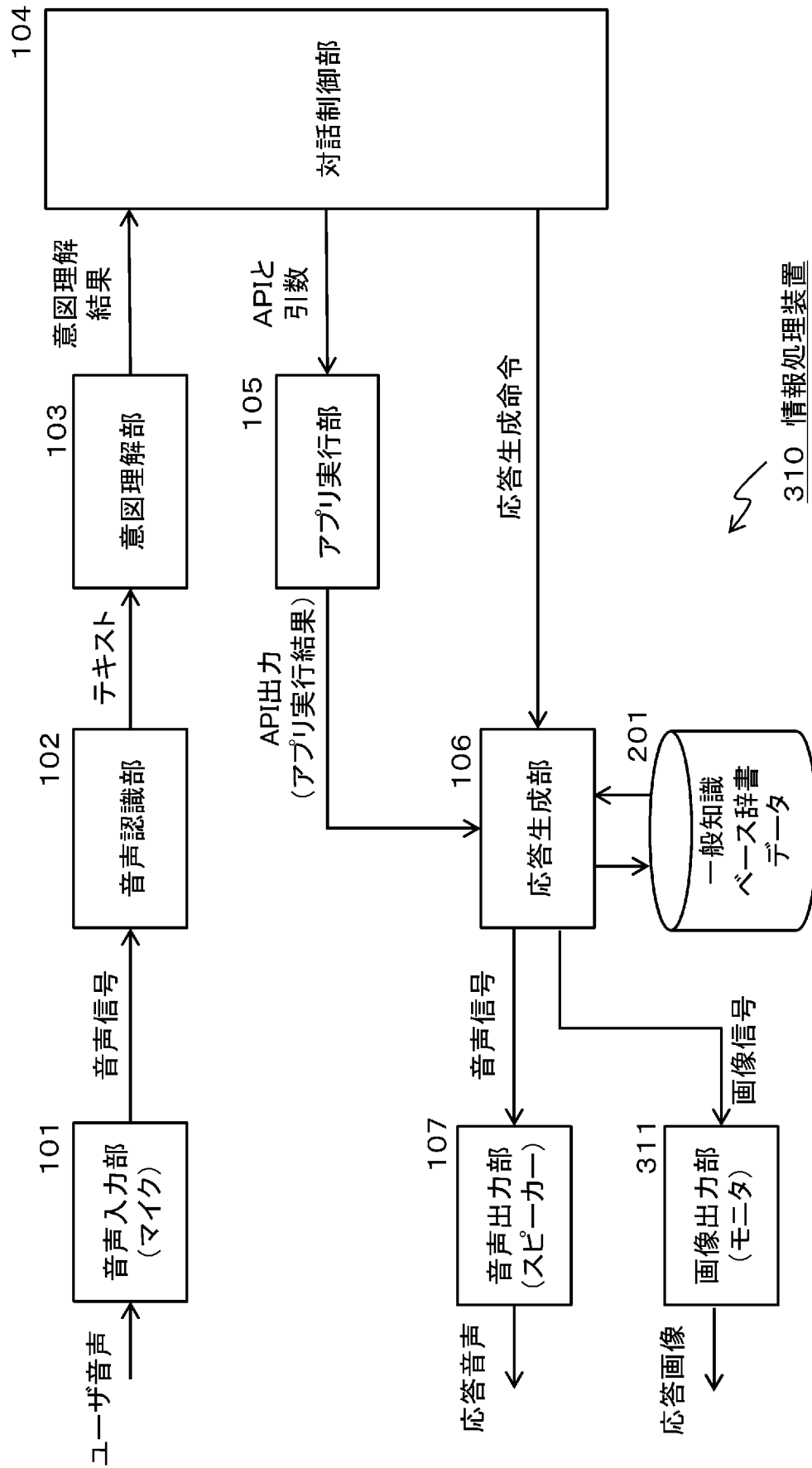


[図15]

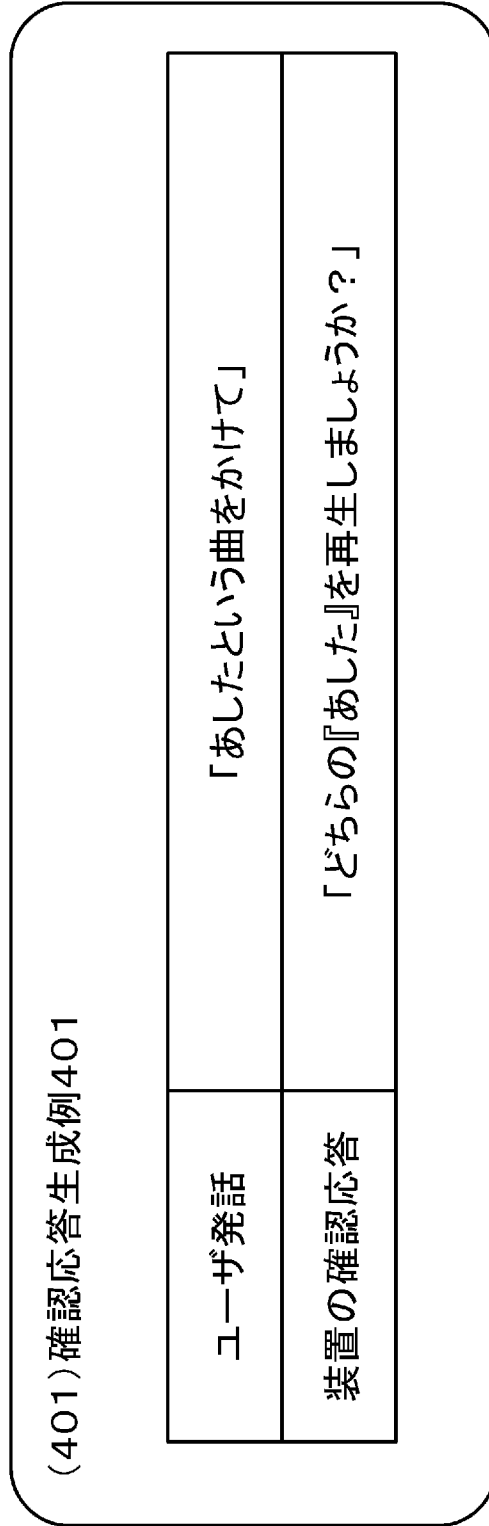
(308) 確認応答生成例308

ユーザ発話	「飲み物をください」
装置の確認応答	「ホットコーヒーでよろしいですか？ それともお茶がよろしいでしょうか？」

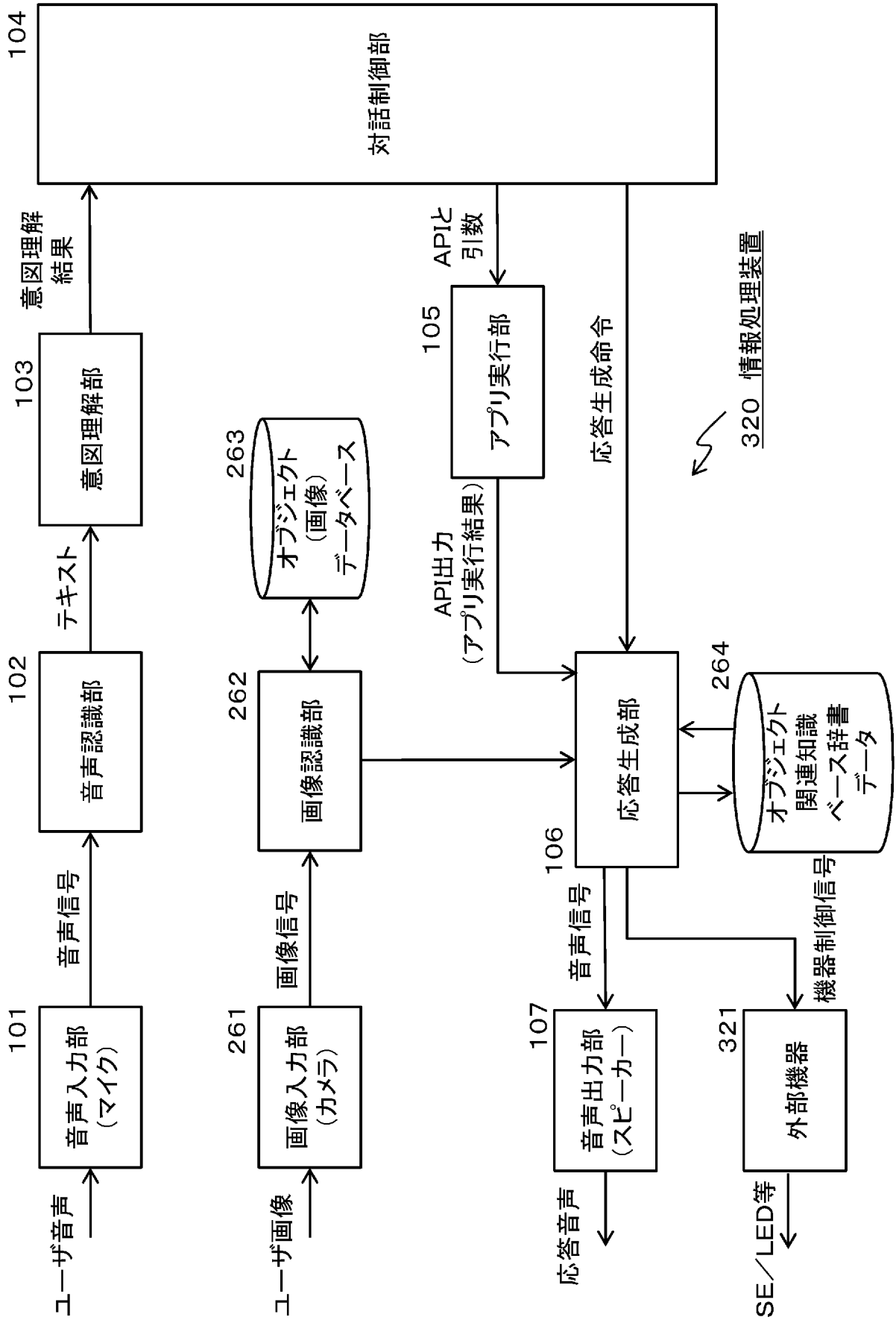
[図16]



[図17]



[図18]



[図19]

(402) 確認応答生成例402

ユーザ発話	「スピーカーから音楽を再生して」
装置の確認応答	「こちらのスピーカーでよろしいですか？」

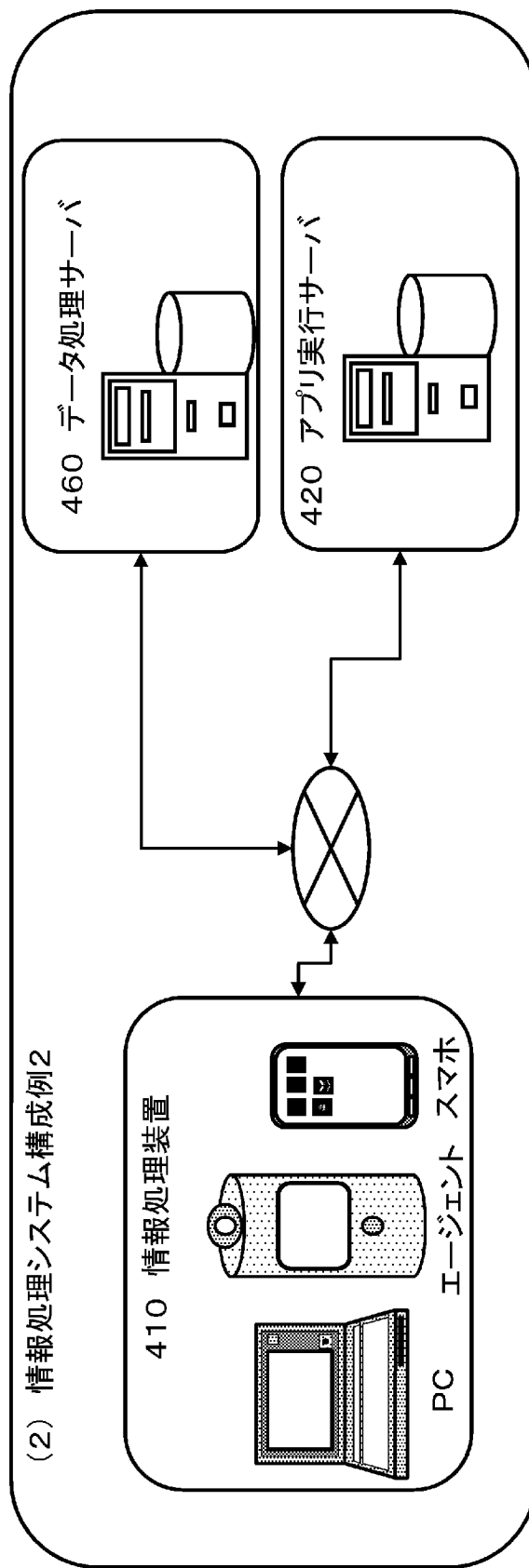
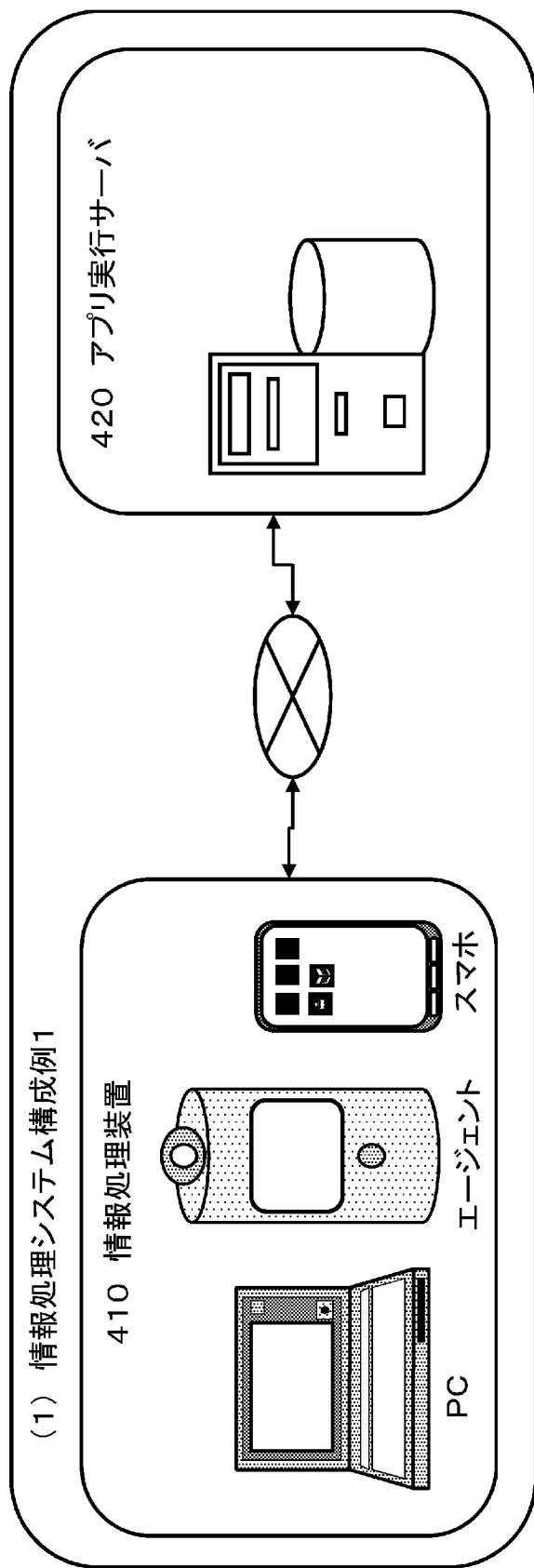
[図20]

(501) 確認応答生成例501

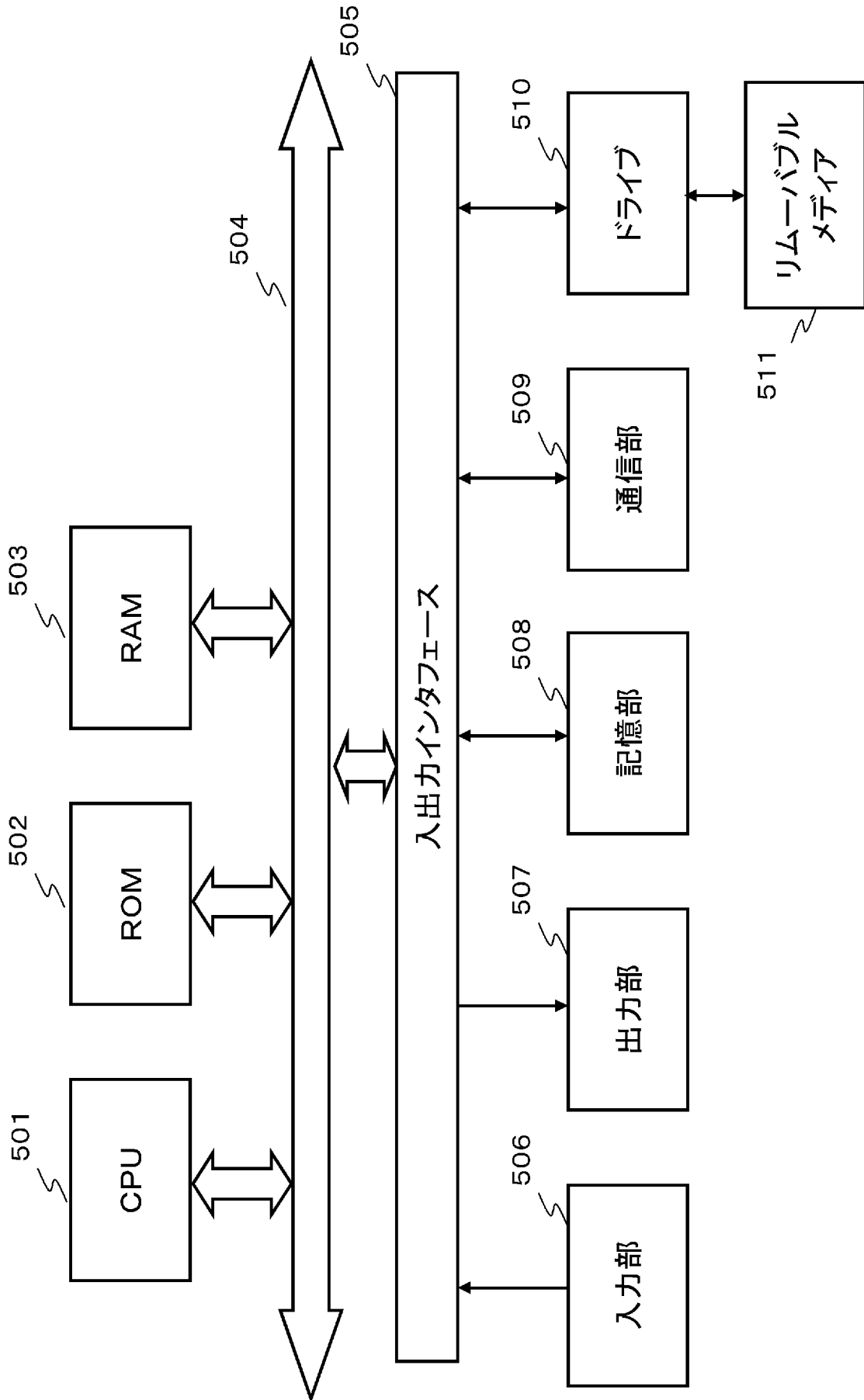
ユーザ発話	「横浜で遊べるところを教えてください」
装置の応答	「横浜みなとみらいに横浜コスモワールドという遊園地がありますよ」
ユーザ発話	「横浜みなとみらいのレストランを探して」
装置の応答	「例えば、ランドマークタワーの中いくつかレストランがありますよ。詳細をお知らせしましょうか？」
ユーザ発話	「港南区の地図を見せて」
装置の確認応答	「横浜の港南区でよろしいですか？」

↕ 対話履歴 ↕

[図21]



[図22]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/037347

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl. G10L15/22(2006.01)i, G06F3/16(2006.01)i, G10L13/00(2006.01)i, G10L13/08(2013.01)i, G10L15/10(2006.01)i, G10L15/30(2013.01)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) Int.Cl. G10L15/22, G06F3/16, G10L13/00, G10L13/08, G10L15/10, G10L15/30 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2018 Registered utility model specifications of Japan 1996-2018 Published registered utility model applications of Japan 1994-2018 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 2010-197709 A (NEC CORPORATION) 09 September 2010, paragraphs [0044]-[0072], fig. 1-6 (Family: none)	1-2, 4, 18, 20 8-12, 16-17, 19
X Y	JP 2006-18028 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 19 January 2006, paragraphs [0021], [0026]-[0031], fig. 9 (Family: none)	1-3, 15, 18, 20 8-12, 16-17, 19
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 20.12.2018		Date of mailing of the international search report 08.01.2019
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/JP2018/037347

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	WO 2014/024751 A1 (ADC TECHNOLOGY INC.) 13 February 2014, paragraphs [0010]-[0014], [0100] & JP 2018-36653 A	1, 5-6, 18, 20 8-12, 16-17, 19
X Y	JP 11-205430 A (KOBE STEEL, LTD.) 30 July 1999, paragraphs [0018]-[0021], fig. 9 (Family: none)	1, 5, 7, 18, 20 8-12, 16-17, 19
X Y	JP 11-249689 A (MITSUBISHI ELECTRIC CORPORATION) 17 September 1999, paragraphs [0040], [0041], fig. 6-8 (Family: none)	1, 13-14, 18, 20 8-12, 16-17, 19
Y	JP 2004-310034 A (MATSUSHITA ELECTRIC WORKS, LTD.) 04 November 2004, paragraph [0065] (Family: none)	8-11
Y	WO 2017/199486 A1 (SONY CORPORATION) 23 November 2017, paragraphs [0072]-[0077], fig. 7 (Family: none)	8, 12
Y	JP 2011-54088 A (NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY) 17 March 2011, paragraphs [0010], [0011], [0151]-[0156] (Family: none)	16-17, 19

<p>A. 発明の属する分野の分類（国際特許分類（I P C））</p> <p>Int.Cl. G10L15/22(2006.01)i, G06F3/16(2006.01)i, G10L13/00(2006.01)i, G10L13/08(2013.01)i, G10L15/10(2006.01)i, G10L15/30(2013.01)i</p>												
<p>B. 調査を行った分野</p> <p>調査を行った最小限資料（国際特許分類（I P C））</p> <p>Int.Cl. G10L15/22, G06F3/16, G10L13/00, G10L13/08, G10L15/10, G10L15/30</p>												
<p>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</p> <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2018年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2018年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2018年</td> </tr> </table>			日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2018年	日本国実用新案登録公報	1996-2018年	日本国登録実用新案公報	1994-2018年		
日本国実用新案公報	1922-1996年											
日本国公開実用新案公報	1971-2018年											
日本国実用新案登録公報	1996-2018年											
日本国登録実用新案公報	1994-2018年											
<p>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）</p>												
<p>C. 関連すると認められる文献</p> <table border="1"> <thead> <tr> <th>引用文献の カテゴリー*</th> <th>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</th> <th>関連する 請求項の番号</th> </tr> </thead> <tbody> <tr> <td>X Y</td> <td>JP 2010-197709 A（日本電気株式会社）2010.09.09, 段落 [0044] - [0072], 図1-6（ファミリーなし）</td> <td>1-2, 4, 18, 20 8-12, 16-17, 19</td> </tr> <tr> <td>X Y</td> <td>JP 2006-18028 A（日本電信電話株式会社）2006.01.19, 段落 [0021], [0026] - [0031], 図9（ファミリーなし）</td> <td>1-3, 15, 18, 20 8-12, 16-17, 19</td> </tr> </tbody> </table>			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	JP 2010-197709 A（日本電気株式会社）2010.09.09, 段落 [0044] - [0072], 図1-6（ファミリーなし）	1-2, 4, 18, 20 8-12, 16-17, 19	X Y	JP 2006-18028 A（日本電信電話株式会社）2006.01.19, 段落 [0021], [0026] - [0031], 図9（ファミリーなし）	1-3, 15, 18, 20 8-12, 16-17, 19	
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	JP 2010-197709 A（日本電気株式会社）2010.09.09, 段落 [0044] - [0072], 図1-6（ファミリーなし）	1-2, 4, 18, 20 8-12, 16-17, 19										
X Y	JP 2006-18028 A（日本電信電話株式会社）2006.01.19, 段落 [0021], [0026] - [0031], 図9（ファミリーなし）	1-3, 15, 18, 20 8-12, 16-17, 19										
<p><input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。</p>												
<p>* 引用文献のカテゴリー</p> <table border="0"> <tr> <td>「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</td> <td>「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</td> </tr> <tr> <td>「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</td> <td>「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</td> </tr> <tr> <td>「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）</td> <td>「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</td> </tr> <tr> <td>「O」 口頭による開示、使用、展示等に言及する文献</td> <td>「&」 同一パテントファミリー文献</td> </tr> <tr> <td>「P」 国際出願日前で、かつ優先権の主張の基礎となる出願</td> <td></td> </tr> </table>			「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	「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	「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	「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	「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）	「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	「O」 口頭による開示、使用、展示等に言及する文献	「&」 同一パテントファミリー文献	「P」 国際出願日前で、かつ優先権の主張の基礎となる出願	
「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	「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											
「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	「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											
「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）	「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											
「O」 口頭による開示、使用、展示等に言及する文献	「&」 同一パテントファミリー文献											
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願												
<p>国際調査を完了した日</p> <p>20.12.2018</p>	<p>国際調査報告の発送日</p> <p>08.01.2019</p>											
<p>国際調査機関の名称及びあて先</p> <p>日本国特許庁（I S A / J P）</p> <p>郵便番号100-8915</p> <p>東京都千代田区霞が関三丁目4番3号</p>	<p>特許庁審査官（権限のある職員）</p> <p>上田 雄</p> <p>電話番号 03-3581-1101 内線 3591</p>	<p>5Z 5095</p>										

C (続き) . 関連すると認められる文献		
引用文献の カテゴリ*	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	WO 2014/024751 A1 (エイディシーテクノロジー株式会社) 2014.02.13, 段落 [0010] - [0014], [0100] & JP 2018-36653 A	1, 5-6, 18, 20 8-12, 16-17, 1 9
X Y	JP 11-205430 A (株式会社神戸製鋼所) 1999.07.30, 段落 [001 8] - [0021], 図9 (ファミリーなし)	1, 5, 7, 18, 20 8-12, 16-17, 1 9
X Y	JP 11-249689 A (三菱電機株式会社) 1999.09.17, 段落 [0040] - [0041], 図6-8 (ファミリーなし)	1, 13-14, 18, 2 0 8-12, 16-17, 1 9
Y	JP 2004-310034 A (松下電工株式会社) 2004.11.04, 段落 [006 5] (ファミリーなし)	8-11
Y	WO 2017/199486 A1 (ソニー株式会社) 2017.11.23, 段落 [0072] - [0077], 図7 (ファミリーなし)	8, 12
Y	JP 2011-54088 A (独立行政法人情報通信研究機構) 2011.03.17, 段 落 [0010] - [0011], [0151] - [0156] (ファ ミリーなし)	16-17, 19