

プレゼンタ動作を再現・再構成する ロボットプレゼンテーションシステム

後藤 充裕^{*1}, 石野 達也^{*2}, 稲澤 佳祐^{*2}, 松村 成宗^{*1}, 布引 純史^{*1}, 柏原 昭博^{*2}

^{*1} NTT サービスエボリューション研究所

^{*2} 電気通信大学 大学院情報理工学研究科

A Presentation System with Robot for Recomposing Presenter Behavior

Mitsuhiro Goto, Tatsuya Ishino, Keisuke Inazawa,
Narimune Matsumura, Tadashi Nunobiki, Akihiro Kashiara

^{*1} NTT Service Evolution Laboratories

^{*2} The University of Electro-Communications

様々な場面で実施されるプレゼンテーションでは、聴衆の注意を誘導したり、聴衆の内容理解を促進するプレゼンタの非言語動作が非常に重要である。しかし、これら動作を適切に実施することは経験の浅いプレゼンタだけではなく、熟練のプレゼンタにとっても難しい。そこで、プレゼンタの代わりにロボットがプレゼンテーションを代行し、分かりやすいプレゼンテーションを聴衆に向けて実施するロボットプレゼンテーションシステムを開発した。本システムは、プレゼンタによるプレゼンテーションを収録し、不十分・不適切な非言語動作を検出・修正することで、聴衆の注意誘導や内容理解を促進させる適切な非言語動作を再現する。こうした再現を可能とするために、本研究ではプレゼンタの伝達意図と非言語動作の対応関係を表現したプレゼンテーション動作モデルをデザインした。本モデルを基盤に、プレゼンタの伝達意図からその意図を聴衆へ伝えるための適切な非言語動作を絞り込み、それらを伴ってプレゼンテーションが実施されているかを診断し、必要に応じて再構成を行う。

キーワード: ロボットプレゼンテーション, プレゼンテーションスキル, ロボット制御, 非言語動作, モーションキャプチャ

1. はじめに

近年、ビジネスにおける企画説明や商品紹介、大学における講義や e-Learning など様々な場面においてプレゼンテーションを実施する機会が増加している。一般に、プレゼンテーションとは、プレゼンタが事前に作成したスライド（以下、スライドコンテンツ）を投影し、口頭での説明（以下、オーラルコンテンツ）を通して、聴衆の理解・納得を得るための情報伝達手段の1つである。このようなプレゼンテーションを分かりやすく実施するためには、上述した2種類のコン

テンツを提示するだけではなく、聴衆の注意をプレゼンタが意図しているコンテンツに集め、意図しているコンテンツ内の情報の重要度や論理展開などの情報の関係性を正しく伝えることが重要となる。これらを効果的に実施するには、プレゼンタがアイコンタクトやジェスチャといった非言語動作を活用して、意図しているコンテンツへ聴衆の注意を誘導したり、重要性や関係性といった説明しようとしている情報の特徴を伝達する必要がある。

しかしながら、プレゼンテーション中にこれらの非言語動作を適切に用いることは学生や新入社員といっ

た経験の浅いプレゼンタだけではなく、大学講師や企業の管理職といったある程度経験を持ったプレゼンタにとっても難しい。加えて、プレゼンタが内向的な性格の場合、手元の PC や投影しているスクリーンを凝視し、聴衆への非言語動作がおろそかになってしまう。この結果、内容理解を促せず、聴衆への情報伝達が不十分なまま、プレゼンテーションを終えてしまうことがある。

そこで、本研究では、プレゼンタの代わりにロボットが内容理解を促進する非言語動作を伴わせてプレゼンテーションを行うロボットプレゼンテーションシステムを提案する。本システムは、プレゼンタによるプレゼンテーションを収録し、聴衆の注意誘導や内容理解を促進する非言語動作が適切に用いられているかを診断する。そして、この診断結果にもとづき、不十分・不適切な動作が存在する場合にはプレゼンテーションを再構成して、適切な非言語動作を伴ってロボットがプレゼンテーションを代行することを実現する。このプレゼンテーションの再構成には、プレゼンタが「オーラルコンテンツに注意を向けて欲しい」や「スライド中のあるキーワードが重要であることを認識して欲しい」といったプレゼンテーション中のある時点における聴衆へ伝えたい意図（以下、伝達意図）を加味して、注意誘導や内容理解を促進するプレゼンタの非言語動作を明らかにする必要がある。このような観点から、本研究では、プレゼンタの伝達意図とロボットのプレゼンテーション動作を対応づけるプレゼンテーション動作モデルをデザインし、本モデルに基づいてシステム開発を行っている。

本稿では、プレゼンテーションにおける非言語動作活用の必要性と問題点について整理し、その解決アプローチとして、プレゼンテーション動作モデルに基づくプレゼンテーション再構成手法について提案する。さらに、開発を進めているロボットプレゼンテーションシステムについて述べる。

2. 関連研究

本章では、プレゼンテーションにおける非言語動作活用やロボットによるプレゼンテーション代行に関する従来研究について述べる。

2.1 プレゼンテーションにおける非言語動作活用

Ekman⁽¹⁾や McNeill⁽²⁾は、非言語動作が人同士のコミュニケーションを円滑にすることを明らかにし、実際にコミュニケーションで用いられるジェスチャ動作の形態や機能をもとにした分類観点を提案している。この分類観点の詳細については、3章に詳述する。本研究では、この分類観点を参考にプレゼンテーション実施時に円滑な情報伝達を促すジェスチャ動作を明らかにしていく。

Kopf ら⁽³⁾は、プレゼンテーション実施時のプレゼンタの手の位置や視線方向を Microsoft 社の Kinect⁽⁴⁾で取得し、リアルタイムにフィードバックするシステムを提案している。これらの情報は、プレゼンタ自身が自らのプレゼンテーションをレビューする指標となるが、聴衆から見た振る舞いの不自然さを直すことが主であり、聴衆にとって分かりやすいプレゼンテーションを実現するための非言語動作を明らかにするという本研究とは目的が大きく異なっている。

2.2 ロボットによるプレゼンテーション代行

Kamide ら⁽⁶⁾の研究では、ロボットによるプレゼンテーションにおいて、ロボットの視線方向の制御が聴衆の注意をスライドコンテンツへ誘導することができることを示している。しかしながら、この研究においては、視線方向をスライドもしくは聴衆に向けるよう制御した場合の聴衆への注意誘導効果を検証しており視線方向以外の非言語動作の制御は行っていない。Chien-Ming ら⁽⁷⁾は、動的ベイジアンネットワークを用いて、16人のプレゼンタの動作を学習し、オーラルコンテンツに基づき、視線方向とジェスチャを自動付与する手法を提案している。この手法はオーラルコンテンツのみをロボットに与えるだけで、視線方向とジェスチャを自動的に決定できるという特徴があるが、付与できるジェスチャや視線方向が学習するデータに依存してしまうという問題がある。実験の結果、不要な腕のジェスチャが付与されて、プレゼンテーションを聞くのに影響があったことも示唆されている。

以上述べてきた従来研究を踏まえて、筆者らは、ロボットによるプレゼンテーション代行の枠組みとして、プレゼンテーション動作モデルをデザインしてきた⁽⁶⁾。本モデルにより、視線制御以外の要素も含めたロボッ

トの動作制御をプレゼンタの伝達意図に基づいて付与し、プレゼンテーションの理解を妨げる不十分・不適切な非言語動作の診断・再構成を可能とする。

3. プレゼンテーションにおける非言語動作

本章では、プレゼンテーションにおける非言語動作の必要性や問題点を詳述した後に、分かりやすいプレゼンテーションの実施に必要な非言語動作について整理する。

3.1 非言語動作の必要性

通常、プレゼンテーションの実施には、プレゼンタが聴衆に伝えたい情報を、キーワードや図表などで端的に表現したスライドコンテンツと、そのスライドコンテンツを補足説明するオーラルコンテンツを用いて行う。プレゼンタは「説明している情報の重要性」や「ある情報とある情報との関連性」などの伝達意図を聴衆に認識してもらうために、アニメーションやテキスト効果を用いてスライド中の一部をハイライトしたり、オーラルを話す際の発話速度や声量を変化させるといった表現の強調を行うことが多い。しかしながら、プレゼンタがこれらのコンテンツ内で伝達意図を強調したプレゼンテーションを実施しても、聴衆がその強調に気付けないことがある。その結果、聴衆に伝達意図が正しく伝わらず、情報伝達が不十分なままプレゼンテーションを終えてしまうことになる。また、聴衆がプレゼンタによる表現の強調に気付いた場合にも、「どの程度重要であるのか」や「どういった関係性があるのか」といったように、伝達意図の詳細を理解する情報がスライドコンテンツやオーラルコンテンツに陽に表現されていない場合には、情報の理解が十分に進まないことがある。このようにスライドコンテンツやオーラルコンテンツのみで、プレゼンタの伝達意図自体に気付いてもらうことや、その伝達意図の詳細を理解してもらうのは、聴衆にとって難しい場合がある。

そこで、プレゼンタは、アイコンタクトやジェスチャなどの非言語動作を用いて、プレゼンタの意図したコンテンツに注意誘導を行うことで、「オーラルに注意を向けて情報を聞いて欲しい」という伝達意図を伝えたり、聴衆に提示しているある情報とある情報のそれぞれの大きさをジェスチャで表現することで、「2つの

情報が大小関係を持っていることを認識して欲しい」といった伝達意図の詳細を伝えながら、プレゼンテーションを実施することが求められている。

3.2 問題点

上述した通り、非言語動作により、プレゼンタは伝達意図を聴衆へ十分に伝え、分かりやすいプレゼンテーションを実施することが可能となる。しかしながら、多くのプレゼンタにとって、プレゼンテーション中に非言語動作を適切に行うことは難しい。例えば、学生や新入社員といったプレゼンテーションの経験が浅いプレゼンタは、オーラルコンテンツを正確に話すことに終始してしまい、非言語動作を伴いながらプレゼンテーションを実施する余裕がないことが多い。大学講師や企業の管理職といったある一定のプレゼンテーション実施機会があり、経験を持ったプレゼンタにおいては、非言語動作を伴ったプレゼンテーションを行うことができるが、アイコンタクト時の顔向きやジェスチャ実施時の腕の角度が不十分となり、聴衆に伝達意図を伝えることのできる非言語動作までに至らないことが多い。また、プレゼンテーション経験の大小だけではなく、プレゼンタの性格に起因して非言語動作が十分に用いられないケースもある。プレゼンタが内向的な性格の場合、手元の PC や投影しているスクリーンを凝視し、聴衆へのアイコンタクトなどの非言語動作がおろそかとなってしまいうこともある。この結果、聴衆への情報伝達が不十分なままプレゼンテーションを終えてしまうケースが多々発生する。

3.3 伝達意図を伝える非言語動作

前節で述べた問題の解決に向けて、本節ではプレゼンタの伝達意図を伝えるための非言語動作について整理する。まず、聴衆の状態に合わせて、プレゼンタの考える伝達意図がどのように変化するかをまとめる。

図 1 に示す通り、聴衆の状態は大きく分けて、以下の 4 状態に分類できる。

状態 1: プレゼンテーションを聞いていない。

状態 2: プレゼンテーションに耳を傾けている。

状態 3: プレゼンテーション中の伝達箇所に気付く。

状態 4: プレゼンテーション中の伝達箇所の詳細を理解する。

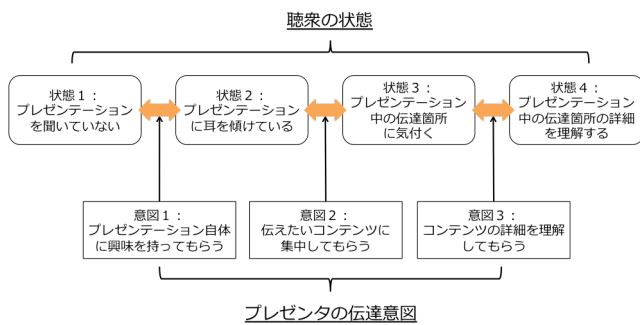


図1 聴衆の状態に応じたプレゼンタの伝達意図

これらの聴衆の状態に対して、プレゼンタの伝達意図は以下の通りに分類できる。

意図1（状態1→状態2の変化を促す）：

聴衆にプレゼンテーション自体への興味を持ってもらう。

意図2（状態2→状態3の変化を促す）：

プレゼンタの伝えたいコンテンツに集中してもらう。

意図3（状態3→状態4の変化を促す）：

聴衆にコンテンツを詳細に理解してもらう。

本研究では、プレゼンタがこうした伝達意図を検討する機会として、(1)プレゼンテーションの準備（リハーサル）中、(2)プレゼンテーションの実施中、があると考えている。まず、(1)ではプレゼンタが発表内容や発表場所、対象となる聴衆の属性などを基に、聴衆の状態変化を仮定して、伝達意図を設定する。例えば、学会発表のような聴衆のモチベーションが高い場合には、聴衆の状態は状態3以降が多くなると仮定することができるため、より情報の詳細を理解してもらうような伝達意図を中心に設定することになると考えられる。一方、街頭での商品紹介のようなプレゼンテーションでは、あまり興味を持っていない聴衆が中心になると考えられるため、聴衆の状態は状態1や状態2になると仮定して、そもそもプレゼンテーションを聞いてもらうための伝達意図を設定すると考えられる。(2)では、プレゼンテーション実施時における聴衆のリアクションから推察される状態に応じて、インタラクティブに伝達意図を設定・変更する。本研究では、現在のところ(1)の文脈で、伝達意図を設定することを想定している。

以降では、どのような非言語動作が各意図の伝達に

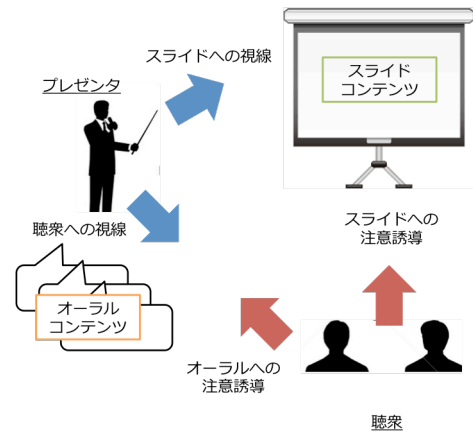


図2 プレゼンタの視線・顔向きによる聴衆の注意誘導

活用できるかを整理する。

3.3.1 意図1：聴衆の興味を促す非言語動作

聴衆がプレゼンテーションへ興味を持ち、耳を傾けてもらうためには、プレゼンタが聴衆に対して語りかけているような印象や何か面白そうだとインパクトを与える必要がある。これらを実現するには、聴衆の頷きに合わせて発話タイミングやジェスチャの開始タイミングを合わせたり、ビデオや音楽といったマルチメディアの利用、大げさなジェスチャによるオーバーアクションを用いて、聴衆にプレゼンテーションへの興味を持ってもらうことが考えられる。

3.3.2 意図2：聴衆の注意を誘導する非言語動作

聴衆がプレゼンテーションを聞きながら、プレゼンタが伝達したい箇所に注意を向けてもらうためには、プレゼンタがある時点で、スライドコンテンツとオーラルコンテンツのうち、どちらのコンテンツに注力して説明をしているかを聴衆に伝える必要がある。これを実現するのに、プレゼンタの顔や視線の向きを活用する。図2に示す通り、プレゼンテーション実施時にプレゼンタがスライド方向へ顔向きや視線を移すことにより、聴衆もスライドに視線を移し、スライドコンテンツに対して注意が向けられる。逆に、プレゼンタが聴衆方向へ顔向きや視線を移すことにより、聴衆もスライドから視線を外し、プレゼンタに視線を向けることになる。これにより、プレゼンタの発話しているオーラルコンテンツに対して注意が向けられる。

3.3.3 意図3：聴衆の理解を促進する非言語動作

聴衆がプレゼンテーションを聞きながら、プレゼンタが特に伝達したい箇所の詳細を理解させるためには、

プレゼンタの説明している情報をさらに補足する情報を聴衆に伝える必要がある。これを実現するのに、情報をどのように補足するかという観点を踏まえて、身振り・手振りのジェスチャやパラ言語（発話速度や声量など）を活用する。ジェスチャには、以下の通り McNeill²⁾が分類した 3 タイプのジェスチャが存在する。本研究では、これらを補足したい情報に合わせて使い分けることで聴衆の理解を促す。

- (a)Deictic: 重要箇所など注目して欲しい対象を指定する、指差しなどのポインティングジェスチャ
- (b)Iconic: 説明している情報の大きさや長さを手の動きで表現するジェスチャ
- (c)Metaphoric: 説明している複数の情報の間にある順序関係を立てた指の数で表現したり、大小関係を手の動きにより表現するジェスチャ

これらのジェスチャは、プレゼンタの説明している情報の詳細を聴衆に伝え、内容理解の促進に寄与すると考えられる。

4. ロボットによるプレゼンテーション

前章では、プレゼンタの伝達意図とその意図を正しく聴衆に伝える非言語動作との対応について述べた。これら動作を無闇に行うと、本来聴衆に伝えなかった伝達意図が正確に伝わらず、分かりにくいプレゼンテーションになってしまう恐れがあり、非言語動作の活用範囲を適切に設定する必要がある。そこで、本章では、まず適切な非言語動作を決定するプレゼンテーション動作モデルについて述べ、本モデルを用いたロボットによるプレゼンテーション再構成手法について詳述する。

4.1 プレゼンテーション動作モデル

本モデルは、伝達意図と動作カテゴリ、プレゼンテーションの基本動作の 3 層で構成され、これらの対応関係でプレゼンテーション時の非言語動作の組み合わせを決定するモデルである (図 3)。具体的には、プレゼンタの思い描く「伝達意図」と、ロボットが利用可能な「プレゼンテーションの基本動作」を入力として、ロボットが分かりやすいプレゼンテーションを実施す

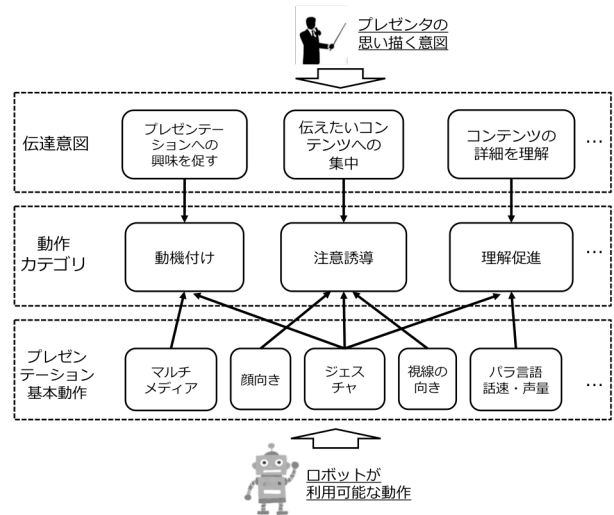


図 3 ロボットのプレゼンテーション動作モデル

る上で、必要となる非言語動作の組み合わせを「動作カテゴリ」として出力する。例えば、図 3 に示す通り、プレゼンタが「伝えたいコンテンツへ集中して欲しい」という伝達意図を持っていた場合に、顔向きや視線、Deictic なジェスチャを組み合わせることにより、ロボットによる注意誘導を実現する。

プレゼンタが「コンテンツの詳細を理解して欲しい」という意図を持ち、それが重要点の強調を意図したのなら、パラ言語の声量の変化と Deictic なジェスチャの組み合わせで説明している情報の重要性を聴衆に伝える。また、それが関係性の強調を意図したのなら、パラ言語と Iconic もしくは Metaphoric なジェスチャを組み合わせることで、ある時点で説明している情報の順序関係や大小関係を分かりやすく聴衆に伝え、ロボットによる理解促進を実現する。

4.2 動作モデルにもとづくプレゼン再構成

本研究ではロボットによるプレゼンテーションとして、プレゼンタの動作を忠実に再現するのではなく、プレゼンテーション動作モデルと照合しながら、不十分・不適切な動作を診断・再構成し、必要に応じてプレゼンタ自身による非言語動作の再修正を行いながら、プレゼンテーションを代行する。再構成は以下で 3 つのステップで実行される。

ステップ 1：伝達意図の推定

ステップ 2：モデル照合による診断・再構成

ステップ 3：再構成結果のフィードバック

各ステップの詳細は以下の通りとなる。

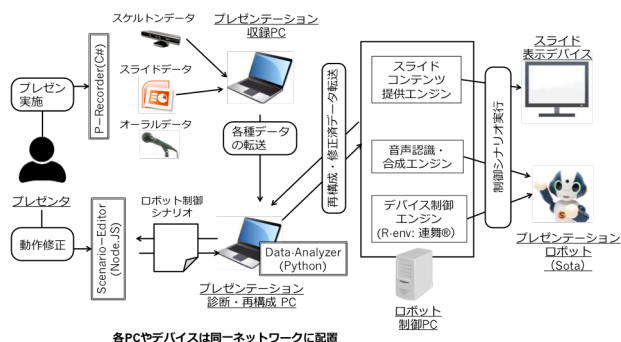


図4 プレゼンテーションシステム概略図

4.2.1 ステップ1：伝達意図の推定

プレゼンタによるプレゼンテーションや聴衆の状態を入力として、どのような伝達意図を持っているかを推定するステップである。現状では、プレゼンタが聴衆を前にせず実施したプレゼンテーションのみを入力とし、その動作やスライドコンテンツ、オーラルコンテンツ内のキーワードを分析して伝達意図の推定を行っている。そのため、現在推定できる伝達意図は「プレゼンタの伝えたいコンテンツへの集中」と「コンテンツの詳細を理解（重要点の強調）」に限られている。一方で、本来のプレゼンテーションは聴衆を前にして実施するものであり、3.3節でも述べたとおり、聴衆の状態変化に合わせて、プレゼンタが設定・変更する伝達意図を推定する手法も必要となる。

4.2.2 ステップ2：モデル照合による診断・再構成

推定した伝達意図をプレゼンテーション動作モデルと照合しながら、不十分・不適切な動作を診断・修正するステップである。「プレゼンタの伝えたいコンテンツへの集中」の意図では、プレゼンタの意図した方向にしっかりと顔や視線を向けているかを診断し、向きが不十分な場合には、それらを再構成する。「コンテンツの詳細を理解（重要点の強調）」の意図では、プレゼンタが意図している情報をしっかりとポイントしているかを診断し、ポイント時の腕の角度や顔向きが不十分な場合には、それらを再構成する。

4.2.3 ステップ3：再構成結果のフィードバック

ステップ2での動作の再構成結果をプレゼンタへフィードバックし、不適切な修正があれば、プレゼンタ自身で再度修正を行い、最終的なプレゼンテーション動作を決定するステップである。プレゼンタが全く意図していない方向へ顔や視線を向けるような不適切な

修正がされている場合には、それらをプレゼンタが手動で修正して、ロボットによるプレゼンテーション代行を実現する。なお、ステップ1の最後に述べた聴衆の状態変化に合わせた伝達意図の切替え時には、本ステップでのプレゼンタによる手動修正は実施せず、ステップ2での再構成結果をリアルタイムにプレゼンタへ反映することを考えている。

5. ロボットプレゼンテーションシステム

本研究で開発中のロボットプレゼンテーションシステムの概略図を図4に示す。本システムは、プレゼンテーション収録PCとプレゼンテーション診断・再構成PC、ロボット制御PC、プレゼンテーションロボット、スライド表示デバイスにて構成される。各PC・デバイスは同一のネットワークに配置されており、各種データの転送をネットワーク経由で行う。プレゼンテーションロボットには、8自由度(胴体1軸、腕2軸、肩2軸、首3軸)を持ったコミュニケーションロボットであるVstone製のSota⁽⁸⁾を採用する。音声認識・合成エンジンにはNTTテクノクロス製のVoice-Series⁽⁹⁾、デバイス制御エンジンには、日本電信電話株式会社製のR-env:連舞⁽¹⁰⁾を使用している。本システムを用いてロボットによるプレゼンテーション代行を実施するには、プレゼンテーションの収録、収録データの診断・再構成、再構成データのユーザ修正という流れでデータ処理を進めながら、ロボット制御シナリオを作成していく。制御シナリオは、ロボットの発話内容や非言語動作と、スライド表示デバイスで利用するスライドコンテンツの実施タイミング時系列順に記述して作成する。以降では各データ処理の詳細を本研究にて作成したアプリケーションと合わせて記していく。

5.1 プレゼンテーションの収録

プレゼンタの顔向きや各関節の座標といったスケルトンデータをプレゼンテーション収録PCに接続したKinectを用いて収録する。スケルトンデータと合わせて、オーラル音声データを外部マイクにより録音し、スライドデータとして、各スライドのアニメーション動作を含めたスクリーンキャプチャ動画(解像度: 1280×960, fps:10, AVI形式)やスライドの遷移タイミングをPowerPointのAPIを介して取得する。そし

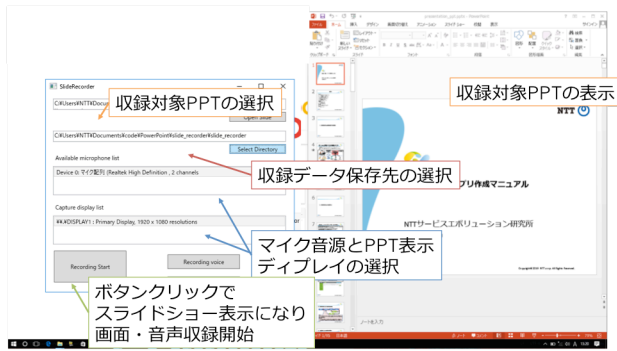


図5 P-Recorder のユーザインタフェース

て、これらデータを一括して取得するアプリケーションを P-Recorder と呼称し、Kinect SDK や MS-Office SDK などを用いて C# で実装している (図 5)。本インタフェースにて、収録対象の PowerPoint ファイルを指定し、プレゼンテーション記録ボタンをクリックすると、各データを収録時のタイムスタンプと共に記録する。このタイムスタンプを利用することで、スケルトンデータやオーラルデータ、スライドデータを正確に同期して、プレゼンテーション代行に利用する。

5.2 収録データの診断・再構成

前節で収録したスケルトンデータやオーラルデータ、スライドデータなどの各種データをプレゼンテーション診断・再構成 PC を用いて解析し、ロボット制御シナリオに利用できる形に変換する。スケルトンデータやオーラルデータから伝達意図が推定される、そして、プレゼンタが実際に行った非言語動作をプレゼンテーション動作モデルと照合しながら、顔向きやジェスチャ動作を診断し、不十分・不適切な動作があればそれらを修正・再構成する。また、スケルトンデータはロボットの関節数に合わせる形に変換される。具体的には、Kinect で取得できる関節数 25 のスケルトンデータを Sota の関節数 8 に変換しながら、非言語動作を診断、再構成される。オーラルデータは、音声認識エンジンを用いて全てテキスト化する。スライドデータは、スライドコンテンツ提供エンジンに合わせて MP4 ファイルを埋め込んだ HTML ファイルに変換され、ロボット制御 PC にアップロードされる。この各種データの変換や顔向きやジェスチャ動作の診断・再構成を行うのが、Data-Analyzer というアプリケーションであり、Python で実装している。Data-Analyzer は GUI を持たないコマンドライン実

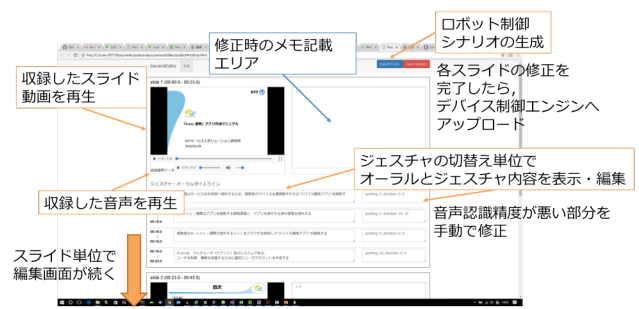


図6 Scenario-Editor のユーザインタフェース

行型アプリケーションで各収録データをバッチ的に処理する。

5.3 再構成データのユーザ修正

プレゼンタは、前節で診断・再構成した非言語動作やテキスト化したオーラルデータ、HTML ファイルに埋め込まれたスライドデータをタイムスタンプで同期し、時系列順に並べたロボット制御シナリオを確認して、必要に応じて再構成結果を手動で修正していく。ここでは、プレゼンテーション動作モデルにもとづき、再構成が適切に行われなかった場合やプレゼンタの伝達意図に合わせた非言語動作が付与されていない場合などに対応し、簡単に制御シナリオを修正できるようになっている。この制御シナリオのオーサリングツールが Scenario-Editor であり、Web ブラウザ上で動作する Web アプリとして実装している (図 6)。Scenario-Editor では、収録したスライドデータやオーラルデータ、再構成により付与された非言語動作をスライド単位で、一覧できるようになっている。また、本システムでデバイス制御エンジンに利用する R-env : 連舞®では、ロボットの発話内容や動作、スライドコンテンツ表示を状態遷移図の形で記述するフォーマット(図 7)となっているため、画面右上青色の「ロボット制御シナリオの生成」をクリックすることで、Scenario-Editor 上で修正した結果を R-env : 連舞®に合わせた形に変換し、ロボット制御 PC にアップロードされる。そして、このプレゼンタによる修正の履歴を利用した伝達意図の細分化や、動作カテゴリを構成するプレゼンテーションの基本動作の組み合わせパターンの変更といったプレゼンテーション動作モデルの詳細化を考えている。

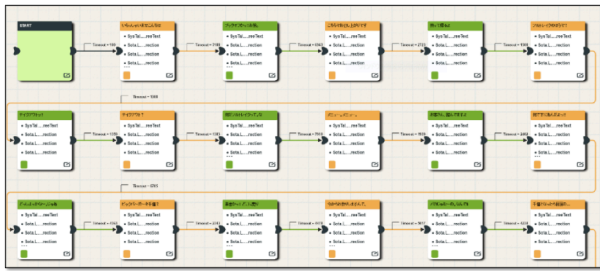


図 7 R-env:連舞®でのロボット制御シナリオ例

5.4 ロボットによるプレゼンテーション代行

5.1 節～5.3 節でのデータ処理を経て作成したロボット制御シナリオを用いたロボットによるプレゼンテーション代行の例について図 8 に記す. 図 8 上部では, 聴衆方向とスライド方向への顔向き制御により, 聴衆の注意誘導を行っている例である. 図 8 下部では, Deictic なポインティングジェスチャにより, スライド中の重要点を指し示し, スライド上部と下部を指し分けて, 理解促進を行っている例である. このように, ロボットによるプレゼンテーション代行において, 非言語動作を適切に活用することにより, 聴衆に対して分かりやすいプレゼンテーションの実施が可能となる.

6. まとめ

本稿では, プレゼンテーションでのプレゼンタの伝達意図と非言語動作の対応関係を表現したプレゼンテーション動作モデルを提案し, プレゼンタの代わりに, ロボットが分かりやすいプレゼンテーションを実施するための, プレゼンテーション再構成手法を提案した. そして, この再構成手法にもとづいて開発を進めているロボットプレゼンテーションシステムについて述べた. 本システムは, プレゼンタの伝達意図を推定し, プレゼンテーション動作モデルとの照合を行いながら, プレゼンテーションを再構成する. さらに, 再構成結果をプレゼンタにフィードバックし, 簡単にロボット制御シナリオの修正が行えるオーサリングツールを提供することで, プレゼンタの思い描くプレゼンテーションをロボットに代行させることができる.

今後は, 開発したシステムの有効性を検証する評価実験を実施して, ロボットによる分かりやすいプレゼンテーションの代行に必要な非言語動作を明らかにしていく. この評価実験で得られるデータを用いて, プレゼンテーション動作モデルにおける伝達意図の細分

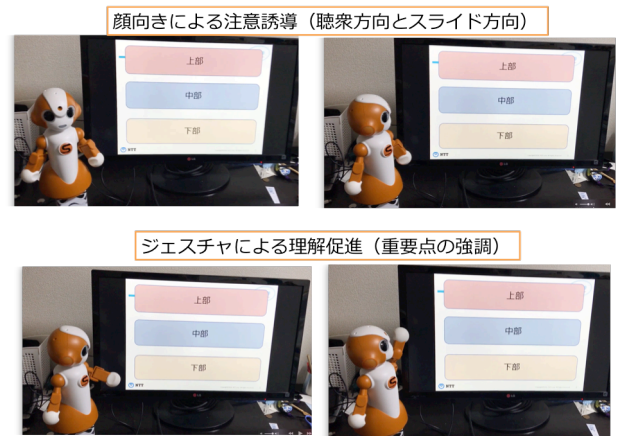


図 8 プレゼンテーション代行時の非言語動作例

化や動作カテゴリを構成するプレゼンテーション動作の組み合わせパターンの変更といったモデルの詳細化を行っていく. また, 聴衆の状態変化に合わせた動的な伝達意図の推定と推定結果に基づくインタラクティブな非言語動作の切替えについても検討を進めていく.

参考文献

- (1) Ekman, P.: "Movements with Precise Meanings", *Journal of Communication*, Vol. 26, pp.14–26(1976).
- (2) McNeill, D.: "Hand and Mind. The University of Chicago Press", (1992).
- (3) Kopf, S., Schön, D., Guthier, B., et al.: "A Real-time Feedback System for Presentation Skills", *Proceedings of EdMedia 2015*, pp. 1686-1693, (2015).
- (4) Microsoft, : "Kinect – Windows アプリの開発", <https://developer.microsoft.com/ja-jp/windows/kinect>.
- (5) 石野 達也, 後藤 充裕, 柏原 昭博: "ロボットを用いたプレゼンテーション代行による講義理解支援", *HAI シンポジウム 2017*, (2017 in press).
- (6) Kamide, H., Kawabe, K., Shigemi, S., et al.: "Nonverbal behaviors toward an audience and a screen for a presentation by a humanoid robot", *Artificial Intelligence Research*, Vol. 3, No. 2, pp. 57-66, (2014).
- (7) Chien-Ming, H., Mutlu, Bilge : "Learning-based modeling of multimodal behaviors for humanlike robots", *Proc of HRI 2014*. pp. 57-64, (2014).
- (8) VStone. ; "Sota", <https://sota.vstone.co.jp/home/>.
- (9) NTT テクノクロス. : "音声ソリューション Voice-Series", <https://www.ntt-tx.co.jp/products/voice-mall/v-series.html>.
- (10) 日本電信電話株式会社. : "ヒト×モノ×ロボットをつなげる R-env:連舞®", <https://www.r-env.jp/>.