

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

Lecture Understanding Support by Presentation Substitution using Robot

石野 達也^{1*}, 後藤 充裕², 柏原 昭博¹

Tatsuya Ishino¹, Mitsuhiro Goto², and Akihiro Kashihara¹

¹電気通信大学 情報理工学研究科

¹Graduate School of Informatics and Engineering, The University of Electro-Communications

²NTT サービスエボリューション研究所

²Service Evolution Laboratories, NTT

概要: 大学の対面講義や e-Learning で行われるプレゼンテーションでは、学習者の視線誘導や理解促進のための講師による非言語動作が非常に重要である。しかし、講師によってはこれらの動作をうまく行えず、伝えたい講義内容がうまく学習者に伝わらないという問題が生じる。そこで、本研究では、講師の代わりにロボットが講義プレゼンテーションを代行し、学習者の講義理解を支援するシステムの開発を行っている。本ロボットの主目的は、学習者の視線制御や理解を促進するための非言語動作を適切に再現することである。そのために、講師のプレゼンテーション動作から不適切・不十分な点を検出し、修正を施す。その際、講師の伝達意図と非言語動作を関係付けたプレゼンテーション動作モデルを用いて講師の動作を診断し、プレゼンテーションの再構成を行う。また、学習者の注意をより引きつけるために、オーバーアクションや、音や光などのマルチメディア情報を適宜組み込むことで、プレゼンテーションにエンターテインメント性を加味することも可能としている。このようなプレゼンテーション代行によって、学習者はより効果的に講義を受けることができると期待される。

1. はじめに

大学などの高等教育機関では、対面講義や e-Learning においてプレゼンテーションが盛んに行われており[1]、講師が受講者（以下、学習者）に対して、講義内容を伝達するためにプレゼンテーションが実施される。そこでは、学習者の講義内容の理解を促すために、講師による視線誘導や理解促進のための非言語動作が重要となる。具体的には、講師が伝達しようとしている内容（スライドコンテンツやオーラルコンテンツ）に注目させるためには、講師の視線や指差しなどの動作によって伝達内容に学習者らの視線を誘導することが必要である。Kamideら[2]は、プレゼンタが聴衆の視線を誘導することによって、プレゼンテーションの内容理解を促すことに成功している。また、より学習者の理解を促すた

めに、講師が学習者の注目を誘導することに加えて、ジェスチャなどで講義内容の特徴を表現したり、重要な箇所を強調することなどが必要といえる。

しかしながら、講師がプレゼンテーションを通して、必ずしも学習者らの視線を誘導したり、講義内容の理解を促すことができるとは限らない。例えば、手元の PC を常に見ながら講義を行ったり、ジェスチャを使わない内向的な講師も存在し、講義内容がうまく学習者に伝わらず、講義への集中が途切れて学習者の意欲や興味が講義途中から低下することが考えられる。その結果、学習者の理解を促すことができない場合も多い。

そこで、本研究では講師の代わりに、ロボットが講義プレゼンテーションを代行し、学習者の講義理解を支援する手法とシステムを提案している[3]。本システムでは、講師が行ったプレゼンテーションを撮影し、視線誘導や講義内容の理解促進のために必要となる講師の動作について診断を行い、不適切・不十分な動作を検出・修正した後、プレゼンテーションを再構成してロボットが代行することを目的

*連絡先：電気通信大学 大学院情報理工学研究科情報学専攻
〒182-8585

東京都調布市調布ヶ丘 1-5-1 電気通信大学西 3 号館 302
E-mail : ishino.tatsuya@uec.ac.jp

としている。このようなプレゼンテーション動作の再構成を行うためには、プレゼンタの伝達意図や聴衆の状況を加味して、聴衆の視線誘導や理解促進に必要なプレゼンテーション動作を明らかにする必要がある。このような観点から、本研究では現在伝達意図とプレゼンテーションの動作を対応づけるプレゼンテーション動作モデルのデザインを進めている。本モデルでは、伝達意図、動作カテゴリ、プレゼンテーション基本動作の3層に分け、プレゼンテーション動作として必要となる非言語動作を表現する。開発を進めているプレゼンテーション代行システムでは、このモデルに基づき、講師の動作からプレゼンテーション中の視線誘導や理解促進動作として、不適切・不十分な点を診断し、修正することで、内向的な講師による講義でもロボットを通して伝えたい講義内容をうまく学習者に伝えることが可能となり、学習者の理解を促すことができる。また、オーバーアクションや、音楽・光などのマルチメディアを提示するなど、人間のプレゼンタが実施したいことをロボットに行わせることで学習者の興味をより引きつけて講義を行うことが可能である。

以下、本稿では、講義プレゼンテーションにおける現状と課題を明らかにし、その解決アプローチとして講師のプレゼンテーション動作モデルを示す。そして、視線誘導や講義内容の理解促進のために必要となる講師のプレゼンテーション動作について考察した上で、ロボットによるプレゼンテーション代行の枠組みを示し、モデルに基づく講師の不適切・不十分な点の診断・修正方法、および誇張再現の方法について論じる。最後に、現在開発を進めているロボットによるプレゼンテーションシステムについて述べる。

2. 講義でのプレゼンテーション

ここでは、講義プレゼンテーションの現状と課題を述べ、その課題解決アプローチとしてプレゼンテーション動作モデルを説明する。

2.1 講義プレゼンテーションの現状と課題

通常、大学での講義では、講師が講義内容を表現したスライドを用いてプレゼンテーションを行う場合が多い。スライドには、イラストやグラフ、キーワードが散りばめられており、講師は口頭でこれらを説明するとともに、スライドには陽に表現されていない事項についても口頭で補足説明を行う。例えば、日本経済の講義では、スライドで貿易輸出額について表やグラフを用いて説明されるが、その際に使わ

れる用語であったり他国との関係性については口頭のみで補足説明される場合がある。このように講義でのプレゼンテーションでは、講師が伝達しようとするスライドコンテンツとオーラルコンテンツといった2種類の講義コンテンツが存在する。講師は、学習者に対してどちらのコンテンツに注目して欲しいかという伝達意図を踏まえて、プレゼンテーションを遂行することが望まれる。

しかしながら、講師によっては手元のPCを常に見てプレゼンテーションを進めたり、ジェスチャをあまり用いず、淡々と口頭説明をするなど重要な箇所をうまく伝えられずに進めてしまうことがある。このような場合、学習者の意欲や興味が講義途中で無くなって、講義への集中が途切れてしまう場合もある。その結果、伝えたい講義内容がうまく学習者に伝わらずにプレゼンテーションを終えてしまうことも少なくない。

2.2 視線誘導・理解促進動作

前節で述べた課題を解決するために、プレゼンタにどのような動作が求められるかを整理することが不可欠である。特に、講師がプレゼンテーション時に意図したコンテンツへ学習者の注目を集めるために、学習者の視線を誘導するような動作が必要である。また、伝達しようとするコンテンツの理解を促すような動作も必要である。さらに、聴衆である学習者の受講状況も加味した上でプレゼンテーション動作をとることが望ましい。

本稿では、視線誘導、理解促進のために求められる動作に焦点を当て、以下ではプレゼンテーションに必要な非言語動作について述べる。

2.2.1 講師による聴衆の視線誘導

講師が学習者の注目を集める一つの方法として、学習者の視線誘導が考えられる。Kamideら[2]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が聴衆の視線誘導を促す可能性が示されており、注目させたいコンテンツへの視線誘導を行うことで聴衆のプレゼンテーション内容への理解が高まることが示唆されている。この知見に基づけば、図1に示すように、講義プレゼンテーションにおいても講師がスライドへ視線を移すことで学習者の視線をスライドに誘導したり、スライドから学習者に視線を向けることで学習者の注意を講師に向けさせることができると考えられる。

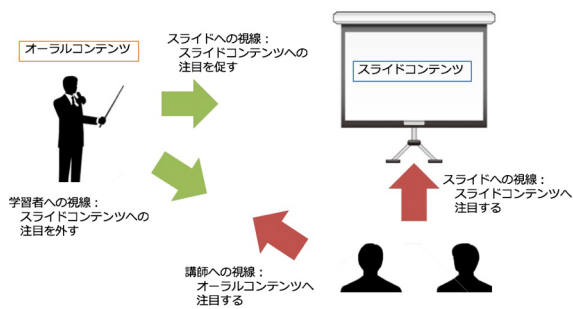


図1 講師による視線誘導

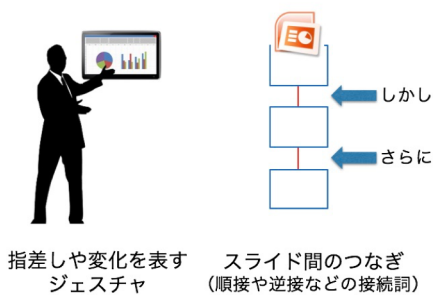


図2 理解促進のための講師の動作

2.2.2 理解を促すためのジェスチャ

視線誘導に加えて、講義プレゼンテーションでは、重要な箇所を強調しながらコンテンツを伝達することで、学習者の理解をより促進することができると考えられる。そのためには、スライドコンテンツやオーラルコンテンツをわかりやすく表現することに加えて、効果的にコンテンツの理解を促すようなジェスチャなどの非言語動作を交えたプレゼンテーションを行う必要がある[4]。

- McNeill[5]は、こうした観点から人のジェスチャを、
- (a)Deictic: 強調部分を指差すなどのポインティングジェスチャ、
 - (b)Metaphoric: 3つの主張がある場合に指を3本立てて表現したり、増減を表現するために手を上下に動かすような概念を表現するジェスチャ、
 - (c)Iconic: 両手で大きさや長さを表現するジェスチャ、

に分類している。図2に示すように、講義でもこのようなジェスチャ分類に基づき、伝達しようとするコンテンツに合わせて適切なジェスチャを行えば、講義内容が学習者によりわかりやすく伝わり、印象に残るプレゼンテーションを実施することができると考えられる。

また、講義ではスライド間の遷移において学習者

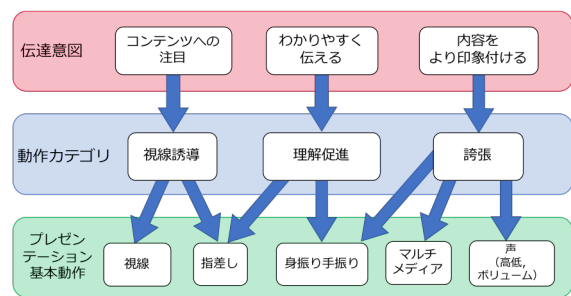


図3 プレゼンテーション動作モデル

の理解が途切れることのないように、スライドのつながりをわかりやすく伝えることも重要である。そこで、図2に示すように、スライド間をつなぐ接続詞を口頭で強調したり、接続詞に合わせたジェスチャを施すことで理解を促すことも必要である。例えば、「ところが」という接続詞で話を転換させる際には、手を裏返すといったジェスチャが考えられる。

2.3 プレゼンテーション動作モデル

ここまで視線誘導・理解促進のために必要となる動作について述べてきた。他にも、プレゼンテーション動作にはマルチメディアの使用やパラ言語（声の高低、ボリューム）の活用などがある。しかし、これらの動作は無闇に使うものではなく、プレゼンターの伝達意図や聴衆の状況を考慮して、実施することが重要である。

そこで、本研究では図3に示すように、伝達意図、動作カテゴリ、プレゼンテーション基本動作の3層に分けて、これらの対応関係でプレゼンテーションに必要な動作を表現するモデルのデザインを進めている。本モデルでは、講師の伝達意図を考慮した上で動作カテゴリを決定しており、その動作カテゴリに基づいてプレゼンテーション基本動作を組み合わせ、伝達意図に見合う動作を表現するモデルである。

例えば、コンテンツへの注目を実施する場合には、学習者の視線誘導を行う必要がある。視線誘導を成功させるためには講師の視線や指差しを組み合わせた動作が望まれる。また、コンテンツをわかりやすく伝えたい場合には、対象を指差した上で2.2.2で述べたジェスチャを行うことが考えられる。さらに、伝達内容をより印象付ける場合には、誇張カテゴリの身振り手振りの動作を大きくしたり、マルチメディアの使用やパラ言語（声の高低、ボリューム）を効果的に使うことが望まれると考えられる。

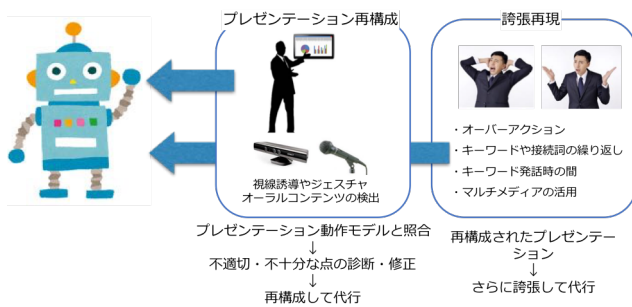


図 4 ロボットによるプレゼンテーション

3. ロボットによる

プレゼンテーション代行

ロボットによるプレゼンテーションとして、本研究では図 4 に示すようにプレゼンテーション再構成と誇張再現を検討している。プレゼンテーション再構成は講師の動作を忠実に再現するわけではなく、プレゼンテーション動作モデルと照合しながら不適切・不十分な点を診断・修正し、再構成してプレゼンテーションを代行する。誇張表現は内容をより印象付けて理解を促進させるために再構成されたプレゼンテーションをさらに誇張して代行するものである。

3.1 プレゼンテーションの再構成

プレゼンテーション再構成では、講師によるプレゼンテーションを入力として、入力された情報とプレゼンテーション動作モデルを照合しながら、不適切・不十分な点を診断・修正し、プレゼンテーションを再構成して代行する。例えば、スライドコンテンツを口頭で説明している際にスライドへの視線誘導ができていない場合や、スライドの補足説明をしている際に講師の視線が学習者に向けられていない場合などを図 5 に示すように、プレゼンテーション動作モデルと照合しながら視線誘導で不十分な点を診断・修正し、ロボットが視線誘導のための動作を代行する。理解促進の動作についても、同様に診断し、キーワードに指差しのジェスチャが行われていない場合や、スライド遷移時に接続詞が使われていない場合、ロボットが対応するジェスチャや口頭説明を自動的に代行することを検討している。

これらの視線誘導や講義内容の理解促進について、不適切・不十分な動作を診断・修正し、プレゼンテーションの再構成が行われることで内向的な講師も伝えたい講義内容をうまく学習者に伝えることが可

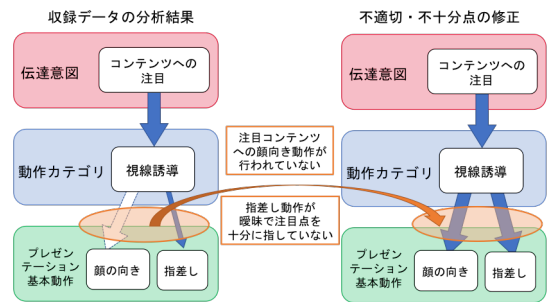


図 5 プレゼンテーションモデルとの照合

能となる。また、講師はロボットによって再構成されたプレゼンテーション動作を内省することで、自らのプレゼンテーションのあり方を改善することも可能となる。

3.2 プレゼンテーションの誇張再現

次に、誇張再現では、プレゼンテーションにおける視点誘導や理解促進のための動作に対してエンターテイメント性を付与し、学習者に講義内容をより印象付けることを目的としている。そのために、プレゼンテーション動作モデルの誇張カテゴリに基づいて、

- (i)身振り手振り：ジェスチャの動きを大きくする、
- (ii)マルチメディア：音や光などを伴わせる、
- (iii)声の大小や高低：キーワードやスライド接続を強調するために声を大きくする・高くする、

などを組み込んだプレゼンテーション代行を実現する。つまり、ジェスチャの動作を大きくすることや、音楽や光などのマルチメディアを提示するなど、人間のプレゼンタが実施しがたいことをロボットに行わせることがプレゼンテーション代行における誇張再現の意義である。

4. ロボットによる

プレゼンテーションシステムの設計

本研究では、図 6 に示すとおり、ロボットを用いたプレゼンテーション代行システムの開発を行っている。

本システムでは、プレゼンテーション代行ロボットとして Vstone 製の Sota[6]を使用する。Sota は、発話や身振り手振りが利用できる小型のコミュニケーションロボットである。合計 8 自由度(胴体 1 軸、腕

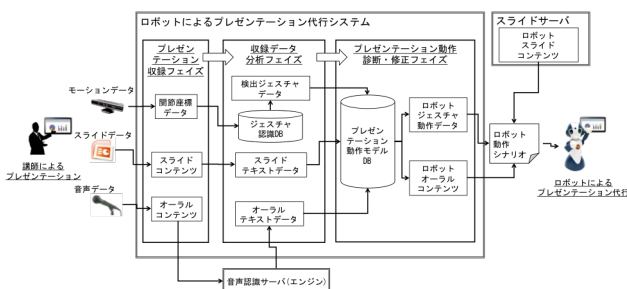


図 6 プレゼンテーション代行による講義理解支援システム

2軸, 肩2軸, 首3軸)を持っており, Javaによる制御が可能である.

ロボットがプレゼンテーションを代行するにあたり, プレゼンテーションの収録, 収録データの分析, プレゼンテーション動作の診断・修正の3つのフェイズに分けて開発を進めている. プレゼンテーション収録フェイズでは, 講師によるプレゼンテーション時の骨格座標データやスライドコンテンツ, オーラルコンテンツをそれぞれ収録する. 収録データ分析フェイズでは, 収録した各種データを分析し, 骨格座標データと後述するジェスチャ認識データベースを利用してプレゼンテーション中のジェスチャを検出する. また, スライドコンテンツ中のテキストをスライド単位で抽出したり, オーラルコンテンツの音声認識により, テキスト化する. プレゼンテーション動作診断・修正フェイズでは, プレゼンテーション動作モデルと照合しながら, 不適切・不十分な動作を診断・修正し, ロボットジェスチャ動作データとロボットオーラルコンテンツとして, プレゼンテーションを再構成する. この3フェイズを経て, Sotaの動作シナリオを作成して適用する. 以下に, 本システムの詳細と講義での利用方法を説明する.

4.1 プレゼンテーション収録フェイズ

プレゼンテーション収録フェイズでは, 顔向きや指差しといったプレゼンタの骨格座標データをMicrosoftのKinectを用いて収録する. 骨格座標データの収録と合わせて, オーラルコンテンツを外部マイクを介して録音し, スライドコンテンツとして, 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. スライドコンテンツは, ロボットがネットワーク経由で接続可能なスライドサーバ上にアップロードされ, ロボットによるプレゼンテーション時に利用される. また, 各データは収録時のタイムスタンプを保持しており,

このタイムスタンプを利用して, 骨格座標データやオーラルコンテンツとスライドコンテンツの表示タイミングを正確に同期できる.

本システムの利用に際しては, まずプレゼンテーション中に使用できるジェスチャを講師に提示し, 確認後プレゼンテーションを行ってもらいたいことを想定している.

4.2 収録データ分析フェイズ

収録データ分析フェイズでは, 骨格座標データからプレゼンテーション中に利用されているジェスチャを認識する. この認識にはVisual Gesture Builderを用いる. これはMicrosoftが提供しているKinect v2用ツールで, 機械学習を使ってジェスチャを認識するための識別器が作成される[7]. 従来手法では, ジェスチャの判定に要するコードが多く, 開発の際の負担も大きくなるという問題があったが[8], 作成された識別器を読み込むことで簡単にジェスチャを認識するプログラムを作成することができる. 本研究では, このツールで作成した識別器をもとにジェスチャ認識データベースを構築し, 骨格座標データの分析に利用する. また, スライドコンテンツから, PowerPointのAPIを用いて各スライドに記述されているキーワードをスライドテキストデータとして抽出する. この抽出時には, 文字サイズや文字色, 太字などのフォント情報を取得し, 重要度の判定に利用する. そして, オーラルコンテンツを外部の音声認識サーバを用いてテキスト化する.

4.3 動作診断・修正フェイズ

プレゼンテーション動作診断・修正フェイズでは, 前フェイズで分析したスライドテキストデータとオーラルテキストデータを比較し, 同時刻におけるスライドコンテンツとオーラルコンテンツとでキーワードの対応が取れない場合や, 講師の顔向きや指差しといったジェスチャ動作が適切かつ十分に行われていない場合などを診断し, 修正点を検出する. この検出には, 講師の伝達意図に基づいてプレゼンテーション動作の組み合わせを決定するプレゼンテーション動作モデルを用いる. この検出結果に基づき, 修正・再構成したロボットジェスチャ動作やロボットオーラルコンテンツと, 4.1節でスライドサーバにアップロードしたロボットスライドコンテンツの実行タイミングを記載したロボット動作シナリオを講師に提示する. 本研究では, 講師自身がこのロボット動作シナリオ中のジェスチャやオーラルコンテンツを簡単に編集・修正できるオーサリング機能の実

装を予定している。そして、この講師によるプレゼンテーション動作の編集・修正の履歴を利用して、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考えている。

また、動作モデルに基づき、不適切かつ不十分な点の修正だけではなく、よりプレゼンテーション内容を印象付けるための誇張再現を動作シナリオに追加する機能の検討も進めている。

4.4 Sota による講義プレゼンテーション

Sota による講義プレゼンテーション実施には、プレゼンテーション収録、収録データ分析、プレゼンテーション動作診断・修正の3つのフェーズを経て作成されたロボット動作シナリオを利用する。実際の講義で使用する際には、Sotaはこのロボット動作シナリオに従いながら動き、指定されたタイミングでスライドサーバとの通信を行い、サーバ内に格納されているロボットスライドコンテンツとSotaのジェスチャ動作やオーラルコンテンツとを同期しながらプレゼンテーションを代行する。このようなロボットによる講義プレゼンテーション代行は、大学の講義におけるプレゼンテーションだけでなく、e-Learningとしてのプレゼンテーションコンテンツと連動への活用も考えられる。自宅で学習する際に画一的な講義プレゼンテーションを聴講するのではなく、学習者自身が好む講師のプレゼンテーションスタイルをロボットに反映させて学習するといった新しい学習環境の実現も可能となる。

5. まとめ

本稿では、プレゼンテーションでの講師の伝達意図とプレゼンテーション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、講師の代わりに、ロボットが講義プレゼンテーションを実施し、学習者の講義理解を支援する手法を提案した。さらに、講師のプレゼンテーションをロボットに代行させる上で不適切・不十分な動作を診断し、修正する方法を示し、誇張再現方法についても述べた。そして、現在開発を進めている講義理解支援のためのロボットを用いたプレゼンテーション代行システムについて、プレゼンテーションの収録、収録データの分析、プレゼンテーション動作の診断・修正の3つのフェーズに分けて紹介した。

今後は、これら3つのフェーズの未実装部分を開発し、ロボットによる講義プレゼンテーションの評

価実験を行い、その有用性や視線誘導、ジェスチャによる講義理解への効果について検証していく予定である。さらに、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。現状のプレゼンテーション動作モデルでは、伝達意図とプレゼンテーション動作が関係付けられており、このモデルに基づき、ロボットから聴衆への情報伝達を実施しているが、聴衆の状況についてもプレゼンテーション動作モデルに組み込み、一方向的な情報伝達からインタラクティブな情報伝達としてプレゼンテーション動作を行えるようなシステムにすることを目指していく。

参考文献

- [1] 後藤 充裕, 柏原 昭博: スライド間の接続関係の可視化に基づくプレゼンテーションドキュメント理解支援, 電子情報通信学会技術研究報告, Vol. 115, No. 492, pp. 273-278, (2016)
- [2] Kamide, H., Kawabe, K., Shigemi, S., and Arai, T.: 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)
- [3] 石野 達也, 後藤 充裕, 柏原 昭博: 代講を目的としたロボットによるプレゼンテーション, 人工知能学会第81回先進的学習科学と工学研究会 (2017.11 in press).
- [4] Chien-Ming, H., Mutlu, Bilge.: Learning-based modeling of multimodal behaviors for humanlike robots. *Proc of HRI 2014*. pp. 57-64(2014)
- [5] McNeill, D.: *Hand and Mind*. The University of Chicago Press, 1992.
- [6] Sota
<https://sota.vstone.co.jp/home/>
- [7] 中村 薫, 杉浦 司, 高田 智宏, 上田 智章: KINECT for Windows SDK プログラミング Kinect for Windows v2 センサー対応版, 秀和システム, (2015)
- [8] Buu: Kinect v2 × Visual Gesture Builder でファイトだよっ!, <http://sv.buu0528.com/d/c87vvgb.pdf>, (2014)