

# 3次元仮想空間における エージェントの連続的動作を対象とした視点の変化

大谷 章      黒澤 義明      目良 和也      竹澤 寿幸

広島市立大学大学院 情報科学研究科

{otani,kurosawa,mera,takezawa}@nlp.its.hiroshima-cu.ac.jp

## 1. はじめに

近年、CG技術の発達により、仮想空間上で物体の複雑な動作が表現できるようになっている。また、音声認識技術の向上に伴い、音声入力を用いて仮想空間上のエージェントを対話的に操作する研究が試みられている[1], [2]。

音声によるエージェントの操作では、解決しなければならない問題がある[3]。それは入力発話に含まれる「左」、「右」などの空間表現が持つ曖昧性である。例えば、図1.1における仮想空間内のエージェント（中央に配置されているキャラクター）を矢印の方向に動かす場合について考える。ユーザがエージェントに対して「左に移動して」と言った場合、システムはこの発話に含まれる「左」をユーザから見て左と解釈しなければならない。一方、「右に動いて」と言った場合、「右」をエージェントから見て右と解釈しなければならない。このように音声入力による仮想空間操作システムでは、「左」、「右」などの空間表現がユーザ視点とエージェント視点、どちらの視点から述べられているのかを決定する必要がある。そこで本研究では、音声入力を用いたエージェント操作における空間表現の曖昧性について扱っていく。特にその空間表現が頻繁に使われるエージェントの移動動作（「左に移動する」、「奥に移動する」）について言及する。本論文では以後、指定がない限り、「左」、「右」、「手前」、「奥」をユーザ視点として記述する。

しかし、図1.1のように操作対象のエージェントがユーザに対して静止して前を向いているとき、「左に移動して」など左、右を用いた指示に対して、ユーザ視点、エージェント視点を決定付けることは難しい。そこで本研究では、エージェントの状態に応じて、ユーザが特定の視点から発すると推測する。例えば、ユーザは動いている物体について話すとき、その物体を基点にして話すと思われる。そのため、ユーザが動き続けている状態のエージェントに対して指示を与える場合、エージェント視点から指示をすることを考えられる。そこで本研究では、エージェントが静止している状態と動き続けている状態を比較し、指示を与える際にユーザが採用する視点の傾向について検討する。

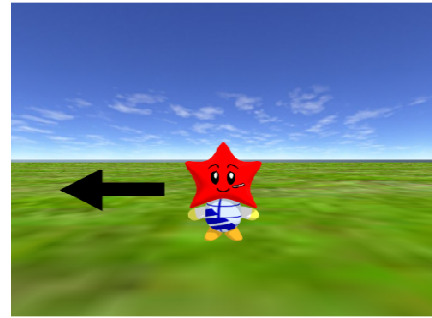


図 1.1 : エージェントの移動方向

## 2. 研究目的

本研究と同様に、空間表現の曖昧性を扱っている研究がある[4]。Imaiらの研究では、図2.1のように、2つの物体の位置関係を表す空間表現「前、後、左、右」が持つ曖昧性の解決を試みている。なお図2.1の画像はImaiらの研究とは関係ない。図2.1の観察者は球を指定する場合、基本的に「球はエージェントの右にある」、または「球はエージェントの左にある」と表現する。前者の表現は観察者視点から、後者の表現は参照されている物体(参照物体)の視点から述べられている。この研究では、物体位置の指定時に観察者が採用する視点の傾向について実験を行っている。

Imaiらの実験結果では、図2.1のような状況において、観察者視点（右）、参照物体視点（左）はほぼ同じ割合で採用されている。この結果になった原因として、Imaiらの実験方法に問題があると考えられる。Imaiらは実験を行う際に、図2.1のような静止画を提示することで、被験者に空間表現を選択させている。この手法により、被験者は仮想空間内の物体に対して十分な感情移入をすることができず、左右の選択が割れたと考えられる。また、この研究では、参照物体としてエージェントだけでなく、向きのあるイスやテレビも用いている。Imaiらの結果では、参照物体がエージェントのとき、他の参照物体に比べてエージェント視点からの選択が多いということがわかっている。このような結果になったのは、参照物体に対する感情移入の度合いによって決まったものだと考えられる。

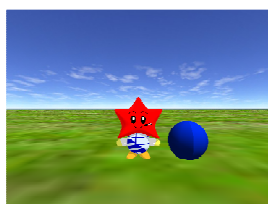


図 2.1 : エージェントと球の位置関係

そこで本研究では、参照物体が動き続けている様子を用いて、観察者の物体への感情移入を促すことにする。本節で記述した観察者視点と参照物体視点は本研究で言及するユーザ視点とエージェント視点と同義であると言える。つまり、エージェントが動いている状態において、ユーザが指示をする場合、エージェント視点での指示が多くなると考えられる。エージェントの動きにおいて、空間表現が頻繁に使われる動作として移動動作（右に移動）が挙げられる。

そこで本研究では、この移動動作における空間表現に含まれる曖昧性の解決を行っていく。その中で、エージェントが動き続けている状態が視点の変化に影響を与えるかを検討する。そのため、エージェントが動き続けている状態と静止している状態について比較を行う。なお、本論文では、この要因を動作の連続性と呼称する。

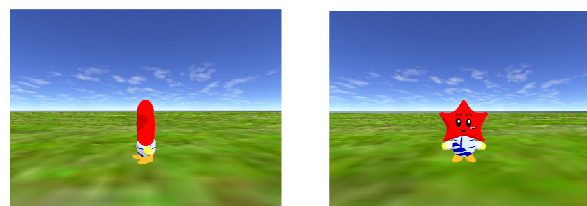
### 3. 動作の連続性

本研究では、動作の連続性がユーザの採用する視点に影響を与えるかについて検討する。この要因は2つの水準を持つ。一つはエージェントが静止している状態である。もう一つはエージェントが動き続けている状態である。本論文では前者の水準を不連続的動作（図 3.1）、後者の水準を連続的動作（図 3.2）と記述する。なお、不連続的動作は動作終了時、必ずエージェントは手前を向くようになっている。この挙動の理由として、これら2つの水準について比較する上で、最も差が出ると思われるからである。

本研究では、各水準に関する2つの実験に分け、それぞれの実験について異なる被験者を割り当てた。各実験では、複数の被験者に対して同時に行った。そのため、音声入力による収集でなく、エージェントが仮想空間内を移動する動画を用いて、命令表現に含まれる視点情報を収集した。また、エージェントの移動のために、移動経路を設けた。

### 4. エージェントの移動経路

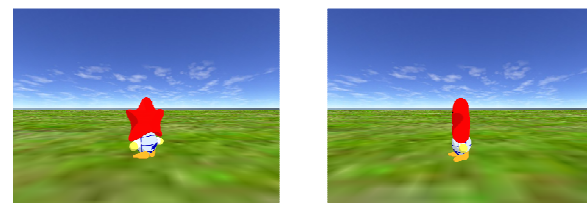
本実験では、エージェントが移動できる方向として、左、右、奥、手前の4方向を選択した。連続的動作において、



a : 次動作

b : 静止

図 3.1 : 不連続的動作（左へ移動）



a : 次動作

b : 動き続けている

図 3.2 : 連続的動作（奥へ移動）

各移動方向は、一つ前の移動方向によって命令表現の内容が変わってくる。例えば、エージェントが奥へ移動する場合について考える。この動作前の移動方向が左のとき（図 3.2）、ユーザは「奥へ移動して」（ユーザ視点）、「右に曲がって」（エージェント視点）の2通りの指示を行うと考えられる。また、動作前の移動方向が奥のとき（進み続ける）、「奥へ移動して」（ユーザ視点）、「直進して」（エージェント視点）の指示が想定される。このように、連続的動作では、エージェント視点の命令表現が前の移動方向によって変化する。そのため、前の動作を考慮に入れた16通りの移動（4×4）を行う必要がある。

本研究では、エージェントが16通りの全ての移動を行うことができるような経路を考えた（図 4.1）。そして、

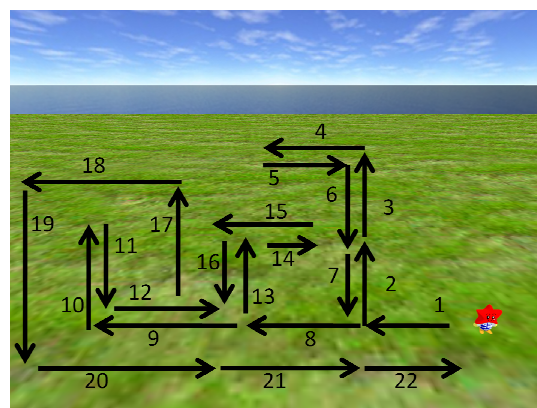


図 4.1 : エージェントの移動経路

各水準におけるエージェントが図 4.1 の経路を辿って移動する動画を MikuMikuDance [5]により作成した。なお、それぞれの動画において、エージェントの最初の状態は手前を向いて静止している状態である。また、各移動動作の時間は約 4 秒間であり、それぞれ同じ距離を移動する。本研究では、この作成した動画を用いて、各水準における視点情報を収集するための実験を行った。

## 5. 実験

本研究では、動作の連続性によって、ユーザが採用する視点の傾向について検討する。そのため、各水準における視点情報を収集するための実験を行った。

本実験では、不連続的動作に関する実験と連続的動作に関する実験を分けて行った。また、各実験には、それぞれ異なる被験者を割り当てた。その結果、不連続的動作に関する実験を大学生 38 名に、連続的動作に関する実験を大学生 72 名に対して行った。以下の節では、実験の方法、収集したデータの解析法、実験結果について述べる。

### 5.1. 実験方法

各実験では、エージェントが図 4.1 の経路を移動する動画を用いた。各実験において、被験者には動画を見せ、その動画内の各移動に対して、その移動動作を実現するための命令表現を記述してもらった。それぞれの実験において、命令表現を記述するタイミングについて説明する。

不連続的動作では、移動終了時に必ずエージェントが手前を向く。そのため、手前を向いたと同時に動画を停止した。停止後、被験者には再生中の移動動作について記述してもらった。以後の動作も同様の方法で行った。

一方、連続的動作では、エージェントは止まらず、次の移動動作を行う。そのため、図 4.1 の移動 2 では、図 3.2-a のような動作を行った 1 秒後に停止した。また、移動 3 のように進み続ける場合、再生してから 3 秒後に停止した。停止後、被験者には再生中の移動動作について記述してもらった。以後の動作も同様に行った。

### 5.2. 対象とするデータ

本実験で収集した命令表現において、以下の記述のみを行った被験者に対して、本研究のデータとして扱わないことにした。

- (1) 英語による記述  
例：Go left, Go ahead
- (2) 矢印による記述  
例：←, ↑

本研究では、日本語音声入力による仮想空間操作について考えている。そのため、上記 2 つの命令表現を使用した被験者は対象外とした。

その結果、対象となる被験者の数は不連続的動作に関する実験では 28 名、連続的動作に関する実験では 62 名であった。本研究ではこれらの被験者が記述した命令表現について扱っていく。

### 5.3. 命令表現の解析

対象とする被験者の各命令表現に対して、視点情報に関するタグを付与した。タグは全 5 種類である。表 5.1 はエージェントが手前を向いているときのタグの付与例を示している。

「ユーザ視点」はユーザ視点からの命令表現に付与されるタグである。「エージェント視点」はエージェント視点からの命令表現に付与されるタグである。「視点一致」はユーザ視点、エージェント視点の両方の視点から解釈可能な命令表現に付与されるタグである。「複合」はユーザ視点とエージェント視点に関する 2 つの空間表現を含む命令表現に付与されるタグである。「判別不可」は視点の判別ができない命令表現に付与されるタグである。

表 5.1：タグの付与例

タグ	命令表現	移動方向
ユーザ視点	左に歩け	左
エージェント視点	右に歩け	左
視点一致	後ろに歩け	奥
複合	左を向いて直進	左
判別不可	歩け	左

### 5.4. 実験結果

本論文では、左右に関してユーザ視点からエージェント視点に変化するかどうかについて議論する。そのため、「左」、「右」をエージェント視点の解釈から命令できる移動動作に着目した。不連続的動作では図 4.1 における [1,4,5,8,9,12,14,15,18,20,21,22] 番の移動について扱った。また、連続的動作では図 4.1 における [1,2,6,8,10,12,13,16,19,20] 番の移動について扱った。

上記で示した移動において、命令表現に付与されたタグの数を計算した。そして、最も多い数のタグをその被験者の視点情報とした。そのときの結果を表 5.2 に示す。なお、表 5.2 には判別不可、複合の被験者は載せていない。

表 5.2：各視点を採用した被験者の平均

	不連続的動作	連続的動作
ユーザ視点	0.476 (10/21)	0.491 (29/59)
エージェント視点	0.523 (11/21)	0.508 (30/59)

表 5.2 より、不連続的動作と連続的動作を比較すると、「左」、「右」が影響を受ける移動において、被験者が採用する視点の傾向はほぼ同じになった。しかし、連続的動作において、エージェント視点を採用した被験者の中に、移動 1、2 でユーザ視点を採用している傾向が見られた。

## 5.5. 連続的動作がもたらす視点の変化

各水準において、エージェント視点を採用した被験者について着目した。その中で、ユーザ視点からエージェント視点に変化した被験者の数を求めた。そのときの結果を表 5.3 に示す。なお、不連続的動作は移動 1 を、連続的動作は移動 1、2 のどちらかがユーザ視点の場合、変化したと定義した。表 5.3 の「ユーザ→エージェント」はユーザ視点からエージェント視点に変化した被験者を意味する。一方、「エージェント→エージェント」は最初からエージェント視点のまま記述した被験者を意味する。また、表 5.2 と表 5.3 の不連続的動作におけるエージェント視点の被験者数が合わない理由として、移動 1 における命令表現が「判別不可」のタグが付けられているからである。

表 5.3 より、連続的動作の方がユーザ視点からエージェント視点へ変化しているように見える。しかし、この差は統計学上有意のある差かどうかかわからない。そこで、 $\chi^2$  検定を用いて、表 5.3 の結果が有意のある差かどうかを確かめた。なお、このときの自由度は 1 である。その結果、 $\chi^2$  値は 5.76、有意確率は 0.0163 となった。この有意確率は有意水準 0.05 を下回っている。つまり、表 5.3 の結果は有意な差があるといえる。そのため、不連続的動作に比べて連続的動作の方がユーザはエージェント視点からの指示を多用すると考えられる。

表 5.3：エージェント視点に変化した被験者の平均

	不連続的動作	連続的動作
ユーザ → エージェント	0.100 (1/10)	0.533 (16/30)
エージェント → エージェント	0.900 (9/10)	0.466 (14/30)

しかし、全体の実験結果である表 5.2 では、ユーザ視点、エージェント視点の採用される割合はほぼ同じであった。この結果はエージェントが最初に左に移動したことが原因であると考えられる。このとき、被験者がユーザ視点として「左に移動して」と記述した場合、以後の左右に関する動作も同様にユーザ視点からの命令表現で記述したと考えられる。そのため、エージェントの最初の動作を手前、奥から移動し始めた実験も行う必要があると考えられる。

## 6. おわりに

本研究では、音声を用いたエージェントの操作におけ

る空間表現の曖昧性について扱った。特にエージェントの移動動作について考えた。その中で、エージェントが動き続けている状態と、静止状態のときと比べて、ユーザが採用する視点の傾向が変わるかどうか検討した。その結果、動き続ける状態の方が、ユーザ視点からエージェント視点へと変化しやすいことがわかった。

本研究では、収集した命令表現へのタグの付与に「右」、「直進」などの空間表現を表す名詞のみに着目した。視点の特定に利用できる空間表現は名詞だけでなく動詞でも特定することができる[6]。例えば「来て」という命令表現はユーザ視点から指示しているということがわかる。今後、名詞だけでなく動詞にも着目することで、幅広い命令表現に対応できると考えられる。

## 謝辞

本研究において、実験に御協力してくださった被験者の皆様に心より感謝致します。

本実験で使用した動画を作成するにあたって、以下のものをらせて頂きました。

MikuMikuDance：樋口優様

地面テクスチャ：Bo Hammarberg 様

空、海テクスチャ：おやぶん L 様

はちゅね用ボーン：朽葉様

上記の作者の皆様に心より感謝致します

## 参考文献

- [1] Otani, S., Kurosawa, Y., Mera, K., and Takezawa, T., “Classification of Utterances into Command Expressions and Conversational Expressions in Virtual Space Operation Task,” eaSDS2009, 2009.
- [2] 新山祐介, 徳永健伸, 田中穂積, “自然言語を理解するソフトウェアロボット：傀儡,” 情報処理学会論文誌, vol.42, no.6, pp.1359-1367, 2001.
- [3] 徳永健伸, 田中穂積, “ロボットにおける言語理解,” 日本音響学会論文誌, vol.63, no.1, pp.35-40, 2007.
- [4] Imai, M., Nakanishi, T., Miyashita, H., Kidachi, Y., and Ishizaki, S., “The meanings of front/back/left/right,” Cognitive Studies, pp.207-255, 1999.
- [5] <http://www.geocities.jp/higuchuu4/>
- [6] 澤田治美, “視点と主観性 -日英語助動詞の分析,” ひつじ書房, 1993.