

# 隠喩的ジェスチャの分析とジェスチャ自動付与に向けた検討

## Analyzing Metaphoric Gestures towards Automatic Gesture Generation

門野 友城<sup>\*1</sup>  
Yuki Kadono

高瀬 裕<sup>\*1</sup>  
Yutaka Takase

中野 有紀子<sup>\*1</sup>  
Yukiko Nakano

<sup>\*1</sup> 成蹊大学理工学部  
Seikei University

Virtual agents are expected to allow more natural human-computer interaction because they can express facial expressions and eye gaze using animations. Therefore, in such interfaces, expressing appropriate nonverbal signals is indispensable, and especially agent's gestures contribute to improve comprehensibility of the agent's utterances. With a goal of automatic gesture generation in virtual agents, this study, first, collects and analyzes hand gestures. We focus on metaphoric gestures, which express abstract concept, and categorize the collected metaphoric gestures. Then, analyzing the relationship between the concepts and the gesture shapes, we propose gesture decision rules for metaphoric gestures. Finally, we implement a gesture generation mechanism that receives a sentence as an input, and assigns gesture using the proposed gesture decision rules, and produce speech and agent animation through a text-to-speech system and a game engine.

### 1. はじめに

アニメーションによるバーチャルエージェントは、表情や視線等、様々な非言語行動を表出することができ、人とのより自然なコミュニケーションを実現するインタフェースとして期待されている。このようなバーチャルエージェントによる身体的表現を利用した機能の 1 つにジェスチャ(ハンドジェスチャ)がある、人同士のコミュニケーションと同様、エージェントがジェスチャを行うことにより、発話内容の理解が向上することがわかっている[Rogers 1978]。しかしその一方、システムによるジェスチャの自動生成では、状況に応じた適切なジェスチャをシステムが自動的に決定するのが難しく、さらに研究が必要な課題である。そこで本研究では、適切な場面で適切なジェスチャを表出するエージェントを実現するためのジェスチャの自動付与機構を提案することを目的とする。

[McNeill 1992]ではジェスチャを以下の 4 種類に分類している。

図像:ものの形や大きさを表現する

隠喩:抽象的な概念を表現する

拍子:発話の強調点で、手を上下に動かす動作

指示:指差し動作

[Lücking et al. 2012]では、この分類の中で、図像のジェスチャに着目し、音声発話と共に起したジェスチャをエージェントアニメーションにより表現するシステムを開発している。まず、仮想空間での街において道案内をする対話データを収集し、物体の形や空間的な関係を伝える際に使用された様々なジェスチャを収集している。次に、収集したジェスチャをその形態の観点から分類し、物体を空間に置いたり並べたりする placing, 物体の形を手で表現する Shaping, 形を軌跡で描く Drawing, 距離や大きさを示す Sizing 等、8 種類の形態に分類している。さらにこれらのジェスチャを付与すべきか否かを決定するルールを提案している。

そこで本研究では、[Lücking et al. 2012]では取り上げられていなかった隠喩的ジェスチャに焦点を当て、ジェスチャ自動生成の実現を目指す。具体的には、テキストを入力すると、隠喩的

ジェスチャを共起させるべき単語を同定し、ジェスチャタグをテキストに付与する機構を実装する。さらに、ジェスチャタグ付きテキストを入力するとアニメーションエージェントによる、音声とジェスチャの同期したコンテンツが出力されるシステムを実現する。本研究と同様、隠喩的ジェスチャに着目した研究として、[塩入ら 2014]では、講義映像を観察し、そこで発生する隠喩的ジェスチャについて分析をし、内容により頻出するジェスチャが異なること、品詞や統語的關係により、ジェスチャは共起頻度が異なることを報告している。この研究は、ジェスチャ発生が言語的情報によって変化することを示唆しており、本研究でも言語的情報からジェスチャを規定する方法を検討した。

### 2. 対話データの収集とジェスチャの注釈付け

対話データとして、インタビュー番組から 10 人分のインタビュー、計 98 分 21 秒の対話データを収集し、その中で観察された 254 のジェスチャを分析対象とする。

ビデオ分析ツール ELAN を使用して対話データに注釈付けを行った。まず、[McNeill 1992]による図像・隠喩・指示・拍子の 4 種類のジェスチャ分類について注釈付けを行った。次に、隠喩的ジェスチャのみに着目し、それらが発生した発話の区間と共に起している単語をさらに注釈として対話コーパスに付与した。

図 1 に注釈の一例を示す。この例では、「学校から帰って」の、「帰って」という言葉のところで、ジェスチャが行われている。「帰



図 1. ビデオ分析ツールによる注釈づけ

って」という言葉は抽象的な概念なので、これは隠喩的ジェスチャに分類される。

### 3. 隠喩的ジェスチャの分類

収集したジェスチャのうち、隠喩的ジェスチャに分類されたもののみを対象とし、その意味や形態に関して観察を行った。その



(a) 「行く」 (b) 「そのとき」

図 2. 隠喩的ジェスチャの例

例を図 2(a), 図 2(b)に示す. 図 2(a)は「行く」という単語と共起して表出されたジェスチャである. 「行く」という概念は具体的な形を持つものではないので, これは抽象的な概念を表す隠喩的ジェスチャであり, 腕の動きにより場所を移動している様子が表現されている. また, 図 2(b)は「そのとき」という言葉と共起していたジェスチャであるが, これも「とき」は具体的な形をもつ概念ではないので, 隠喩的ジェスチャである. 左手を用いて, ある時間の 1 点を示していると解釈できる.

観察の結果, 収集した隠喩的ジェスチャを表 1 に示す 12 種類に分類した. 「明日」「1 年間」などの時間を表すジェスチャは, さらに, 一定の時間間隔を表す「期間」のジェスチャと, 時間軸上の 1 点を表す「時点」のジェスチャに分類された. 行動は「行く」「帰る」など, 人や物などがとる行動のジェスチャを表す. 物は「ガス」などの形のないものを表す. 考えは「嫌だ」などの人の気持ちなどについてのジェスチャを表す. 関係は「年齢が上」などの上下などの見えない位置関係などを表す. 様子は「ちよつと」などの程度や, 「グーつと」などの擬音をジェスチャで表す. 場所は「うち」などの細かく形を定めることが難しいジェスチャを表す. そしてそれらの概念を 12 種類に分類した.

表 1. 隠喩ジェスチャの細分類

対象概念	小分類	説明
時間	期間	「1 年間」などの一定の時間
	時点	「明日」などの時間の一時点
行動	移動	「行く」「帰る」などの移動
	その他	その他の行動を表す概念
物	物	形の定まらない物の概念
考え	否定	「いいです」などの否定
	その他	その他の考え
関係	関係	物と物の関係性
様子	程度	「ちよつと」「いっぱい」などの物事の程度
	擬音	「くらつと」「グーつと」などの擬音・擬態
	その他	程度と擬音・擬態に該当しないもの
場所	場所	形で表すことのできない場所

#### 4. ジェスチャ付与ルール

ジェスチャが発生した単語について, シソーラスを使って上位語を調べた. シソーラスとは言葉を同意語や意味上の類似関係, 抱合関係などによって分類した辞書, あるいはデータベースのことである[小林 2010]. 上位概念が共通である単語は, 類似し

た概念を表すとみなせ, 従って, 同じジェスチャを付与できると考えた. これにより, 単語ごとにジェスチャの形態を登録するのではなく, より汎用的なジェスチャルールが設定可能になる. 図 3 に示すように, 日本語 Wordnet を使用して分析を行った結果, 「年間」と「過去」は共通の上位概念として **period** を持ち, どちらの単語にも「期間」のジェスチャを付与可能であることが確認できた.

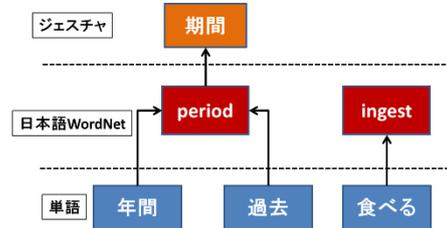


図 3. ジェスチャ付与の例

同様の分析を行い, 表 1 のジェスチャの分類の中で, 「否定」「移動」「期間」「時点」の 4 種類のジェスチャについて, それらを付与すべき単語を決定するルールを作成した. 表 2 にジェスチャ付与ルールを示す. 例えば, 上位概念が **dislike, interrupt** である単語には「否定」のジェスチャルールが適用し, 前で小さく手を振るジェスチャを付与する. しかし, 「食べる」という単語の上位概念を分析した場合, 上位概念は図 3 に示すように **ingest** となり, 表 3 のいずれのジェスチャルールにも当てはまらないので, ジェスチャは付与しない.

表 2. ジェスチャ付与ルール

小分類	上位概念	ジェスチャの形態
否定	dislike, interrupt	前で小さく手を振る
移動	go	両手で物を移動させるような動作
期間	period, time unit	両手で長さを示すような動作
時点	point	手を置くような動作

#### 5. ジェスチャ自動付与機構の実装

4 章で提案したジェスチャ付与ルールを実装することにより, テキストを入力すると, 発話音声と同期してエージェントがジェスチャを行うアニメーションコンテンツが自動生成されるシステムを実現する.

##### 5.1 ジェスチャ決定モジュール

ジェスチャ決定モジュールでは発話する文中のどの単語でジェスチャを発生させるかを決定する. まず, 入力文を形態素解析し, 単語に分割する. 本システムでは, 形態素解析器 Mecab を使用した. 次に, シソーラスを用いて, 分割した各単語の上位概念を調べ, ジェスチャ付与ルールにマッチすれば, 該当するジェスチャタグを XML 形式で付与する. 例えば, 「今日私は京都に行きます。」という入力文があった場合, 「行き」に移動のジェスチャルールがマッチし, go の属性値を持つ **gesture** タグが付与され, 図 4 に示す XML 形式のファイルが出力される.

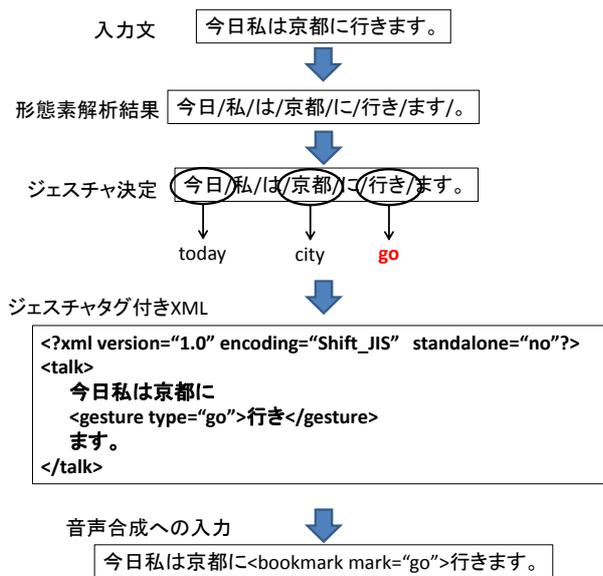


図 4. 処理の流れ

## 5.2 音声とアニメーションの出力

音声合成モジュールでは、発話をするために入力した文章を発音記号に変換する。アニメーションエージェントのリップシンクは、Microsoft Speech API(SAPI)に準拠した音声合成エンジンから出力される、発音に応じた口の形を表す visimeID と呼ばれる発話記号をモジュール間通信プラットフォーム OpenAirServer を通じてアニメーションの描画を行うゲームエンジン Unity に送信することにより実現している。また、ジェスチャ決定モジュールで決定されたジェスチャに応じてジェスチャ ID を Unity に送信することによりエージェントのアニメーション命令が実行される。ジェスチャを発生させたい単語の前に bookmark タグを付与して SAPI に送ることにより、指定された単語の音声出力される時間がフィードバックされ、それをジェスチャ命令のトリガとすることにより、音声とジェスチャを同期させることができる。

## 6. 実行例

「私は京都に行きます。」という文章で例を示す。処理の流れを図 4 に示す。まず、発話させたい文章「私は京都に行きます。」という文章を形態素解析で単語に分割する。分割した各単語に対し、シソーラスで上位概念を調べ、表 2 のジェスチャ付与ルールに示す上位概念に当てはまった場合は、その単語に gesture タグを付与し、XML 形式で出力する。この例では、図 4 に示すように、「gesture type=go」のタグが付与された XML が生成される。

次に、入力文が音声合成器に送信されるが、その際、XML 中でジェスチャタグが付与されている単語の直前に bookmark タグを挿入しておく。これにより、音声出力が bookmark の位置に到達したときに、該当するジェスチャ、この例では「go」ジェスチャの実行命令がアニメーション描画部である Unity に送信され、アニメーションエージェントによるジェスチャが出力される。「行く」のアニメーションが実行されている様子を図 5 に示す。



図 5. アニメーションエージェントの実装

## 7. おわりに

対話データの観察をすることでジェスチャの発生について分析し、4 種類の隠喩的ジェスチャに関するジェスチャ付与ルールを提案した。さらに、テキストを入力すると、アニメーションエージェントにジェスチャを自動付与するシステムを実装した。発話とジェスチャを共起させることによって、アニメーションエージェントのコミュニケーションをより自然なものにする手助けになると考える。今後の課題として、ジェスチャ付与ルールを拡充するとともに、様々なジェスチャを実装したい。また、必要のない部分でジェスチャが発生してしまうことがあるので、ルールの改良も必要である。

## 謝辞

第 5 章で用いた OpenAirServer は、本学知能インタフェース研究室の高坂氏の制作によるものです。氏のご尽力によりアニメーションエージェントへの隠喩的ジェスチャの付与が行えました。深く感謝いたします。

## 参考文献

- [Rogers 1978] Rogers, W., The Contribution of Kinesic Illustrators towards the Comprehension of Verbal Behavior within Utterances. *Human Communication Research*, 5: pp. 54-62, 1978.
- [McNeill 1992] McNeill, D.: Hand and mind: what gestures reveal about thought, 1992.
- [Lücking et al. 2012] Andy Lücking, Kirsten Bergman, Florian Hahn, Stefan Kopp, Hannes Rieser: Data-based analysis of speech and gesture: the Bielefeld Speech and Gesture Alignment corpus (SaGA) and its applications, *Journal of Multimodal User Interfaces*, Springer, DOI 10.1007/s12193-012-0106-8, 2012.
- [小林 2010] 小林 暁雄, 増山 繁, 関根 聡: Wikipedia と汎用シソーラスを用いた汎用オントロジー構築手法, 電子情報通信学会論文誌. D, J93-D(12), 2597-2609, 2010.
- [塩入 2014] 塩入直哉, 塙俊樹, 長谷川大, 白川真一, 佐久田博司, 大原剛三: 講義映像における暗喩的ジェスチャーと品詞及び統語的情報の関連性, 第 76 回情報処理学会全国大会, 2014.