

# 言語・空間情報に基づくメタバースアバターへの ジェスチャ自動付与機構

塚本 剛生<sup>†</sup> 中野 有紀子<sup>‡</sup>

成蹊大学大学院理工学研究科理工学専攻<sup>†</sup> 成蹊大学工学部情報科学科<sup>‡</sup>

## 1. 背景

セカンドライフ(SL)に代表されるメタバースアプリケーション上で相手とコミュニケーションをする主な方法は、テキスト入力によるチャットであり、対面場面での音声や身振りによる表現は、非常に限定的にしか使うことができない。一方、対面コミュニケーションの研究では、ジェスチャ等の非言語行動は言語行動を補う重要な情報であることがわかっている。特に、道案内を行う対話では言語表現を補足するために、空間的なジェスチャが豊富に用いられるのが特徴的である。

そこで本稿では、入力テキストに応じて仮想空間のアバターにジェスチャを自動付与し、音声合成と同期したジェスチャアニメーションを生成する機構を提案する。[Tsukamoto, 11]では、道案内対話の収集実験と収集されたジェスチャデータの分析を行い、会話参加者の立ち位置によって、ジェスチャの形態が異なることを示した。本稿では、さらに道案内において用いられる言語表現とジェスチャ形態との関連性を分析し、言語情報がジェスチャ形態の決定に有用であることを示す。また、これらの分析結果を基に、SL上の空間的情報とチャットテキストとして入力された道案内の言語的情報を統合し、ジェスチャの形態を決定する機構を提案するとともに、SL上に道案内アバターシステムを実装する。

## 2. 言語表現とジェスチャ形態との関連性

### 2.1 ジェスチャコンセプトの分類

まず、会話収集実験で得られた被験者発話の書きおこしデータから、ジェスチャが共起している言語表現を取り出し、それらを以下の5種類に分類した。

- (i) 方向：方向についての言及。方向を表す単語(右, 左等)+助詞により表現される。
- (ii) 相対的な位置関係：2つの対象物の相対的な位置関係への言及。助詞「の」/「その, そのの」+方向を表す単語(右, 左, 奥, 後ろ等)により表現される。
- (iii) 動作：動作, および動作の方向についての言及。方向を表す単語(右, 左, まっすぐ等)+助詞の文節が動作を示す動詞にかかっている場合と、動作を示す動詞のみで表現される場合がある。
- (iv) ランドマーク：ランドマークとなる対象物についての言及。名詞+「がある, が見える」の表現形式をとり、ランドマークを表す名詞の初出現時にのみジェスチャを伴う場合が多い。
- (v) 建物の特徴：ランドマークとなる建物の特徴に関する言及。

観測された右腕のジェスチャ 229 個のうち、44 個は beat ジェスチャと呼ばれる手の上下動のみのジェスチャ

Gesture Generation for Metaverse Avatars using Linguistic and Spatial Information

Takeo Tsukamoto<sup>†</sup>, Yukiko Nakano<sup>‡</sup>

<sup>†</sup>Graduate School of Science and Technology, Seikei University

<sup>‡</sup>Faculty of Science and Technology, Seikei University

であり、概念によって形態が異なるものではないため、今回の分析からは除外した。残りの 185 個のジェスチャのうち、(i)から(iv)の4種類のカテゴリで、全体の95%を占めることから、道案内において特徴的に表れるジェスチャの大部分は、これら4種類の概念のいずれかであると言える。そこで、これらのカテゴリをジェスチャコンセプトと呼び、これらとジェスチャ形態との関係をさらに詳細に分析する。

### 2.2 ジェスチャコンセプトとジェスチャ形態との関係

次に、収集したジェスチャのビデオデータを注意深く観察し、ジェスチャの形態的特徴を記述するために以下の4種類のジェスチャ形態パラメータを設定した。

- (a) 手の移動方向：ジェスチャのストロークにおける右腕の移動方向や軌跡
- (b) 手のひらの向き：ストローク時の右手のひらの向き
- (c) 手の形態：ジェスチャストローク時の手の開き具合
- (d) 反復：1つのジェスチャ中に、同じ動きの反復が含まれるか否か

前節で提案した4種類のジェスチャコンセプトを表現するジェスチャをそれぞれ分析し、“相対的な位置関係”を表現するジェスチャを、上記4種類のパラメータの組み合わせにより、15種類のジェスチャ形態に分類することができた。同様に、“方向”のジェスチャは10種類、“動作”のジェスチャは13種類、“ランドマーク”のジェスチャは19種類のジェスチャ形態に分類された。

そこで、各ジェスチャコンセプトにおいて、どのような要因がジェスチャ形態パラメータの決定に影響しているかを調べるために、ジェスチャ形態パラメータと言語表現との関係を詳細に分析した。その一例として“相対的な位置関係”のジェスチャコンセプトにおけるジェスチャパラメータの分布を表1に示す。このジェスチャコンセプトの場合、方向を表す単語が「左」や「右」

「下」を意味する場合、必ず手の移動方向はそれと一致し、方向を表す単語が「中」の場合、手の移動方向は上方向となる。方向を表す単語が「隣」か「横」の場合、手の移動方向は左か右のどちらかの方向となり、方向を表す単語が「裏」「後ろ」等の場合は、手の移動方向は前方向となる。また、方向を表す単語が「右」である場合のみ手のひらの向きが「右」となる場合があるが、それ以外の場合は、手のひらを右に向けることはない。また、手の開き具合についても、指差し型の手の形をとるときは、手のひらの向きは必ず下向きであった。最後に、反復については、多くの場合、反復はされないことが分かった。その他のジェスチャコンセプトについても分布を分析し、同様の結果を得ることができた。

以上、我々の収集したデータにおいて、50種類以上のジェスチャ形態が観察され、4種類のジェスチャ形態パラメータの組み合わせによりこれらを定義可能であることを示した。また、方向を表す単語やパラメータ間の依

表 1: ジェスチャパラメータの分布 (相対的な位置関係)

手の移動方向						手のひらの向き		手の形態		反復
方向を表す単語: 右	方向を表す単語: 左	方向を表す単語: 下	方向を表す単語: 上	方向を表す単語: 隣、横	方向を表す単語: その他	方向を表す単語: 右	方向を表す単語: その他	手のひらの向き: 下	手のひらの向き: その他	無(90%), 有(10%)
右(92%), 前(8%)	左(100%)	下(100%)	上(100%)	左(50%), 右(50%)	前(100%)	右(40%), 左(20%), 下(20%), 前(20%)	左(63%), 右(0%), 下(24%), 前(13%)	指差し型(18%), 半開き(18%), 全開(64%)	半開き(13%), 全開(87%), 指差し型(0%)	

関係により、ジェスチャ形態パラメータの値を明確に限定できる部分を見出すことができた。

### 3. ジェスチャ形態決定手法

本章では、前章の分析結果と[Tsukamoto, 11]を統合し、入力された道案内発話の言語情報と SL 中の空間情報からジェスチャ形態を決定する方式を提案する。ジェスチャ決定部では、以下の処理によりジェスチャを付与する。

(1) ジェスチャコンセプトの付与: まず、チャット画面から入力されたテキストを形態素解析 JUMAN, 構文解析 KNP による言語タグ付与機構に送り、形態素情報・係り受け情報を付与する。次に、これらの言語情報に基づき、ジェスチャコンセプトが付与可能な文節には、ジェスチャコンセプトを付与する。

(2) 立ち位置タイプの判定: 次に、アバタの立ち位置タイプを判定する。[Tsukamoto, 11]では、案内者役と訪問者役それぞれの前方の空間と訪問者役の視野の3つが重なる領域をジェスチャ表示空間と定義し、このジェスチャ表示空間の中心と各被験者との距離に基づき、立ち位置を4種類に分類し、立ち位置によってジェスチャの分布範囲が異なることを示した(図1)。この知見を利用し、SLより得られたジェスチャ表示空間の中心の座標から訪問者ユーザのアバタ(DRアバタ)、および案内者ユーザのアバタ(DGアバタ)との距離に応じて4種類の立ち位置タイプを判定する。

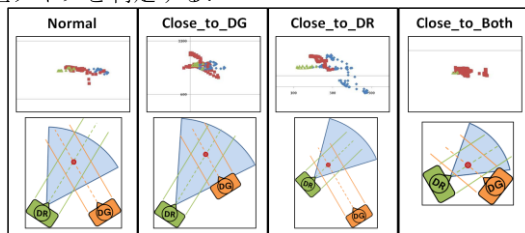


図 1: 立ち位置(下)とジェスチャ分布範囲(上)

(3) ジェスチャ形態の決定: 次に、アバタの立ち位置タイプを考慮しながら、各ジェスチャコンセプトを表現するためのジェスチャ形態パラメータの値を決定する。各ジェスチャコンセプトについて作成したジェスチャパラメータの分布表(例、表1)に基づき、各事象の発生確率に従って、ジェスチャ形態パラメータ値を決定した。

尚、ランドマークについては、概念によっては、四角や円形等、複雑な軌跡をとるため、ランドマークごとに

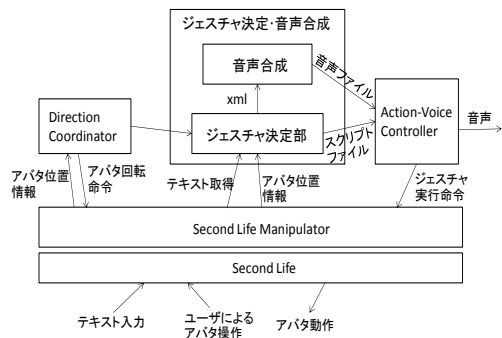


図 2: 道案内アバタシステムの構成

個別にジェスチャ形態パラメータセットを用意した。

### 4. システム実装

前章で提案したジェスチャ形態決定手法を組み込み、SL上に道案内アバタシステムを実装した。システム構成を図2に示す。尚、プラットフォームには SL のオープンソース版である OpenSim を使用している。

(1) Second Life Manipulator (SLM): SL ではプリムと呼ばれるオブジェクトに、SL 独自のスクリプト言語であるリンデンスクリプト言語(LSL)を組み込み、それをアバタに装着させることでアバタの制御ができる。SLMはこの機能を用いて、アバタの制御、チャットテキストの取得、位置情報の取得を行う。

(2) Direction Coordinator: Direction Coordinator では、SLM から得られた DG アバタと DR アバタ、および目的地の座標情報と DG アバタと DR アバタの体の向きを用いて、それぞれのアバタを回転させる角度を計算する。ここで計算された角度を用いてアバタ回転の命令を SLM 経由で SL に送信することにより、両アバタが自動的に目的地方向へ回転する。

(3) ジェスチャ決定・音声合成: ジェスチャ決定部は CAST[Nakano, 04]を拡張することにより実装した。CASTは、テキストを入力として受け取ると、日本語形態素解析と係り受け解析により得られた言語的特徴からジェスチャを付与すべきタイミングを決定し、アニメーションのタイムスケジュールを自動生成するツールである。本研究では、3章で提案した手法を CAST に実装することにより、道案内で必要となる特定の形態を持つ空間的ジェスチャを生成できるシステムを実装した。

(4) Action-Voice Controller (AVC): AVC では作成した日本語合成音声を再生すると同時に、SL中の DG アバタに対し、スクリプトファイルに示されたタイムスケジュールに従ってジェスチャ実行命令を送信する。

以上の機構により、SL上で道案内の文章を入力すると、日本語音声合成が再生されると同時に、その文章に適した道案内のジェスチャが自動的に案内役アバタ(DGアバタ)に付与され、アニメーションとして出力される。

### 5. まとめ

本研究では、メタバースにおける道案内アバタを実現するため、メタバースの空間情報とチャットテキストの言語情報から、道案内アバタによる空間的なジェスチャのタイミングと形態を決定する機構を提案した。本稿で提案したジェスチャ決定手法の汎用性の検証や、生成されたジェスチャの妥当性の評価については今後の課題として取り組む予定である。

#### <参考文献>

[Tsukamoto, 11]Tsukamoto.T et al.," Collection and Analysis of Multimodal Interaction in Direction Giving Dialogues: Towards an Automatic Gesture Selection Mechanism for Metaverse Avatars" Proc. of the Workshop on AAMAS '11(2011)  
 [Nakano, 04]Nakano.Y et al.," Converting Text into Agent Animations: Assigning Gestures to Text", HLT-NAACL2004, Companion Volume, Boston (2004)