

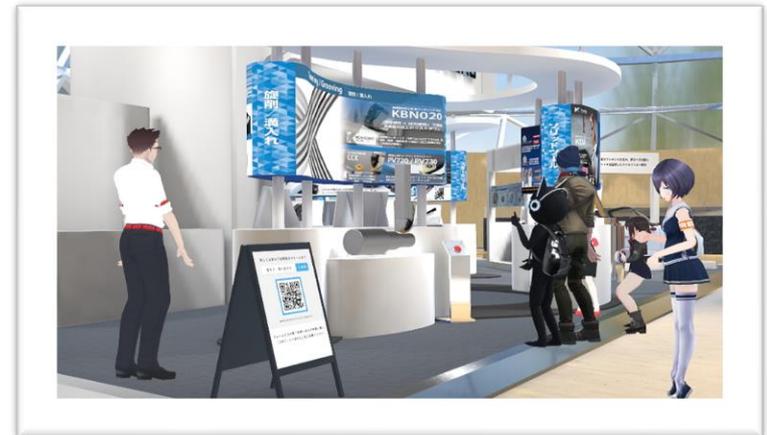
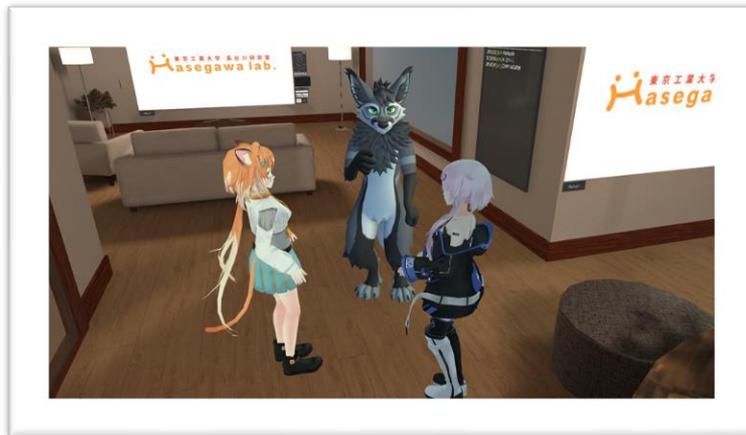
ソーシャルVRにおいて多人数会話に参加する  
会話エージェントのためのLSTMを用いた  
自然な身体動作の生成

2B1-02

加藤圭悟, 長谷川晶一 (東京工業大学)

# ソーシャルVRの利用シーン

- ユーザ同士の交流の場
- 求められるサービスの多様化 (Ex.展示会で, グループごとに質問対応できる存在)
  - サービス提供を行える人材が必要
  - コスト面から, 安易な人材追加は難しい



- サービス提供が可能なNPCの存在が注目されている
- ユーザの行動を理解した応答による会話の促進 [3]
  - 動作の出力モデル

身体をもつ会話エージェント(ECA) [2]



[2] Advanced Media, Inc : <https://www.advanced-media.co.jp/newsrelease/23695>

[3] M.Iwasaki et al. 2019. "That Robot Stared Back at Me!": Demonstrating Perceptual Ability Is Key to Successful Human-Robot Interactions

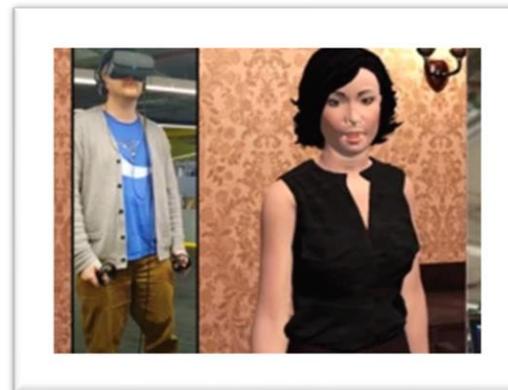
# 動作の出力モデル

- 機械学習モデル 
  - 自然の反応のための細かいルールを学習により機械が作成

動作(連続)を注視方向(離散)に変換[4]



動作(連続)をエンゲージメント  
レベル(離散)に変換[5]



離散値として捉えているため,決まった動作の検出・生成に限定

[4] S.Zou et al. 2017. Conversational Agent Learning Natural Gaze and Motion of Multi-Party Conversation from Example

[5] G.Cristina Dobre et al. 2022. Immersive machine learning for social attitude detection in virtual reality narrative games

- 研究の目的
  - VRの多人数会話において、動作を連続値として捉えて多様な動作を検出・生成できる人間らしいECAの実現

多人数会話 [6]



## 研究対象

- ① 三者会話
- ② 傍参与者 [7]

[6] CNET Japan : <https://japan.cnet.com/article/35173551/>

[7] Goffman E. 1981. Forms of talk

# 提案モデルの作成方法

- 作成の流れ

Step1: 会話データの収集

Step2: ラベリング

Step3: LSTMによる学習

Step4: 動的な動作決定



# Step4: 動的な動作決定

- 実現された動作
  - 学習データにある注視の仕方,細かな動作(頭の微小な揺れ)を実現
  - 提案モデルの動作は小さく,ノイズがある

● 生データ



● 提案した連続モデル



● 従来 of 離散モデル



※HMMによる離散的な注視方向の生成

.....

※注視方向が変化する動作は確認できたものの、これは稀にのみ発生して、かつ大概は小さい動作

# 動作が小さくノイズがある原因

- 動作の小ささ
  - R2が上がりきっていない、中間の予測値ばかり取るため
  - 対策: より多くのデータを収集
    - 本研究: 約120万フレーム
    - 同手法で動作する現実の二者会話の好印象なECA: 約62万フレーム [8]
- 動作のノイズ
  - 予測値にノイズがあるため
  - 対策: **細かな動作表現を損なわない程度のフィルターの適用**

	3分割交差検証
平均R2(訓練)	0.634
平均R2(検証)	0.553

※この計算は間違えている。上記より低い値となる。

.....