

# 講演者の視線と発話に基づく積極的傾聴キャラクターを用いた 没入型講演体験システムの開発

石坂 諒任<sup>†</sup> 渡辺 富夫<sup>‡</sup> 石井 裕<sup>‡</sup>

岡山県立大学大学院 情報系工学研究科<sup>†</sup> 岡山県立大学 情報工学部<sup>‡</sup>

## 1. はじめに

対面コミュニケーションでは、話し手の語りかけに対し、聞き手がうなずくなどの反応動作が行われ、両者の身体的リズムが共有されることで円滑なコミュニケーションが行われる。これまでに著者らは、発話音声からうなずきや身振り・手振りなどの身体引き込み動作を自動生成する iRT を開発している<sup>[1]</sup>。また、これを用いた音声駆動型身体引き込み CG キャラクター InterActor を教師・学生キャラクターとして仮想教室内に複数体設置した身体引き込み集団コミュニケーションシステムとして SAKURA<sup>[2]</sup>を開発している。しかし、これらのシステムには話者の視線情報に起因するインタラクションの生成は考慮されていなかった。

そこで本研究では集団での対面コミュニケーションの一つである講演に着目し、講演者の視線移動に対して観客としての反応を行う InterActor を用いることで、より臨場感のある講演を体験できる没入型講演体験システムを開発した<sup>[3]</sup>。本稿では、講演者の発話情報と視線情報を併用した観客キャラクターの傾聴動作制御を行うことで、より質の高い講演体験ができるシステムを開発している。

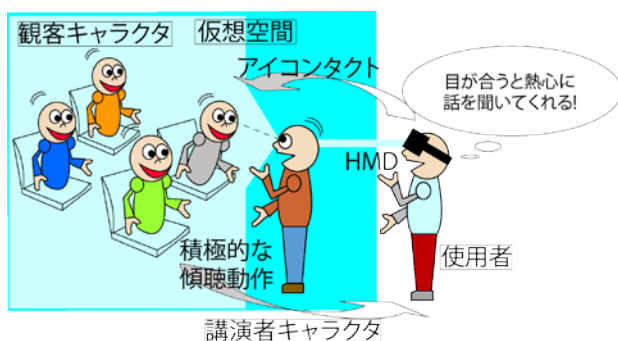


図1 コンセプト

Development of an Immersive Presentation Experience System Using Active Listening Audience Characters Based on Lecturer's Gaze and Utterance

<sup>†</sup>Ryoto Ishizaka · Graduate School of Systems Engineering, Okayama Prefectural University

<sup>‡</sup>Tomio Watanabe, Yutaka Ishii · Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

## 2. システム概要

本システムのコンセプトを図1に示す。講演やプレゼンテーションにおいて顔や視線を相手に向けることは大変重要である。そこで、講演者が観客と視線が合うことにより、観客が講演者の話により関心を持って聞く態度を示すことで講演者が視線への意識を強め、円滑なコミュニケーションを実現させる。また、使用者の音声と視線から得られたデータにより、観客キャラクターが深いうなずきを行うようになり、前傾姿勢になるなどの、積極的な傾聴動作を行う。これにより、観客への視線を意識した講演を行うことができる。

## 3. システム構成

本システムにおいて、HMDはOculusRiftを用いており、マイクとコントローラにより音声入力やボタン入力を行う。また、仮想空間の構築・システムの制御には3DゲームエンジンUnity、キャラクターの制作には3DモデリングソフトBlenderを用いている(図2)。

## 4. 観客の聞き手反応

コミュニケーションにおいて、視線は重要な意味を持ち、意思表示・感情表現・関心表明などの機能を持つことが明らかになっている<sup>[4]</sup>。そこで、本システムでも使用者の視線を用いることでインタラクションを生成する機能の開発を行っている。観客キャラクターの傾聴動作は、通常傾聴状態と積極的傾聴状態の2つにより構

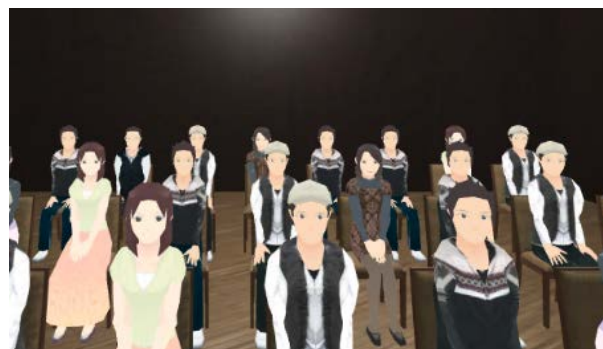


図2 没入型講演体験システム

成されており (図 3), まず通常傾聴状態の観客キャラクタは, 使用者に視線を向け, 発話に対して iRT を用いて自動生成されたうなずき反応を行う. それに対し, 積極的傾聴状態の観客キャラクタは, 通常傾聴状態の反応に加え, 上体を使用者の方へ傾け, 通常より深いうなずき反応を提示する. 観客キャラクタは視線情報と音声情報からこれらの状態遷移を行う. 本システムでは, 使用する HMD(OculusRift)に内蔵されている慣性計測装置や赤外線センサから使用者の頭部の角度情報を取得し, これを使用者の視線情報としてシステムに反映している. 観客キャラクタの頭部にそれぞれ球形の判定範囲を配置し, 講演者キャラクタの頭部から前方方向に球形の判定範囲をもった透明なオブジェクトを常に発射しており, それが観客キャラクタの判定範囲と接触した際に観客キャラクタとの視線一致の判定を行う.

5. 視線と発話による制御

従来のシステムでは, 使用者が観客キャラクタと視線が一致したときに観客キャラクタは通常傾聴状態から積極的傾聴状態に遷移し, 視線が外れた際には即座に通常傾聴状態に戻るという制御であった<sup>[3]</sup>. 実際の講演場面では視線が外れたとしても観客が即座に姿勢を変えることはないため, 現実の講演場面との違和感が生じていた. 本システムでは, 積極的傾聴状態から通常傾聴状態に遷移する際に, 視線情報だけで



(a) 通常傾聴状態 (b) 積極的傾聴状態  
図 3 傾聴状態

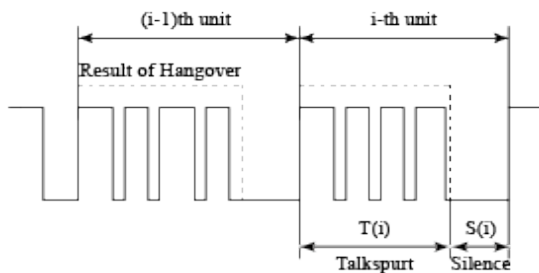


図 4 発話音声とユニット区間

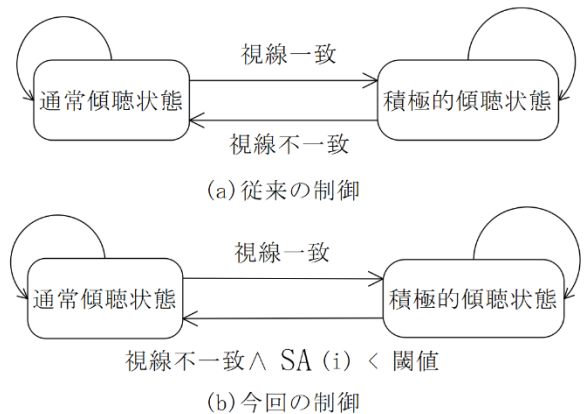


図 5 観客キャラクタの状態遷移図

ではなく式(1)のユニット区間  $(i - 1)$  から  $(i - K)$  における発話時間率  $SA(i)$  (Speech Activity) を因子として採用した (図 4). 発話時間率は, 一般的に発話速度と相関が高いことが明らかになっている<sup>[5]</sup>. よって視線が外れており, かつ発話時間率がある一定の閾値を下回った際には, 話が落ち着いたとみなし, 通常傾聴状態への遷移を行うことで, 実際の講演場面に即した制御を可能とした (図 5).

$$SA(i) = \frac{\sum_{k=1}^K T(i-k)}{\sum_{k=1}^K \{T(i-k) + S(i-k)\}} \quad (1)$$

$K$ : 予測次数 (6)

$T(i)$ :  $i$  番目ユニットにおける ON 区間長

$S(i)$ :  $i$  番目ユニットにおける OFF 区間長

6. おわりに

本研究では, 音声情報と視線情報から積極的傾聴と通常傾聴の状態遷移を決定する身体引き込みキャラクタを用いることで, より質の高い講演・プレゼンテーションの体験ができる没入型講演体験システムを開発した.

参考文献

[1] Watanabe, T.Okubo, M., Nakashige, M. and Danbara, R.: InterActor : SpeechDriven Embodied Interactive Actor; International Journal of Human-Computer Interaction, Vol.17, No.1 (2006), pp.43-60.  
 [2] 小畑淳, 渡辺富夫, 大久保雅史: 音声駆動型身体引き込みコミュニケーションシステム"SAKURA", ヒューマンインタフェースシンポジウム2000論文集, (2000) pp.331-334.  
 [3] 田中 一也, 渡辺 富夫, 石井 裕, “音声駆動型身体引き込み観客キャラクタを用いた没入できる講演体験システムの開発”, ヒューマンインタフェースシンポジウム2016 DVD論文集, pp.187-190(2016).  
 [4] 黒川隆夫, 電子情報通信学会 編(1994)「ノンバーバルインタフェース」pp. 52, 原島 博 監修, オーム社.  
 [5] Tomio Watanebe, "The Adaptation of Machine Conversational Speed to Speaker Utterance Speed in Human-Machine Communication", IEEE Trans of Systems, Man, Cybernetics. SMC-20-2, 502/507(1990).