

令和元年6月25日現在

機関番号：12608

研究種目：若手研究(B)

研究期間：2017～2018

課題番号：17K17713

研究課題名(和文) 会話事例からの動作制御モデル自動獲得による自然な会話動作の再現

研究課題名(英文) Natural Conversation Motion Generation by Control Model Acquisition from Examples

研究代表者

三武 裕玄 (Mitake, Hironori)

東京工業大学・科学技術創成研究院・助教

研究者番号：30613939

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：物理エンジンを用いた身体制御モデルにより会話動作を生成する手法を実現すると同時に、会話動作生成モデルを逆に用いることで実会話動作を分析し、複雑な会話動作の生成モデルを獲得する手法の実現を目指した。身体の動力学を再現し自然な動作を生成する手法として、物理エンジンを用いて近似的なトルク変化最小軌道を生成する簡便な動作生成手法を実現した。実会話動作を基に動作生成手法の振る舞いを人力で調整するための開発環境の実現も行った。表情や抑揚について会話データを基にした学習手法の実現も行った。また、見る位置によって視線の向きが異なって見え、特定方向にのみアイコンタクトを生じさせる安価なディスプレイを実現した。

研究成果の学術的意義や社会的意義

物理エンジンによる簡便なトルク変化最小軌道の生成手法は、計算の不安定性等の課題はあるが、力学的に理にかなった動作を自動的に行うキャラクタ動作生成手法の基盤に将来なりうる手法である事が確認できた。また、物理エンジンを利用した動作デザイン環境を構築したことで、会話動作データを利用しつつ、キャラクタらしい動作調整を行って、自然な身体動作を持つ実用的な対話エージェントを実現するためのソフトウェア環境を実現することができた。加えて、選択的な視線提示を可能にするディスプレイは安価な平面ディスプレイで従来不可能だった正しいアイコンタクトを実現するもので今後の普及が見込まれる。

研究成果の概要(英文)：We aimed to realize motion generation for embodied conversational agents utilizing physics engine for natural synthesis of motion.

Firstly, we focused on "Forward Inverse Relaxation Model (FIRM)" of human hand reaching motion. We created physics engine based FIRM computation method, which is easy to use and generate smooth and natural motion. Also, we created motion design environment for creating characteristic behaviors of virtual agents utilizing physics engine based motion generation. On the other hand, facial expressions or voice pitch cannot be generated with physical simulation. Therefore, we also researched on HMM based machine learning method of conversation model based on real conversation examples.

Moreover, we also created new display for appropriate gaze expression of agents. With 2 layered LCD, we realized low-cost and high resolution display which can show eye-contact only for limited view position.

研究分野：バーチャルクリーチャ、キャラクタ動作生成、対話エージェント

キーワード：対話エージェント 非言語コミュニケーション 物理エンジン 動作生成 視線提示 モナリザ効果  
ディスプレイ

## 様式 C-19、F-19-1、Z-19、CK-19 (共通)

### 1. 研究開始当初の背景

自然な会話による UI は、それと意識させずに生活に溶け込む情報環境の理想像の一つとされると同時に、会話の楽しさを人工物で再現し楽しみを得る目的でも実現が試みられている。情報ネットワークの発達とともに、情報環境が人同士をつなげ社会を活性化する役割を担うようになった今、情報環境の端末となるロボットや会話エージェントもまた、人々の交流の場に溶け込み、コミュニティを活性化することが期待されている。

しかし、会話エージェントが多人数に溶け込んで自然な会話を実現するには、ジェスチャや視線によるターンテイキング等の身体動作など、自然な会話に必要な身体動作が必要となる。身体性を持つ会話エージェントは研究されているが、多くは動作はアニメーションの切替・合成によるもので切替えが不自然、同じ動作の繰返しが目立つ、自在なタイミングで動作できない等の問題がある。身体の動力学と制御を用いて自然な動作を実時間生成する研究もあるが、多くは歩行動作や物体操作が対象で、自然な会話動作は複雑で原理が不明なため身体制御モデルの実現には至っていない。

また、自然な会話の再現には実会話事例の利用が効果的であるが、最も有効な分析手段が人力のアノテーションであるため、扱う特徴を限定して多人数会話を分析・再現する研究が多く、会話制御における複雑な身体動作の相互作用の解明には至っていない。一方で会話事例を直接機械学習しようとする、同種の動作でも話者の位置関係によって変わる・力学の影響で毎回微妙に異なるなど、生の会話データは同種の動作同士でも差異があり、共通性を見いだしにくい。

### 2. 研究の目的

本研究では、物理エンジンを用いた身体制御モデルにより会話動作を生成する手法を実現すると同時に、会話動作生成モデルを逆に用いることで実会話動作を分析し、複雑な会話動作の生成モデルを獲得する手法の実現を目指した。

また、非言語対話において視線は重要な役割をもつが、一般的なディスプレイはモニリザ効果（平面映像の顔をどの方向から見ても自分が注視されていると感じる現象）により視線の表現に大きな制約が生じることから、正確な視線表現を可能にするディスプレイの研究も行った。

### 3. 研究の方法

- (1) 身体の動力学を再現し自然な動作を生成する手法として順逆緩和モデルに着目し、物理エンジンを用いて順逆緩和モデルの計算を行い近似的なトルク変化最小軌道を生成する簡便な動作生成手法を実現した。
- (2) 非言語会話は動作以外にも表情や抑揚の重要度が高いが、これらは動力学モデルでは生成できないため、会話データを基にした学習手法の実現を行った。
- (3) 実用的な対話エージェントを実現するために、機械学習だけでなく、実会話動作を基に動作生成手法の振る舞いを人力で調整するための開発環境の実現を行った。
- (4) 液晶パネルを二層に配置し顔と目を別々に描画することで、見る位置によって視線の向きが異なって見え、特定方向にのみアイコンタクトを生じさせる安価なディスプレイを実現した。

### 4. 研究成果

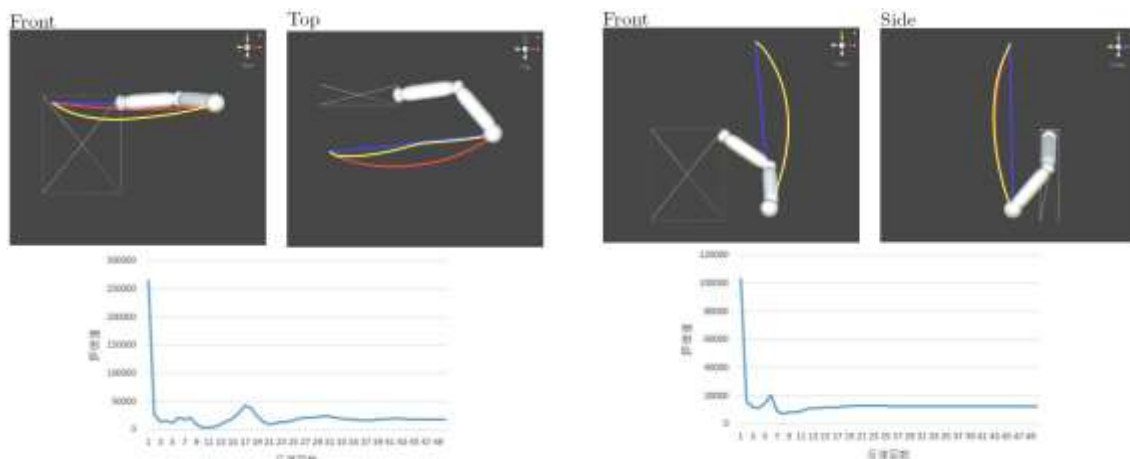


図1 物理エンジンを用いた順逆緩和モデルにより生成された手先軌道の例

(1) 会話動作生成のための手先軌道制御に用いていた躍度最小軌道は計算が簡便であるが軌道が直線的で不自然な場合があった。対して関節の慣性や粘性を考慮するトルク変化最小軌道は人の手先到達軌道が描く曲線を再現できるとされ、比較的簡易な計算手法として身体動力学的順モデルと逆モデルを交互に繰返し計算して動作を得る順逆緩和モデルが知られていた。

従来、順逆緩和モデルを利用するには身体力学モデルの定式化が必要であったが、物理エンジンは動力学的順モデルとしてだけでなく、計算過程で関節トルクが求まるため逆モデルとしても利用できる。この特徴を用いて物理エンジンによる汎用な身体モデルに対する順逆緩和計算を実現した。

結果として、特に移動距離の大きな時に、躍度最小軌道よりも人間の動作により近い滑らかな曲線を描く動作を生成することに成功した(図1)。一方で、いくつかの条件では計算が発散し軌道生成に失敗するなどの課題も残った。

(2) 表情については、特に人同士の会話では笑顔の同調が見られるという観察が知られていることから、連続値HMMにより笑顔同調の推定モデルを学習し、これを用いて同調的的表情を行う対話エージェントの実現を行った。

笑顔の同調は音声にも依存すると考えられるため、話し手・聞き手の音声と、表情認識により得た笑顔の度合いを組み合わせた時系列データを連続値HMMに学習させた。実現したエージェントとのインタラクション実験により、学習した会話事例に含まれるような笑顔の同調パターンがいくつか生成され、体験者も提案手法による笑顔の推移が他の手法よりも人間らしいと感じたことが示唆された。



図2 連続値HMMによる表情変化モデル、及びインタラクションの例

また、会話データの観察から、聞き手の相槌の音量・音高が、話し手の音量・音高に影響を及ぼすことが明らかになった。そこで、エージェントが話し手側となるような対話において、聞き手である人間の傾き音声の音量・音高を入力とし、エージェントの音量・音高の変化を移動平均モデルに基づいて決定する話し手エージェントの実現も行った。

(3) 生活に溶け込み人を楽しませる対話エージェントには、親しみやすく人気のあるキャラクターの外観が用いられることが多い。こうした対話エージェントに見た目に即した個性的で多様な対話動作をさせるには、物理エンジンによる身体動作生成手法に加え、人がキャラクターらしさを与えるための動作デザインを行う必要がある。



図3 手先位置キーフレーム(四角形で図示)と動作の例

物理エンジンによる身体動作生成手法と親和性が高く、一例をデザインしただけで類似する



多数の動作を実現でき、従来の関節角アニメーションと同等の直感的な調整が可能な手法として、手先目標位置ベースのキーフレームアニメーションを主体とする動作データの表現と、そのための動作デザイン環境の実現を行った。

実現した環境により実際に複数のインタラクティブキャラクターを実現した。展示会等で来場者に自然な動作により親しまれ、高い評価を得た。

(4) 従来、モナリザ効果を起こさず特定の相手へのみ被注視感を生じさせられるディスプレイは、コスト・解像度・動作の自由度の少なくとも一つにおいて問題のあるものであった。そこで、液晶ディスプレイ2枚を5mm程度の間隔をあけて層状に重ねるという安価に実現できる機器を用い、奥の層に黒目を除いた顔画像、手前の層に黒目を描画することにより、視点位置により白目と黒目の相対位置関係が変化して見えることで特定の方向から見た場合にのみ黒目が白目の中心に見える(=被注視感が生じる)手法を実現した。

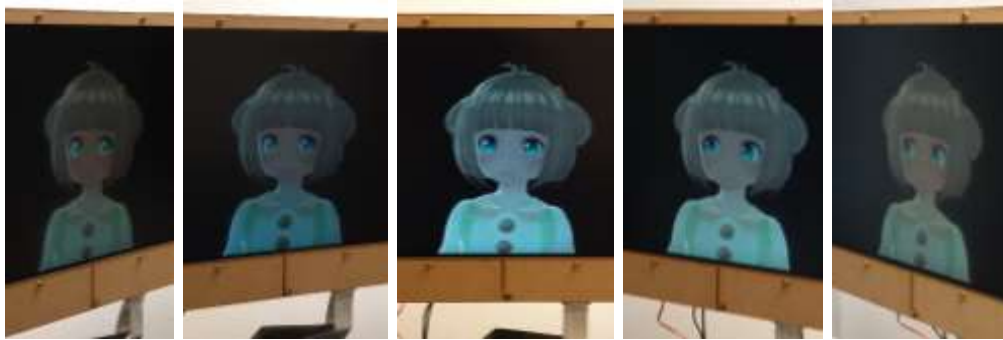


図4 視点位置によって視線方向が異なって見えるディスプレイ

## 5. 主な発表論文等

[雑誌論文] (計 1 件)

① 江添正剛, 佐藤裕仁, 三武裕玄, 長谷川晶一, ” 動力学指標の最適化により静止姿勢を身体負荷のより少ない姿勢に改善するバーチャルデッサン人形”, 画像電子学会誌, 査読有, 47巻, 2018, pp382-390

[学会発表] (計 9 件)

① 佐藤裕仁, 三武裕玄, 長谷川晶一 “トルク変化最小軌道に基づくキャラクターの自然な動作の自動生成”, エンタテインメントコンピューティング 2017

② Shuai Zou, Kento Kuzushima, Hironori Mitake, Shoichi Hasegawa, “Conversational Agent Learning Natural Gaze and Motion of Multi-Party Conversation from Example”, Proc. Of 5<sup>th</sup> International Conference on Human Agent Interaction, 2017

③ 藪下剛史, 三武裕玄, 長谷川晶一, ” 対話事例に基づく機械学習を用いた同庁的表情を提示する対話エージェント”, 情報処理学会第47回 EC 研究発表会, 2018

④ 平沼英翔, 三武裕玄, 長谷川晶一, ” 人の相槌に対する韻律の同調制御と発話進行制御により親和的な情報提供を行う対話エージェント”, 情報処理学会第47回 EC 研究発表会, 2018

⑤ 三武裕玄, Hsueh-Han Wu, 長谷川晶一, ” キャラクターを用いたデジタルサイネージが通行人の注意を惹きつけるための視線制御”, エンタテインメントコンピューティング 2018

⑥ 一居太郎, 三武裕玄, 立石和也, 長谷川晶一, ” 広い視聴範囲で美観を損ねずにモナリザ効果を排除できる二層式ディスプレイを用いた平面映像提示手法”, 情報処理学会インタラクティブセッション 2019

⑦ 杉森健, 三武裕玄, 佐藤裕仁, 長谷川晶一, ” 物体とのインタラクション時に身体の硬さを自在に変えられるVTuberアバター”, 情報処理学会インタラクティブセッション 2019

⑧ 佐藤裕仁, 三武裕玄, 杉森健, 長谷川晶一, ” 視線と身体を用いたインタラクティブキャラクターシステムとその編集インタフェースの構築”, 情報処理学会インタラクティブセッション 2019

⑨ Hironori Mitake, Taro Ichii, Kazuya Tateishi, Shoichi Hasegawa, “Wide Viewing Angle Fine Planar Image Display without the Mona Lisa Effect”, IEEE-VR 2019

〔産業財産権〕

○出願状況（計 1 件）

名称：視線表示方法及び視線表示装置

発明者：三武裕玄

権利者：同左

種類：特許

番号：特願 2019-10703

出願年：2019 年

国内外の別： 国内

〔その他〕

ホームページ等

<http://haselab.net/~mitake/>

## 6. 研究組織

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。