

交替潜時と韻律情報に基づく会話同調制御方式と情報収集を目的とした会話エージェントへの実装

速水 達也^{1,a)} 佐野 睦夫² 向井 謙太郎¹ 神田 智子²
宮脇 健三郎² 笹間 亮平³ 山口 智治³ 山田 敬嗣³

受付日 2012年11月16日, 採録日 2013年5月18日

概要: 本論文では, 交替潜時と韻律情報に基づく非線形振動子による連続的な引き込み制御とインタラクションルールによる離散的な引き込み制御により, 人とエージェントが協調的にコミュニケーションを行うための会話制御方式を提案する. また, 提案手法を情報収集を目的とする会話エージェントへ実装し, 会話同調制御の評価実験を行った. その結果, 直前の状態量の局所的制御を行う方式や交替潜時のみの単純な同調方式の従来手法に比べて, エージェントがどの程度親しみやすいかの評価尺度である親身度に対して有意差を示すことができ, 同時にエージェントの会話の自然さ(ストレスが少ない)に対しても有意差が認められた. これらの結果により, より親しみが持てストレスの少ない会話のもとで親しみやすく持続的な情報収集が行える会話システムの実現可能性を示すことができた.

キーワード: 交替潜時, 韻律情報, 引き込み制御, 会話同調制御, 会話エージェント

A Synchrony Control Method Based on Switching Pauses and Prosodic Information in Conversation and Implementation of a Conversational Agent for Gathering Information

TATSUYA HAYAMIZU^{1,a)} MUTSUO SANO² KENTAROU MUKAI¹ TOMOKO KODA²
KENZABURO MIYAWAKI² RYOHEI SASAMA³ TOMOHARU YAMAGUCHI³ KEIJI YAMADA³

Received: November 16, 2012, Accepted: May 18, 2013

Abstract: In this paper, we propose a conversation synchrony method which consists of continuous entrainment control based on non-linear oscillator using switching pauses with prosodic information and discrete entrainment control based on interaction rules for communicating humans with an conversational agent coordinately. We implemented a conversational agent system for gathering information and conducted evaluation experiments of the proposed method. Consequently, the experimental results show that our proposed method has statistical significant difference to both kindness and natural communication compared to locally controlled methods and conventional synchrony methods based on only switching pauses. And we could show feasibility of conversational agent system to gather information sustainably with friendly and natural communication.

Keywords: switching pause, prosodic information, entrainment control, conversational synchrony control, conversational agent

¹ 大阪工業大学大学院
Graduate School of Osaka Institute of Technology, Hirakata,
Osaka 573-0196, Japan

² 大阪工業大学
Osaka Institute of Technology, Hirakata, Osaka 573-0196,
Japan

³ NEC C&C イノベーション推進本部
NEC C&C Innovation Initiative, Ikoma, Nara 630-0101,
Japan

a) t.hayamizu@outlook.com

1. はじめに

近年, ロボティクスの分野では, 人とのコミュニケーションや介護を目的としたホームロボットや介護ロボットなどが開発され, ロボットが一般の人々にも身近な存在となりつつある. ホームロボットや介護ロボットのように, ロボットが人々の生活により密接にかかわる場合, 人と持

統的かつ自然にコミュニケーションをとる必要がある。そのためには、人とロボットが協調的に対話するための対話の流れを解明し、より自然でストレスの少ない親和性のあるコミュニケーション制御が必要である。

人間同士の会話では、言語情報とともに非言語情報とパラ言語情報がやりとりされる。伝達される情報の多くは非言語情報およびパラ言語情報であるが、その中でも発話時間・交替潜時・イントネーション・アクセント・音圧レベル・発話速度などのパラ言語情報が、話者に対する印象や対話の流れに大きな影響を与えていることが確認されている。同時に、パラ言語情報は、話者間で互いに同調することも確認されている [1], [2], [3], [4], [5], [6], [7]。

長岡ら [2], [3] は、協調的な音声対話では、対話者同士がほぼ同じような発言長や交替潜時で発言や相槌をする引き込み現象を確認している。また、対話者の交替潜時が円滑なコミュニケーションを促すことも示唆している。交替潜時とは、一方の話者が話し終わって次の話者が話し始めるまでの時間間隔のことである [2]。三宅ら [4] は、機械の発話特徴（発話長・発話間隔）が一定であるならば、人間は自らの発話特徴を機械に同調させ、機械が途中で発話特徴を変化させた場合には人間は変化後の発話特徴に向けて同調させていく、ことを確認している。小松ら [5] は、人間同士の音声対話において、発話速度の引き込み現象が存在することを示し、人間と人工物との対話においても同様に観察されることを示した。垣田 [6] は、質問応答形式の実験において、一方の話者の基本周波数が高ければもう片方の話者の基本周波数が高くなることを指摘している。西村ら [7] は、協調的な音声対話システムを実現するために、人間同士の対話の印象と韻律変化との間にどのような関係性があるのかを分析し、人間同士の対話における韻律変化には、同調傾向の相互作用があることを示した。また、同調傾向が強い対話は盛り上がっている印象を持っており、フランクで親しみのある盛り上がった対話では、対話の韻律特徴（F0、パワー、話速）が同調していることを示した。

パラ言語に基づく引き込み制御による会話システムの実現方式としては、(1) 最初に最適解を求めておき会話に対する状態量を静的に適用するもの、(2) 直前の状態量に着目し動的に局所的制御を行うもの、(3) 会話全体のリズムに着目し同調させ動的に制御を行うもの、に分類ができる。(1)の問題点としては、状態量が固定されているため会話中のコンテキストの変化に対して対応ができない [8]。(2)については、傾きや身体動作タイミングの予測モデルに基づき引き込み制御を行い、局所的なコンテキストの変化には対応することはできる [9] が、会話全体のリズムに関わるような大局的な状態量の変化には対応はできない。(3)の会話全体にリズムがあるような会話例としては、聞き取りのように会話をテンポよく進行させるようなケースや応答・共感などが繰り返され会話に一定以上のダイナミ

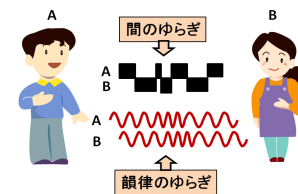


図 1 会話における状態量のゆらぎとダイナミクス

Fig. 1 Fluctuation of state and dynamics in conversation.

クスがあるようなケースがあげられる。(3)の会話全体のリズムに着目し同調させながら引き込み制御を行う方式としては、林ら [10]の研究があげられる。しかし、交替潜時や発話速度の時間的リズムに基づく引き込み制御を行うにとどまっておき、韻律情報による同調効果 [7]については考慮されていない。会話全体をより自然で親しみやすくするためには、会話全体の「間」のゆらぎを反映した時間的リズムに加え、会話全体の「トーン」のゆらぎを反映した韻律的なリズムを同時に制御する必要があるが、このような会話全体のゆらぎに着目し引き込み制御を行い、会話システムをデザインしている研究例はない。本研究では、図 1 に示すように、会話の「間」の変動と韻律の変動を対話の流れにおける「ゆらぎ」ととらえモデル化を行う。ゆらぎのリズムが共振したときに同調現象が起こることが知られている [11]。コミュニケーションにおける同調現象は、協調的会話の状態や、会話が自然に盛り上がっているときなどに観察される [3], [7]。本研究では、このような会話全体のゆらぎに基づいた引き込み制御モデルを特徴とする会話エージェントを提案する。会話エージェントの効果としては、「ストレスが少なく自然」で、より「協調的である」ことが期待される。具体的には、非線形振動子に基づき引き込み制御モデルの定式化を行い、交替潜時と韻律情報に基づいた大局的な同調制御を行う。

この会話全体の時間的ゆらぎと韻律的ゆらぎに着目した同調制御方式を情報収集会話エージェントとして実装し、提案する同調制御方式の有効性の検証を行う。

2. 提案する同調制御方式と会話エージェント

2.1 非線形振動子の引き込み現象とパラ言語

非線形振動子には、外部の振動に合わせて振動数を同調させる引き込み現象が見られる。

引き込み現象は、コミュニケーションを行っている二者間でも起こり、自らの表出するパラ言語と相手のパラ言語が同調する現象が観察されている。本研究では、この引き込み現象を Van der Pol 方程式を拡張することで、実現する [11]。

Van der Pol 方程式は、式 (1) で与えられる。

$$\ddot{x} + a(x^2 - 1)\dot{x} + x = 0 \quad (a > 0) \quad (1)$$

ここで、 a は粘性係数であり、振動子の挙動の強度を制御

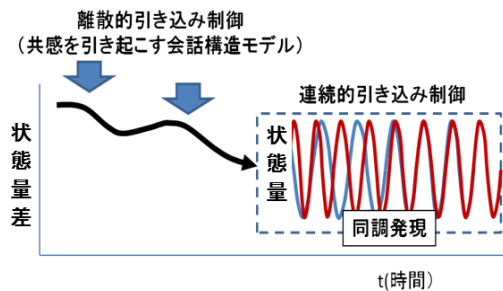


図 2 同調現象発現のための引き込み制御

Fig. 2 Entrainment control for realizing synchronization phenomenon.

できる. 式 (1) は, 2つの振動子が相互に影響を与えるように, 式 (2) に拡張可能であり, これにより, 相互引き込みを発現させることができる [11].

$$\begin{aligned} \ddot{x}_1 &= \varepsilon(1 - x_1^2)\dot{x}_1 - x_1 + \lambda(x_2 - x_1) \\ \ddot{x}_2 &= \varepsilon(1 - x_2^2)\dot{x}_2 - x_2 + \lambda(x_1 - x_2) \end{aligned} \quad (2)$$

ここで, λ は結合係数であり, 相互作用の強さを表す. ε は粘性係数である.

2.2 提案する同調制御方式の概要

人とエージェントが協動的にコミュニケーションを実現するためには, 対話者とエージェント間の対話の流れを考慮してコミュニケーション制御方式を設計しなければならない. この対話の流れをモデル化する一般的な方法論として, 制御すべき状態量を, 状態方程式で表現する方法が考えられる.

本研究で検証するコミュニケーションの状態量は, 先行研究で人間同士の同調傾向が確認されている交替潜時 (時間的引き込み制御パラメータ), および基本周波数 F_0 (韻律的引き込み制御パラメータ) とする. 本論文では, 基本周波数 F_0 を以下, ピッチと呼ぶ. これらの状態量の同調現象を非線形振動子に基づく状態方程式によりモデル化し, 対話者とエージェント間の連続的で大局的な引き込み制御を行う.

しかし, 状態方程式で記述される連続的な引き込み現象は, 状態量差が大きいと発現しない. したがって, エージェントと人間の会話がまったく成立していない状況では, このような引き込み制御を行うことは困難であることに留意しなければならない. 本研究では, 共感を引き起こす会話構造モデルの下で, 頷きなどの身体的インタラクションなどの離散的な引き込み制御を行うことによって, 応答可能な会話を成立させ, 同調現象を引き起こす条件を調整する戦略をとる. すなわち, 図 2 に示すように, 離散的引き込み制御によって, 同調現象を発現するだけの状態量差まで調整し, 同調現象を発現させる. いったん, 同調現象が発現されると, 安定したコミュニケーションを維持することが可能となる.

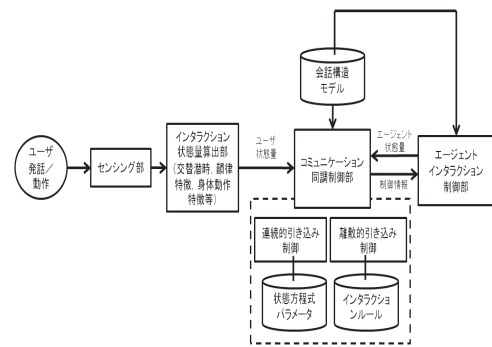


図 3 コミュニケーション同調制御方式の構成図

Fig. 3 System configuration diagram of communication synchronization control method.

図 3 に, 本研究で提案する制御方式を示す. 人間の発話や動作をセンシングし算出されたユーザ状態量と, インタラクション制御部から生成されるエージェントの状態量に対して, 状態方程式による連続的で大局的な引き込み制御とインタラクションルールによる離散的で局所的な引き込み制御に基づきコミュニケーション同調制御を行い, 制御情報をエージェントのインタラクション制御 (バーバル制御, ノンバーバル制御) に反映させ, 共感を引き起こす会話構造モデルのもとで会話状態を動的に遷移させ, コミュニケーションを促進する. 2.3 節および 2.4 節でその詳細を示す. このように, 会話エージェントによる同調発現を系統的にとらえデザインした研究例はない.

2.3 コミュニケーション同調制御

2.3.1 連続的引き込み制御

本研究では, 引き込み制御の数値モデルとして, Van der Pol 方程式を用いる. Van der Pol 方程式は, 引き込み特性を持つ非線形振動子である. 非線形振動子は, ここで, 大局的な時間的リズムを示す交替潜時の引き込み制御と, 大局的な韻律的リズムを示すピッチの引き込み制御は独立であると仮定し, それぞれ定式化を行う. エージェントの時刻 (t) における交替潜時の変位を式 (3), ピッチの変位を式 (4) で定義する.

$$\ddot{x}(t) = \mu(1 - x(t)^2)\dot{x}(t) - \omega^2x(t) + \alpha c(t) \quad (3)$$

$$\ddot{y}(t) = \mu(1 - y(t)^2)\dot{y}(t) - \omega^2y(t) + \beta p(t) \quad (4)$$

ここで, ω, μ はそれぞれエネルギー保存係数, 粘性係数である. $c(t), p(t)$ はエージェントに同調させる対話者のパラ言語情報であり, 対話者の交替潜時およびピッチのリズムを近似した \sin 波上の時刻 t におけるそれぞれの変位を表す. $\alpha (0 \leq \alpha \leq 1), \beta (0 \leq \beta \leq 1)$ はその際に用いる引き込み係数で, これらを変化させることで対話者のパラ言語に対する同調のしやすさを調節する. $\mu, \omega, \alpha, \beta$ のパラメータの値は, パラメータの値を刻み幅 0.1 で変更していくことで Van der Pol 方程式においてリミットサイクルが発現する値を求めた. そして, 同調を通して変化し

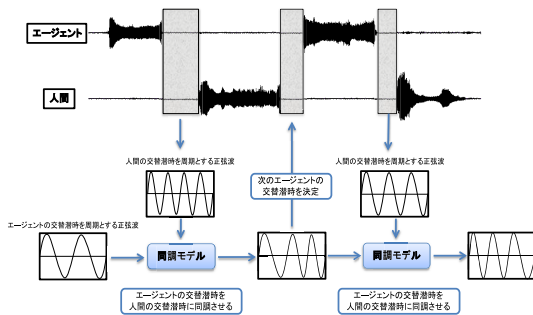


図 4 交替潜時の同調制御のフロー

Fig. 4 Synchronization control flow of switching pauses.

た周期情報を用いてエージェントが次に表出する交替潜時やピッチを制御する。一般的に、人間の脳波や心拍・呼吸なども一定の生体リズムを有しており、調和振動（非線形振動）が起こることが確認されている [12], [13]. 調和振動（非線形振動）は、微分方程式で表され、一般解は、正弦波となり振動する。筆者らは、これらの知見から、交替潜時や韻律情報などの会話パラメータも、大局的には、非線形振動の挙動の中に支配されていると仮説をたてた。そこで、対話者の時刻 (t) における交替潜時の変位 $c(t)$ を式 (5)、ピッチの変位 $p(t)$ を式 (6) に定義する。

$$c(t) = \sin\left(\frac{2\pi t}{ks}\right) \quad (5)$$

$$p(t) = \sin\left(\frac{2\pi t}{kr}\right) \quad (6)$$

ここで、 s は対話者から観測される交替潜時長である。 r はピッチの変化率であり現在のターンの発話区間の基本周波数と 1 ターン前の発話区間の変動値として定義する。 k は波の位相のスケールングファクタで定数である。 k の値は、事前実験の結果より決定する。対話者の交替潜時やピッチの正弦波は、交替潜時長 s やピッチ変化率 r を k 倍した周期を持つ波として表現できる。

交替潜時の同調制御のフローを図 4 に示す。人の交替潜時を周期とした正弦波とエージェントの交替潜時を周期とした正弦波を同調モデルを介して同調させる方式をとる。

次に、同調モデルによるピッチの制御方法を図 5 に示す。ピッチの制御は、交替潜時の制御と同じく人のピッチ変化率とエージェントのピッチ変化率を正弦波で近似し同調モデルを介して同調させる方式をとる。

2.3.2 離散的引き込み制御

人間同士の対話では、適切なタイミングで相槌や頷きを返している。聞き手に相槌や頷きなどの反応がなければ、話し手は相手が自分の発話を理解しているのか不安になってしまう。したがって、相槌や頷きなどの現象は自然対話によるマンマシインタフェースを考えるうえで非常に重要なインタフェースであると考えられる [14]. そこで、本研究ではエージェントに、発話開始および発話終了を意味する身体的インタラクションとして頷きを導入する。エー

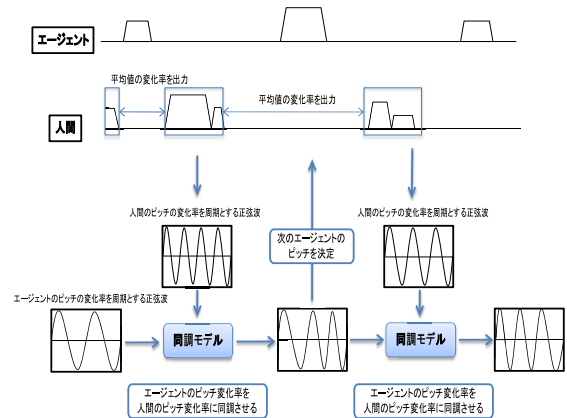


図 5 ピッチの同調制御のフロー

Fig. 5 Synchronizing control flow of pitch.

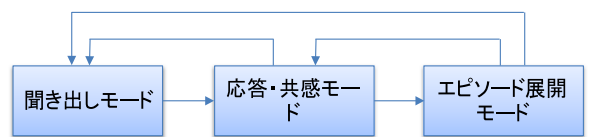


図 6 エージェントの会話構造モデルと引き込みの促進

Fig. 6 Conversation structure model of agent and promotion of entrainment.

ジェントは、話者の発話内容について「話者に了解を求める」あるいは「話者の発話に対して賛同する」場合において頷くとする。エージェントには、人間の発話終了後に浅い頷きを 2 回行い、エージェントの発話終了後に深い頷きを、行うようにインタラクションルールを設定する。これにより、会話を成立させ、同調現象を引き起こす条件を調整する。

2.4 情報収集を目的とした会話構造モデルと引き込み制御

会話を成立させるには、適切な会話構造モデルが必要となる。本研究では、対話を通して問診を行う情報収集タスクを想定して、会話構造モデルを設計した。図 6 に示すように、エージェントは、まず相手から情報を聞き出し（聞き出しモード）、聞き出した情報に対して何らかの応答を行い必要に応じて共感し（応答・共感モード）、共感した情報に対して深く掘り下げる（エピソード展開モード）ことを繰り返す。このように、対話者の行動やエピソードに共感し、エージェントへの引き込みを促進しながら、情報収集を円滑に進めていく。各モードの発話内容例を表 1、表 2、表 3 に示す。先行研究 [7] では、会話の共感時に話者間で韻律パラメータの相関があることが確認されており、本研究では、ここで提案する会話構造が、韻律パラメータによって非線形振動子による同調現象を引き起こす前提条件であると仮定する。また、共感し、情報をさらにテンポよく聞き出すことを繰り返すことで、交替潜時についても同調現象を引き起こすことにつながると想定する。なお、応答内容は人間の発話に関係なく質問内容に応じてランダムに決定される。

表 1 聞き出しモードの発話内容

Table 1 Utterance contents of hearing mode.

聞き出しモード	
今日の気分はどうですか？	最近、急激な体重の増減はありましたか？
定期的な運動をしていますか？	医師に高血圧と診断されたことはありますか？
規則正しい生活をおくっていますか？	アレルギーがありますか？
自分に適切な睡眠時間を取っていますか？	朝起きるのが辛いと思いませんか？
お酒を飲みますか？	最近、イライラすることが増えましたか？
たばこを吸いますか？	肩の凝りはありますか？
定期検診を受けていますか？	手足のしびれを感じたことはありますか？
体力に自信がありますか？	食欲が無いことがありますか？
体は軽いですか？	腰の調子は悪くないですか？
食生活に気を付けていますか？	最近、身体の冷えを感じることはありますか？
現在、身体に関する悩みはありますか？	疲れを慢性的に感じていませんか？
急に気分が沈んだりすることはありますか？	視力が落ちてきたと感じますか？
現在服用している薬はありますか？	生活においてストレスを感じる事が多いですか？
過去に大きな病気になったことがありますか？	冷え性ですか？
生活環境に不安はありませんか？	毎日30分以上は歩いていますか？

表 2 応答・共感モードの発話内容

Table 2 Utterance contents of response and sympathy mode.

応答・共感モード
私も～に、チャレンジしたくなってきました。
やっぱり、自分のペースでやるのが一番いいですね。
嫌いなものはなかなか好きになれませんね。
私も応援しています。これからも続けていってくださいね。
最近は特に、色々と考えさせられることが多いですね。
私もやってみたいと思います。
私も一緒にチャレンジさせてください。
私も好きです。好きが一緒ですね。食事の中に好きなものが入っているととても嬉しくなりますよね。
いいですね。是非私も行ってみたいです。
私も、～に行く事があれば、楽しみたいです。
わかります。私もこれから見てみたいです。
そうですか。

表 3 エピソード展開モードの発話内容

Table 3 Utterance contents of episode expansion mode.

エピソード展開モード
～さんは～が趣味だとお伺いしましたが、～の良さを教えてください。
趣味の～は、どれくらいの頻度でされているのですか？
嫌いな食べ物～は～ですが、なぜ嫌いなんですか？
～さん、日課の～の楽しさを教えてください。
～の気になるニュースは、～ですが、なぜか教えて頂けますか？
～さんは何かチャレンジしたい事はありますか？
～さんは～が好きだということですが、～がなぜ好きなのか理由を教えてください。
～さんは～に旅行に行きたいとのことでしたが、～に行きたい理由を教えてください。
～さんは出身地が～ということですが、～の良いところを教えてください。
～さんが好きなテレビ番組は～ということですが、どこが好きなのか教えてください。

3. 提案手法の実装

提案方式は、図 3 に示すようなモジュール構成として実装した。インタラクション状態量算出部、コミュニケーション同調制御部、エージェントインタラクション制御部の各部の動作について示す。

3.1 インタラクション状態量算出部

インタラクション状態量算出モジュールは、マイクから入力された音声のパワーとピッチを用いて、人間の発話開始/終了時間とピッチの平均値を求め、その値を用いてエージェントの発話開始時間、ピッチの平均値を算出する。発話末の変化が多い日本語では、発話末の制御が必要であるが、本研究では詳細な引き込み制御ではなく大局的な引き込み制御を主眼に置いている。したがって、発話区間の代表値を用いて制御を行う。発話区間のピッチの代表値としては平均値と中央値が考えられる。ここで、事前実験で

ピッチの平均値と中央値の残差平均は $\pm 3.0 \sim \pm 3.5\%$ に収まっており、ピッチの変動が 4% 以内であれば、実験参加者は変化を感じないことが確認されている。本研究では、中央値より演算コストの低い平均値で十分であると判断した。また、音声パワーの基準値を 20×10^{-6} [Pa] とした。次式を用いて実験参加者の発話音声のパワーを算出する。

$$L_p = 20 \times \log_{10}(p/p_0) \quad (7)$$

(p : 音圧, p_0 : 基準値)

発話開始・終了判定にはパワーを用いた。約 0.02 秒ごとにパワーを算出し記録する。そのデータを用いて約 1 秒間の単純移動平均を算出しその時点でのパワーとする。閾値を超えると発話と見なし、約 0.25 秒間発話し続けると発話開始とした。閾値を下回ると無音と見なし、約 0.75 秒間無音が続くと発話終了とした。これらの発話開始および発話終了の時間間隔は、高杉らの研究 [15] を基に事前実験を行い、0.75 秒とした。また、発話音声から基本周波数を求めピッチとして算出した。算出方法には音声の基本周期を音声データから読み取り、その逆数をとる方法を用いた。発話開始から発話終了までの区間の音声データを用いて、約 0.02 秒ごとにピッチを算出し、区間内のその値の平均をその発話のピッチとした。算出した現在のターンのピッチと 1 ターン前のピッチの変動値をピッチ変化率とし、引き込みモジュールへは人間の発話開始/終了時間とピッチの変化率を送信する。ピッチ変化率の算出式を式 (8) に示す。

$$\text{ピッチの変化率} = \frac{\text{今回の発話ピッチ}}{\text{前回の発話ピッチ}} \quad (8)$$

3.2 コミュニケーション同調制御部

コミュニケーション同調制御部は、インタラクション状態量算出部から人間の発話開始時間と終了時間およびピッチ変化率を受け取る。エージェントインタラクション制御部の音声合成モジュールからエージェントの発話終了時間を受け取り、受け取った時間から人間の交替潜時を計算し、Van der Pol 方程式によりエージェントの交替潜時を決定する。同様に、人間のピッチ変化率を受け取り、Van der Pol 方程式によりエージェントのピッチ変化率を決定する。交替潜時の終了後、エージェントインタラクション制御部の音声合成モジュールへ次の質問番号を送信し、エージェントインタラクション制御部の動作生成モジュールに口の動作 ID を送信する。また、エージェントの発話終了タイミングと人間の発話終了タイミングで、エージェントインタラクション制御部に頷きの動作 ID を送信する。

3.3 エージェントインタラクション制御部

音声合成モジュールと動作生成モジュールから構成される。

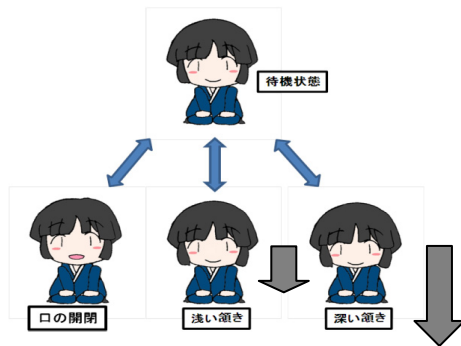


図 7 会話エージェントの動作例
 Fig. 7 An example of conversational agent.

3.3.1 音声合成モジュール

音声合成モジュールは、コミュニケーション制御同調部から受信したパラメータをもとに発話内容やピッチを制御し、再生を行うモジュールである。音声の再生終了時には、その終了時刻をコミュニケーション同調制御部へ送信する。音声合成エンジンとしては、ピッチ、パワー、スピードなどの韻律情報、アクセントなどが調整可能な「Voice Text」(HOYA 社)の女性音声版 (HARUKA) を用いた。ピッチの変動幅については、違和感を与えないという条件下で事前実験により算出し、基準値の±20%とした。また、音声合成による遅延は約 30 ms であり対話タイミングに影響がないことを確認した。会話内容は、会話構造モデル (図 6) の各モードに対応した発話パターン (表 1~表 3) に基づき生成する。エピソード展開モードの発話内容は、事前に獲得済みのプロフィールデータを参照することにより生成する。

3.3.2 動作生成モジュール

会話エージェントの表示制御を行うとともに、引き込みモジュールから受信した動作 ID をもとにエージェントの動作を制御するモジュールである。動作の種類は以下の 4 つとした。

- (1) 各動作が終了した後、停止する。
- (2) 音声合成モジュールが発話を開始した際に口を開閉させる。
- (3) 人間が発話終了した際に浅く 2 回頷く。
- (4) 音声合成モジュールが音声の再生を終了した際に深く 1 回頷く。

会話エージェントの動作例を図 7 に示す。

4. 評価実験

4.1 実験目的

本実験では、提案する会話同調制御方式を会話エージェントへ実装し、エージェントの交替潜時およびピッチを、人の交替潜時およびピッチと同調させることにより、従来方式よりも、実験参加者が、「自然でストレスが少なく」かつ「協調的である」ことがいえるかどうかを検証することを目的とする。

表 4 実装した 3 種類の引き込み制御方式

Table 4 Three types of implemented entrainment control methods.

方式	種類	パラメータ	説明
①	非同調方式	交替潜時 ピッチ	人間の交替潜時・ピッチを模倣したエージェント(比較方式)
②	同調方式	交替潜時 ピッチ	人間の交替潜時・ピッチと同調したエージェント(提案方式)
③	同調方式	交替潜時	人間の交替潜時と同調したエージェント(比較方式)

4.2 比較制御方式

本研究では、提案する同調制御方式と比較検証する制御方式として、引き込み制御方式や引き込み制御パラメータが異なる 3 種類のエージェントを作成した。実験参加者とこの 3 種類のエージェントを会話させることで、提案方式に関する評価を行った。作成した 3 種類の引き込み制御方式の詳細を表 4 に示す。

インタラクション状態量算出、会話構造モデル、離散的な引き込み制御などの条件は 3 方式とも同じとした。提案方式は ② である。

比較方式 ① は、模倣という最も基本的な引き込みモデルで、人間の直前の交替潜時・ピッチという局所的な状態量に着目し直前の状態量を模倣するという制御方式とした。具体的には、人のピッチ・交替潜時をそれぞれ $P_H \cdot C_H$ 、エージェントのピッチと交替潜時をそれぞれ $P_A \cdot C_A$ とし、現在のターンを t とすると、エージェントの状態量 $P_A \cdot C_A$ は、式 (9) で算出できる。

$$P_A(t) = P_H(t - 1) \tag{9}$$

$$C_A(t) = C_H(t - 1)$$

コミュニケーション全体のダイナミクスを反映しておらず、直前の状態量のみに着目しているということで非同調方式と命名した。このように引き込み制御方式のまったく異なる非同調方式と対話の流れを反映した提案方式の効果の差を検証する。

比較方式 ③ は、交替潜時のみの同調制御方式であり、提案方式 ② とは、同調に適用するパラメータが、交替潜時のみであるという違いがある。交替潜時のような時間的引き込み制御パラメータのみで同調させるという考え方 [10] に基づく。時間的引き込み制御に韻律的引き込み制御を付加する効果を検証する。

4.3 実験方法

大学生 10 名を実験参加者とした。実験参加者は、対話エージェントが表示されるモニタに向かい合うように着席させた。実験中、実験参加者らはマイクを装着し、側面と後方に設置したビデオカメラにより実験風景を撮影した。本実験では音声認識は用いておらず、音声はピッチの平均値を算出するために使う。実験参加者からは、事前に、趣味や特技などエピソードに関係する情報を取得し、会話エージェントシステムに実験参加者のプロフィールデータとし

表 5 アンケートの項目

Table 5 Questionnaire items.

分類項目	質問内容	質問番号
親身度	個人的な話ができた	1
親身度	エージェントが私に興味を持っていると感じた	2
ストレス	ストレスなく話すことができた	3

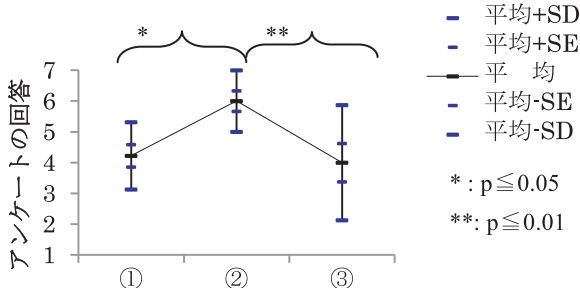


図 8 質問 1 の項目に対する各方式の平均値と標準偏差

Fig. 8 Average/standard deviation for questionnaire item 1.

表 6 質問 1 の項目に対する各方式の平均値と標準偏差

Table 6 Average/standard deviation for questionnaire item 1.

因子A	平均	標準偏差(SD)	平均-SD	平均+SD	標準誤差(SE)	平均-SE	平均+SE
①	4.22	1.09	3.13	5.32	0.36	3.86	4.59
②	6.00	1.00	5.00	7.00	0.33	5.67	6.33
③	4.00	1.87	2.13	5.87	0.62	3.38	4.62

て組み込んだ。実験参加者には、問診システムであるので、質問されたことにのみ答えるように指示をした。エージェントの実験を行う順序はランダムとし、実験時間は3分とした。実験終了後、表5に示すようなアンケートを実施した。会話エージェントに期待される効果である「協動的であるか」という評価項目については、エージェントの親身度ととらえ、「個人的な話をすることができたか?」と「エージェントが私に興味を持っていると感じたか?」の2項目を設定した。また、「ストレスが少ない(会話そのものに対するストレス)」という項目に関しては、直接、「ストレスなく話すことができたか?」という問いかけとした。以上のアンケート項目に対して、7段階(7が最も良い)のレベルを設定した。

4.4 実験結果

上記3項目(質問1, 質問2, 質問3)のアンケート結果を分析し、多重比較検定を行った結果、3つの項目ともに有意差が認められた。以下、分析結果の詳細を示す。

4.4.1 親身度の項目における分析

質問1「個人的な話ができた」の項目に対する各制御方式の平均値と標準偏差を、図8と表6に示す。図8と表6から、最もアンケート評価の平均値が高いのは提案方式②であることが分かる。

次に、質問1に対する分散分析表を表7に示す。分散分析の結果から提案方式②は、比較方式①と③に対して

表 7 質問 1 の項目に対する分散分析表

Table 7 Analysis of variance table for questionnaire item 1.

因子	平方和	自由度	平均平方	F 値	P 値	判定
因子A	21.63	2.00	10.81	5.70	0.01**	
誤差	45.56	24.00	1.90			
全体	67.19	26.00				

**危険率1%有意 *危険率5%有意

表 8 質問 1 の項目に対する多重比較検定の結果

Table 8 Results of multiple comparison test for questionnaire item 1.

水準1	水準2	平均1	平均2	差	統計量	P 値	判定
②	①	6.00	4.22	1.78	2.74	0.01*	
②	③	6.00	4.00	2.00	3.08	0.01**	

**危険率1%有意 *危険率5%有意

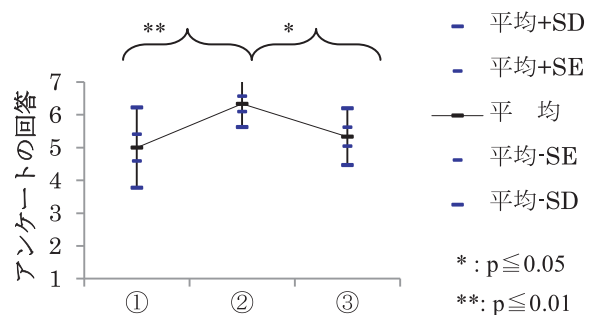


図 9 質問 2 の項目に対する各方式の平均値と標準偏差

Fig. 9 Average/standard deviation for questionnaire item 2.

表 9 質問 2 の項目に対する各方式の平均値と標準偏差

Table 9 Average/standard deviation for questionnaire item 2.

因子A	平均	標準偏差(SD)	平均-SD	平均+SD	標準誤差(SE)	平均-SE	平均+SE
①	5.00	1.22	3.78	6.22	0.41	4.59	5.41
②	6.33	0.71	5.63	7.04	0.24	6.10	6.57
③	5.33	0.87	4.47	6.20	0.29	5.04	5.62

危険率1%で有意差があった。さらに、提案方式②と比較方式①、および提案方式②と比較方式③との多重比較検定をFisherのLSD(最小有意差検定)を用いて実施した結果を、表8に示す。提案方式②と比較方式①の間に危険率5%で有意差があり、提案方式②と比較方式③の間に危険率1%で有意差があった。以上のことから各方式間の評価は、② > ①, ② > ③であるといえる。これにより、親身度に関しては、比較方式よりも提案方式の方が、優位性があるといえる。

質問2の項目に対する各制御方式の平均値と標準偏差を、図9と表9に示す。図9と表9から、最もアンケート評価の平均値が高いのは提案方式である②であることが分かる。

次に、質問2に対する分散分析表を表10に示す。分散分析の結果から提案方式②は、比較方式①と③に対して危険率5%で有意差があった。さらに、提案方式②と比較方式①、および提案方式②と比較方式③との多重

表 10 質問 2 の項目に対する分散分析表

Table 10 Analysis of variance table for questionnaire item 2.

因子	平方和	自由度	平均平方	F 値	P 値	判定
因子A	8.67	2.00	4.33	4.73	0.02*	
誤差	22.00	24.00	0.92			
全体	30.67	26.00				

表 11 質問 2 の項目に対する多重比較検定の結果

Table 11 Results of multiple comparison test of questionnaire item 2.

水準1	水準2	平均1	平均2	差	統計量	P 値	判定
②	①	6.33	5.00	1.33	2.95	0.01**	
②	③	6.33	5.33	1.00	2.22	0.04*	

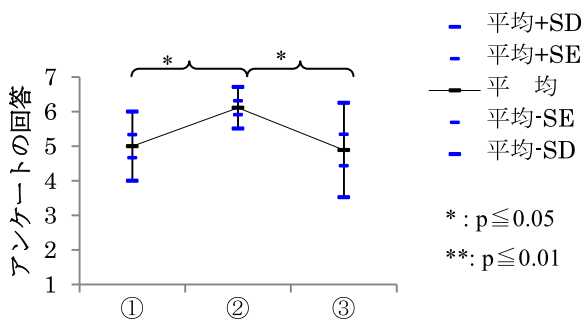


図 10 質問 3 の項目に対する各方式の平均値と標準偏差

Fig. 10 Average/standard deviation for questionnaire item 3.

表 12 質問 3 の項目に対する各方式の平均値と標準偏差

Table 12 Average/standard deviation for questionnaire item 3.

因子A	平均	標準偏差(SD)	平均-SD	平均+SD	標準誤差(SE)	平均-SE	平均+SE
①	5.00	1.00	4.00	6.00	0.33	4.67	5.33
②	6.11	0.60	5.51	6.71	0.20	5.91	6.31
③	4.89	1.36	3.52	6.25	0.45	4.43	5.34

比較検定を Fisher の LSD を用いて実施した結果を表 11 に示す。提案方式 ② と比較方式 ① の間に危険率 1% で有意差があり，提案方式 ② と比較方式 ③ の間に危険率 5% で有意差があった。以上のことから各方式間の評価は，② > ①，② > ③ であるといえる。これにより，親身度に関しては，比較方式よりも提案方式の方が，優位性があるといえる。

4.4.2 ストレスの項目における分析

質問 3 の項目に対する各制御方式の平均値と標準偏差を図 10 と表 12 に示す。図 10 と表 12 から，アンケート評価の平均が最も高いのは提案方式 ② であることが分かる。

次に，質問 3 の項目の分散分析表を表 13 に示す。分散分析の結果から提案方式 ② は，比較方式 ① と ③ に対して危険率 5% で有意差があった。さらに，提案方式 ② と比較方式 ①，および提案方式 ② と比較方式 ③ との多重比較検定を，Fisher の LSD を用いた結果を表 14 に示す。

表 13 質問 3 の項目に対する分散分析表

Table 13 Analysis of variance table for questionnaire item 3.

因子	平方和	自由度	平均平方	F 値	P 値	判定
因子A	8.22	2.00	4.11	3.83	0.04*	
誤差	25.78	24.00	1.07			
全体	34.00	26.00				

表 14 質問 3 の項目に対する多重比較検定の結果

Table 14 Results of multiple comparison test for questionnaire item 3.

水準1	水準2	平均1	平均2	差	統計量	P 値	判定
②	①	6.111	5.000	1.111	2.274	0.032*	
②	③	6.111	4.889	1.222	2.502	0.020*	

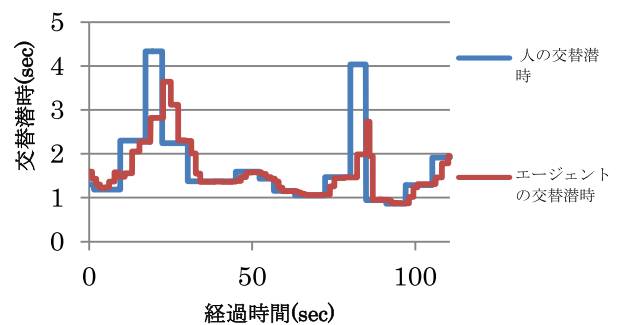


図 11 人とエージェントの交替潜時の変化

Fig. 11 Switching-pause changes of both human and agent.

表 14 から，提案方式 ② と比較方式 ① の間に危険率 5% で有意差があり，提案方式 ② と比較方式 ③ の間に危険率 5% で有意差が確認された。以上のことから各方式間の評価は，② > ①，② > ③ であるといえる。これにより，ストレス性に関しても，比較方式よりも提案方式の方が，優位性があるといえる。

4.5 考察

実験結果 (提案方式 ② と比較方式 ① の比較) より，対話者の直前のコミュニケーション状態量を模倣する方式 (非同調方式) と比べて，提案の同調方式の方が，親身度を向上させ，ストレスを軽減させる効果があることが確認された。提案の同調方式は，対話者のコミュニケーション状態量を模倣するのではなく，会話全体のコミュニケーションのゆらぎに着目し，そのゆらぎの中で引き込みを起こし同調現象を発現させることにより，より協調的で自然でストレスのない会話を実現できたと解釈できる。図 11 および図 12 は，エージェントと実験参加者の会話開始時点から 110 秒までの交替潜時とピッチ変化の推移を図示したものである。図 11 と図 12 から，エージェントの交替潜時とピッチ変化が，実験参加者の交替潜時とピッチ変化に合わせて同調しているのが確認された。また，図 11 の 20 秒付近と 80 秒付近は実験参加者が考え込んでしまった場面で，交替潜

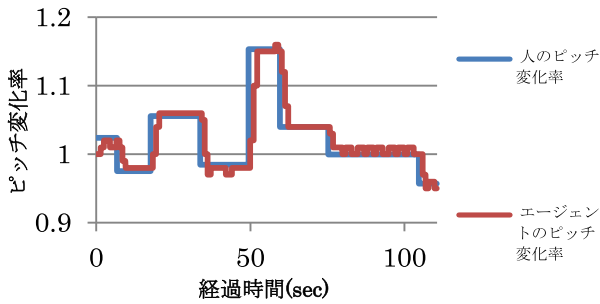


図 12 人とエージェントのピッチ変化

Fig. 12 Pitch changes of both human and agent.

時が長くなっているが、交替潜時が長くなってもエージェントは人の交替潜時に動的に追従していることが確認できる。また、図 12 のピッチ変化率では、50-60 秒付近にかけて、ピッチ変化が大きくなっており、この付近では、ビデオ観察により人とエージェントの会話が盛り上がっていた。

このように、本提案方式により、人とエージェント間の対話の流れの中で、エージェントが人のコミュニケーション状態量の変化に適応的に振る舞うことができ、かつ、いったん同調を発現すると、多少の変動があっても、安定的にコミュニケーション状態を維持することが可能となる。これは、システムの挙動がいったんアトラクタの中に入れば、多少のノイズがあっても頑健に振る舞うことができるという非線形振動子の挙動特性を裏付けている [12], [16]。

さらに、提案方式 ② と比較方式 ③ を比較した実験結果により、本同調方式の引き込み制御パラメータとして、交替潜時だけでなく、韻律情報（ピッチ）が、協調的で自然でストレスのない会話の実現に寄与していることが裏付けられた。このように、韻律的なリズムの同調は会話全体の調整に大きな役割を果たしているといえる。

親身度およびストレス軽減の評価について、提案方式に基づく会話エージェントとして、両方とも平均 6 点（7 点満点）以上と高い評価を得ることができた。このことは、図 3 で示した同調制御方式に基づく会話エージェントシステムが有効に機能したといえる。さらに、3 分間の会話において、平均 10 個程度の問診情報を安定的に獲得できており、情報収集エージェントとしても有効に動作した。

5. おわりに

本論文では、非線形振動子による連続的で大局的な引き込み制御とインタラクションルールによる離散的で局所的な引き込み制御に基づきコミュニケーション同調制御を行い、共感を引き起こす会話構造モデルのもとで会話状態をダイナミックに遷移させ、コミュニケーションを促進するエージェント会話制御方式を提案した。さらに、提案した会話同調制御方式を情報収集を目的とした会話エージェントへ実装し、有効性の検証を行った。その結果、

(1) 提案方式は、対話者のコミュニケーション状態量を模

倣する制御方式と比べて、より協調的でストレスのない引き込み制御ができる、

- (2) 会話エージェントは、人のコミュニケーション状態量の変化に適応的に振る舞うことができる、
- (3) 交替潜時だけでなく、韻律情報（ピッチ）は、協調的でストレスの少ない会話の実現に寄与し、韻律的なリズムの同調は会話全体の調整に大きな役割を果たしている、ことが明らかとなり、これらを実現するシステム仕様について明確化した。

また、本実験を通して、より親しみが持てストレスの少ない会話のもとで持続的な情報収集が行える会話システムの実現可能性を示すことができた。

今後は、本会話制御方式に基づき、強力な音声認識機能を付加し、遠隔（医師と患者間）で問診可能な「遠隔問診エージェント」の実現に向けて取り組んでいく予定である。

参考文献

- [1] 木村昌紀, 余語真夫, 大坊郁夫: 感情エピソードの会話場面における同調傾向の検討—擬似同調傾向実験パラダイムによる測定, 対人社会心理学研究, Vol.4, pp.92-99 (2004).
- [2] 長岡千賀, Maria Draguna, 小森政嗣ほか: 音声対話における交替潜時が対人認知に及ぼす影響, ヒューマンインタフェースシンポジウム予稿集, Vol.22, pp.171-174 (2002).
- [3] 長岡千賀, 小森政嗣, Draguna Raluca Maria ほか: 協調的対話における音声行動の 2 者間の一致—意見固持型対話と聞き入れ型対話の比較, ヒューマンインタフェースシンポジウム 2003 論文, Vol.2003, pp.167-170 (2003).
- [4] 三宅美博, 辰巳勇臣, 杉原史郎: 交互発話における発話長と発話間隔の時間的階層性, 計測自動制御学会論文集, Vol.4, No.6, pp.670-678 (2004).
- [5] 小松孝徳, 森川幸治: 人間と人工物との対話コミュニケーションにおける発話速度の引き込み現象, 情報処理学会研究報告—知能と複雑系, Vol.2004, No.105, pp.71-78 (2004).
- [6] 垣田邦子: 簡単な“質問—答”形式の対話における F0 の話者間相互作用, 日本音響学会研究発表会, Vol.1995, No.2, pp.305-306 (1995).
- [7] 西村良太, 北岡教英, 中川聖一: 音声対話における韻律変化をもたらず要因分析, Journal of the Phonetic Society of Japan, Vol.13, No.3, pp.66-84 (2009).
- [8] 志和敏之, 神田崇行, 今井倫太ほか: 対話ロボットの反応時間と反応遅延時における問投詞の効果, 日本ロボット学会誌, Vol.27, No.1, pp.87-95 (2009).
- [9] 渡辺富夫: 身体的コミュニケーションにおける引き込みと身体性—心が通う身体的コミュニケーションシステム E-COSMIC の開発を通して, ベビーサイエンス, Vol.2, pp.4-12 (2003).
- [10] 林 貴宣, 加藤昇平, 伊藤英則: パラ言語に基づいた会話ロボットの精神リズム同調モデル, 人工知能学会全国大会論文集, Vol.23, pp.ROMBUNNO.1H2-4 (2009).
- [11] 伊藤宏司: 身体知システム論, 共立出版 (2005).
- [12] 蔵本由紀: リズム現象の世界, 非線形・非平衡現象の数理 ①, 東京大学出版会 (2005).
- [13] ハーマン・ハーケン: 脳機能の原理を探る—非平衡協同現象としての脳神経活動・行動・認識, シュプリンガー・フェアラーク東京 (2000).
- [14] 市川 熹, 堀内靖雄, 西田昌史: マルチモーダル音声対話コーパスの収録とうなずきの分析, 広瀬・峯松研究

室 (オンライン), 入手先 (http://www.gavo.t.u-tokyo.ac.jp/tokutei_pub/houkoku/corpus/ichikawa.pdf) (参照 2013-01-06).

- [15] 高杉将司, 山本知仁, 武藤ゆみ子ほか: コミュニケーションロボットとの対話を用いた発話と身振りのタイミング機構の分析, 計測自動制御学会論文集, Vol.45, No.4, pp.215-223 (2009).
- [16] 多賀厳太郎: 脳と身体の動的デザイン—運動・知覚の非線形力学と発達, 金子書房 (2002).



速水 達也 (正会員)

2010年大阪工業大学大学院情報科学研究科情報科学専攻博士前期課程修了。現在, 大阪工業大学大学院情報科学研究科情報科学専攻博士後期課程在籍。



佐野 睦夫 (正会員)

1983年京都大学大学院修士課程修了。同年電電公社(現NTT)電気通信研究所入所。2002年大阪工業大学情報科学部教授, 現在に至る。ロボットビジョン, ロボットコミュニケーション, 料理メディア, 認知リハビリテーションの研究等に従事。



向井 謙太郎

2011年大阪工業大学大学院情報科学研究科情報科学専攻修士課程修了。2011年から, 株式会社日立ソリューションズに勤務。大阪工業大学大学院修士課程在学中に本研究に従事。



神田 智子 (正会員)

1996年マサチューセッツ工科大学メディアラボラトリー修士課程修了。2003年京都大学大学院情報学研究科社会情報学専攻博士後期課程修了。博士(情報学)。現在, 大阪工業大学情報科学部情報メディア学科教授。ヒューマン・エージェント・インタラクション(HAI)や擬人化エージェントを用いた異文化コミュニケーション研究に興味を持つ。



宮脇 健三郎 (正会員)

2008年大阪工業大学大学院情報科学研究科博士後期課程修了。博士(情報学)。現在, 同大学情報科学部情報メディア学科特任講師。ユビキタス調理支援等コンテキストウェアサービスの研究に従事。



笹間 亮平 (正会員)

2007年京都大学大学院修士課程修了。同年日本電気株式会社入社。2012年日本電気株式会社C&Cイノベーション推進本部所属, 現在に至る。コミュニケーションロボット, ヒューマンインタフェース等の研究を経て, イノベーション事業創出の研究に従事。



山口 智治

1990年広島大学大学院博士課程(前期)修了。同年日本電気株式会社入社。1997~1998年スタンフォード大学ロボティクス研究所客員研究員。2008年京都大学大学院情報学研究科博士課程(後期)単位取得認定退学。2012年日本電気株式会社C&Cイノベーション推進部主任研究員, 現在に至る。自然言語処理, ヒューマンインタフェース, コミュニケーションロボット等の研究を経て, イノベーション事業創出の研究に従事。



山田 敬嗣 (正会員)

1987年京都大学大学院工学研究科情報工学専攻博士後期課程修了。同年日本電気株式会社入社。現在, 日本電気株式会社理事兼C&Cイノベーション推進本部本部長。奈良先端科学技術大学院大学客員教授を兼任。専門分野は, 情報理論, パターン情報処理, ユビキタスシステム, ヒューマンコミュニケーション。情報処理学会理事, 電子情報通信学会理事, 国際パターン認識学会日本代表理事を歴任。電子情報通信学会, IEEE, 等会員。