

科学研究費補助金研究成果報告書

平成 24 年 5 月 17 日現在

機関番号：14301
 研究種目：基盤研究(S)
 研究期間：2007～2011
 課題番号：19100001
 研究課題名(和文) 会話エージェント研究共有プラットフォームの構築と利用技術の研究
 研究課題名(英文) Studies on Construction and Utilization of a Common Platform for Embodied Conversational Agent Research

研究代表者

西田 豊明 (NISHIDA, TOYOAKI)
 京都大学・大学院情報学研究科・教授
 研究者番号：70135531

研究成果の概要(和文)：

本研究の目的は、高度な会話エージェントシステム開発のためのさまざまなチャレンジが円滑にできるようにするための研究基盤と方法論を確立することである。研究成果は、会話エージェントシステム構築用プラットフォーム、コンポーネント技術、没入型WOZを用いた会話エージェントシステム開発環境、会話コーパスに基づく会話行動モデル開発方法論の開発、コンテンツ制作支援システム、評価手法の8項目から構成される包括的なものである。

研究成果の概要(英文)：

The goal of this research is to develop a platform and methodology for enabling researchers to conduct challenging researches on advanced embodied conversational agent systems. The result is comprehensive, covering a couple of generic platforms for the development of embodied conversational agent systems, component technologies, an immersive WOZ environment for the development of embodied conversational agent systems, a methodology for corpus-based development of embodied conversational agents, a content production assistance system, and a method for evaluating embodied conversational agents.

交付決定額

(金額単位:円)

	直接経費	間接経費	合計
2007年度	19,100,000	5,730,000	24,830,000
2008年度	17,600,000	5,280,000	22,880,000
2009年度	17,600,000	5,280,000	22,880,000
2010年度	17,600,000	5,280,000	22,880,000
2011年度	17,600,000	5,280,000	22,880,000
総計	89,500,000	26,850,000	116,350,000

研究分野:総合領域

科研費の分科・細目:情報学・知能情報学

キーワード:エージェント, 会話エージェント, 音声言語, 非言語コミュニケーション

1. 研究開始当初の背景

実用に資する会話エージェントシステムを開発するためには、音声処理、言語処理、画像処理、談話管理など多数のコンポーネントを統合し、並列実行するための強力な

計算基盤、会話エージェントのインタラクション行動制御モデルの開発支援環境、多様なコンテンツの制作支援環境が一体的に提供される必要がある。

2. 研究の目的

高度な会話エージェントシステム開発のためのさまざまなチャレンジが円滑にできるようにするための研究基盤と方法論を確立する。このために、次の目標を設定した。

- (1) 会話エージェントシステム構築用プラットフォームの開発
- (2) 基本コンポーネントセットの開発
- (3) コーパスに基づく会話エージェント行動モデル開発方法論の確立
- (4) 会話エージェント評価尺度の構成と実証実験実施

以上の研究開発は、会話エージェントシステムに関わる標準化活動とも連携を取りながら進める。

3. 研究の方法

研究期間前半では、ブラックボードモデルアーキテクチャモデルに基づき、並列動作するコンピュータ上で動作する会話エージェントシステム構築用プラットフォームの開発、コンポーネント最小セット開発、会話分析支援システムの開発、会話コーパスに基づく会話行動モデル開発方法論の構築、評価用パイロットシステムの開発、視聴覚情報に基づく評価手法開発を行う。研究期間後半では、より先進的な会話エージェントシステム構築用プラットフォーム、コンポーネントセット拡充、会話エージェントシステム開発支援環境、コミュニケーションモデル自律構築方式、コンテンツ制作支援システム、先進的な評価用パイロットシステムの開発、生理指標を用いた評価手法の開発を行う。最終年は、成果を取りまとめる。

4. 研究成果

研究成果は、会話エージェントシステム構築用プラットフォーム、コンポーネント技術、没入型 Woz を用いた会話エージェントシステム開発環境、会話コーパスに基づく会話行動モデル開発方法論の開発、コンテンツ制作支援システム、評価手法の 8 項目から構成される。本研究の成果全体の構成は、図 1 に示すように、会話エージェントシステム技術と会話エージェントシステム開発方法論に大別される。

(1) 会話エージェントシステム構築用プラットフォーム

(1a) GECA

並列分散環境・異種 OS 上の異種ツールで作られた部品を協調させて会話エージェントシステムとして統一性のあるインタラクションの実現を可能にする GECA を開発した。GECA は、分散計算環境における異種 OS とプログラミング言語への対応、分散環境上での負荷分散、リアルタイム性の高い応答実現に適した単層コンポーネント構成、多様なコンポーネント構成をもつ擬

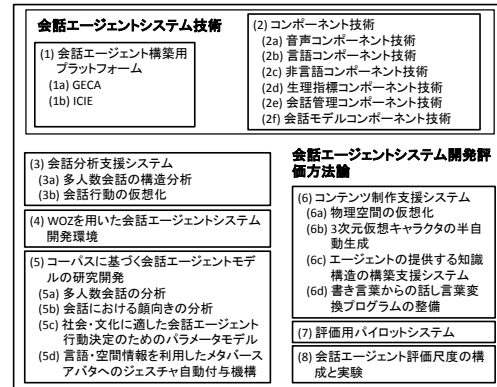


図 1: 成果全体の構成

人化会話エージェントシステムへの対応、タイミング制御と同期の実現、オンラインのコンポーネント・スイッチングとシステムの構造変更、低負荷のメッセージ転送、複数のプログラムによるコンポーネント並列開発のサポートが可能であるという特色がある。

MINDMAKERS.ORG に、中野と黄が Founding Member として参加した。研究協力者の黄が会話エージェントの行動記述標準化を行っている BML の会合に参加し、動向調査を行った。

会話エージェントの意図表現と感情表現を会話の高次機能を記述する FML を拡張することにより具体的に提案し、BML に変換する方法を示した。

(1b) ICIE

ユーザを取り囲むように配置された高さ 2m 程度の液晶モニタによる全方位視聴覚提示環境と、複数の非接触センサから構成される没入型インタラクション環境 ICIE (Immersive Collaborative Interaction Environment) の開発を行った。現実空間の仮想化ツール、キャラクタ生成・アニメーションツール、姿勢計測システムなどの開発を行い、ICIE の上に、仮想協調作業空間、テレプレゼンスシステム、およびハイブリッド協調作業空間を構築した。

(2) コンポーネント技術

音声コンポーネント技術、言語コンポーネント技術、非言語コンポーネント技術、生理指標センシングコンポーネント技術、会話管理コンポーネント技術、会話モデル学習コンポーネント技術から構成される。

(2a) 音声コンポーネント技術

音声言語認識機能をもつコンポーネントの開発とその高度化に関わる研究を行った。Julius と SynGraph を GECA システムに追加し、外部システムの深い部分に変更を加えることなく、ネットワーク経由で GECA に接続するための GECA-Proxy コンポーネントを開発した。Julius のラッパーコンポーネントを改良し、キーワードス

ポッシング機能を実現した。音声認識エンジン Julius, 自然言語解析パーザ KNP, 音声合成エンジンをモジュール間通信で結合した形で音声対話システムを実装し, 本ソフトウェアは, 共同研究等を通して, 複数の企業に供与した。

音声対話エンジンの高度化に関して次の成果を得た。人間どうしの自然な音声会話を収録して分析を行い, 3つのあいづち形態が, 韻律パターン(ピッチ・パワー・持続時間長など)の変動が大きく, それぞれ異なる韻律パターンによって話の内容に対する興味・関心を示していることを明らかにした。さらに, 有声休止に着目した韻律的特徴ならびに音声認識に基づく音韻的特徴を統合することによって, これらのあいづちを音声の中から検出する方法を実装し, 70%程度の検出率を実現した。

(2b) 言語コンポーネント技術

ネット上の言語資源を活用して言語コンポーネントを開発する技術を開発した。WWW コーパスの機械解析結果から計算される文要素関の共起尺度を観測変数として用い, 変数間の相関を因子分析モデルによって分析し, 多くの抽出因子について何らかの言語現象として解釈が可能であることが確認できた。文末に含まれるモダリティの同定技術の改良をおこない, 従来, ルールベースの解析器を利用することで行われてきたモダリティ解析に統計的な文脈知識を利用する手法を導入することによって, 同定精度が改善することを示した。

話し言葉の解析・生成において重要となるモダリティ表現の自動推定を行なった。「同じモダリティである表現は同じような機能語と共起しやすい」という分布仮説に基づくコーパスベースの手法により, 約 85%の精度でモダリティを自動推定できることを確認した。

会話エージェントの知識源として言語知識を利用可能とするために, 国語辞典の定義文からの語彙知識獲得と, 事典に現れる名詞句の自動分割を行う手法を開発した。

(2c) 非言語コンポーネント技術

非言語的なインタラクションに必要なコンポーネントを開発するとともに, それにより, 会話の雰囲気や推定できることを示した。

GECA 上の基本的なコンポーネントとして, 頭部追跡, 加速度センサー入力, データグローブ入力, モーションキャプチャ入力, CG アニメーションプレイヤーのためのコンポーネントを開発した。複数の Microsoft Kinect による身体動作計測サブシステムも開発した。圧力センサを用いた歩行行動の認識のためのコンポーネントも開発した。

顔向きによるユーザ間会話の状態推定モデルを構築するため, アの顔向き変化と発話区間, 及びそれらに付随する時間情報から, 決定木学習によりユーザペア間の会話状態を推定するモデルを構築した結果, ペア間での相談, 雑談, およびエージェントへの質問の3つの状態を約 70%の精度で区別することができた。ヘッドトラックから得られる顔向きの情報に加え, 韻律情報として, ピッチ(F0), パワー, 話速に着目し, これらと顔向きの情報を組み合わせることにより, 受話者が他のユーザである場合の F 値が 0.807, エージェントである場合は 0.752, そして全体の予測精度を 78.3% にまで向上させることができた。

(2d) 生理指標センシングコンポーネント技術

映像と生理指標を用いて会話中の場の盛り上がりと個人の盛り上がりの関係を分析し, 映像による盛り上がりの判定と生理指標による盛り上がりとの関係を明らかにした。エージェントによる場の盛り上がりと個人の盛り上がりの自動判別手法を実装し, 予備実験を行った。得られた結果はまだ不十分であるが, ユーザセンシングを改善することにより, 盛り上がりの自動判別手法を開発できるという見通しが得られた。生理指標を用いたユーザの内部状態推定手法を発展させて, 対話中にユーザの選好構造を形成していく手法の開発を行った。ロボットの組み立てタスクについて人間同士で対話を行い, 映像と生理指標を計測する予備実験を行った結果, 新しい要因の発見や優先順位の変化と, 言語情報, うなずき, SCR の上昇, 心電の LF/HF 成分の上昇との関係が確認できた。この結果をふまえ, ユーザの言語情報, うなずき, SCR, 心電, 皮膚温, 提案の二者選択を入力として, ユーザの重視要因やその変化を推測して次の提案を決定するエージェントの提案手法を考案し評価を行った結果, 重視度の大きな変化が確認でき, 重視度の推測結果, 提案の満足度等のほとんどの項目において統計的に有意によい結果が得られた。

意識的な要因と潜在的な要因を提案に反映させる手法を複数検討した。知識構造に基づいて潜在的な部分を複合的に変化させた場合とさせなかった場合を比較したところ, 提案の自然さ, 提案内容の参考度において, 複合的に変化させなかった場合の方が有意によい結果を示した。複合的に変化させた場合の方が, 提案回数は少ない傾向にあり, すべての選択肢の中から実験参加者が選んだ提案と本研究の手法が選んだ提案の一致度も高かった。予備実験により, 意見の対立を避けるか, 結果の質を上げるかというスタンスの違いによって, 自らの

重視要因をグループの意見に反映させようとする場面が異なることが明らかになった。グループ全体の重視要因を、それぞれ異なる推定手法を使い分けて推定する手法について検討した結果、2つのグループ重視度動的推定法を使い分けることで、高い確率でグループの選択結果とメンバそれぞれが最良だと思ふ選択を一致させることが出来た。また、満足度に関して意見の対立の有無と2つのグループ重視度動的推定法の間有意傾向があり、対立を避けるか結果を優先するかで使い分けることが妥当であることがわかった。

(2e) 会話管理コンポーネント技術

エージェントと会話している多人数ユーザの状態を計測して、中心人物や盛り上がりの状態を推定する手法を開発した。映像からでも、人間が考える最小の盛り上がりまで視野に入れれば場に引張られる小さな盛り上がりを検出できることがわかった。

音声認識、顔認識、会話プランニング、エージェントアニメーション、コーパスからの人間の会話行動推定モデルを統合し、多人数会話における対話制御方式を示した。それに基づき、ユーザ同士の会話に介入する会話エージェントのプロトタイプを実装した。本システムでは、ユーザの顔向きから会話の状態を推定し、ユーザが相談中であると推定された場合は、ユーザペアに情報提供を申し出たり、雑談をしていると判定された場合には、本題に戻るよう促すといったように、ユーザペアの会話にエージェントが介入することができる。

プロアクティブな会話エージェントを実現するために音声対話エンジンの高度化を行った。情報検索・質問応答に基づく対話の枠組みを拡張し、述語構造解析に基づく部分的なマッチングを導入することにより、検索要求／質問に合致するものがなくても関係する情報を推薦できるようにした。述語項構造に着目した Web 上のテキストからの情報抽出に基づく雑談的な対話生成手法を開発し、音声対話システムに組み込んだ。さらに、複数の応答生成モジュールでの発話の候補から最適なものを選択することで対話を実現し、強化学習によって対話戦略を自動獲得するようにした。

(2f) 会話モデル学習コンポーネント技術

教師なし学習と教師あり学習を組み合わせ、非言語的なインタラクションの模倣学習を実現するための基本アルゴリズムを構成した。この手法は、相互行為のバブリング、相互行為の構造学習、相互行為中の適応の3段階から構成される。この枠組みを用いて一定の自然さを持つ対話的視線行動を実現可能であることを確認した。

モーションキャプチャのデータを自動的にセグメンテーションし、頻出パターンを抽出する方法と、追加学習の途中で段階的な特徴抽出を行い、入力データに応じた部分特徴空間を動的に抽出する手法を開発した。

従来開発していた変化点発見とモチーフ発見プログラムを核とし、さらに評価用信号生成ルーチン、評価ルーチンを含むツールパッケージを開発し公開した。

(3) 会話分析支援システム

(3a) 多人数会話の構造分析

ポスターを参照しつつ複数の人が会話するポスター会話を題材として会話を収録し、指差し、視線、発話について自動的にアノテーションを行い、そこから N-gram を生成することにより、非言語的なインタラクションの時間構造を抽出する手法を開発した。

3人会話データから特徴的な会話プロトコルの抽出を試みた。自由会話には存在するが、ポスター発表には少ない会話構造や、ポスター発表会話には存在するが応答場面には少ない会話構造の存在を見出した。

(3b) 会話行動の仮想化

Kinect を用いて、6m × 6m 程度の空間における会話行動を背景とともに仮想化してさまざまなアングルからの分析を可能にする分析支援環境を構築した。環境としてスタティックなものは事前に取り込んだ距離情報と RGB 画像から計算することで、2cm 程度の粒度で取り込むことができる。人間の情報を取得する上では、空間的・時間的にキャリブレーションを行う手法を開発することで、オクルージョンやデータ同期の問題に対応した。人間の詳細な表面情報を取得するシステムは別のシステムとして開発した。骨格情報と、人間の表面情報を組み合わせることで、違和感のないインタラクション場面の再生が可能になる。このシステムと時系列分析のシステムを組み合わせる実際の複数の人間におけるインタラクションを分析するシステムを開発した。

システムに蓄積されている教師の情報と、システムからリアルタイムに得られる生徒の情報をマッチングし、適切な助言を行ったり、タスクを達成する上での Tips を蓄積するシステムを開発した。実際に社交ダンスという難しい動作の習得に利用して評価したところ、教師データと生徒データの差が大きいところを発見して生徒に提示できており、ダンス習得の役に立っていることを確認した。また、本システムをチュートリアルのように使い、繰り返し練習することによってどの生徒役の被験者も各フェーズごとに上達していくことが観察された。

(4) WOZを用いた会話エージェントシステム開発環境

インタラクションにおける人間の行動データの取得やヒューマンエージェントインタラクションにおける問題の分析を行うために ICIE の上にテレプレゼンス機能を開発し、没入型 WOZ 環境を開発した。実験者が会話エージェントまたはロボットの内側に入り込んだ着ぐるみ感覚で、普通の人間とインタラクションを行い、人間と、人間に相当する能力を持ったエージェントがインタラクションする WOZ 実験を実施し、インタラクションログの記録と分析ができる。没入型全方位投影システム、複数台のレンジセンサなどを組み合わせて、人間への行動制約と認知的負荷を軽減している。

遠隔地の実世界センシングと ICIE 内の操作者の身体動作計測を同時に考慮し、コミュニケーションインターフェースとなっているロボットが、半自律的にコミュニケーション行動を生成するコンポーネントを開発した。実際の遠隔コミュニケーションタスクに対して評価実験を行ったところ、特に非言語行動による意図の伝達において有用であることが示された。

(5) コーパスに基づく会話エージェント行動モデルの研究開発

(5a) 多人数会話の分析

情報提供を題材にして、複数の実験協力がエージェント(実験協力者に知られていない状態で人間が操作する)と会話をする実験で収録されたユーザの発話のタイミングと顔の動きなどの非言語情報を基に決定木を生成し、発話対象を 60%程度の精度で推定できた。

(5b) 会話における顔向き分析

ショッピングモールの情報提供エージェントを想定した 60 会話を収録し、OKAO vision を用いて会話中の被験者(来客者役)の顔の位置と回転角度のデータを収集した。収集した OKAO vision のデータを用いて、被験者の顔向き(正面, 左, 右, 下)を判別する決定木学習を行った結果 93.5%の精度が得られ、顔向きの自動ラベリングが可能であることを確認した。さらに、得られた顔向きデータと会話内容との関係を分析し、被験者がエージェントに話しかけているのか、被験者同士で話しているのか、またそれが課題についての相談であるのか雑談であるのかによって、二者の顔向きのパターンや顔向き変化に要する時間が異なることを明らかにした。この成果は非言語コンポーネント開発に活用された。

(5c) 社会・文化に適した会話エージェント行動決定のためのパラメータモデル

アウグスブルグ大学と共同で収集した 2 カ国間での対照可能な CUBE-G コーパス

の分析を進めた。初めて出会う学生同士の会話と、彼らの課題遂行結果を評価する立場の人との会話とを比較した結果、後者の状況において、姿勢変化の頻度が減少し、より小さく、かたく緊張した姿勢が頻出することが明らかになり、社会的関係によって、姿勢の出現特徴が異なることがわかった。心理学の関連研究を参考にし、社会的関係を特徴づける 4 つのパラメータを提案した。さらに、これらのパラメータをノードとし、ベイジアンネットワークモデルに追加することにより、文化、社会的関係、そして非言語行動との相互関係を表現するモデルを構築した。

(5d) 言語・空間情報を利用したメタバースアバターへのジェスチャ自動付与機構

メタバースアプリケーションにおいて、道案内を行うアバターの発話に適切なジェスチャを自動付与する機能の実現を目指し、道案内対話の収集実験の実施、モーションキャプチャによるジェスチャデータの収集を行った。収集データの予備的な分析では、会話参加者の立ち位置によって、適切なジェスチャの形態が異なることが分かった。

ジェスチャの形態を規定する空間情報として、会話参加者の立ち位置の情報が有用であることを見出した。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、方向、位置関係、動作、ランドマークを表す言語表現がジェスチャ形態の決定に有用であることがわかった。これらの分析結果を統合し、セカンドライブ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、セカンドライブ上に道案内アバターシステムを実装した。

(6) コンテンツ制作支援システム

(6a) 物理空間の仮想化

通常の写真から、屋外の広い空間を、奥行き情報付きで没入型環境へ投影できる仮想空間構築システムを開発した。比較的広い範囲(20m×20m 程度)の実空間を 1200 枚程度の写真から、4 日間程度の時間で復元するシステムを開発し、HAI における背景的部分を実世界情報に基づいてコンテンツ化することを実現した。計算を高速化することでさらに広範囲(50m×50m 程度)の実空間を同程度の時間で復元する手法も開発し、実用化を進めた。

(6b) 3次元仮想キャラクターの半自動生成

実在人物の特徴を維持した 3次元仮想キャラクターの半自動合成手法を開発した。Kinect を利用し静止した 4 点から表面情報を取得する手法と、Kinect を人物の周りで動かしながら表面情報を取得する手法を実装・評価した。

(6c) エージェントの提供する知識構造の構築支援システム

エージェントの提供する知識構造の構築を補助するシステムを開発した。知識構造における人間の感覚的な部分を可視化し、他人と共有することで、知識の円滑な伝達と再構築が可能になる。知識の継続的な追加に対応するために、クラスタリングの手法を適用して、ユーザの知識構造の再構築や目的の知識の円滑な閲覧を助けるシステムを実装した。

(6d) 書き言葉からの話し言葉変換プログラムの整備

論文や報告書などの書き言葉で書かれたテキストを会話エージェントで読み上げることを目的とし、GECAプラットフォーム上で利用できるように書き言葉からの話し言葉変換プログラムの整理を行なった。

(7) 評価用パイロットシステムの開発

GECAを用いた多人数会話システムを2008年8月にパリで開催されたeINTERFACE'08において開発した。独立行政法人が開催している一般公開の来訪者が実際に体験できるオンサイト食品安全クイズエージェントを開発した。没入型多人数インタラクティブプラットフォームVISIEを開発した。軽症の認知症患者や健全な高齢者の日常生活を支援するシステムの研究開発を進めた。

(8) 会話エージェント評価尺度構成と実験

IATを使って評価し、有用性の評価を行った。アンケート調査とGNATを組み合わせて会話エージェントを評価する方法も開発した。潜在的態度の実験計画記述言語ExPML, ExPML記述を実行し、データを採取するインタープリタを実装した。実験方法を改良し、IAT評価が誤りの影響を受けにくくした。また、誤りを提示するために生じる実験参加者の記憶効果を打ち消すための手法とスコア計算方法を考案し、実験によりその有効性を確認した。

5. 主な発表論文等

[雑誌論文] (計 37 件)

- ① Matthias Rehm, Yukiko I. Nakano, Elisabeth André, Toyoaki Nishida, Nikolaus Bee, Birgit Endraß, Michael Wissner, Afia Akhter Lipi, Hung-Hsuan Huang: From observation to simulation: generating culture-specific behavior for interactive systems. *AI & Society* 24(3): 267-280 (2009)
- ② R. Gomez and T. Kawahara: Robust speech recognition based on dereverberation parameter optimization using acoustic model likelihood. *IEEE Trans. Audio, Speech & Language Process* 18(7): 1708-1716, 2010.

[学会発表] (計 188 件)

- ① Birgit Endrass, Elisabeth Andre, Yukiko Nakano, Matthias Rehm, Afia Akhter Lipi: Culture-related differences in aspects of

behavior for virtual characters across Germany and Japan, in the proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011), pp. 441-448, 2011.

- ② Hung-Hsuan Huang, Takuya Furukawa, Hiroki Ohashi, Toyoaki Nishida, Aleksandra Cerekovic, Igor S. Pandzic, Yukiko I. Nakano: How multiple concurrent users react to a quiz agent attentive to the dynamics of their game participation. AAMAS 2010: 1281-1288

[図書] (計 2 件)

- ① 西田豊明, 角康之, 松村真宏: 社会知デザイン, 人工知能学会編, 知の科学シリーズ, オーム社, 2009.
- ② Toyoaki Nishida, Lakhmi Jain, Colette Faucher (eds.): Modelling Machine Emotions for Realizing Intelligence: Foundations and Applications, Smart Innovation, Systems and Technologies Series, Springer, 2010.

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

- 受賞 8 件 (Finalists, プロジェクト所属学生単独受賞を含む)
- 国際会議基調講演/招待講演 19 件
- 分担著 4 件
- 報道 4 件
- 特記事項 8 件

6. 研究組織

(1) 研究代表者

西田豊明 (NISHIDA, TOYOAKI)
京都大学・大学院情報学研究科・教授
研究者番号: 70135531

(2) 研究分担者

河原達也 (KAWAHARA, TATSUYA)
京都大学・学術情報メディアセンター・教授
研究者番号: 00234104

黒橋禎夫 (KUROHASHI, SADA0)
京都大学・大学院情報学研究科・教授
研究者番号: 50263108

中野有紀子 (NAKANO, YUKIKO)
成蹊大学・理工学部・准教授
研究者番号: 40422505

角康之 (SUMI, YASUYUKI)
はこだて未来大学・複雑系知能学科・教授
研究者番号: 30362578

大本義正 (OHMOTO, YOSHIMASA)
京都大学・大学院情報学研究科・助教
研究者番号: 90511775

黄宏軒 (HUANG, HUNG-HSUAN)
立命館大学・情報理工学部・助教
研究者番号: 00572950