

機関番号：14301

研究種目：特定領域研究

研究期間：2006 ～ 2010

課題番号：18049045

研究課題名（和文） 実世界インタラクションの分析・支援・コンテンツ化

研究課題名（英文） Analysis, Support, and Materialization of Real World Interaction

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）：われわれの日常会話は、会話の場に含まれる参加者同士および事物の間の精巧な相互作用(実世界インタラクション)から構成される。本研究では、実世界インタラクションを計測し、分析することにより、コミュニケーションの豊かさをもたらす要因を解明するとともに、実世界インタラクションを増強し、そこから教育や訓練のための新たなコンテンツを自動獲得するための新しい手法の開発に取り組んだ。成果は、(a) 脳・生理計測を用いた実世界インタラクションの分析、(b) 実世界インタラクションのデータ収集と分析、(c) WOZ (Wizard of Oz)によるインタラクションロボット開発環境、(d) 実世界インタラクションの行動記録からのロボットの行動自動生成、(e) 実世界インタラクションの支援・記録・再利用からなる。

研究成果の概要（英文）：Our daily conversation consists of a sophisticated interaction – real world interaction – among participants and the environments. This research addresses uncovering principles underlying real world interactions through quantitative measurement and analysis and developing a computational method for augmenting real world interactions to elicit interactive contents reusable across varying conversational environments. The outcome comprises (a) analysis of real world interactions by brain and physiological measurements, (b) data collection and analysis of real world interactions, (c) the WOZ-based interactive robot development environment, (d) automatic generation of communicative behaviors of robots from real world interaction log, (e) reusing real world interaction log.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2006 年度	29,600,000	0	29,600,000
2007 年度	26,300,000	0	26,300,000
2008 年度	21,000,000	0	21,000,000
2009 年度	9,100,000	0	9,100,000
2010 年度	9,100,000	0	9,100,000
総計	95,100,000	0	95,100,000

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：社会知デザイン、会話情報学

1. 研究開始当初の背景

情報システムの多機能化と高機能化によってもたらされる複雑さを克服し、人間が安

心して利用できる快適な情報環境を構築するには、人間と情報システムとの間に存在する様々なレベルでのギャップを、実世界にお

けるマルチモーダルインタラクションによって解消することが必要である。先行する学術創成研究「人間同士の自然なコミュニケーションを支援する智能メディア技術」(平成13年度~17年度)では、人間にコンピュータの存在を意識させず、人間のもつ豊富なコミュニケーションスキルを活用した人間同士の自然なコミュニケーションを可能にする智能メディア技術を研究開発することを目的とし、コンピュータを環境の背景に埋め込んでコミュニケーションをさりげなく支援する環境メディア技術、逆にコンピュータを人間と普通に会話できる社会的主体として人間同士の会話を媒介させる会話エージェント技術、両者を実現するための技術基盤となる会話のコミュニケーション・モデルの開発に焦点をあてた取り組みが行われた。この取り組みの中で、環境メディア技術、会話エージェント技術、会話のコミュニケーション・モデルからなる会話情報学の枠組みが提案され、その中心概念が実現された。会話的なインタラクションを支援しつつ、活性化された会話から別の文脈で使用できるコンテンツを抽出する技術を開発することが次の課題として同定された。

2. 研究の目的

本研究では、実世界インタラクションの理解と支援技術の研究開発を目的とし、(a) ユビキタスセンサ技術を利用した実世界インタラクションの計測と分析、(b) 人間と相互適応的に身体インタラクションを行う会話ロボットを用いた人間の行為の見守りと状況に応じた知識提供による会話支援、(c) 意味あるインタラクションを捉え場面を越えて適用可能なコンテンツとして結晶化させるコンテンツ化技術を中心とする研究開発を行う。これらの研究は、いずれも従来の研究で一般化しつつあるマルチメディアデータに加えて、生理指標まで含めた統合的なデータをリアルタイムで取り扱えるレベルまで進めることを目標とする。これらの研究を統合的に進めるために、支援班で開発する実世界インタラクション計測分析環境(IMADE)を基盤とするシステム開発を行う。複数人数の会話や共同作業を題材とする実世界インタラクションを共通題材として研究分担者間の研究成果の統合を図る。

3. 研究の方法

本研究では、図1に示すように、センサ、ディスプレイ、ロボットなどで拡張された実世界環境 IMADE (real world Interaction, Measurement, Analysis and Design Environment)と没入型協調的インタラクション環境 ICIE (Immersive Collaborative Interaction Environment)を接続し、実世



図1. 実世界インタラクションの分析と活用のための技術の枠組。

界と仮想世界を連携させた実験環境を構成することにより、実世界インタラクションの分析・支援・コンテンツ化の課題に取り組んだ。

4. 研究成果

(1) 脳・生理計測を用いた実世界インタラクションの分析

実世界インタラクションを人間の脳・生理活動から説明し、インタラクションに活用する試みを行った。主な成果として、事象関連電位によるアニメシー知覚の生理指標の確立、時系列変化点検出アルゴリズム (RSST) を使ったインタラクションの自然さの評価法、生理指標に対して信号処理と筋電の特性を生かしたインタフェース設計論「EMGUI」の提案があげられる。

対象が生物であるときと非生物であるときとでリーチング行動に伴う事象関連電位に現れる違いを図2に示す。これは、アニメシーの知覚には少なくとも二つの下位メカニズムが関与している可能性を示唆している。左前頭部での違いは、行動開始時に起こることから、実験参加者の主観的な生物としての認識、すなわち対象に生物らしさを帰属させる認識との関係が推察される。また、右側頭部での違いは、リーチング中に起こることから、生物の運動の予測との関係が推察される。

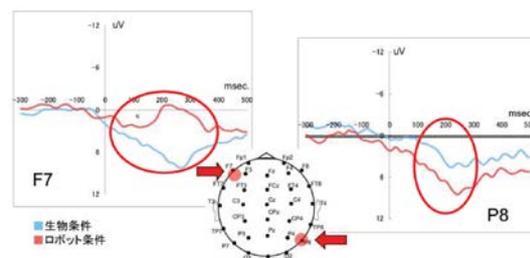


図2. 条件間に差がみられた電極(F7, P8)における事象関連電位を示す。横軸は時間を示し、リーチング行動の開始地点を0とした。赤丸で囲まれた部分に差が見られたことを示している。

人間・ロボット・インタラクションにおけるエージェントのふるまいの自然さを、コミュニケーションパートナーである人間に与える負の感情（ストレス、認知的負荷、心配、フラストレーション、…）の小ささと規定し、皮膚抵抗、脈波、呼吸数から得られる8個の生理指標（皮膚コンダクタンス：SCL (Skin Conductance Level), GSR (Galvanic Skin Response), 脈波：HH (Heart Rate), HRV (Heart Rate Variability), P (raw Pulse data), 呼吸数：RR (Respiration Rate), RRV (Respiration Rate Variability), R (raw Respiration pattern))を用いて評価する手法を開発した。信号処理によって外れ値を除去し、我々が開発した時系列変化点検出アルゴリズム RSST (Robust Singular Spectrum Transform)を適用して、時系列の背後にある内部状態のダイナミクスの変化を検出して、特徴量に組み入れてから、判別を行う。全体

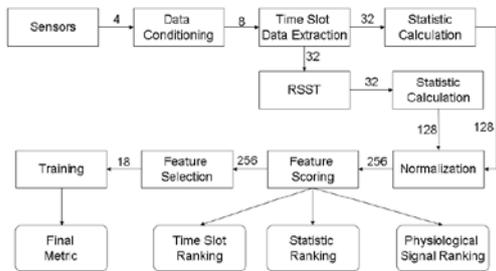


図3. 生理指標によるインタラクション判別法。

の処理過程を図3に示す。

(2) 実世界インタラクションのデータ収集と分析

多人数会話を記録・分析するための環境 IMADE(図4)を構築した。この研究では、ハードウェア面として、時間同期した複数映像記録を実現するためにネットワークカメラ、各自の発話を独立して録音するためのワイヤレスマイク、複数人の身体移動や身振り・

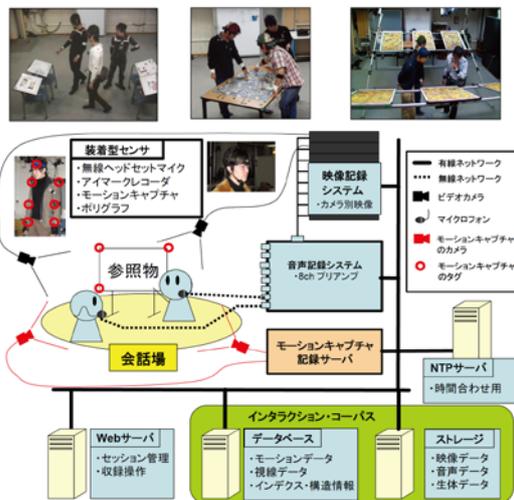


図4. 多人数会話の計測環境 IMADE.

手振りを計測するためのモーションキャプチャシステム、視線移動を計測するための眼球運動計測装置を導入し、それらを統合的に記録・閲覧するためのソフトウェア環境を実現した。

多人数会話における言語・非言語の時間構造分析を可能にするためのソフトウェア環境 iCorpusStudio を構築した(図5)。iCorpusStudio は、複数映像や音声の同期再生、ラベリング付与とラベル間の演算が可能であり、モーションデータや視線データの可視化、それらから解釈された注視対象変化やジェスチャに関するラベルデータを一括し

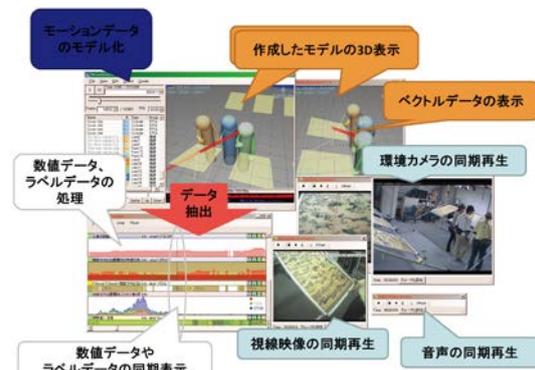


図5. マルチモーダルデータ分析環境 iCorpusStudio.

て扱うことができる。

ポスター発表、自由会話、ボードゲーム等の様々な種類の多人数会話データを収録し、それらの会話状況に共通あるいは特有の会話プロトコルを分析するために、非言語インタラクションの時間構造を解析する手法(インタラクションマイニング)を開発した。

(3) WOZ (Wizard of Oz)によるインタラクションロボット開発環境

人と持続的にインタラクションを発展させていくことが可能なエージェントを実現するために、相互適応現象に注目して研究を進めた。人間同士のプロトコル分析、人間・手動操作ロボットのインタラクションのプロトコル分析の2段階を経て、相互適応の詳細に関する知見を得たうえで目標とする自律ロボットを構築するアプローチ(3段階アプローチ)を取った。

第1段階については、人間同士の経路案内タスクを観察し、お互いに意思疎通を効率的に行うためのプロトコルを持たない場合に、人間同士がどのようにプロトコルを成立させていくのかを分析した。

第2段階については、人間・手動操作ロボットのインタラクションのプロトコル分析をするために図6のような実験環境を構築した。人間の指示者に経路を与え、その経路に沿ってロボットを走行させ、パフォーマンス

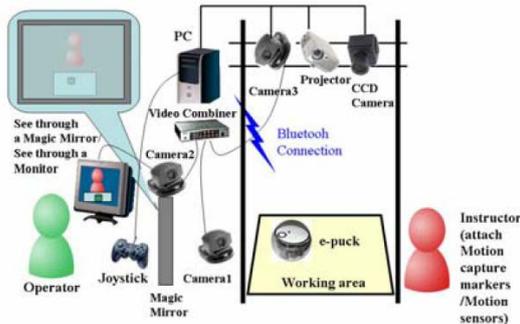


図 6. 人間・手動操作ロボットのインタラクションプロトコル計測環境.

を評価するというタスクを設定した。人間の指示者（被験者）にはロボットはジェスチャを知覚する能力があると伝える一方で、被験者には指示者に気付かれないよう指示者のジェスチャとロボットの現在位置をロボット操作者（被験者）に提示し、指示者の意図をくみ取ってロボットを走行させるよう指示した。指示者ロボットが小さくて操作しにくかったので、ロボット操作者に対する情報提示方法や操作方法を工夫した。予備実験の結果、指示者のジェスチャが途中から平易なものに置き換えられるなど、相互適応の兆候が観察された。

以下の5つを相互適応現象の誘発条件として提案した。

- ① インタラクション相手から、自分の行動に対する直接的な反応もしくは報酬を受け取ることができる。
- ② 自発的な行動によって自らの内部状態（行動意図など）をインタラクション相手に表現することができる。
- ③ タスクを遂行するに当たって必要な基本的なプロトコルを理解できる。
- ④ インタラクション相手がどのような適応能力を持っているかが理解できる。
- ⑤ インタラクション参加者個別の能力や情報だけでは、タスクを達成できないか、著しく効率が悪い。

実際にこれらの実験環境とエージェントを用いて実験を行った。実験では、実験参加者（25名が参加）は店のマネージャの立場から、エージェントに接客の指示（席への案内、注文取り、テーブルの片付け）を与え、エージェントは、マネージャである参加者の指示と仮想店舗の席の状態、および、自らが受け取ることのできる報酬に従って適応学習を行った。エージェントが前述の相互適応の誘発条件を満たしていれば、人間同士の場合のように、人間はエージェントの学習状態に応じて出す指示の頻度を変化させるなどといった適応を行い、エージェントは指示の変化に適応して過去の学習状態を考慮して新しい

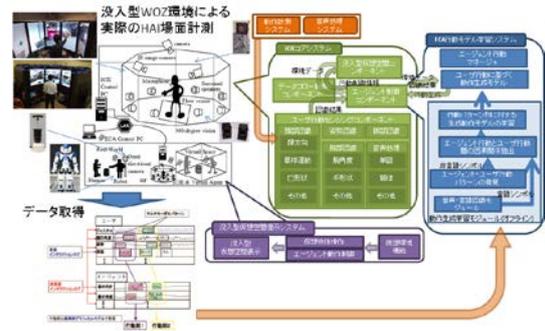


図 7. 没入型協調的インタラクション環境 ICIE (Immersive Collaborative Interaction Environment)を用いた HAI 開発環境.

適切な学習を行うと考えられる。このように、前述の実験環境とエージェントを作成することによって、相互適応現象の誘発条件を実験的に検討することが可能になった。

実験環境とエージェントの学習、および実験を行った結果から、前述の誘発条件を検討した。誘発条件①、③、⑤は、実験環境と学習の条件によって達成されるように設計されていた。このことは、実験時のアンケートからも確認された。また、44%の参加者がエージェントの適応に気がつくことができ、しかも、エージェントの振る舞いを予想することができた。これは、誘発条件②、④をある程度満たしていたことを示している。これらによって、本実験環境とエージェントが、提案した相互適応の誘発条件を満たしていたことが確認された。

一方、人間による指示のログとエージェントの行動ログを、Change Point Detection と Motif Discovery を用いてセグメンテーションとクラスタリングを行なった。それらのデータを分析したところ、エージェントの適応状態によって人間の指示が変わり、その人間の指示の変化に応じて過去の学習を壊さずにエージェントの適応が変わるといふ、相互適応的な行動を機械的に発見できた。このような行動は学習手法の種類によらずに出現した。従って、エージェントの適応学習の種類によらず、提案した発生条件によって相互適応現象が誘発されることを明らかにできた。

WOZ による強力なインタラクションロボット開発環境を構築するために、没入型協調的インタラクション環境 ICIE (Immersive Collaborative Interaction Environment) を用いて、ロボット操作者 (Wizard) がロボットの置かれている状況を認知的な負荷なしに直感的に把握し、特別なデバイスを用いることなく直感的にロボット操作ができる HAI 開発環境 (図 7) を試作した。没入型インタラクション環境を用いて、マッチングやキャリブレーションを高速化するための技法をいくつか組み込んでいる。プロトタイプ

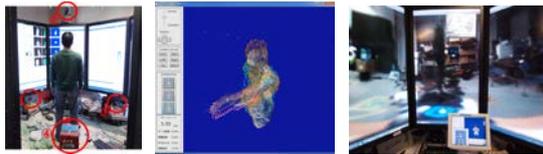


図 8. 没入型インタラクション環境における姿勢認識とロボットの身体動作への反映。(a) レンジセンサ 4 台を用いたセンシング, (b) 距離データの Voxel 化と人体モデルのフィッティング, (c) WOZ 操作環境

の動作例を図 8 に示す。距離情報を利用した姿勢認識を行っているため、操作者の背後のディスプレイの映像が切り替わっても認識に影響を与えない。

また、360 度のパノラマ画像とそれに対応する奥行き情報および存在するオブジェクトの領域情報を生成し、これらの情報を没入型スクリーンへ提示する手法を開発した。本手法では、撮影後の画像処理によってカメラパラメータを推定するため、撮影時点において大がかりな機材を必要とせず、レンジセンサなどを利用する手法に比べて実世界での作業時間を短縮することができる。また、一般的なカメラが進입可能であれば復元が可能であるため、狭く、入り組んだ場所や車両が進入できない場所を復元することもできる。さらに、ある程度の奥行き情報が復元されていれば、マンハッタンワールド仮説に従わないオブジェクトに関しても復元可能である。

(4) 実世界インタラクションの行動記録からのロボットの行動自動生成

実世界インタラクションを計測して得られたログから模倣学習によって行動パターンを自動生成するための手法 (図 9) を開発した。この手法では、収集された指示系列と行動系列から頻繁に繰り返されるセグメント (モチーフ) を抽出し、因果関係を推定して対応付けを行うことにより、行動者のモデルを推定する。この手法を用いて撮影ロボットなどを試作した。

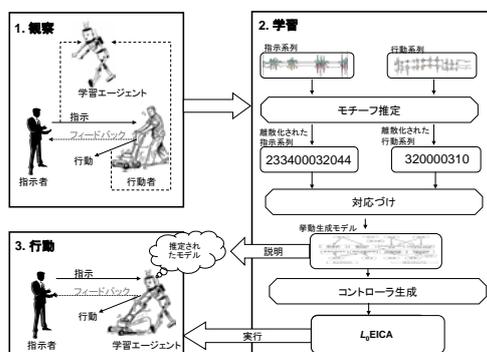


図 9. 模倣学習によるインタラクションパターン獲得.

(5) 実世界インタラクションの支援・記録・再活用

TV 番組などで説明者を補助する人間のアシスタント (アナウンサーやインタビューなど) の機能を模倣し、説明者との様々なインタラクションを活用し、説明者が話しやすくなり、説明を視聴者にとってわかりやすくしたり、その場の雰囲気を良くしたりする機能をもつ仮想アシスタント (図 10) を試作した。仮想アシスタントが補助することによって、視線方向の改善など視聴者を意識した振る舞いが増えること、撮影中の説明行動が増えること、説明者がそれをあまり負担に感じないことなどがわかってきた。また、人間を取り巻く情報システムがその「場」や「状況」を記録し、利用者の必要に応じてそれらを提示する機能を「環境記憶」とし、それを用いて利用者の行動支援を行うモデルを提案し、

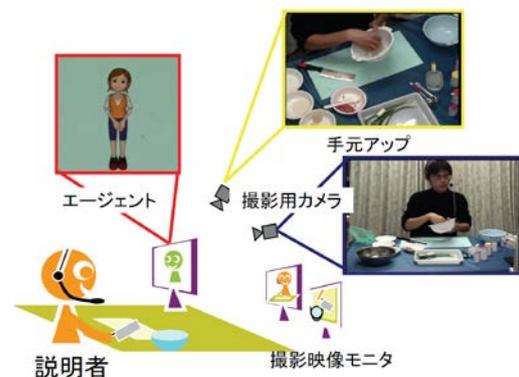


図 10. 仮想アシスタントを用いた知識コンテンツ取得.

その有効性を検証した。

会議の参加者やその場の状況を自動認識し、それを議事録として実時間で提示して議論の支援とするための基礎的研究を行った。参加者の頭部ジェスチャなどの非言語行動から「コメント」、「了承」、「質問」などの形式的役割を 80%以上の精度で認識できることがわかった。また、会議のコンテンツ化に関しては、会議の進行状況を可視化するツールを作成した。

(6) 総括

本研究では、実世界インタラクションを計測し、分析することにより、コミュニケーションの豊かさをもたらす要因を解明するとともに、実世界インタラクションを増強し、そこから教育や訓練のための新たなコンテンツを自動獲得するための新しい手法の開発に取り組んだ。成果は、(a) 脳・生理指標を用いた実世界インタラクションの分析、(b) 実世界インタラクションのデータ収集と分析、(c) WOZ によるインタラクションロボット開発環境、(d) 実世界インタラクションの

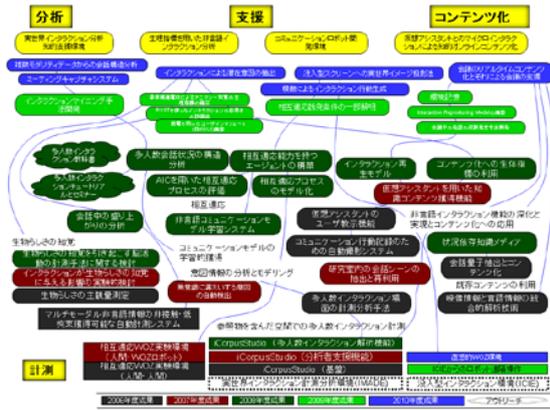


図 11. 本研究の展開の詳細.

行動記録からのロボットの行動自動生成, (e) 実世界インタラクションの支援・記録・再活用から成る. 研究展開の詳細を図 11 に示す.

上記の研究成果を礎とする今後の展開はいろいろ考えられる. 第一に, さらに全体を統合し, 計測からコンテンツ生成までの一貫したモデリングとシステム化のための方法論にまとめることが考えられる. 第二には, 扱っている現象の拡大である. 現在の段階では, 見通しの良い研究の進展を行うために自然言語を意図的に避けてきた. 次の段階では, 自然言語処理と音声言語処理を本格的に導入してさらに包括的な技術を構築することが考えられる. また, 感情コンピューティングや文化コンピューティングを導入して, 感情情報をより本格的に扱うとともに種々の応用に適用することが考えられる. 第三には, 生理指標や脳計測を用いた研究手法とより本格的な融合を行い, 認知科学的な基盤をより強固にすることが考えられる.

5. 主な発表論文等

[雑誌論文] (計 20 件, すべて査読有)

1. 中田 篤志, 角 康之, 西田 豊明. 非言語行動の出現パターンによる会話構造抽出, 電子情報通信学会論文誌, Vol.J93-D, No.1, 2011 年 1 月.
2. Yasser F. O. Mohammad, Toyoaki Nishida. Using physiological signals to detect natural interactive behavior. *Applied Intelligence* 33(1): 79-92 (2010).
3. Fukuda, H., Ueda, K. (2010). Interaction with a Moving Object Affects One's Perception of Its Animacy. *International Journal of Social Robotics*, 2(2), 187-193.
4. 大本 義正・植田 一博・大野 健彦 (2010). 複数の非言語情報による自由なコミュニケーション中の嘘の自動判別の可能性の検討. 電子情報通信学会論文誌, J93-D(6), 848-856.
5. Y. Xu, K. Ueda, T. Komatsu, T. Okadome, T. Hattori, Y. Sumi and T. Nishida. WOZ experiments for understanding mutual adaptation. *AI & Society* 23(2): 201-212, 2009.
6. M. Ozeki, S. Maeda, K. Obata, Y. Nakamura. Virtual Assistant: Enhancing Content Acquisition by Eliciting Information from Humans, *Multimedia Tools and Applications*, Vol.44, No.3, pp.433-448, 2009.
7. 青山秀紀, 尾関基行, 中村裕一. ユーザの状態に適応した支援を行うためのインタラクションモデ

[学会発表] (計 46 件, すべて査読有)

1. Yasser Mohammad and Toyoaki Nishida. Learning Interaction Protocols using Augmented Bayesian Networks Applied to Guided Navigation, Taipei, Taiwan, IROS 2010.
2. Zhiwen Yu, Zhiyong Yu, Hideki Aoyama, Motoyuki Ozeki, Y. Nakamura. Capture, Recognition, and Visualization of Human Semantic Interactions in Meetings, *IEEE International Conference on Pervasive Computing and Communications*, pp.107-115, 2010.
3. Fukuda, H., Ueda, K. (2010). An Even-Related Study of Animacy Perception: Turtle or Robot. *Proceedings of Cognitive Neuroscience Society 2010 Annual Meeting*, 73.

[図書] (計 3 件)

1. T. Nishida, L. Jain, C. Faucher (eds.) *Modelling Machine Emotions for Realizing Intelligence: Foundations and Applications*, Smart Innovation, Systems and Technologies Series, Springer, 2010.
2. 坊農真弓, 高梨克也 (編著). 多人数インタラクションの分析手法, 知の科学シリーズ, オーム社, 2009 年 9 月.
3. 西田豊明, 角康之, 松村真宏. 社会知デザイン, 知の科学シリーズ, オーム社, 2009 年 6 月.

[産業財産権]

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

[その他]

6. 研究組織

(1) 研究代表者

西田 豊明 (TOYOAKI NISHIDA)
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

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

(3) 研究分担者

中村 祐一 (NAKAMURA YUISHI)
京都大学・学術情報メディアセンター・教授
研究者番号: 40227947

(4) 研究分担者

植田 一博 (UEDA KAZUHIRO)
東京大学・総合文化研究科・准教授
研究者番号: 60262101

(5) 研究分担者

角 康之 (SUMI YASUYUKI)
京都大学・大学院情報学研究科・准教授
研究者番号: 30362578

(6) 研究分担者

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