

論文 / 著書情報
Article / Book Information

題目(和文)	さりげないヒューマン・ロボットコミュニケーションにおける深層的状况理解
Title(English)	Deep Level Situation Understanding in Casual Communication between Humans and Robots
著者(和文)	湯永康
Author(English)	Yongkang TANG
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第9579号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:廣田 薫,寺野 隆雄,室伏 俊明,長谷川 修,小野 功
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第9579号, Conferred date:2014/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	湯 永康	
論文審査 審査員		氏名	職名	氏名	職名
	主査	廣田 薫	教授	小野 功	准教授
	審査員	寺野 隆雄	教授		
		室伏 俊明	准教授		
長谷川 修		准教授			

論文審査の要旨 (2000 字程度)

本論文は、ヒューマン・ロボット共存社会におけるヒューマン・ロボット間のコミュニケーションに関する研究をまとめたもので、「Deep Level Situation Understanding in Casual Communication between Humans and Robots (さりげないヒューマン・ロボットコミュニケーションにおける深層的状況理解)」と題し、英文4章から成っている。

第1章「Introduction」では、ヒューマン・ロボット共存社会において、ロボット側から人間の行動を正しく理解して適切に返答することが重要な課題であることを述べ、ロボットが人間の行動を誤認識することが多いことや、コミュニケーションが必ずしも円滑に行えないことを指摘し、その解決方法として、ショケ積分に基づくマルチモーダルシステムで認識率を向上し、深層的状況理解でコミュニケーションを円滑化させることが重要であると述べた上で、各章の構成に言及している。

第2章「Multimodal Gesture Recognition based on Choquet Integral」では、ジェスチャの認識率を向上するために、視覚センサと加速度センサを用いたショケ積分に基づくマルチモーダルジェスチャ認識システムを提案している。提案システムは、視覚センサによるジェスチャ認識ユニットと加速度センサによるジェスチャ認識ユニットで構成され、それぞれが類似度を計算してショケ積分によって各ユニットの類似度を統合した最終認識結果を出力している。実験には、申請者所属研究室のマスコットロボットシステムプロジェクトで使われている代表的な8つのジェスチャを、6名の被験者にそれぞれ5回繰り返してもらった計240のジェスチャデータを用いている。山登り法による各ジェスチャ認識ユニットにおいて、8ジェスチャのファジィ測度を探索して有効性を検証している。そして、視覚センサユニット認識結果の76.7%、及び加速度センサユニット認識結果の70%に比して、約20%のジェスチャ認識性能向上を実現している。さらに、菅野積分によるジェスチャ融合方法及び逆菅野積分によるジェスチャ融合方法と比較して、提案するショケ積分によるジェスチャ融合方法が、ロバスト性を保ちながらより有効的であることも示している。

第3章「Deep Level Situation Understanding」では、ヒューマン・ロボットコミュニケーションのための音声理解、ジェスチャ・ポスチャ理解、感情理解、意図理解及び雰囲気理解の先行研究を概観し、これらのみでは、ヒューマン・ロボット間のさりげないコミュニケーションを実現するのに不十分であることを指摘したうえで、話者の隠している本音や真の感情などを推定するために、深層的状況理解の概念を提案している。さらに、深層的状況理解において円滑なさりげないコミュニケーションを行うための会話相手に関連する特化した知識や思いやりを考慮する必要性を指摘している。深層的状況理解をするために必要な5パラメータ（発話内容情報、一般的な知識と特化された知識、感情情報、雰囲気情報及びジェスチャ・ポスチャ情報）について言及している。実験では、主に発話内容のテキスト情報に着目し、発話内容の意味を解析して、解析された意味によって知識データベースへ問い合わせを行い、発話者の意図を推定して、その意図に合わせる思いやりのある応答を話者に返している。会社で秘書ロボットと社員間で会議予約をするシナリオを想定し、社員の発話内容をシステムの入力として提案システムによる推論を行い、思いやりのある返事を秘書ロボットの発話として出力する実験を行っている。提案システムを評価するために、12名の被験者により、同じ背景知識を承知している知人からの返事と比較してシステムの出力がどの程度自然らしいかを比較してもらい、そのアンケートによって評価を行っている。評価指標を「自然」、「少し自然」、「普通」、「少し不自然」、「不自然」の五段階に分け、それぞれに「1」、「0.75」、「0.5」、「0.25」、「0」の数値を割り当てた評価結果は、「自然である」と「少し不自然」の間である「0.84」という結果で、提案する深層的状況理解が、ヒューマン・ロボット間のコミュニケーションをより円滑に実現することが可能であると結論づけている。

第4章「Conclusions」では、ヒューマン・ロボット間でさりげないコミュニケーションを行うための、視覚センサと加速度センサを用いたショケ積分に基づくマルチモーダルジェスチャ認識システムと、深層的状況理解概念及びその推論システムを総括し、さらに、深層的状況理解における応用の広さに言及した上で、その実用性およびその将来展望などを述べている。

以上を要するに、本論文では、ヒューマン・ロボット間のさりげないコミュニケーションを実現するために、視覚センサと加速度センサを用いてショケ積分に基づいたマルチモーダルシステムと、対話相手の本当の感情、気持ちなどを推定する深層的状況理解の概念及び推論システムを提案し、実証実験を通して提案の有効性を確認している。従って、本論文は博士(学術)の学位論文として、十分に価値があるものと認められる。