

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)	論文要旨
Type(English)	Summary

論文要旨

THESIS SUMMARY

専攻： Department of	知能システム科学	専攻	申請学位 (専攻分野)： 博士 (学術) Academic Degree Requested Doctor of
学生氏名： Student's Name	湯 永康		指導教員 (主)： 廣田 薫 教授 Academic Advisor(main)
			指導教員 (副)： Academic Advisor(sub)

要旨 (和文 2000 字程度)

Thesis Summary (approx.2000 Japanese Characters)

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、シヨ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ヨ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シヨ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ケ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% に比べて、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、提案するシヨ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・ボスチャ理解、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・ボスチャ情報) 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 人の被験者により、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 5 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ヨ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ヨ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。

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

専攻： 知能システム科学 専攻
Department of
学生氏名： 湯 永康
Student's Name

申請学位(専攻分野)： 博士 (学術)
Academic Degree Requested Doctor of
指導教員(主)： 廣田 薫 教授
Academic Advisor(main)
指導教員(副)：
Academic Advisor(sub)

要旨 (英文 300 語程度)

Thesis Summary (approx.300 English Words)

The concept and inference system of deep level situation understanding are proposed to realize human like natural communication among agents (e.g., humans and robots/machines). The human robot communications based on visible and audible information are called surface level communications, such as gesture/posture understanding, facial expression understanding, and speech/voice understanding. The deep level situation understanding is characterized as unifying the surface level understanding, emotion understanding, intention understanding, and atmosphere understanding by applying universal and customized knowledge of each agent.

Gesture communication is an important communication way in surface level understanding. A multimodal gesture recognition method for Mascot Robot System is also proposed based on Choquet integral by fusing camera and 3D accelerometer data. By calculating two fuzzy measures in the training phase for camera based and accelerometers based units, the proposed system obtains enough recognition rate of 96.0% for 8 types of gestures by improving the recognition rate approximate 20% compared with that of each unit.

The proposed deep level situation understanding aims to smooth the communications between human and robot, to realize harmonious communication by excluding unnecessary troubles or misunderstandings among agents, and finally to create a peaceful, happy, and prosperous humans-robots society.

A simulated experiment is established to implement the proposed deep level situation understanding system where meeting-room reservation in a company is done between a human employee and a secretary-robot. Twelve subjects are asked by questionnaires to evaluate the response of the proposed inference system comparing to the responses from familiar people. The proposed deep level inference system achieves a naturalness value of 0.84 which is between the ranks of “natural (=1.0)” and “a little natural (=0.75)” comparing to communicate with familiar people.

The proposed deep level situation understanding may be applied in robot systems for casual communications such as restaurant service robot systems, secretary robot systems, domestic robot system, and therapy robot systems.

備考：論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 1 部提出してください。

Note：Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 1copy of 800 Words (English).