

論文 / 著書情報
Article / Book Information

題目(和文)	ロボット聴覚とプラン認識に基づく環境理解のためのPlan-Intention-Eventフレームワーク
Title(English)	Plan-Intention-Event Framework for Scene Analysis Based on Robot Audition and Plan Recognition
著者(和文)	小島 諒介
Author(English)	Ryosuke Kojima
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10554号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:井村 順一,天谷 賢治,中島 求,早川 朋久,篠田 浩一,中臺 一博
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10554号, Conferred date:2017/3/26, 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

(博士課程)

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	小島 諒介
論文審査 審査員		氏 名	職 名	
	主査	井村 順一	教授	篠田 浩一
	審査員	天谷 賢治	教授	中臺 一博
		中島 求	教授	
		早川 朋久	准教授	

論文審査の要旨 (2000 字程度)

本論文は「Plan-Intention-Event Framework for Scene Analysis Based on Robot Audition and Plan Recognition」(ロボット聴覚とプラン認識に基づく環境理解のための Plan-Intention-Event フレームワーク)と題し、英文全 7 章から構成されている。

第 1 章「Introduction (序論)」では、本論文で扱う環境理解、すなわち、シーン分析について、研究背景を述べ、その課題を整理している。シーン分析における主要課題として、1) センサ情報から、対象シーン内で起きた各イベントを、「いつ、だれが、どこで、何が」という属性と共に抽出すること(イベント抽出)、および 2) イベント間の関係性を考慮してシーンの構造を再構成すること(シーン再構成)を設定し、議論を行っている。

第 2 章「Literature Review (関連研究)」では、関連研究の紹介・議論を行い、研究の位置付けを行っている。シーン分析のさらなる応用のためには、複数の属性に同時に注目したイベント抽出、また、シーン再構成は、イベント間の関係性の抽出一つをとっても挑戦的な課題であるため、限定的な状況に限った研究報告が多く、様々な状況に対応できる柔軟な手法が必要であることを主張している。本論文はこれらに必要な技術の開発を行い、シーン分析システム設計の汎用的枠組みとして、整理することに意義があると位置付けている。

第 3 章「Proposed Framework (提案フレームワーク)」では、イベント抽出時に複数属性抽出を考慮できることに加え、プランや意図モデルを事前知識として組み込むことにより 2 つの主要課題を扱うことができるシーン分析システム設計の汎用的枠組みを「Plan-Intention-Event (PIE) フレームワーク」として提案している。この有効性を示すため、以降の章で、調理認識、鳥の歌分析、Web セッションログ分析の 3 つの応用を試みている。

第 4 章「Cooking Recognition (調理認識)」では、マイクとカメラを用いたマルチモーダル調理認識システムを構築し、評価している。このシステムでは、PIE フレームワークに基づき「切る、煮る、焼く」といった調理イベント抽出とシーン再構成のためのプラン・意図認識を行っている。さらに、ウェブ上のレシピから階層隠れマルコフモデルを構築する手法を提案している。マイクもしくはカメラの一方の情報しか利用できない場合は、PIE フレームワークの利用により欠落した情報が補完でき、認識性能が顕著に向上することを実験により示している。調理認識システムへの適用を通じて、提案した PIE フレームワークの有効性を実証している。

第 5 章「Bird Song Analysis (鳥の歌分析)」では、シーン分析の主要課題であるイベント抽出問題において、複数の属性を同時に抽出する問題に焦点を当て、これを解決する Spatial-Cue-Based Probabilistic Model を PIE フレームワークの一部として提案・評価している。従来のロボット聴覚技術では、音源の検出、定位、分離、識別といった処理を直列に接続するため、誤差が蓄積するという問題があったが、提案法では音源の位置情報に着目した統合によりこの問題を緩和している。この手法を評価するため、鳥の歌の分析システムを構築し、実験を行っている。同時に 4 羽程鳴いている実環境で 5 種類の鳥の歌を識別するタスクにおいて、最大で従来法より 5 ポイント高い 40% の正答率を達成するなど、定量的に提案法の有効性を実証している。

第 6 章「Web Session Log Analysis (Web セッションログ分析)」では、全イベントが観測できない状況でのシーン再構成という PIE フレームワークを実問題に適用する際に避けられない問題に挑戦している。階層隠れマルコフモデルより柔軟なモデリングが可能な確率文脈自由文法を用いたプラン・意図認識を例にあげて、その文法上の接頭辞確率を導入する手法を提案している。Web サイト来訪者の意図をオンラインで推定する実験により提案法の有効性を実証している。

第 7 章「Conclusion (まとめ)」では、PIE フレームワークの適用事例から得られた知見を整理し、シーン分析システム設計の汎用的なフレームワークとしての学術的な貢献、および今後の課題について議論している。

以上を要するに、本論文は、環境理解の実現のために、ロボット聴覚とプラン・意図認識を統一的に扱う PIE フレームワークを提案している。また、シーン分析における複数属性の同時抽出、および全イベントが観測できない状況でのシーン再構築という課題について、それぞれ PIE フレームワークに基づいた独自の手法を提案し、その解決を図っている。また、これらを 3 つの応用例を用いて評価し、その有効性を定量的に実証している。こうしたフレームワークの提案と実証は、環境理解の発展に、工学上貢献するところが大きい。よって本論文は博士(工学)の学位論文として十分な価値があると認められる。

注意:「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。