OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Robot Motion Generation and Auditory Scene Manipulation for Tele-presence Systems
Author(s)	劉, 超然
Citation	大阪大学, 2015, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/52067
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

様式３

論　文　内　容　の　要　旨

氏　名　（　劉　超然　）	
論文題名	Robot Motion Generation and Auditory Scene Manipulation for Tele-presence Systems （遠隔存在感システムのためのロボット動作生成及び音場再構築）

論文内容の要旨

This thesis presents a tele-operation robot system that focused on tele-presence for both robot and tele-operator. In proposed system, robot's motion was generated by a rule-based model based on dialog act function tags, auditory environment in robot side is augmented and reproduced to tele-operator for increasing tele-presence.

Head motion occurs naturally and in synchrony with speech during human dialogue communication, and may carry paralinguistic information, such as intentions, attitudes and emotions. Therefore, natural-looking head motion by a robot is important for smooth human-robot interaction. Based on rules inferred from analyses of the relationship between head motion and dialogue acts, this paper proposes a model for generating head tilting and nodding, and evaluates the model using three types of humanoid robot (a very human-like android, "Geminoid F", a typical humanoid robot with less facial degrees of freedom, "Robovie R2", and a robot with a 3-axis rotatable neck and movable lips, "Telenoid R2"). Analysis of subjective scores shows that the proposed model including head tilting and nodding can generate head motion with increased naturalness compared to nodding only or directly mapping people's original motions without gaze information. We also find that an upwards motion of a robot's face can be used by robots which do not have a mouth in order to provide the appearance that utterance is taking place. Finally, we conduct an experiment in which participants act as visitors to an information desk attended by robots. As a consequence, we verify that our generation model performs equally to directly mapping people's original motions with gaze information in terms of perceived naturalness.

In a tele-operated robot system, the reproduction of auditory scenes, conveying 3D spatial information of sound sources in the remote robot environment, is important for the transmission of remote presence to the tele-operator. We proposed a tele-presence system which is able to reproduce and manipulate the auditory scenes of a remote robot environment, based on the spatial information of human voices around the robot, matched with the operator's head orientation. In the robot side, voice sources are localized and separated by using multiple microphone arrays and human tracking technologies, while in the operator side, the operator's head movement is tracked and used to relocate the spatial positions of the separated sources. Interaction experiments with humans in the robot environment indicated that the proposed system had significantly higher accuracy rates for perceived direction of sounds, and higher subjective scores for sense of presence and listenability, compared to a baseline system using stereo binaural sounds obtained by two microphones located at the humanoid robot's ears. The effect of variable ITD (interaural time differences) in perceived distance expression of near field sound source has been investigated. Result of localization experiment showed that compare to HRTFs measured from a dummy head, calculated impulse response which implied ITD's change in near field can provide higher sense of distance with same accuracy in perceiving direction of sound in horizontal plane. We also proposed three different user interfaces for augmented auditory scene control. Evaluation results indicated higher subjective scores for sense of presence and usability in two of the interfaces (control of voice amplitudes based on virtual robot positioning, and amplification of voices in the frontal direction).

様式7

論文審査の結果の要旨及び担当者

氏　名　（　劉　超然　）		
	（職）	氏　名
論文審査担当者	主 査　教 授 副 査　教 授 副 査　教 授	石黒　浩 西田　正吾 飯國　洋二

論文審査の結果の要旨

近年、メディアとしての遠隔操作コミュニケーションロボットが多数開発されており、これらのロボットを用いて円滑で自然な遠隔コミュニケーションの実現が期待されている。しかし、従来のビデオコミュニケーションシステムでは、これらに寄与する存在感や臨場感が必ずしも十分に伝達されていないという問題がある。この問題に対して、本論文では、遠隔操作コミュニケーションロボットを介して操作者の存在感を遠隔地へ伝達すること、および、遠隔地の臨場感を操作者へ伝達することを目的として、ロボットの自然な動作生成および空間的音環境の再構築手法を提案している。

ロボットの自然な動作生成では、対面対話における人の頭部動作を詳細に分析し、その結果に基づいた頷き・首傾げ生成モデルが提案された。複数の種類の人型ロボットを用いた被験者実験の結果、本論文で提案されたモデルは、人によるオリジナルの頭部・眼球動作を再現するのと同等の自然さ評定が得られることが示されている。即ち、提案モデルを用いることによって、操作者の存在感が効果的に遠隔地へ伝達可能であることが実証された。

空間的音環境の再構築手法では、ロボット周囲の音源位置情報に基づいて音環境を３次元に再現可能なシステムが提案された。複数のマイクロフォンアレイとヒューマントラッキング技術を用いて音源の定位と分離を行い、さらに、操作者の頭部回転トラッキング情報に基づいて操作者の動きを補正することで、遠隔地の空間的音環境を操作者側に再構築している。被験者実験の結果、本論文で提案されたシステムは、従来法よりも高い定位精度と強い臨場感、および、聴き取り易さが得られることが示されている。これによって、提案システムは遠隔地の臨場感を操作者に有効に伝達するものであることが明らかとなっている。

以上のように、本論文では従来のビデオコミュニケーションシステムが有する存在感・臨場感の伝達上の問題点に着目し、ロボットを介した遠隔地への操作者の存在感の伝達、および、遠隔地の空間的音環境の操作者への伝達について、堅実な分析結果に基づいた具体的な提案手法が示されており、その評価についても被験者実験によって検証されている。本論文で提案された手法は、特定のロボットや環境に依存しないため汎用性が高く、ロボットを介した遠隔コミュニケーション研究にとって意義あるものである。よって，本論文は博士（工学）の学位論文として価値のあるものと認める。