

ジェスチャ及び音声指示によるロボットの制御

篠塚 晃希[†] 田村 仁[‡]日本工業大学工学研究科機械システム工学専攻[†]日本工業大学先進工学部ロボティクス学科[‡]

1. はじめに

近年では工場から日常生活に至るまで様々な場面にてロボットの活用が想定されている。その中でロボットの操作を人間が行う場合、コントローラを用いて行うほかに音声やジェスチャをロボットに認識させることで指示を行う方法が注目されている。ジェスチャなど非言語でのコミュニケーションを用いたロボットの操作ができれば、コントローラ等を用いることのできない環境においても活躍が期待できる。例えば、商店等でロボットが特定の位置に商品を運ぶ際に指差しによる指示ができれば手軽にロボットが操作可能である等が挙げられる。

2. 提案手法

本研究では、先行研究である[1]によって指差し指示を行った際に生じた誤差を、音声によって補正することで任意の地点へロボットを移動させることを目的とする。

移動ロボットから生じる誤差を音声によって補正する手法としては[2]などが挙げられる。音声での補正はロボットの視点から行われているため、ロボットの向き等によっては指示者が指示しづらい可能性がある。このため、本研究は指示者の視点基準での音声指示が可能な機能の実装を行う。

平面を前後左右に移動可能なロボットに搭載することを想定し、ロボットに360度カメラを搭載する。図1(a)のように360度画像から指示者の写るRGB画像を展開し、そのRGB画像によって取得した指示者の骨格座標から、右手で指している地点を認識し移動を行い、生じた誤差に応じて指示者による声掛けによって位置を修正することで、目的地点までの移動を行う。システムの概要を図1に示す。

誤差の補正は、ロボットが移動を完了したのち、生じた誤差に応じて決められたいくつかの

キーワードを指示者が発声することでそのキーワードに応じてロボットが位置の微調整を行う。キーワードは表1に示す通り、合計8種類と設定した。表1の移動距離は、入力された場合のロボットの移動距離を表し、方位は図2のようにそれぞれ指示者から見たロボットの前後左右に当てはまる方向へ移動する。例として、「もっと右」という指示を与えた場合、ロボットと指示者を通る直線上をy軸とし、指示者から見て右、図2の場合4の方向にロボットが0.5[m]移動することになる。

指示者視点でロボットの誤差修正を行う場合、指示者から見たロボットの進行方向をロボットが認識する必要がある。そのため、音声指示が入った際、360度カメラの画像から指示者の立っている方位を測定し、図1(b)に示すように、それを基準とした座標系を用いることで移動する方位を決定する。

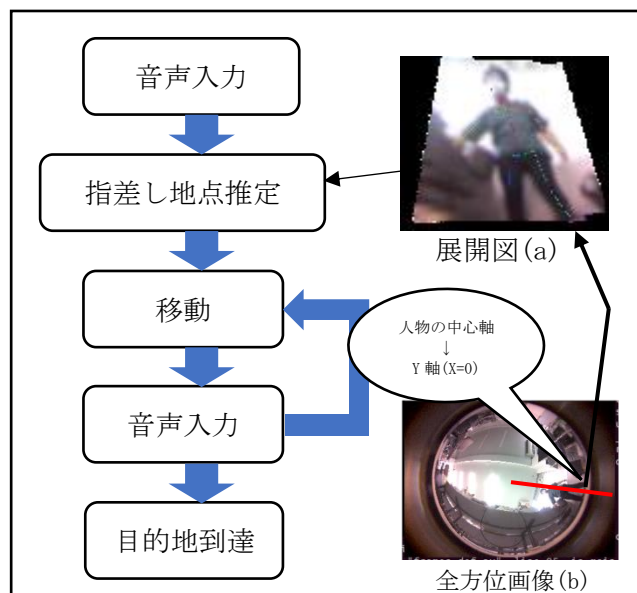


図1 システム概要

Robot control by gestures and voice instructions

[†]Koki Shinotsuka, Graduate School of Eng., Dept. of Mechanical Systems Eng., Nippon Institute of Technology.[‡]Hitoshi Tamura, Dept. of Robotics, Faculty of Advanced Eng., Nippon Institute of Technology.

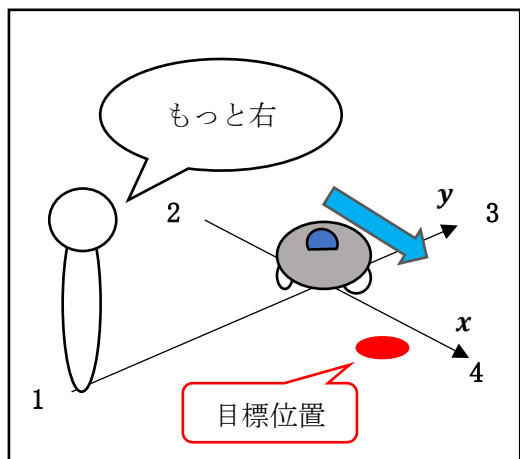


図 2 音声補助

表 1 キーワード

キーワード	移動距離[m]
ちょっと手前	0.1
ちょっと奥	
ちょっと右	
ちょっと左	
もっと手前	0.5
もっと奥	
もっと右	
もっと左	

3. 音声指示の流れ

指差し推定地点への移動を終えたロボットに対し音声指示を行うと想定した場合、例として[1]の実験のように、ロボットの右側 0.5[m]の地点を指さした際、x 軸方向に-0.4[m]、y 軸方向に 0.9[m]の誤差が生じたとする。この場合図 3 のような音声指示を与えることで、誤差を 0.2[m]以内に抑えることが可能であると考えられる。

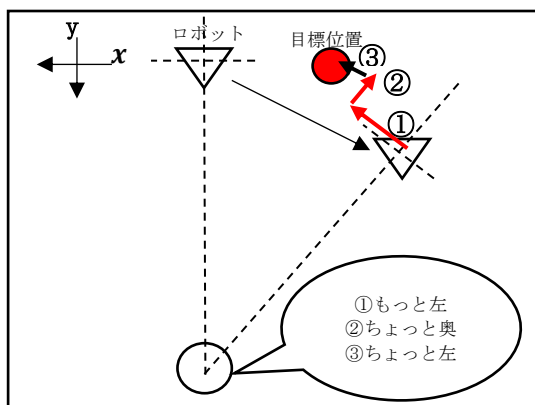


図 3 音声指示による移動

4. 実装方法の検討

本研究には、画像および音声情報の処理に Jetson nano を使用する。

撮影に用いる 360 度カメラは視野角 220 度の、Pi VR220 Camera を使用する。

指示者の骨格座標取得には Open Pose[3]の機械学習モデルである tf-pose-estimation[4]を用いた。

音声補助のための音声の認識に、Google から発表されているライブラリの、Speech Recognition[5]を使用する。

5. 今後の予定

実験として、実装を行ったうえで特定地点へ指差し、音声による補正によって誤差をどの程度減少できるかの検証を行う。

また、「〇〇の近く」といったキーワードと物体認識を用いることで指差し地点の補正を行う手法の検討も行う予定である。

5. おわりに

本研究では、ロボットへ指差しのジェスチャの指示によって生じた誤差を音声によって補正し移動させることを目的とし、システムの実装を試みた。今後は、音声の補助の実装を行い総合的な検証を行いたい。

参考文献

- [1] 篠塚 晃希, 田村 仁, “指差しジェスチャと音声指示によるロボットの制御手法”, 情報処理学会第 83 回全国大会 講演論文集, 2021, 分冊 4 pp143 - 144
- [2] 佐藤 隼, 中野 幹生, Antoine Raux, 船越 孝太郎, 竹内 誉羽, “移動ロボットに対する指示音声の状況依存逐次理解”, 情報処理学会第 74 回全国大会 講演論文集, 2012, 1 pp199 - 200
- [3] CMU-Perceptual-Computing-Lab /openpose, <<https://github.com/CMU-Perceptual-Computing-Lab/openpose>>2021 年 1 月 8 日閲覧
- [4] ildoonet/tf-pose-estimation, <<https://github.com/ildoonet/tf-pose-estimation>>2021 年 1 月 8 日閲覧
- [5] SpeechRecognition 3.8.1, <<https://pypi.org/project/SpeechRecognition/>>2022 年 1 月 7 日閲覧