

사람 손 포즈 인식을 통한 로봇-휴먼 사물 전달 시스템

최용혁, Ismael Espinoza, 오지현, 정단비, 이강산, 김태성
경희대학교 전자정보융합공학과

{hope5090, inespinoza24, dhwlgjs3, dbingsu, kangsanl, tskim} @khu.ac.kr

Robot-Human Handover of Objects via Human Hand Pose Recognition

Yong-Hyeok Choi, Ismael Espinoza, Ji-Heon Oh, Danbi Jung, Kangsan Lee, Tae-Seong Kim

Dept. of Electronics and Information Convergence Engineering, Kyung Hee Univ.

요약

본 논문은 인간-로봇 사물 전달(Object Handover)에서 손 포즈 추정을 통해 물체 전달을 수행하는 시스템을 구현하였다. RGB-D 비전을 통해 사물 및 사람 손의 위치와 형태를 인식하고, 이를 기반으로 로봇 팔이 물체를 정확하게 전달하는 시스템을 구현하였다. 물체 인식과 손 포즈 추정을 통합하여 실시간으로 로봇이 인간의 손동작을 이해하고 적절히 반응할 수 있도록 설계하였으며, 로봇 팔이 물체를 효율적이고 정확하게 전달하는 성능을 보였다.

I. 서론

로봇은 산업 분야 및 일상적인 생활 작업 분야에도 적용하기 위한 연구 개발이 활발하게 진행되고 있다. HRI(Human-Robot Interaction), 즉 인간과 로봇이 직접 상호작용하는 환경에서는 로봇의 사람 동작을 인식하고 해석하는 능력이 중요하다[1]. 특히 사물의 Robot-Human Handover 를 위해서는 물체와 사람의 손을 인식하는 로봇 비전과 동작을 제어할 수 있는 시스템이 필요하다.

본 연구에서는 손 포즈 추정과 로봇 조작을 통합하여, 효율적으로 물체를 전달하는 HRI 시스템을 제안한다. 로봇 3D 비전과 딥러닝 모델을 통해 실시간으로 사람의 손 포즈를 추정하고, 로봇 팔이 물체와 손 모양에 적절한 동작을 수행할 수 있도록 구현하였다. 이를 위해 딥러닝 모델의 손 포즈 추정 성능을 평가하여 시스템에 적용하고, 로봇이 물체를 사람 손까지 효과적으로 전달하는 것을 검증하였다.

II. 본론

본 연구에서의 Robot-Human Object Handover 시스템을 그림 1에 도시하였다. 로봇 비전을 담당하는 카메라(USB-D Intel RealSense camera D415)는 RGB-D 영상 정보로부터 FastRCNN 및 주성분 분석(PCA)을 통해 물

체의 검출 위치와 방향 정보를 추론한다. 또한 검출한 사람의 손 영상으로부터, 딥러닝 모델은 손 포즈를 추정하여 관절 좌표로 변환한다. 최종적으로, 로봇 시스템은 물체와 사람 손에 대한 위치 정보를 이용해 물체를 파지하고 사람 손 포즈에 맞추어 전달한다.

A. Object Recognition

로봇 비전에서 사물 및 손 인식을 위해 CNN 기반 객체 탐지 모델인 FastRCNN[2]을 사용하였다. FastRCNN을 통해 물체를 검출하고, 이후 물체의 깊이(depth) 정보를 point cloud로 변환하여 위치와 방향 정보를 추론한다.

B. Hand Pose Recognition

Virtual-View Pose Estimation[3]은 손 포즈를 인식하여 3D 관절 좌표를 추정하는 딥러닝 모델이다. 모델은 손의 point cloud 정보를 여러 각도에서 투사한 virtual view로 증강하여 가려지는 정보를 보완하여 손 포즈를 추정한다. 이 때 virtual view가 포즈 추정에 도움이 되는지를 self-attention을 통해 평가하고, 그 중 평가가 좋은 일부를 사용하여 손 포즈에 대한 16개의 joint point를 추정한다. 손 포즈 인식기의 정확도를 평가하기 위하여 손 포즈 데이터베이스인 ICVL[4]을 이용하여 정확도를 검증하였다.

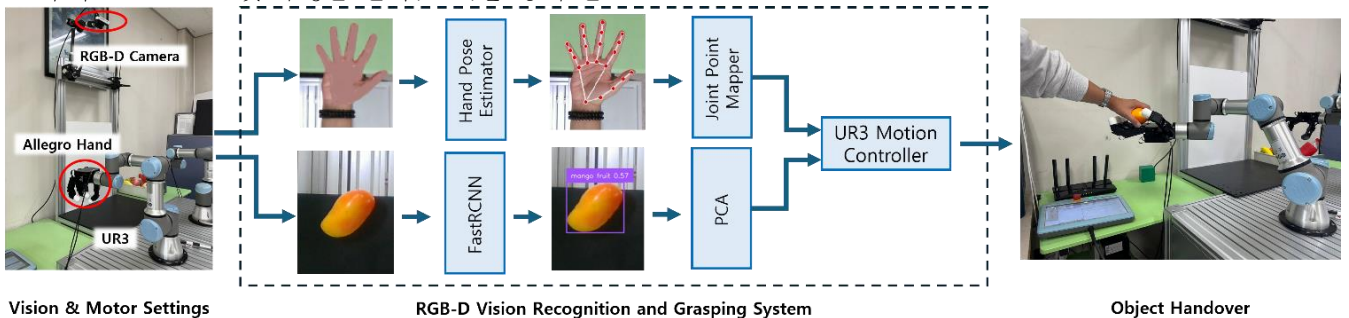


그림 1 Robot-Human Object Handover 시스템 구조

C. Robot-Human Handover System

연구에 사용된 로봇 시스템은 Allegro dexterous manipulation hand 을 장착한 UR3 로봇으로, 각각 16 및 6의 Degree of Freedom을 가진다. 물체에 대한 위치 및 주 방향 정보는 PCA 를 통해 실제 6D 정보로 계산된다. 최종적으로 로봇 제어 소프트웨어인 MoveIt 의 motion planner 와 inverse kinematics[5]을 사용하여, 로봇 팔은 물체를 잡는 위치와 방향 각도에 따라 이동한다. 이후 로봇 팔은 물체를 파지하고, 사람의 손까지 이동하여 물체를 놓아 손에 전달한다.

III. 결과

A. Hand Pose Estimation Results

그림 2 에서 로봇 시스템에서의 손 영상과 딥러닝 모델이 손 포즈를 추정한 관절 좌표에 대한 결과를 보여준다. ICVL 데이터베이스를 통한 모델의 정확도 평가 결과는 관절 좌표의 실제 값과 추정 값에 대한 평균 오차로 4.56mm 으로 계산되었다.

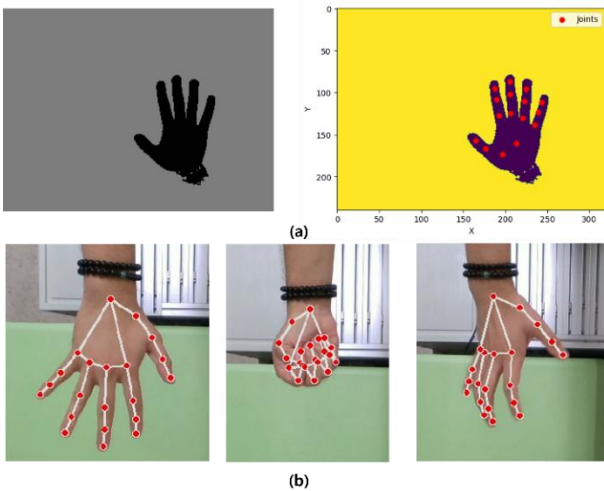


그림 2 딥러닝 모델의 손 포즈 추정 성능 평가 결과. (a) ICVL 깊이 영상 및 검출된 관절 정보, (b) 로봇 시스템의 손 영상 및 추정 손 포즈

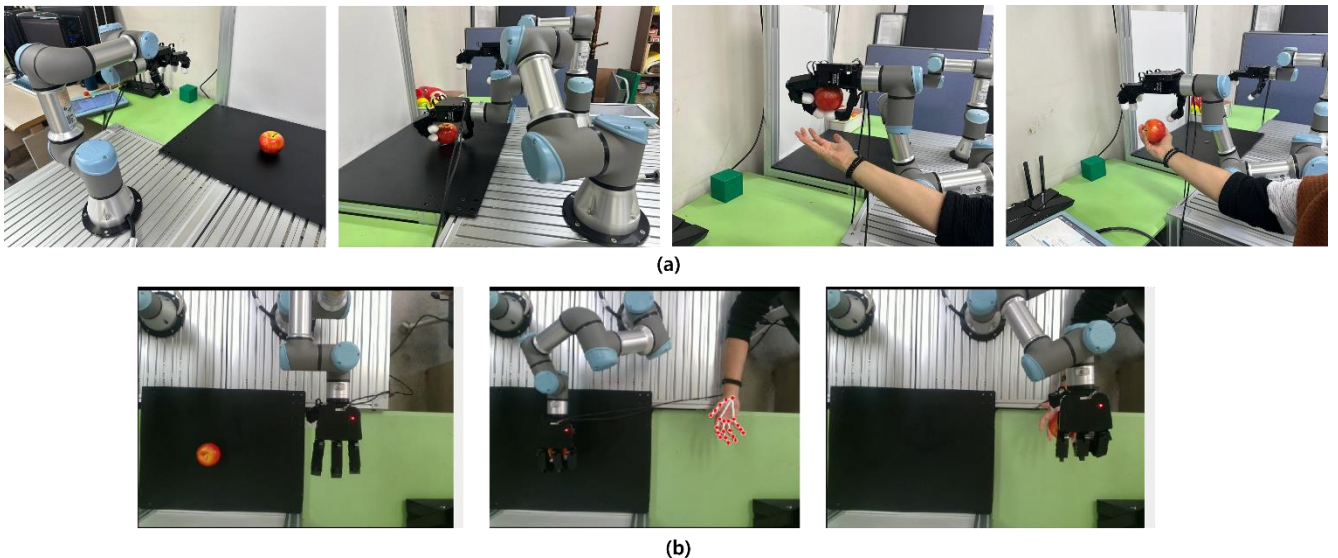


그림 3 (a) Time Series and (b) Top View Images of Robot-Human Handover of an Apple

B. Robot-Human Object Handover Results

그림 3 은 로봇 시스템을 구동한 robot-human object handover 로, 물체에 대해 시간 순의 동작 모습을 보여준다. 로봇 팔은 물체(사과)를 인식하고, 물체를 파지한다. 이후 사람의 손에 전달하기 위해 인식한 사람 손으로 이동하고, 손 포즈에 적절한 방향으로 회전 후 물체를 전달한다.

IV. 결론

본 논문에서는 사람의 손 포즈를 고려한 로봇의 사물 전달 시스템을 구축하고, 실증하였다. 이 시스템은 향후 유용한 HRI 기술로서 활용될 것으로 기대한다.

ACKNOWLEDGMENT

본 연구는 한국연구재단 중견연구사업 (NRF-2023R1-A2C100358511), 산업통상자원부 사회적약자 자립지원로봇기술개발사업 (RS-20-2300232141), 과학기술정보통신부 및 정보통신 기획평가원 대학 ICT 연구센터사업 (IITP-2024-RS-2024-00438239)의 연구결과로 수행되었음.

참고 문헌

- [1] H. Liu and L. Wang, "Gesture recognition for human-robot collaboration: A review," *International Journal of Industrial Ergonomics*, vol. 68, pp. 355-367, 2018.
- [2] J. Huang, V. Rathod, V. Birodkar, A. Myers, Z. Lu, R. Votel, Y.-h. Chen, and D. Chow, "TensorFlow Object Detection API," TensorFlow, Apr. 2020.
- [3] J. Cheng, Y. wan, D. Zuo, and C. Ma, "Efficient Virtual View Selection for 3D Hand Pose Estimation," *arXiv preprint arXiv:2203.15458*, 2022.
- [4] <https://labcvl.github.io/hand.html>
- [5] <https://forum.universal-robots.com/>