

2022 제37회 제어·로봇·시스템학회 학술대회

2022 The 37th ICROS Annual Conference (ICROS 2022)

일시: 2022년 6월 22일(수)~24일(금) 장소: 거제 소노캄

대회장

고광일 (고영테크놀러지)

공동대회장

김진환 (KAIST)

프로그램위원장

박재홍 (서울대학교)

프로그램위원회

최한림 (KAIST)
유선철 (POSTECH)
신동준 (중앙대학교)
오현동 (UNIST)
진상록 (부산대학교)
김아영 (서울대학교)
박종용 (부경대학교)
최진우 (KRISO)
김중혁 (홍익대학교)
박해원 (KAIST)

재무위원회

오세훈 (DGIST)

특별세션위원회

고낙용 (광주전남지부장)
한수희 (대구경북지부장)
김 정 (대전충청지부장)
이경창 (부산경남울산지부장)
박재병 (전북제주지부장)
이종민 (공정시스템연구회장)
서진호 (로보틱스 및 응용 연구회장)
이왕현 (머신비전연구회장)
김윤수 (제어이론연구회장)
최용준 (철강계측제어연구회장)
성상경 (항법유도제어연구회장)
김동원 (진화형자율다체계시스템연구회장)
이세진 (스마트모바일로봇연구회장)
최재순 (의료로봇및시뮬레이션연구회장)
이석재 (국방기술연구회장)
박종오 (바이오로봇연구회장)
박용운 (자율표준화연구회장)
이경수 (자동차제어연구회장)
손형일 (농업및건설로봇연구회장)
한성현 (제조로봇및스마트팩토리연구회장)
김상철 (필드로봇소사이어티회장)

심사위원회

유지환 (KAIST)
우주현 (한국해양대)
이세진 (공주대학교)
김종한 (인하대학교)

출판위원회

이규빈 (GIST)
이동환 (KAIST)
조백규 (국민대학교)

홍보위원회

공경철 (KAIST)
최성록 (서울과학기술대)

총괄간사

윤지영 (ICROS)

학회사무국

이지연 (ICROS)

(사)제어·로봇·시스템학회에서 매년 주최 및 주관하는 국내학술대회인 「2022 제37회 제어·로봇·시스템 학회 학술대회」가 2022년 6월 22일 (수)부터 24일(금)까지 거제 소노캄에서 개최됩니다.

4차 산업혁명을 주도할 제어, 로봇 및 시스템 분야의 최신 연구개발 동향을 공유하며 연구자간에 상호교류 할 수 있는 자리를 마련하였으니 많은 관심과 참여를 부탁드립니다.

논문모집안내

- | | | |
|-----------|--------------|--------------|
| ▪ 제어이론 | ▪ 로봇 지능 | ▪ 스마트 공장 시스템 |
| ▪ 센서 및 계측 | ▪ 로봇 응용 | ▪ 무인이동체 시스템 |
| ▪ 영상처리 | ▪ 로봇시스템 | ▪ 제어시스템 |
| ▪ 제어응용 | ▪ 인간-로봇 상호작용 | ▪ 인공지능 응용시스템 |

논문제출안내

- | | |
|-------------------|--------------|
| ▪ 2022년 4월 22일(금) | 논문 접수 마감일 |
| ▪ 2022년 5월 20일(금) | 세션 일정 공지 |
| ▪ 2022년 5월 27일(금) | 발표자 사전등록 마감일 |

제출방법

소정의 논문양식(홈페이지 참고)에 따라 2페이지로 온라인 제출 (요약문 제출 및 심사과정 생략)

초청세션 구성 및 논문

세션당 논문 5~6편을 모집 후 소정의 구성양식(홈페이지 참고)을 작성하여 논문 접수 마감일까지 사무국 이메일로 제출, 논문은 온라인 제출

- 제출논문은 본 학술대회에 처음 발표되는 것으로 한정
- 국·영문으로 모두 작성 가능
- 발표방법: 구두 또는 포스터 발표
- 우수논문에 대한 시상 및 ICROS 논문지(한국연구재단 및 SCOPUS 등재지) 게재 추천

특별 프로그램

- 나를 감동시킨 논문들: 각 분야의 교수들이 직접 발표
- 우수신진연구자 초청발표
- 응용 논문 발표 및 데모 세션
- 산학포럼: 산업체에 필요한 제어·자동화·로봇·인공지능
- 학부생 논문 경진대회 (포스터 발표): 학부생이 제 1저자로 직접 발표하여야 하며, 우수논문에는 "학부생 논문상" 수여
- 워크샵 세션: 정부과제, 사업단 등 과제수행관련 워크샵

A Real-Time Person Tracking Algorithm Based on Yolact and SIFT for Agricultural Robots

Pablo Vela^{1,3} Jaewon Jung¹ Xiaosa He^{1,3} Yu Zheng^{1,3} Kyeong-Hwan Lee^{1,2,3*}

¹) Department of Convergence Biosystems Engineering, Chonnam National University

²) Agricultural Automation Research Center, Chonnam National University

³) BK21 Interdisciplinary Program in IT-Bio Convergence System, Chonnam National University

*Corresponding author: Kyeong-Hwan Lee (khlee@jnu.ac.kr, 062-530-2156)

Abstract Collaborative robots are becoming more and more important in every field. Thus, this paper proposes a person identification and tracking method based on Yolact and SIFT (Scale Invariant Feature Transform) for an agricultural person-follower robot using a single monocular camera. First, Yolact algorithm was used to detect all the people in the image frame, then their semantic segmentation was obtained and used to separate the people from the background. Secondly, SIFT was used to extract features from all the detected people. Finally, the features obtained are compared with the ones obtained in the last frame or with reference images and matched by calculating the approximate nearest neighbors. The experimental results show that by combining Yolact and SIFT accuracy is increased and robustness against change of illumination and distance is improved even if there are multiple people in the frame.

Keywords Follower Robot, Semantic Segmentation, SIFT, Tracking, Yolact

1. INTRODUCTION

Recently robots work in collaborative environments, where humans perform some tasks with help of robots for example in assembly lines or agricultural fields to load and unload heavy boxes around the field. In this last example, the robot must be able to accurately identify a person and obstacles to follow the farmer.

However, real time object detection and classification is challengeable and computationally expensive work [1]. Moreover, tracking an object also involves several problems such as occlusion when the object is momentarily hidden from the camera and non-stationary camera in which the object and the camera are simultaneously in movement [2].

Some approaches have been done using monocular vision to estimate the height of the target [3], but they do not consider narrow environments where the entire target person cannot fit in the frame. Other approaches consider using face recognition [4] but when following a person walking, it is his back that faces the robot. Also, HSV color extraction methods have been used to extract the colors of the clothes [5] but two people can wear the same outfit or colors can be distorted in different sunlight conditions.

Our proposed method is based on Yolact and SIFT features thus it is scale resilient, it doesn't depend on clothes colors, it doesn't need training time and by matching the features with the features extracted in the previous frame decrease the error.

All these characteristics allow that the target person can be changed dynamically and fast for a new one.

2. METHODOLOGY

Our proposed method is divided into 2 main stages, as shown in Fig. 1. The first stage is the "Searching stage", where the robot is looking for the target person for the first time, i.e., the algorithm has just been initialized or the robot has lost the target. This stage can be triggered when the score obtained in the tracking stage when comparing the current frame with last frame is lower than a previously defined threshold. In the second stage, the tracking stage the algorithm compares the features from the current frame with the ones in the previous frame, in such manner the changes of orientation or scale are not excessively different, as it could happen with the original reference images.

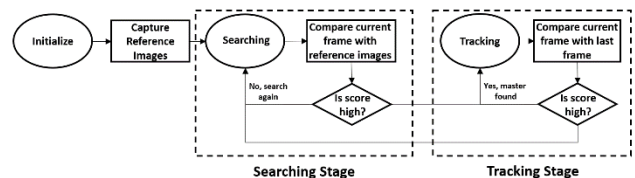


Fig. 1 Proposed algorithm flowchart

2.2 Pre-stage: Master registration

In this stage, the master is alone and multiples photos in different angles are taken, then all the pictures are processed individually using Yolact to obtain their semantic segmentation and remove the background. Finally, the pictures without background are processed with SIFT to extract and save their features.

※ This work was supported by the Defence Acquisition Program Administration (DAPA) (21-SF-GU-01) and the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry (IPET)(12102902).

2.3 Stage 1: Searching stage

In this stage, there is no clear information of whom is the master because the algorithm has just been initialized, thus there is no features extracted from a previous frame to compare the feature of the current frame, or due to a low score in the Tracking stage, that made impossible to identify the master correctly.

First, the semantic segmentation of each person in the frame is extracted using Yolact, then the features of each person are extracted independently from each other using SIFT and then each person features are match with each reference image using FLANN (Fast Library for Approximate Nearest Neighbors) [6]. Finally, the scores obtained in each case are added together. This process is repeated with each person and the one that get the higher score that is above the threshold is identified as the master, as shown in Fig. .

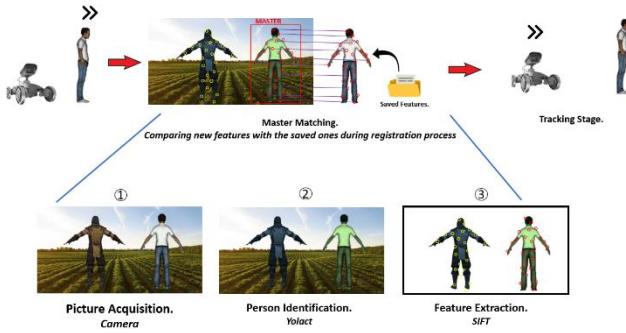


Fig. 2 Searching stage description

2.4 Stage 2: Tracking stage

This stage is similar to the last one, but instead of matching the features with the ones extracted from the reference images, the newly extracted features are compared with the ones extracted in the previous frame, as shown in Fig 3. Because it is reasonable to assume that the master cannot change his position or orientation instantaneously then it better to compare the current features with the previous features. Thus, the orientation and scale changes are not so abrupt, and more matches can be obtained.

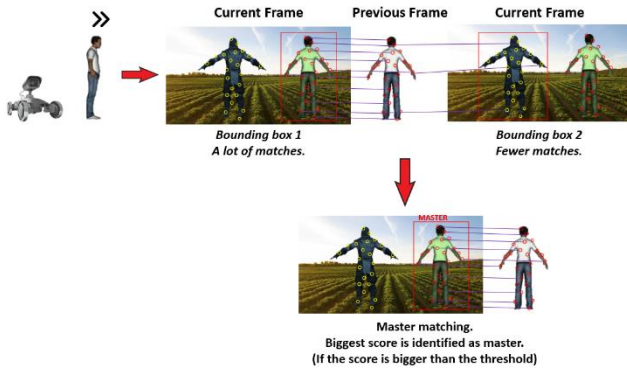


Fig. 3 Tracking stage description

3. RESULTS

The experiments were conducted outside, simulating common scenarios in an agricultural field, as shown in Fig 4. The objective in all of them is to test the capacity of the algorithm to keep tracking the master and the capacity to recover if the master is lost.

The experiment results show that the overall algorithm success rate is above 90%, but when the distance is more than 2 meters between the master and the camera, or the experiments were conducted in the afternoon the success rate decrease. Because for the first case it is harder to extract more features when the images are smaller and for the second case because the light conditions differ with the reference images.



Fig. 4 Field experiments

4. CONCLUSION

Our proposed method was able to identify and track a master person in an open field under different sunlight conditions even though people wore similar plain clothes, without designs and the experiments were performed in different days.

In the future, a dynamic threshold estimator will be implemented, to deal with the dynamic range of matches produced when the target person is far or close to the robot. Furthermore, a target trajectory predictor will be developed to increase the performance of the algorithm when the master is occluded by other people. Also, a local planner will be developed to generate the corresponding waypoints so experiments with a real robot can be conducted.

5. REFERENCES

- [1] M. Beard, B. T. Vo and B. Vo, "A Solution for Large-Scale Multi-Object Tracking," in *IEEE Transactions on Signal Processing*, vol. 68, pp. 2754-2769, 2020.
- [2] Algabri, Redhwan, and Mun-Taek Choi. "Target Recovery for Robust Deep Learning-Based Person Following in Mobile Robots: Online Trajectory Prediction." *Applied Sciences* 11.9 (2021): 4165.
- [3] Koide, Kenji, Jun Miura, and Emanuele Menegatti. "Monocular person tracking and identification with on-line deep feature selection for person following robots." *Robotics and Autonomous Systems* 124 (2020): 103348.
- [4] Chan, Wesley P., et al. "Autonomous Person-Specific Following Robot." arXiv preprint arXiv:2010.08017 (2020).
- [5] Algabri, Redhwan, and Mun-Taek Choi. "Deep-learning-based indoor human following of mobile robot using color feature." *Sensors* 20.9 (2020): 2699.
- [6] Muja, M., & Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. VISAPP.