Poster Abstract: Exploring Rolling Shutter Effect for Motion Tracking with Objective Identification

Xiao Zhang
Michigan State University
zhan1387@msu.edu

Griffin Klevering
Michigan State University
kleveri2@msu.edu

Li Xiao
Michigan State University
lxiao@cse.msu.edu

ABSTRACT

Sensing-based user interfaces hold enormous potential for smart homes, medical equipment, educational systems, AR/VR/MR, etc. However existing hand and body gesture recognition systems are mostly based on frame-level computer vision approaches, which have limitations such as the inability to operate in the environment with low brightness, short detection distance, without the objective identification ability, and coarse-grained tracking when the objectives are in high-speed motion. Therefore, in this paper, we propose to attach active LED elements on objectives and utilize rolling shutter effect to enhance the gesture recognition and achieve the fine-grained motion tracking with objective identification.

CCS CONCEPTS

• **Networks** → *Mobile networks*; **Wireless personal area networks**; • **Human-centered computing** → *Mobile devices*; **Ubiquitous and mobile computing**.

KEYWORDS

rolling shutter camera, visible light sensing, gesture recognition, optical labeling and identification

ACM Reference Format:

Xiao Zhang, Griffin Klevering, and Li Xiao. 2022. Poster Abstract: Exploring Rolling Shutter Effect for Motion Tracking with Objective Identification. In *The 20th ACM Conference on Embedded Networked Sensor Systems (SenSys '22), November 6–9, 2022, Boston, MA, USA*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3560905.3568063

1 INTRODUCTION

Hand gestures and body gestures have become more and more popular Human-Computer Interaction (HCI) across a wide variety of applications due to the fast development of computer technology and artificial intelligence (AI)[1]. Furthermore, hand and body gestures, for example, can be used in smart homes to control IoT devices for a variety of purposes; in interactive video games to provide a user-friendly and immersive gaming experience; and in XR (AR, VR, and MR) enabled mobile applications to provide interactive operations (e.g., navigation)[2]. The vision-based gesture recognition is widely used, which uses the similar processing as human eyes to detect the morphology of objective. However, vision-based approaches have the following drawbacks: (1) they can not be used in dark environments, (2) coarse-grained tracking of objectives with high motion due to their frame-level tracking speed, and (3) high processing cost (e.g., the complex computation with 21 joints of a hand)[1].

Given a real-world use case, remote surgery requires the fine-grained finger tracking and hand identification of different doctors and nurses at the same time. Then the detected fine-grained and identified location variations will be carried out to the remote robot for operation. Thus, *we propose enhanced gesture recognition, which uses some attached LED elements on different parts of objectives and rolling shutter camera to achieve gesture recognition and the fine-grained motion tracking with objective identification.* We implement a prototype of a wearable glove for hand gesture recognition. Low-power LED components with different light waves connected to glove fingers serve as invisible active optical labels. It thus enables the camera to confidently identify each finger from a single hand or even numerous hands. In the end, our system can readily recognize a comprehensive hand gesture of a single or multiple hands with hand identities from a **enhanced** considerable distance: up to 1.5m and **enhanced** finger tracking precision due to higher shutter rate (*10KHz*) than frame rate (*30Hz*). Although higher sampling rate will introduce additional computation overhead, it indeed can improve tracking precision with similar latency of traditional approaches.

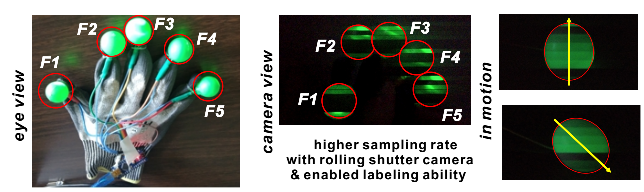

Figure 1: Enhanced gesture recognition with fine-grained tracking and optical labeling via rolling shutter effect.

2 BACKGROUND

2.1 Vision based Gesture Recognition

Traditional vision based gesture recognition use commercial cameras detecting objects via their optical properties such as shape, position, etc., frame by frame. However, cameras update these variations (i.e., sampling rate) at low frame rate (e.g., < 60 FPS) with limit tracking precision. Furthermore, the vision based gesture

SenSys '22, November 6–9, 2022, Boston, MA, USA

ACM ISBN 978-1-4503-9886-2/22/11…$15.00
https://doi.org/10.1145/3560905.3568063

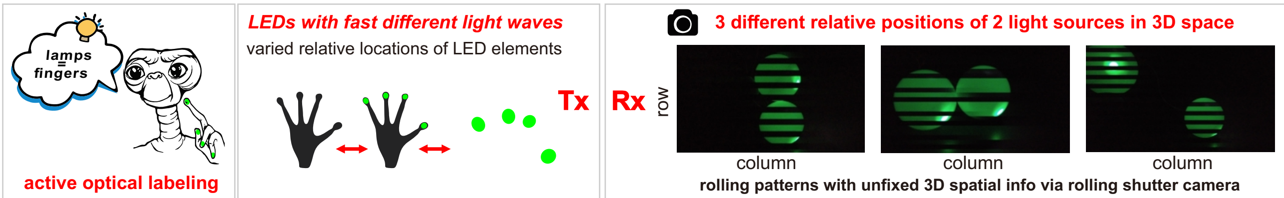

Figure 2: Using attached LED elements at fingers as active optical labels and using rolling shutter camera to parse their labels and 3D locations with fine-grained motion tracking due to the fast shutter rate than the frame rate.

recognition approaches record entire image frame and pose privacy concerns in sensitive circumstances (e.g., home, hospital, etc).

2.2 Rolling Shutter Enhanced Recognition

The global shutter camera records the entire scene at the same time. While the rolling shutter camera works by simultaneously exposing one row of pixels and generating a whole image row by row[3, 4]. When the rolling shutter speed is equal to or slightly faster than the switching speed of the light wave from the transmitter, the distinct stripes are observed, as shown in the right of Figure 2. These clear rolling patterns from different light sources can be utilized as active optical labels for each essential part of the recognized objective. Thus, the different row-level locations variation of each part of the objective can be recorded with identification. Most commercial cameras are based on rolling shutter camera, however, as discussed in Section 2.1, **existing vision approaches did not utilize this rolling shutter effect for enhanced gesture recognition yet.**

3 SYSTEM OVERVIEW

To perform enhanced vision based gesture recognition (i.e., hand gesture, body gesture, etc.,), we implement a prototype for hand gesture recognition which consists of 5 commercial LED elements attached wearable glove and a rolling shutter camera, as shown in Figure 1 and the left of Figure 3. Our goals are (1) optical labeling for each finger for identification, (2) recording motion of each finger from a hand, (3) achieve long distance and finer-grained tracking. The gloves are standard gloves with a low-power LED component connected to each finger and a plastic ball covering it. These LED components are powered by a standard battery and controlled by an Arduino Nano. They produce identical LED waves to represent various fingers. The reader is based on commercial cameras, including smartphones, webcams, cameras placed to drones, and even underwater sports cameras. The lenses on these cameras may be adjusted for focal length, and the rolling shutters have adjustable shutter speeds. After knowing about the finger-grained location variation of each finger and know which finger it is, we can absolutely recognize the current hand gestures.

4 EVALUATION

Setup: To evaluate the performance of the system for enhanced vision based gesture recognition, we use the implemented prototype presented in Section 3 and evaluate its performance in terms of the effective recognition distance, and enhanced location variation ability in both vertical and horizontal motion directions with three different moving paths: (1) →, (2) ↑, and (3) ↗.

Results: ***(1) Sensing Distance:*** As shown in Figure 3, the system can successfully record rolling patterns at up to 2m, which is a significant improvement of sensing distance from previous about 0.5m. Compared with traditional vision based gesture recognition, our rolling shutter enhanced recognition can achieve long-distance recognition with identification and able to be used in private or dark environments. ***(2) Improved Motion Variation Range:*** As shown in the right of Figure 3, our rolling shutter enhanced recognition can improve the motion variation range of more than the average of 20% and 10% for the vertical and horizontal directions for three different motion paths. Theses results demonstrate the rolling shutter enhance gesture recognition can provide finger-grained and long-ranged tracking with objective identification.

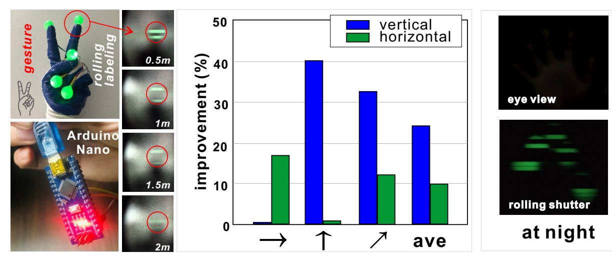

Figure 3: Implemented prototype for rolling shutter enhanced gesture recognition and improved performance.

5 CONCLUSION

In this work, we designed and implemented a rolling shutter enhanced vision gesture recognition prototype for the better HCI. This technology can make gesture recognition have finer-grained tracking, longer sensing distance, objective identification ability as well as privacy protection. In the future, it can also be extended to other different objective recognition with active optical labeling to provide smarter Cyber-Physical interactions.

ACKNOWLEDGMENTS

This work was partially supported by the U.S. National Science Foundation under Grants CCF-2007159 and CNS-2226888.

REFERENCES

[1] Tianxing Li, Xi Xiong, Yifei Xie, George Hito, Xing-Dong Yang, and Xia Zhou. 2017. Reconstructing hand poses using visible light. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–20.

[2] Xiao Zhang, Hanqing Guo, James Mariani, and Li Xiao. 2022. U-Star: An Underwater Navigation System based on Passive 3D Optical Identification Tags. In *The 28th Annual International Conference on Mobile Computing and Networking.*

[3] Xiao Zhang, James Mariani, Li Xiao, and Matt Mutka. 2022. LiFOD: Lighting Extra Data via Fine-grained OWC Dimming. In *2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON).* IEEE, 1–9.

[4] Xiao Zhang and Li Xiao. 2020. RainbowRow: Fast Optical Camera Communication. In *2020 IEEE 28th International Conference on Network Protocols (ICNP).* IEEE, 1–6.