Integrated Sensing and Communication for Optical Camera Communication

Ida Bagus Krishna Yoga Utama, Su Mon Ko, Irzal Zaini, Muhammad Fairuz Mummtaz, Yeong Min Jang
Department of Electronics Engineering
Kookmin University
Seoul, Korea
ibkyutama@ieee.org, sumonko@kookmin.ac.kr, irzal.zaini00@gmail.com, muhammad.fairuz91@ui.ac.id, yjang@kookmin.ac.kr

***Abstract*—The Integrated Sensing and Communication (ISAC) is essential technology for the future 6G networks. By integrating the sensing and communication, more efficient communication systems can be achieved. At present, most ISAC systems are proposed for Radio Frequency (RF). In this paper, we present a novel idea to perform ISAC for Optical Camera Communication (OCC). The proposed ISAC system combines a monocular depth estimation method with the OCC system to perform both communication and sensing in one-shot and end-to-end manner. This method eliminates the need for sensor fusion with LiDAR or RADAR to perform an accurate sensing. The proposed method shows a centimeter-level localization accuracy in indoor environment while maintaining OCC link.**

***Index Terms*—6G Networks, Integrated Sensing and Communication, Monocular Depth Estimation, Optical Camera Communication, Optical Wireless Communication**

I. Introduction

Over the years, demand for hyper-connectivity and ubiquitous intelligence accelerates causing the telecommunications industry to rapidly evolving toward the sixth generation (6G) of wireless networks. A major research direction of this evolution is Integrated Sensing and Communication (ISAC). Traditionally, sensing and communication have been treated as separate system requiring distinct hardware and spectral resources. ISAC represents a paradigm shift by unifying these two functions into a single system, thus enhancing hardware efficiency, spectrum utilization, and overall network intelligence [1], [2].

While the concept of ISAC is gaining significant traction, the majority of current research and implementation efforts are predominantly focused on the Radio Frequency (RF) domain [3]. Numerous research has been published to perform ISAC and optimize its system in RF domain. Differently, in Optical Wireless Communication (OWC) domain, the ISAC method is not as popular as in RF domain. In OWC, the ISAC mainly proposed for Visible Light Communication (VLC) by relying on the passive sensing [4]. In this system, the sensing is achieved by analyzing the changes in light intensity or channel impulse response on the receiver part. Based on the analyzed pattern of light reflections or shadows, the system is able to performhuman activity recognition or object detection and tracking with centimeter-level accuracy.

Another method of ISAC for OWC is applied in a laser communication system. An Optical Phased Arrays (OPA) is utilized to establish the ISAC for OWC [5]. Essentially, an OPA is a steerable lasers where the beam can be controlled to focus only to certain direction, depending on user location, and lock onto them to send the data. By taking the advantages of OPA, the laser communication system can perform sensing and communication in a single device without having to use other device like LiDAR to perform the sensing.

Optical camera communication (OCC) is one type of OWC where the transmitter is using LED and camera is used as the receiver [6]. Technically, by using a camera, the ISAC can be performed directly in OCC systems because the camera provides the location and communication data embedded in a single image frame. However, in the implementation, the image used for decoding in OCC should be processed to filter out noise and remain only the communication data. During this processing steps, the image usually transformed and thresholded to show only the LED lights and remove any other objects. Hence, the sensing is difficult to perform since the resulted image usually only in black and white images with no other objects other than the LED. Moreover, performing sensing and localization from a 2D images usually have a lower accuracy in distance estimation.

That is why several research perform sensor fusion for OCC to enhance the sensing and localization performance. A LiDAR or RADAR that generates a 3D point cloud data can be used to perform the sensing for localizing the presence and position of the LED transmitter [7]. After localization using such sensors, the communication part will start the decoding process to retrieve the original data. Although this method works well, it has several inherent downsides. First, by using sensor fusion, it makes the system more complex due to the usage of multiple sensor. Then, the overall system power consumption might be increased. Finally, sensor like LiDAR or RADAR is costly and usually much more expensive than the camera for receiver part.

In this paper, we present a new method that simplifies the sensing and communication method in OCC system. A novel ISAC framework is proposed specifically designed for OCC usage. Unlike traditional RF-ISAC or complex sensor-fusion models, our proposed system leverages the visual data inherently captured by the OCC camera receiver to perform both communication and sensing simultaneously.

II. Proposed System

The proposed ISAC system is established by integrating monocular depth estimation model in the receiver of OCC system. As such, there are two main components to establish the main systems: OCC and monocular depth estimation model. By integrating the aforementioned system, the ISAC for OCC is established. This section covers the process to establish ISAC for OCC and details of each module used in the proposed ISAC for OCC method.

A. ISAC for OCC

The proposed ISAC for OCC solve issue mainly on the receiver part of the OCC system. The proposed method introduce a steps to perform the ISAC based on the combination of monocular depth estimation and existing system to generate both sensing and communication together at one time, as shown in Fig. 1. The proposed method is started by capturing the RGB image produced by the camera. Every generated image frame is utilized for performing the sensing and communication.

The image frame is copied into two data, one for sensing and another for communication method. The image copy for sensing is inserted into the monocular depth estimation model to generate the estimated depth image of the scene. Based on the generated depth image, each pixel refers to the estimated distance between the object and the camera. Then, an object detection and tracking is applied to detect the location of the LED transmitter. Therefore, the position of the LED is known both in 2D and 3D space, where the object detection and tracking generates the 2D position and monocular depth estimation generates the 3D position.

Then, the second image copy is utilized for the communication module. The image is inserted into the image preprocessing pipeline to change the image properties for extracting the LED features. Then, the bits in the LED are extracted by considering the LED is on or off. After receiving the bits from LED, the demodulation process is continued to recover the original data.

In the end of the method, both information of sensing and communication are generated at the similar time. The retrieved communication data and transmitter location are then utilized to perform further technologies such as routing or topology control in an OCC network. Fig. **??** and Fig. **??** shows the original image captured by the camera and the resulted depth map estimation from the monocular depth estimation model. Based on the pixel value in the depth image, the distance of the LED can be retrieved.

B. Optical Camera Communication

The OCC system considered in this paper is a MIMO-COOK system. A 16×16 LED matrix is employed as the transmitter. Meanwhile a global shutter camera is utilized as the receiver. The modulation strategy is using MIMO-COOK with Manchester coding and Reed-Solomon Forward Error Correction (FEC) to ensure robust and reliable data transmission. The OCC architecture utilized in this work shown in Fig. 4.

1) Transmitter Design and Modulation: The transmitter consists of an 16×16 LED matrix, where each individual LED functions as an independent communication channel. To address the inherent flickering issues and synchronization challenges of standard OOK, the system utilizes Manchester coding. Before modulation, the raw binary data is processed using Reed-Solomon (RS) Forward Error Correction. RS codes are block-based error-correcting codes that are particularly effective at correcting burst errors—a common issue in OCC channels due to motion blur, partial occlusion, or ambient light interference. The RS encoder adds redundant parity symbols to the data stream, allowing the receiver to detect and reconstruct corrupted symbols without requiring re-transmission.

2) Global Shutter Reception and Demodulation: The receiver utilizes a high-speed camera equipped with a Global Shutter sensor. This characteristic eliminates the spatial banding (striping) effect typically seen in rolling shutter OCC. To achieve data transmission, the system relies on frame-by-frame temporal sampling:

- Synchronization: The camera's frame rate (F_{cam}) must be synchronized with the transmitter's symbol rate (R_{sym}). The camera operates at three times higher sampling rate than the transmitter.
- ROI Extraction: For every captured frame, the system detects the LED matrix and isolates the Region of Interest (ROI).
- Manchester Decoding & FEC: The sequence of intensity values from consecutive frames is analyzed to identify the Manchester transitions (Low-to-High or High-to-Low). Finally, the Reed-Solomon decoder processes the demodulated bit-stream to correct errors and recover the original data.

C. Monocular Depth Estimation

Monocular Depth Estimation is a fundamental computer vision task that aims to predict the dense depth value of each pixel from a single 2D RGB image. Formally, given a single input image $I \in \mathbb{R}^{H \times W \times 3}$, the objective is to learn a mapping function f such that $D = f(I)$, where $D \in \mathbb{R}^{H \times W}$ represents the estimated depth map. Monocular depth estimation is inherently an ill-posed problem. The model is developed based on DepthAnything V3 model [8].

In the context of OCC, it is crucial to get the best accuracy value for performing the depth estimation. The model is developed to estimate the distance based on the single training image. The DepthAnything V3 model is chosen for this work because it is suitable for our purpose in ISAAC simulator in Gazebo indoor environment.

A critical challenge in monocular depth estimation is scale ambiguity. Recent SOTA methods, including ZoeDepth and

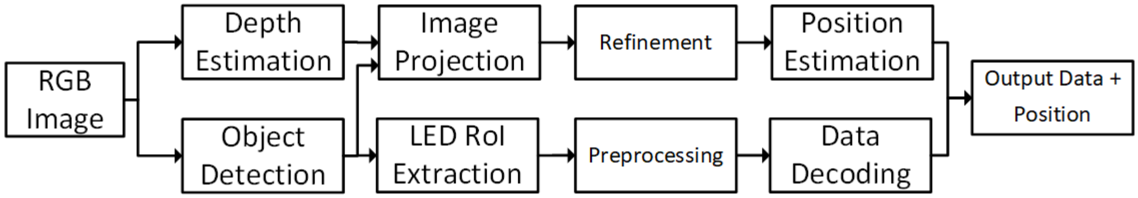

Fig. 1. Proposed method for ISAC for OCC system.

Fig. 2. Original image captured by the camera.

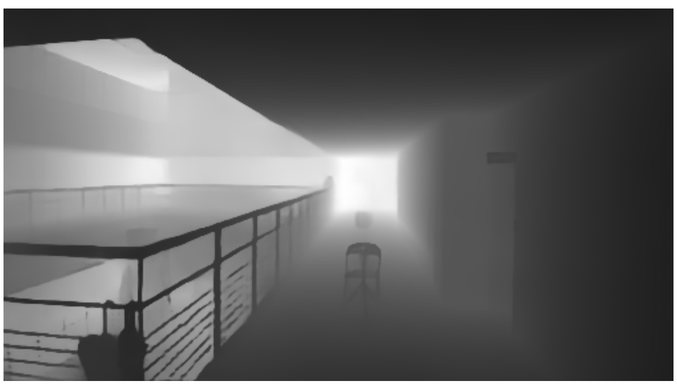

Fig. 3. Depth image from monocular depth estimation model.

Marigold, are focusing on recovering metric depth by incorporating camera intrinsic parameters or using affine-invariant loss functions to align relative predictions with real-world scales.

III. Prospects of ISAC for OCC

The integration of Monocular Depth Estimation (MDE) with Optical Camera Communication (OCC) represents a significant progress for Integrated Sensing and Communication (ISAC) systems, particularly for indoor 6G applications. By leveraging the visual data inherent in the communication channel, this proposed architecture offers unique advantages over traditional RF-based or sensor-fusion approaches.

A. Hardware-Efficient Sensing

The most compelling prospect of this integration is the achievement of sensing capabilities with effectively zero additional hardware cost because only using one camera.

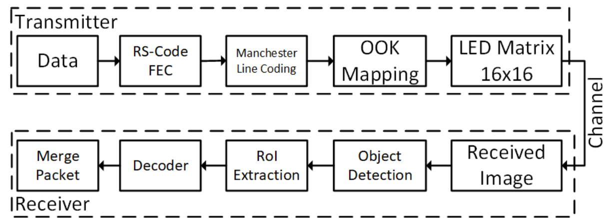

Fig. 4. Architecture of the OCC using OOK modulation.

Current Limitations are caused by traditional robotic or autonomous systems typically require a dedicated sensor stack. The OCC receiver (camera) transforms into a dual-function device. It simultaneously decodes high-speed data from the LED transmitter and calculates the precise distance to it. This eliminates the need for expensive, power-hungry active sensors like LiDAR, drastically reducing the weight, cost, and energy consumption of mobile agents such as UAVs or AGVs.

B. Interference-Free Indoor Localization

RF-based ISAC systems suffer significantly in complex indoor environments due to multipath fading, reflection, and electromagnetic interference (EMI) from other devices. The proposed optical ISAC system operates in the visible light spectrum, which is immune to RF interference.

C. Semantic-Aware Communication

Unlike RF waves, which only detect physical reflections, a camera can see the context clearly by capturing the image. The integration allows for semantic-aware ISAC. The system is smarter when using this method. The semantic allows the model to recognize object other than the LED for knowing the environment results in a better communication technology.

IV. Future Works

While the proposed Monocular Depth Estimation-based ISAC system demonstrates significant potential for indoor applications, several avenues for optimization and expansion remain. Future research will focus on enhancing the system's robustness, computational efficiency, and sensing resolution to meet the rigorous demands of 6G networks.

A. Lightweight Neural Networks for Edge Deployment

The current deep learning models for depth estimation, while accurate, are computationally intensive. To facilitate deployment on resource-constrained edge devices, future work will prioritize the development of lightweight depth estimation model. The goal is to achieve real-time (more than 30 FPS) simultaneous communication and sensing on low-power embedded processors.

B. Robustness to Environmental Variability

Optical channels are susceptible to environmental noise, including motion blur, ambient light interference, and occlusion. Future research will incorporate temporal consistency constraints into the learning framework. By analyzing video sequences rather than individual frames, the system can enforce geometric consistency over time, effectively filtering out transient noise and improving the stability of both the data link and the depth map during high-speed mobility. Anomaly detection also possible to be integrated to the future system to enhance the system robustness and reduce error [9].

C. Integration of Semantic Segmentation

Finally, we aim to evolve the system from purely geometric sensing to semantic-aware ISAC. By integrating semantic segmentation heads into the AI model architecture, the receiver will not only estimate the distance to the transmitter but also classify the surrounding environment. This context-awareness will enable intelligent resource allocation to dynamically adjust and control the LED transmission based on the scene condition.

V. Conclusion

This paper presented a novel framework for Integrated Sensing and Communication (ISAC) specifically designed for Optical Camera Communication (OCC) systems. By integrating a Monocular Depth Estimation model with a MIMO-COOK communication architecture, we demonstrated that a single camera receiver can perform data decoding and precise distance estimation. This approach effectively eliminates the dependency on complex sensor fusion with LiDAR or RADAR, offering a streamlined, cost-effective, and energy-efficient solution for OCC system. The proposed method utilizes the DepthAnything V3 model to achieve centimeter-level localization accuracy while ensuring robust communication through Manchester coding and Reed-Solomon FEC. Ultimately, this work establishes a foundation for pervasive optical ISAC, paving the way for semantic-aware and hardware-efficient mobile agents in future intelligent environments.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No.2022R1A2C1007884).

References

[1] M. U. F. Qaisar, W. Yuan, O. Günlü, T. Riihonen, Y. Cui, L. Zhang, N. Gonzalez-Prelcic, M. D. Renzo, and Z. Han, "The role of isac in 6g networks: Enabling next-generation wireless systems," 2025. [Online]. Available: https://arxiv.org/abs/2510.04413

[2] D. Wen, Y. Zhou, X. Li, Y. Shi, K. Huang, and K. B. Letaief, "A survey on integrated sensing, communication, and computation," *IEEE Communications Surveys Tutorials*, vol. 27, no. 5, pp. 3058–3098, 2025.

[3] K. Wu, Y. Bigdeli, S. A. Keivaan, J. Deng, and P. Burasa, "Integrated sensing and communication (isac) transceiver: Hardware architectures, enabling technologies, and emerging trends," *IEEE Journal of Selected Topics in Electromagnetics, Antennas and Propagation*, vol. 1, no. 1, pp. 37–64, 2025.

[4] R. Zhang, Y. Shao, M. Li, L. Lu, and Y. C. Eldar, "Optical integrated sensing and communication with light-emitting diode," in *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2024, pp. 2059–2064.

[5] Y. Wen, F. Yang, J. Song, and Z. Han, "Optical wireless integrated sensing and communication based on optical phased array: Performance metric and optimal beamforming," *IEEE Transactions on Wireless Communications*, vol. 24, no. 9, pp. 7221–7236, 2025.

[6] M. F. Ahmed, M. K. Hasan, M. Shahjalal, M. M. Alam, and Y. M. Jang, "Design and implementation of an occ-based real-time heart rate and pulse-oxygen saturation monitoring system," *IEEE Access*, vol. 8, pp. 198 740–198 747, 2020.

[7] I. B. K. Y. Utama, O. S. Sitanggang, M. R. A. Nasution, M. I. Joha, J. Yoo, and Y. M. Jang, "Enhancing optical camera communication performance for collaborative communication using positioning information," *IEEE Access*, vol. 12, pp. 11 795–11 809, 2024.

[8] H. Lin, S. Chen, J. H. Liew, D. Y. Chen, Z. Li, G. Shi, J. Feng, and B. Kang, "Depth anything 3: Recovering the visual space from any views," *arXiv preprint arXiv:2511.10647*, 2025.

[9] R. F. Pamungkas, I. B. K. Y. Utama, K. Hindriyandhito, and Y. M. Jang, "A hybrid approach of ConvLSTMBNN-DT and GPT-4 for real-time anomaly detection decision support in edge–cloud environments," *ICT Express*, vol. 10, no. 5, pp. 1026–1033, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405959524000857