Machine Learning-Augmented Architecture for UAV-to-UAV OAM Links

Young Jae Moon, Yeon Ho Chung*
Department of Artificial Intelligence Convergence, Pukyong National University
8068joshua@pukyong.ac.kr, *yhchung@pknu.ac.kr

무인 항공기 (UAV) 간 각 궤도 운동량(OAM) 통신을 위한 기계 학습 알고리즘과 모형

문영제, 정연호*
부경대학교 인공지능융합학과*

Abstract

Orbital angular momentum, a property of electromagnetic waves that allows them to carry multiple data streams on a single frequency, has garnered numerous attention in recent years. However, unmanned aerial vehicles are also used for a wide range of applications with wireless communication. In this paper, we study a UAV-to-UAV free-space optical link that uses orbital angular momentum (OAM) modes and propose a beacon-aided, dual-loop pointing, acquisition, and tracking architecture augmented with lightweight learning. The simulation results show that our system reduces the 95th-percentile lateral miss from 7.0 mm to 4.2 mm and suppresses nearest neighbor crosstalk against misalignment.

Ⅰ. Introduction

In recent years, beams carrying orbital angular momentum (OAM) [1] have attracted interest due to the orthogonality between the beams with different modes [2]. OAM enables multiple orthogonal data streams, thereby enhancing the scalability and system capacity efficiency. With this additional degree of freedom of OAM, unmanned aerial vehicles (UAVs) can use OAM-based optical communication for the efficient data transmission [3]. However, the UAV-to-UAV OAM communication system is yet to be investigated, compared to the quantum key distribution (QKD) link between the UAVs [4]. Therefore, we propose an OAM-based communication system between UAVs, taking into account its pointing error and tracking jitter.

Ⅱ. Background Information

Most of the OAM-based communication systems utilize Laguerre-Gaussian (LG) beams, due to their unique properties. LG beams are a natural solution to the paraxial wave equation, describing how a laser beam propagates in free space. LG beams can be expressed as [5]

$$E_{p,\ell}(r,\phi,z) = C_{p,\ell}\frac{1}{w(z)}\left(\frac{\sqrt{2}\,r}{w(z)}\right)^{|\ell|} L_p^{|\ell|}\left(\frac{2r^2}{w^2(z)}\right) \times \exp\left(\frac{-r^2}{w^2(z)} - ik\frac{r^2}{2R(z)}\right) \times \exp\left[i\ell p - i(2p+|\ell|+1)\zeta(z)\right] \quad (1)$$

where $w(z)$ is the beam waist at propagation distance z, $L_p^{|\ell|}(\cdot)$ is the associated Laguerre polynomial, z_R is the Rayleigh range, $R(z)$ curvature radius, and $\zeta(z)$ is the Gouy phase.

QKD is a point-to-point communication protocol which leverages quantum information to enable security [6]. However, the movements of the UAVs prohibit the alignment between the OAM transmitter and the receiver, which is critical to the communication channel [7]. This returns the communication system a pointing error and tracking jitter. The pointing error with OAM-based channel can be expressed as [7]

$$\eta_{point}(\delta) \approx exp\left[-2\left(\frac{\delta}{w_0\sqrt{1+(R/z_R)^2}\sqrt{|\ell|+1}}\right)\right] \quad (2)$$

where δ is the lateral misalignment, while the nearest-eighbor cross talk with tracking jitter can be expressed as

$$E[P_{\ell\rightarrow\ell\pm1}] \approx \frac{w_\ell(R)}{2}\frac{dP_{\ell\rightarrow\ell\pm1}}{d(\Delta^2)}|_{\Delta=0}\frac{|\mu|^2+2\sigma^2}{w_0\sqrt{1+(R/z_R)^2}\sqrt{|\ell|+1}} \quad (3)$$

where $P_{\ell\rightarrow\ell\pm1}$ is the probability that a symbol with OAM order ℓ is interpreted as adjacent OAM order $\ell\pm1$, $E[\cdot]$ is the average over fast jitter, μ is the lateral bias at the receiver plane from slow misalignment of the beam centroid, σ is the per-axis root-mean-squared lateral tracking jitter at the receiver plane, Δ is the lateral offset, and k_ℓ is the dimensionless constant of order one.

Ⅲ. System Architecture

To mitigate the misalignment, we deploy a beacon-aided dual-loop system with a fast-steering mirror. Moreover, we add a motorized, multi-axis stabilized mount, which searches the partner's beacon and keeps the boresight aligned to the partner UAV. Building on this, the outer loop uses the gimbal guided by a wide-field-of-view acquisition camera and inherital measurment unit (IMU) cues for coarse pointing, while the inner loop drives the mirror at high rate using a position-sensitive detector or a sub-pixel centroid from a lightweight vision head. An extended Kalman filter fuses the IMU to estimate tip and tilt, and a small temporal predictor supplies one-step feed-forward that reduces latency in the fine loop. A supervisory state machine manages search, capture, lock, track, and recover, and quality gates based on beacon confidence and OAM coupling ensure authority handoffs between loops remain stable.

Algorithm 1 Tiny Temporal CNN with TCN Residual Blocks

1: **Input:** Visual frames $V \in R^{B\times H\times W\times T\times 1}$, IMU sequence $I \in R^{B\times T\times d}$
2: **Visual Stream:**
  Reshape $V \rightarrow (B, H, W, T)$
  Conv2D(3×3, 8ch, ELU)
  Residual Block (2 convs, 8ch) + ELU
  MaxPool2D(2×2)
  Residual Block (2 convs, 16ch + 1×1 skip) + ELU
  Global AvgPool$_{H,W}$ $\rightarrow (B, 16)$
  Reshape $\rightarrow (B, T, 16)$ (intended per-time-step features)
3: **IMU Stream:**
  Conv1D(3, 8ch, ELU) $\rightarrow (B, T-2, 8)$
  Global AvgPool$_{time}$ $\rightarrow (B, 8)$
4: **Fusion:**
  Flatten visual seq $(B, T, 16) \rightarrow (B, 80)$
  Concatenate with IMU $(B, 80+8) \rightarrow (B, 88)$
  Reshape $\rightarrow (B, 88, 1)$
5: **Temporal Convolutional Network (TCN):**
  TCN Residual Block: Conv1D(dilation=1, 64ch) → BN + ELU + Dropout
  TCN Residual Block: Conv1D(dilation=2, 32ch) → BN + ELU + Dropout
  Global AvgPool$_{seq}$ $\rightarrow (B, 32)$
6: **Feedforward Head:**
  Dense(16, ELU)
  Dense(2, Linear) → Output $(\hat{y}_{tip}, \hat{y}_{tilt})$
7: **Output:** Predicted tip/tilt vector $\in R^2$

Fig 1. Algorithm table with UAV-to-UAV OAM communication.

We also make an IMU-conditioned convolutional neural network (CNN) to localize the beacon and output a sub-pixel centroid with a confidence score. A tiny temporal CNN predicts the next line-of-sight tip and tilt one step ahead for feed-forward to the fast steering mirror. With the actions of UAV, the residual blocks learn spatial structure at low cost. Then, the temporal fusion uses a lightweight shift over the short clip to share motion cues, while the output shows the next tip and tilt for feed-forward. The overall algorithm is shown in Fig. 1.

Ⅳ. Simulation

We simulate a 623 nm LG beam with p=0 and ℓ=+1, with an initial transmit beam waist of 10 mm. Two UAVs with 100mm apertures fly at a fixed 200m distance while retaining a line-of-sight relationship. However, the atmosphere follows a horizontal Kolmogorov model set of 1.0×10^{-15},3.0×10^{-15} and $1.0\times10^{-14}m^{-2/3}$, while the inner scale is 5 mm and the outer scale is 50mm. Pointing error is scaled as a slow bias of 0 - 10 μrad per axis with zero-mean jitter of 10 μrad per axis per terminal. The OAM sorter assumes −22 dB nearest-neighbor leakage with ≤3 dB mode-dependent loss, and interprets the received OAM beam. For the evaluation method, we calculate lateral miss as

$$\Delta(t) = R\sqrt{\theta_x^2(t)+\theta_y^2(t)} \quad (3)$$

where θ_x and θ_y is the actual respective lateral misalignments for each coordinates. We also compute the nearest-neighbor crosstalk from the different received intensities from (2) while considering the leakage, when the architecture is applied and when it is not.

V. Results

Simulation results demonstrate the algorithm's significant performance improvement as turbulence increases. The proposed method reduces the 95th-percentile lateral miss from approximately 7.0 mm to 4.2 mm at low turbulence. Moreover, our system suppresses nearest-neighbor crosstalk by keeping it low at approximately 3.3% to 4.6%, while it increases from ~20% to ~25% when algorithm is not in place. The results confirm that our proposal, utilizing a CNN, is highly successful

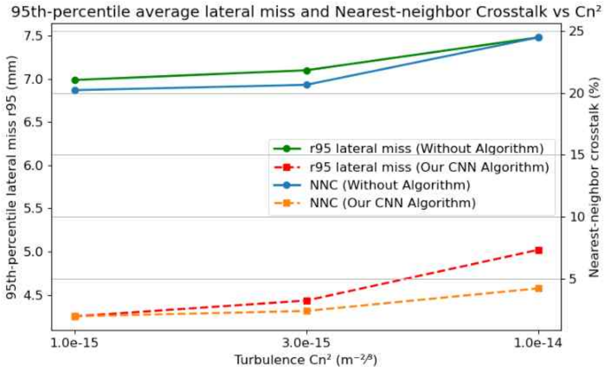

Fig 2. 95% Average lateral miss and nearest-neighbor crosstalk with our algorithm when compared with atmospheric turbulence.

in beacon localizing and predicting necessary line-of-sight corrections.

VI. Conclusion

In our paper, we have studied a UAV-to-UAV free-space OAM link that uses a beacon-aided, tracking architecture augmented with light-weight learning. The architecture utilizes an outer loop with a stabilized gimbal, while the inner loop employs a fast steering mirror and a temporal CNN predictor to reduce latency. Simulation results demonstrate that our method reduces the 95th-percentile lateral miss from 7.0 mm to 4.2 mm and suppresses nearest neighbor crosstalk.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2023R1A2C2006860).

REFERENCES

[1] M. Padgett and R. Bowman, "Tweezers with a twist," Nature Photonics, vol. 5, no. 6, pp. 343 - 348, 2011.

[2] G. Gibson, J. Courtial, M. J. Padgett, et al., "Free-space information transfer using light beams carrying orbital angular momentum," Optics Express, vol. 12, no. 22, pp. 5448 - 5456, 2004.

[3] R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, et al., "Entanglement-based quantum communication over 144 km," Nature Physics, vol. 3, pp. 481 - 486, 2007.

[4] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, et al., "Satellite-to-ground quantum key distribution," Nature, vol. 549, pp. 43 - 47, 2017.

[5] A. E. Siegman, Lasers. University Science Books, 1986.

[6] Z. Qu, G. Xie, L. Li, et al., "Performance of free-space optical communications with misalignment: Spatial mode multiplexing perspective," Optics Express, vol. 25, no. 26, pp. 31 561 - 31 572, 2017.

[7] J. Sun, A. Zhang, and H. Xu, "Image formation for SAR satellites using OAM-carrying beams," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1 - 12, 2022.