A Coarse-to-Fine Approach for Estimating FHSS Parameters of UAVs

Soon-Young Kwon
Dept. of Electrical and Electronics Enginnering
Pusan National University
Busan, Republic of Korea
ysk1680@pusan.ac.kr

Nayun Park
Dept. of Electrical and Electronics Enginnering
Pusan National University
Busan, Republic of Korea
nayun_p@pusan.ac.kr

Hyoung-Nam Kim
School of Electrical and Electronics Enginnering
Pusan National University
Busan, Republic of Korea
hnkim@pusan.ac.kr

***Abstract*—With the increasing deployment of unmanned aerial vehicles (UAVs) in diverse environments, robust signal analysis capabilities are essential for counter-UAV operations. In particular, Accurate estimation of frequency hopping spread spectrum (FHSS) parameters in UAV uplinks is important for link characterization and electronic warfare, especially under low signal to noise ratio (SNR) conditions. However, spectrogram based semantic segmentation methods suffer from limited frequency resolution because the short time Fourier transform uses a fixed FFT size. To overcome this limitation, this paper proposes a coarse to fine FHSS parameter estimation scheme for UAV uplink signals. First, a DeepLabV3+ based semantic segmentation network separates uplink, downlink, drone ID, and noise components on the spectrogram and provides coarse dwell times and frequency bands for each uplink hop. Then, for every dwell time, a high resolution power spectral density (PSD) is computed from the time domain signal, and a region of interest is defined from the coarse frequency band. The PSD outside this region is used to estimate and subtract the average noise power, and the resulting noise calibrated PSD is used to estimate the hop center frequency and occupied bandwidth. Simulation results for OcuSync 2.0 UAV uplink signals show that the proposed method consistently outperforms a coarse only baseline. At high SNR, the RMSE of center frequency estimation is reduced from 0.11 MHz to 0.01 MHz and the RMSE of occupied bandwidth estimation is reduced from 0.86 MHz to 0.05 MHz.**

***Index Terms*—coarse-to-fine, semantic segmentation, FHSS, parameter estimation**

I. Introduction

Recently, unmanned aerial vehicles (UAVs) have been actively used for various military and tactical purposes, such as reconnaissance, surveillance, and target strikes [1]. As a result, the importance of electronic warfare (EW) technologies for detecting and identifying hostile UAVs has significantly increased [2]. In this environment, to characterize UAV communication links, it is essential to precisely estimate frequency hopping spread spectrum (FHSS) parameters [3]. These parameters include the center frequency and bandwidth of each hop. In particular, the ability to stably estimate these parameters even in low signal-to-noise ratio (SNR) conditions is a key element for counter-UAV operations, such as link analysis and jamming strategies [2].

Commercial UAV systems often employ complex signal structures, such as the OcuSync protocol from DJI [4]. In these systems, signals for different purposes, including uplinks using FHSS, downlinks, and Drone IDs, are transmitted and received in a time-division manner across various frequency bands. In such environments, various drone signals and noise are observed in the time-frequency domain at the EW receiver. Therefore, it is crucial to effectively separate each signal component and then precisely estimate the parameters for the FHSS signals [5].

As an approach to address this issue, various methods have been studied that apply the short time Fourier transform (STFT) to the received signal to generate a spectrogram and process the signal directly in the time-frequency domain [6]. For instance, previous research proposed a technique that treats the spectrogram as an image and applies a semantic segmentation model based on DeepLabV3+ [7]. By generating segmentation maps for the uplink, downlink, Drone ID, and noise components, this method effectively separates the drone communication signals. Results from estimating the center frequency using the separated spectrogram showed that this method outperforms the conventional dual sliding window (DSW) based approach in terms of signal separation and parameter estimation [8].

However, these spectrogram-based semantic segmentation techniques have an intrinsic limitation from the perspective of FHSS parameter estimation. In these methods, the center frequency and bandwidth of each hop are estimated directly from the discrete frequency grid of the spectrogram, where the frequency spacing is determined by the fast Fourier transform (FFT) size N_{FFT} used for the STFT. Since it is difficult to set a large N_{FFT} due to computational constraints, spectrograms with relatively coarse frequency spacing are inevitably used in practice. Consequently, even if the semantic segmentation itself is accurate, the estimated values of the hop center frequency and bandwidth are quantized to the coarse frequency grid, resulting in an irreducible root mean square error (RMSE).

Therefore, to more precisely estimate the FHSS parameters of UAV communication signals, a new design paradigm is required to overcome the trade-off between frequency resolution and computational complexity inherent in spectrogram-based semantic segmentation techniques. Specifically, a Coarse-to-

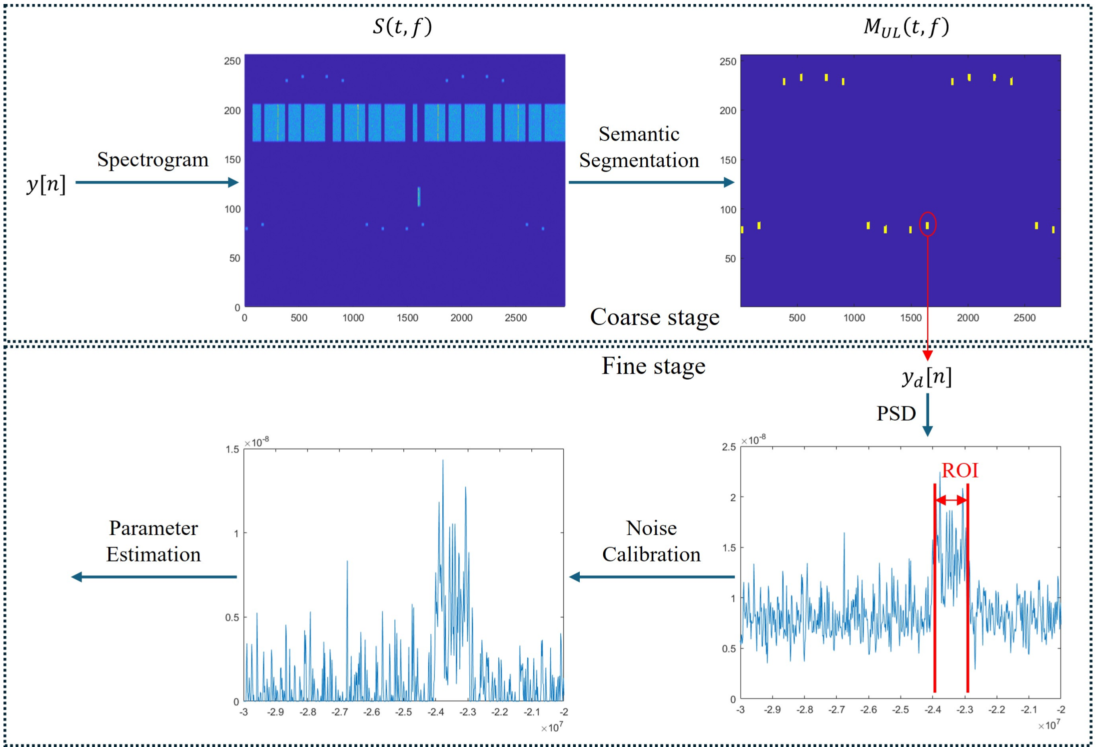

Fig. 1. Overall block diagram of the Coarse-to-Fine framework structure.

Fine architecture is needed to effectively reduce the RMSE. In this structure, deep learning is intensively utilized for the coarse detection of signal presence regions, while parameters requiring high frequency resolution, such as hop center frequency and bandwidth, are processed in the fine stage using spectral analysis with high-resolution FFT within the narrowed regions of interest (ROI). Furthermore, since the areas outside the ROI can be regarded as pure noise intervals, the average noise power can be estimated from these regions and used to calibrate the power spectral density (PSD), thereby further enhancing the accuracy of hopping parameter estimation.

II. Proposed Coarse-to-Fine Framework

A. Overall Framework Structure

The method proposed in this paper builds upon the existing semantic segmentation-based drone signal separation technique by incorporating a parameter estimation stage with a Coarse-to-Fine structure [8]. Fig. 1 illustrates the overall block diagram of the proposed framework.

First, in the coarse stage, the received signal is transformed into a spectrogram via the STFT. Subsequently, a DeepLabV3+ based semantic segmentation model is applied to generate segmentation maps for the uplink, downlink, Drone ID, and noise components. From the uplink segmentation map, the coarse dwell time and the coarse frequency band where the uplink signal resides can be simultaneously extracted.

In the subsequent fine stage, time domain signals are extracted for each uplink dwell time identified in the coarse stage. The PSD is then calculated for each dwell time to precisely estimate the hop center frequency and bandwidth. Furthermore, a frequency ROI is defined based on the coarse frequency range of the uplink obtained in the coarse stage. The spectrum outside this ROI is utilized as a noise-only interval to estimate the average noise power, which is then subtracted from the PSD. Consequently, even if the frequency resolution of the spectrogram used for semantic segmentation is relatively coarse, the RMSE of the FHSS parameters can be effectively reduced through high-resolution PSD analysis and noise calibration in the fine stage.

B. Coarse Stage: Uplink Signal Separation and Information Extraction via Semantic Segmentation

The coarse stage adopts the same architecture as the semantic segmentation-based drone signal separation technique proposed by [8]. The STFT is performed on the received baseband signal $y[n]$ to obtain the complex spectrogram $S(t,f)$, and its magnitude $|S(t,f)|$ is utilized as the input to the DeepLabV3+ model.

The DeepLabV3+ model employs an encoder-decoder structure with ResNet18 serving as the backbone neural network. The trained semantic segmentation model outputs four binary segmentation maps, $M_{\mathrm{UL}}(t,f)$, $M_{\mathrm{DL}}(t,f)$, $M_{\mathrm{ID}}(t,f)$, and $M_{\mathrm{N}}(t,f)$, which indicate whether each time-frequency pixel belongs to the uplink, downlink, Drone ID, or noise class, respectively. Here, $M_{\mathrm{UL}}(t,f) \in \{0,1\}$ denotes whether the corresponding time-frequency pixel belongs to the uplink signal.

a) Coarse Dwell Time: Let $t = 1, \ldots, T$ and $k = 1, \ldots, K$ denote the time and frequency indices of the spectrogram, respectively. The existence of an uplink signal at time index t is defined as $\hat{a}_{\mathrm{UL}}(t)$:

$$\hat{a}_{\mathrm{UL}}(t) = \begin{cases} 1, & \exists k \text{ s.t. } M_{\mathrm{UL}}(t, f_k) = 1 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

By connecting consecutive intervals where $\hat{a}_{\mathrm{UL}}(t) = 1$, the dwell time intervals $[t_s^{(d)}, t_e^{(d)}]$ for $d = 1, \ldots, D$, where uplink hops are present, can be obtained. These intervals represent the coarse dwell time of the corresponding hops.

b) Coarse Frequency Band: For each dwell interval d, the coarse frequency band for the corresponding hop is defined using the minimum and maximum frequency indices k satisfying $M_{\mathrm{UL}}(t, f_k) = 1$ within the time range $t_s^{(d)} \leq t \leq t_e^{(d)}$:

$$f_L^{(c,d)} = \min\{f_k \mid M_{\mathrm{UL}}(t, f_k) = 1\}, \tag{2}$$

$$f_H^{(c,d)} = \max\{f_k \mid M_{\mathrm{UL}}(t, f_k) = 1\}, \tag{3}$$

where the superscript (c, d) indicates that $f_L^{(c,d)}$ and $f_H^{(c,d)}$ are the coarse lower and upper frequency bounds for the d-th uplink hop. Here, since f_k represents a discrete frequency axis with a spacing of $\Delta f_{\mathrm{STFT}} = f_s / N_{\mathrm{STFT}}$ determined by the FFT size N_{STFT} used in the STFT, the frequency information obtained in the coarse stage is an approximation limited to this resolution level.

In summary, the coarse stage simultaneously extracts the dwell time and approximate frequency band of the uplink hops via semantic segmentation, which are subsequently utilized for initial estimation and ROI configuration in the fine stage.

C. Fine Stage: FHSS Parameter Estimation using PSD and Noise Calibration

In the fine stage, the FHSS parameters (center frequency and bandwidth) for each uplink hop are precisely estimated using the dwell time and frequency information obtained from the coarse stage.

First, let $y_d[n]$ denote the time-domain sample interval corresponding to the d-th dwell interval $[t_s^{(d)}, t_e^{(d)}]$. The PSD for this interval is calculated using Welch's method. Given the sampling frequency f_s and the FFT size N_{fine} used in the fine stage, the fine-resolution PSD is obtained as follows:

$$\hat{P}_d(f_\ell), \quad f_\ell = \frac{\ell f_s}{N_{\mathrm{fine}}} \quad (\ell = 0, 1, \ldots, N_{\mathrm{fine}} - 1). \tag{4}$$

Here, N_{fine} is set to be sufficiently larger than N_{STFT} used in the coarse stage STFT. This reduces the frequency axis spacing $\Delta f_{\mathrm{fine}} = f_s / N_{\mathrm{fine}}$, thereby decreasing the quantization error in the estimation of the hop center frequency and bandwidth.

Next, using the uplink frequency range $[f_L^{(c,d)}, f_H^{(c,d)}]$ obtained in the coarse stage, the ROI on the fine frequency axis is defined as:

$$\mathcal{F}_{\mathrm{ROI}}^{(d)} = \{f_\ell \mid f_L^{(c,d)} \leq f_\ell \leq f_H^{(c,d)}\}. \tag{5}$$

This set represents the frequency region where the actual energy of the corresponding hop is expected to be concentrated. Conversely, the region defined as

$$\mathcal{F}_{\mathrm{noise}}^{(d)} = \{f_\ell \mid f_\ell < f_L^{(c,d)} \text{ or } f_\ell > f_H^{(c,d)}\} \tag{6}$$

is assumed to be a noise-only frequency interval where no signal components exist. Using this, the average noise PSD is estimated as

$$\hat{P}_w^{(d)} = \frac{1}{\left|\mathcal{F}_{\mathrm{noise}}^{(d)}\right|} \sum_{f_\ell \in \mathcal{F}_{\mathrm{noise}}^{(d)}} \hat{P}_d(f_\ell) \tag{7}$$

and the noise-calibrated PSD is defined by subtracting this estimate from the total PSD.

$$\tilde{P}_d(f_\ell) = \max\left(\hat{P}_d(f_\ell) - \hat{P}_w^{(d)}, 0\right) \tag{8}$$

Finally, the FHSS parameter estimation is performed using $\tilde{P}_d(f_\ell)$ for frequencies belonging to $\mathcal{F}_{\mathrm{ROI}}^{(d)}$. For instance, the center frequency $\hat{f}_c^{(d)}$ of the d-th hop can be estimated using a power-weighted average as follows:

$$\hat{f}_c^{(d)} = \frac{\sum_{f_\ell \in \mathcal{F}_{\mathrm{ROI}}^{(d)}} f_\ell \, \tilde{P}_d(f_\ell)}{\sum_{f_\ell \in \mathcal{F}_{\mathrm{ROI}}^{(d)}} \tilde{P}_d(f_\ell)} \tag{9}$$

The bandwidth is defined as the 99% occupied bandwidth based on the energy distribution of the noise-calibrated PSD.

III. Simulations

In this section, we validate the performance of the proposed Coarse-to-Fine parameter estimation technique through simulations. The simulated received signals were generated using the DJI OcuSync 2.0 signal model, with the sampling frequency set to 122.88 MHz. The semantic segmentation network employed in the coarse stage utilizes the DeepLabV3+ architecture with ResNet18 as the backbone. The ResNet18 backbone network, known for its strengths in sophisticated feature extraction, has a total computational cost of 2.3 GFLOPs and 1.1×10^7 parameters. The training dataset consists of 31,000 spectrogram-label pairs. For the testing phase, performance was evaluated using 10,000 independent signals generated for each SNR level. Other simulation parameters, including the hopping frequency range and occupied bandwidth are summarized in Table 1.

TABLE I
Simulation Parameters

Parameters	Values
Sampling frequency	122.88 MHz
Center frequency of uplinks	{-50, -49, ..., 49, 50} MHz
Uplink bandwidth	1.4 MHz
FFT size of STFT	256
FFT size of PSD	8192
SNR	{-10, -9, ..., 19, 20} dB

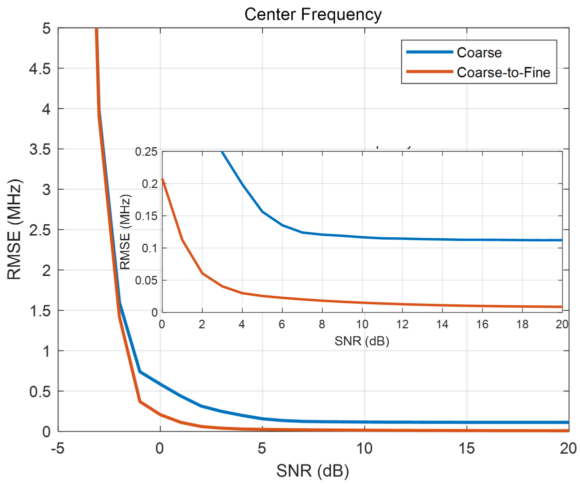

Fig. 2. RMSE comparison of center frequency.

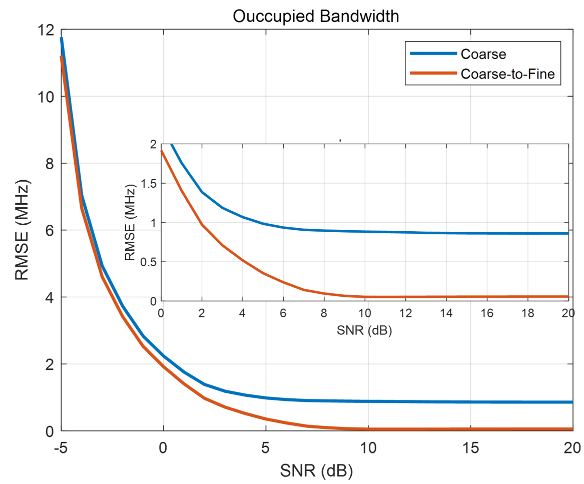

Fig. 3. RMSE comparison of ouccupied bandwidth.

Performance evaluation was conducted by comparing the proposed Coarse-to-Fine method with a baseline Coarse method that utilizes only the results obtained from the coarse stage. The Coarse method estimates the center frequency and occupied bandwidth directly on the grid resolution of the spectrogram by projecting the uplink segmentation map obtained from semantic segmentation onto the frequency axis. In contrast, the proposed Coarse-to-Fine method operates in two stages, as described in Section II. First, it extracts the coarse dwell time and frequency band using the segmentation map and then calculates the PSD for each dwell time using a larger FFT size. Subsequently, an ROI is established based on the frequency information from the coarse stage, and the average noise power is estimated from frequency components outside the ROI to be subtracted from the PSD. Finally, the center frequency and occupied bandwidth of the uplink hop are precisely estimated using this noise-calibrated PSD.

The RMSE for the center frequency and occupied bandwidth is defined for a parameter θ as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\hat{\theta}_i - \theta_i\right)^2} \tag{10}$$

where M represents the number of test signals per SNR.

Fig. 2 illustrates the RMSE comparison results for the uplink hop center frequency with respect to SNR. As observed in this figure, while the RMSE decreases for both methods as SNR increases, the proposed Coarse to Fine method consistently exhibits a lower RMSE compared to the Coarse method across the entire SNR range. In particular, in the high SNR region, the center frequency RMSE of the Coarse method converges to approximately 0.11 MHz, whereas that of the Coarse to Fine method decreases to about 0.01 MHz, confirming that the proposed method provides superior RMSE performance for center frequency estimation.

Fig. 3 shows the RMSE comparison results for the occupied bandwidth. A similar trend is observed: although both methods exhibit high RMSE in the low SNR range, the RMSE of the Coarse to Fine method decreases more rapidly as SNR increases, demonstrating a distinct performance improvement over the Coarse method in the medium to high SNR regions. Specifically, in the high SNR region, the occupied bandwidth RMSE of the Coarse method saturates at approximately 0.86 MHz, whereas the Coarse to Fine method converges to about 0.05 MHz. This indicates that the proposed coarse to fine architecture effectively mitigates the quantization error caused by the spectrogram frequency axis spacing.

IV. Conclusion

In this paper, we proposed a Coarse-to-Fine framework to precisely estimate the center frequency and occupied bandwidth of UAV FHSS uplink signals by extending the semantic segmentation-based drone signal separation technique. The proposed method is designed to extract the uplink dwell time and approximate frequency band from the segmentation map in the coarse stage, while in the fine stage, it calculates a high-resolution PSD for the corresponding interval and estimates hopping parameters using a noise-calibrated PSD, obtained by estimating and subtracting the noise power outside the ROI.

Simulation results based on OcuSync 2.0 demonstrated that the proposed Coarse-to-Fine method exhibited a lower RMSE compared to the conventional Coarse method across the entire SNR range. Notably, in the high SNR region, the RMSE for the center frequency significantly decreased from approximately 0.11 MHz to 0.01 MHz, and for the occupied bandwidth from approximately 0.86 MHz to 0.05 MHz. These results confirm that the proposed technique effectively mitigates the limitations of spectrogram frequency resolution and enhances the performance of FHSS hopping parameter estimation. Future work will focus on further improving practicality by validating the method with actual measured UAV signals and extending it to multi-UAV and multi-link environments.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00557790).

References

[1] H. Kang, J. Joung, J. Kim, J. Kang and Y. S. Cho, ”Protect Your Sky: A Survey of Counter Unmanned Aerial Vehicle Systems,” IEEE Access, vol. 8, pp. 168671-168710, 2020.

[2] F. L. Chiper et al., “Drone Detection and Defense Systems: Survey and a Taxonomy,” Sensors, vol. 22, no. 4, 1453, 2022.

[3] Y. Li et al., “A Fast Estimation Algorithm for Parameters of Multiple Frequency-Hopping Signals Based on Compressed Spectrum Sensing and Maximum Likelihood,” Electronics, vol. 12, no. 8, 1808, 2023.

[4] N. Schiller et al., “Drone Security and the Mysterious Case of DJI’s DroneID,” in Proc. NDSS, 2023.

[5] W. Zhu et al., “Parameter Estimation Algorithm of Frequency-Hopping Signal Based on Time–Frequency Distribution and Frequency Hopping Code Reconstruction,” Sensors, 2023.

[6] R. Uvaydov et al., “Semantic Spectrum Segmentation with Wideband Signal Stitching,” IEEE INFOCOM Workshops, 2024.

[7] L.-C. Chen et al., “Encoder–Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” ECCV, 2018.

[8] D. Park, S.-Y. Kwon, J. Jeong, I. Sim, S. Yun, J. Seo, and H.-N. Kim, ”Semantic segmentation-based drone communication signal separation method,” J-KICS, vol. 50, no. 11, 2025.