# NTR-Gaussian: Nighttime Dynamic Thermal Reconstruction with 4D Gaussian Splatting Based on Thermodynamics

Supplementary Material

## 1. Data Collection

### 1.1. Collection Devices

We use DJI products to collect the required data. Specifically, we employ the DJI Matrice 300 RTK as the drone platform, equipped with the DJI H20T camera to capture infrared data, as shown in Fig. 1. The DJI Matrice 300 RTK is a professional commercial drone with a maximum flight time of 55 minutes and a horizontal positioning error of only 1 cm and vertical positioning error of 1.5 cm in areas with good signal quality. The H20T is a multi-sensor camera, and its long-wave infrared sensor has a temperature measurement range of -40°C to 150°C, which is sufficient to cover most scenarios in typical environments. The camera parameters of the H20T are shown in Tab. 1.

Figure 1. UVA platform and Cameras. DJI Matrice 300 RTK(Left) and DJI H20T(Right).

| Parameter | H20T |
|---|---|
| Sensor | Uncooled VOx Microbolometer |
| Image Resolution | $640 \times 512$ |
| Pixel Pitch | 12 $\mu m$ |
| Spectral Band | 8 $\mu m$ to 14 $\mu m$ |
| Focal Length | 13.5 mm |
| Temperature Range | -40°C to 150°C |

Table 1. H20T Camera Parameters

### 1.2. Data Acquisition

The NTR dataset covers four regions, as shown in Fig. 2. To ensure the drone captures the entire area, we planned the flight path as illustrated, keeping the drone at a constant altitude of 250 meters above the ground and recording the capture time during the flight. Since we aimed to collect nighttime data, the data collection started at sunset and continued until midnight. By utilizing the repeatability feature of the M300 RTK drone, we were able to fly the same flight path multiple times and capture thermal radiation data from the same location at different times, with intervals of approximately 2 to 3 hours. In addition, factors such as GPS signal deviation and wind speed may cause slight positional offsets during the flight.

### 1.3. Data Processing

The data processing consists of two main parts: the generation of the initial point clouds and the processing of thermal infrared images.

First, for the generation of the initial point cloud, we use a 3D textured model with absolute coordinates, created from high-resolution RGB images using photogrammetry software. This model is generated by sampling points from the mesh and performing random down-sampling to create a sparser initial point cloud, with approximately 300,000 points per scene. The reason for generating a sparse initial point cloud is that dense point clouds could lead to memory overflow issues during training.

Second, for the processing of thermal infrared images, the thermal infrared images captured by the H20T are in R-JPG format, containing basic information such as radiation intensity, capture time, and GPS location. However, the internal processing of the H20T applies various image enhancements, such as contrast stretching, to make the images more suitable for human visual perception. To ensure that the constructed 3D thermal radiation field accurately reflects temperature, NTR also provides the raw images recording radiation temperature. Using DJI's Thermal Imaging SDK (TSDK), we converted the R-JPG thermal infrared images into these raw images.

Furthermore, for accurate 3D reconstruction of the thermal infrared images, it is crucial to precisely estimate their absolute orientation. We adopt a rendering-based and matching-based pose estimation method , where the thermal image is registered to a pre-constructed RGB reference 3D model. The overall method includes three stages: (1) Rendering reference images and depth maps from the pre-constructed RGB 3D model using drone sensor priors. (2) Performing cross-modal feature matching between the rendered RGB reference image and the thermal image, establishing a 2D-3D correspondence. (3) Estimating the pose of the thermal image using the Perspective-n-Point (PnP) algorithm . Using the new pose, we render a synthetic RGB image from the high-resolution RGB model with absolute coordinates, which closely matches the thermal infrared image. These images are then used for feature generation with Feature-GS.


Figure 2. Area and route planning.

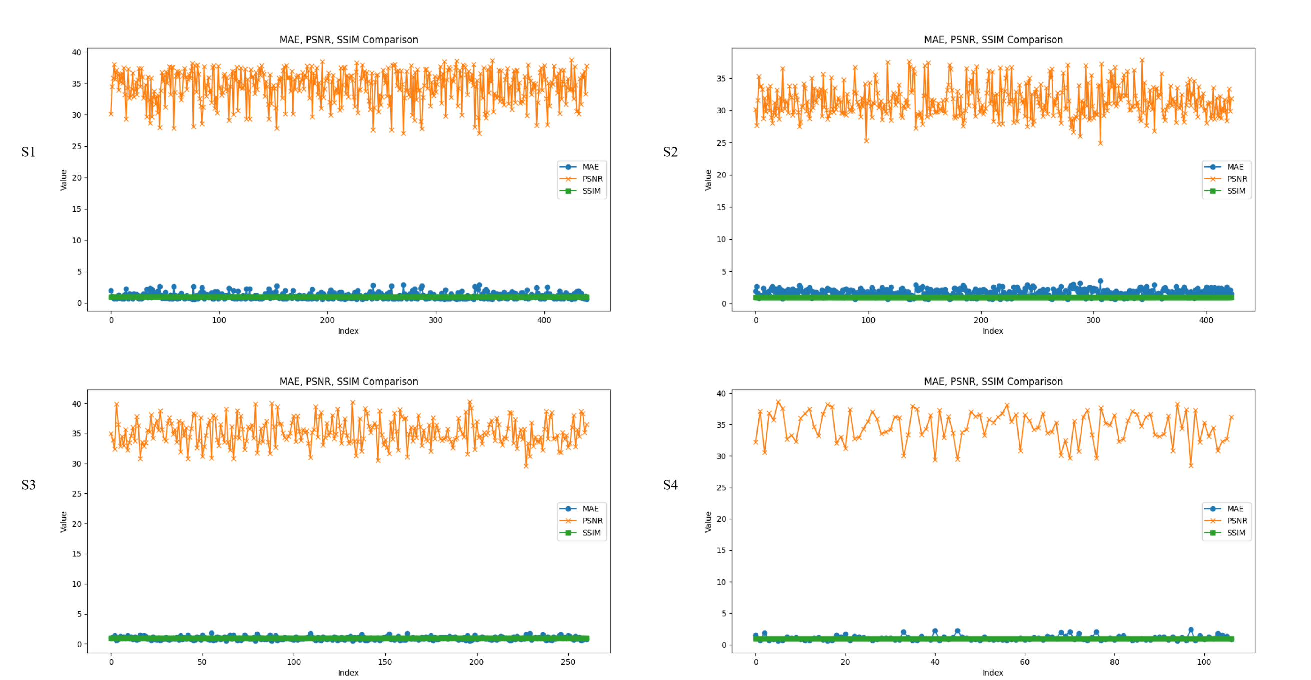


Figure 3. **Quantitative presentation of various scenes.** We presented the PSNR, SSIM and MAE of temperature between the predicted values and the ground truth values for each view in each scenario. Here, the abscissa represents the index of each view. It can be observed that our metrics for each view are relatively good.

## 2. More Experiment Results

Our intermediate process will generate emissivity, convective heat transfer coefficient and heat capacity. Ours is different from the artificial settings. We obtain them through the network and then get the scene radiative temperature through numerical analysis.

We take PSNR , SSIM and temperature MAE (Mean Absolute Error) as the evaluation metrics, and calculate these metrics under all the test views in areas S1, S2, S3 and S4 to illustrate the stability of our method. As shown in Fig. 3, we present the metrics of all views in all regions in the chart. Taking the Mean Absolute Error (MAE) of temperature as a representative, we calculated the metrics of all methods in the test views of the four regions respectively to prove the superiority of our method. As shown in Fig. 4, the MAE of our method is at the lowest level.

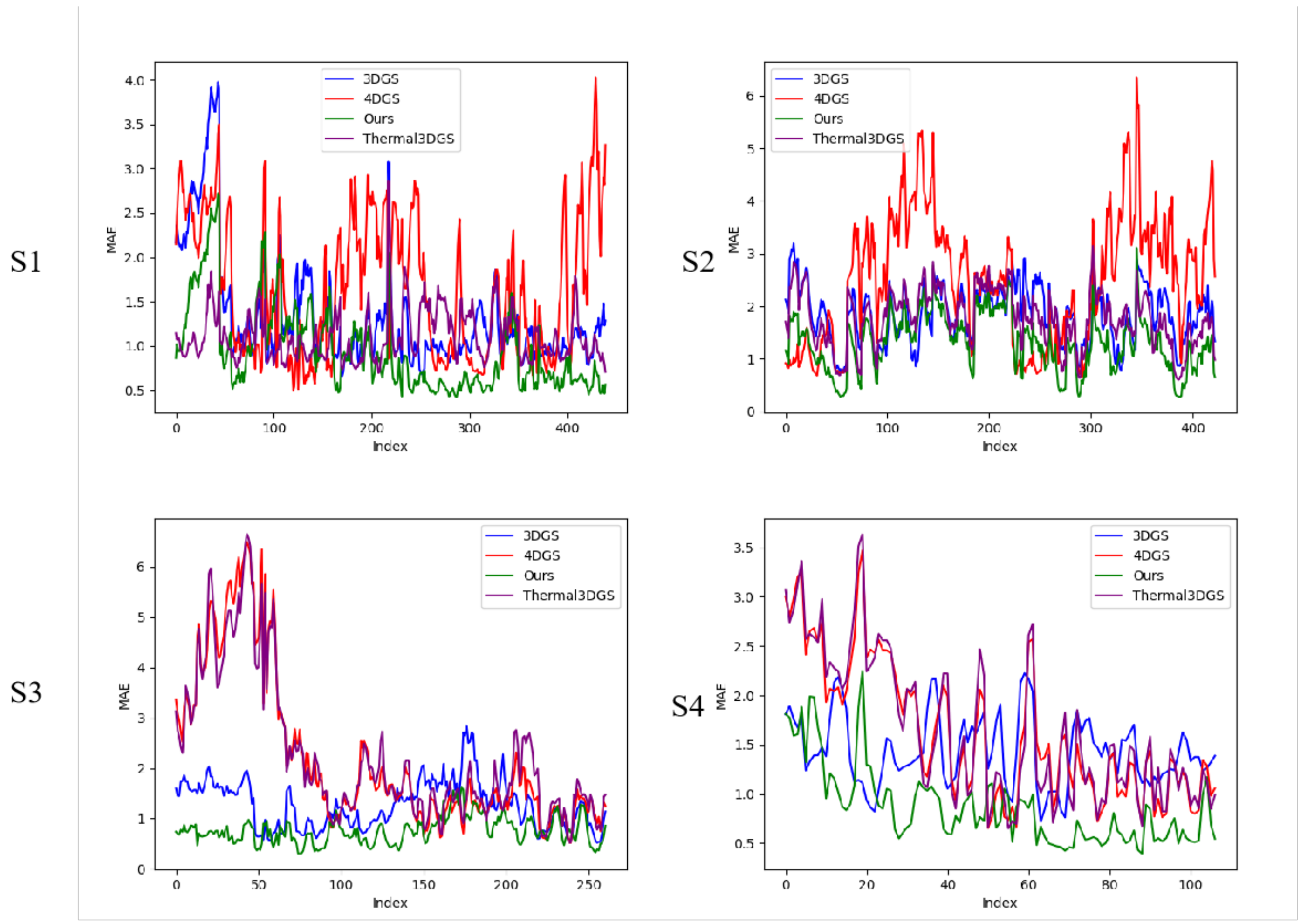


Figure 4. **Quantitative comparison of various scenes.** We present the MAE of temperature between the predicted values and the ground truth values for each view in each scenario. Here, the abscissa represents the index of each view. It can be observed that our method has the lowest MAE among all the methods and outperforms the other methods.