Computer Vision Dataset for Aircraft Taxi Operations

Ryan Horn[1] and Nelson Brown[2]
NASA Armstrong Flight Research Center, Edwards, California, 93523, USA

The development and democratization of computer vision algorithms are contingent on the availability of high-quality datasets. In this paper, we introduce a database of forward-facing videos from taxiing aircraft as well as the time-correlated flight data at approximately 1 to10 Hz containing aircraft state data and environmental conditions. The video data is sourced from the National Aeronautics and Space Administration Airborne Science Program archive and includes over 33 hours of 4k, 1080p, and 720p video from twenty-two airports around the world. This paper describes the method of the database construction and a brief analysis of its contents.

I. Introduction

Computer vision as a profession has witnessed significant development in past decades, in part, due to datasets such as Caltech 101 [1], The MNIST Database of Handwritten Digits [2], and ImageNet [3]. These datasets provide researchers with a benchmark of annotated images, thereby enabling the development, comparison, and advancement of computer vision algorithms. While these datasets are primarily tailored for image classification tasks, this paper presents a computer vision dataset specifically for aircraft ground movement. This dataset is aimed at advancing research in computer vision-based navigation, with potential applications in autonomous aviation systems. Furthermore, domain-specific datasets such as the Lung Image Database Consortium [4] (medical imaging) and Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) [5] (autonomous driving for cars) have demonstrated the importance of datasets tailored to a specific field. Notably, the presence of a large-scale database specifically targeted towards aircraft ground movement is lacking in the public domain. This database gap is particularly relevant for the development of autonomous aircraft, where reliable data for simulating taxiing scenarios is crucial for ensuring safe and robust operations. Much like KITTI played a pivotal role in the development of autonomous driving algorithms for cars, the goal of creating a similar dataset for autonomous aircraft is to facilitate the design, testing, and validation of algorithms that will ensure efficient autonomous aircraft ground operations. To this end, we turn to the NASA Airborne Science Program (ASP).

The ASP focuses on conducting scientific research using aircraft as platforms for data collection. Operating within the NASA Earth Science Division, the ASP primarily focuses on aiding the development, calibration, and validation of aerospace instruments and sensors. To facilitate its mission, the ASP flies a variety of aircraft. This dataset includes data from a McDonnell Douglas (formerly, McDonnell Douglas Corporation, St. Louis, Missouri) DC-8-70 aircraft and a Lockheed (Lockheed Martin Corporation, Bethesda, Maryland) P-3 Orion aircraft in a diverse set of locations. During these missions, flight data and video are often recorded and later uploaded to the ASP archive [6]. The flight data is recorded in IWG1 format [7] and includes the following parameters: latitude, longitude, attitude, ground speed (GS), true airspeed, indicated airspeed, GPS altitude, radar altitude (RA), and pressure altitude. Latitude, longitude, and GPS altitude use the WGS84 geodetic system. The archive contains files from eleven different aircraft across eighty-five campaigns, including thirty-six forward-facing videos.

To address the lack of a dedicated computer vision database for aircraft taxiing, our objective is to assemble a database containing flight logs and videos (synchronized within one second) which were sourced from flights within

---

[1] Student Intern, National Aeronautics and Space Administration Armstrong Flight Research Center.
[2] Aerospace Engineer, National Aeronautics and Space Administration Armstrong Flight Research Center.

the ASP archive. The contents of the database can be accessed at the following URL: https://techport.nasa.gov/projects/157798.

II. Methods

In this study, terminology is used to differentiate between phases of aircraft movement [8], where the term “ground movement” is a general term that describes all aircraft movement on the ground. The term "taxi" refers to ground movement, such as from a gate to the runway, that does not involve acceleration or deceleration on the runway to begin or end flight operations. "Takeoff roll" refers to the acceleration of the aircraft on the runway prior to takeoff. "Landing roll" describes the decelerating movement immediately after landing from the runway to the taxiway or gate. Although the taxi and takeoff roll, or landing roll and taxi, often occur in close succession, the algorithm described below typically identifies them as separate ground movements. This separation is due to the significant periods of time the aircraft remains stationary between the taxi and roll events. The Beautiful Soup Python library [9] was used to download all flight logs and forward-facing videos on the ASP archive. Metadata such as the campaign name, campaign year, and aircraft tail number were also recorded and stored with its respective flight log or video.

The aircraft ground speed and radar altitude were used to determine whether it was engaged in ground movement. For each row in the flight log file a Boolean “is_ground_movement” value was assigned based on aircraft movement at a typical taxi speed and location, which was below an altitude threshold: the criteria values were defined arbitrarily through manual experimentation. The comparisons in Table 1 must be true for the timestamp to be considered a ground movement:

Table 1 The logical comparisons to determine if a timestamp occurs during ground movement.

Flight parameter	Relational operator	Threshold
Ground speed (GS)	>	0 m/s
Ground speed (GS)	<	100 m/s
Radar altitude (RA)	<	25 m

The equation to determine the value of “is_ground_movement” is therefore given in Eq. (1).

$$\text{"is\_ground\_movement"} = \begin{cases} 0, \text{if } 0\frac{\text{m}}{\text{s}} < \text{GS} < 100\frac{\text{m}}{\text{s}} \textbf{ and } \text{RA} < 25\text{ m} \\ 1, \text{otherwise} \end{cases} \quad (1)$$

Due to noise in the flight logs, the “is_ground_movement” parameter often fluctuates between “True” and “False” during ground movement. To address this problem, a one-minute averaging convolution kernel was applied to the “is_ground_movement” parameter, treating True and False as ones and zeros, respectively. The resulting array was rounded and converted back to Boolean values, creating a continuous sequence of True values throughout the entire ground movement event. This blurring effect extends coverage to a brief period just before and after the initial sequence of True values, ensuring that the entire duration of aircraft takeoffs and landings is captured in the database. This process is visualized for the takeoff roll of a mission during the 2023 NASA Student Airborne Research Program (SARP) campaign, as shown in Fig. 1.

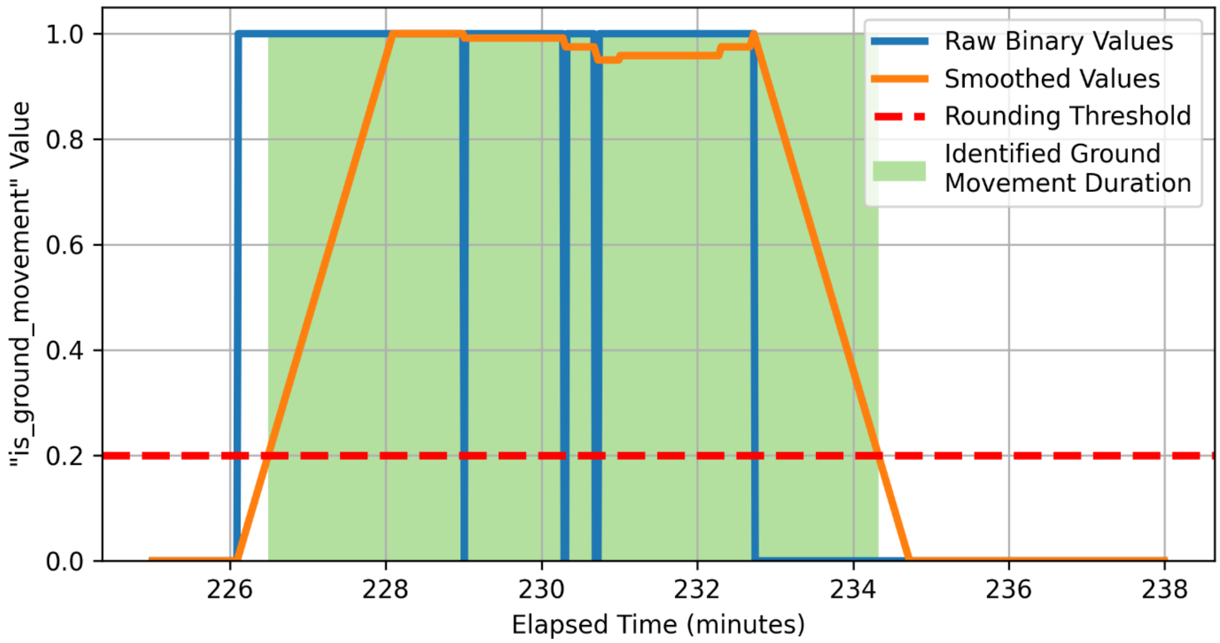

Fig. 1 Ground movement identification visualization.

Figure 1 shows the visualization of how noise in the flight logs is mitigated. The blue line represents the raw binary values, as calculated by Eq. (1). Due to noise in the data, the binary series incorrectly alternates between zero and one, even though the entire segment should be recognized as a single, continuous ground movement event. The orange line shows the result after applying the smoothing kernel. The duration where the orange line exceeds the rounding threshold (red-dashed line) is considered ground movement for the final "is_ground_movement" parameter (highlighted in the green region). This process eliminates the noise, ensuring the entire ground movement event is accurately identified as one uninterrupted segment.

Figure 2 illustrates the radar altitude and ground-speed data for the same 2023 SARP campaign: Fig. 2(a) shows the entire flight duration; Fig. 2(b) provides an expanded view of the taxi phase (from the hangar to the runway) and the takeoff roll; and Fig. 2(c) illustrates the landing roll and taxi phase (from the runway back to the hangar).

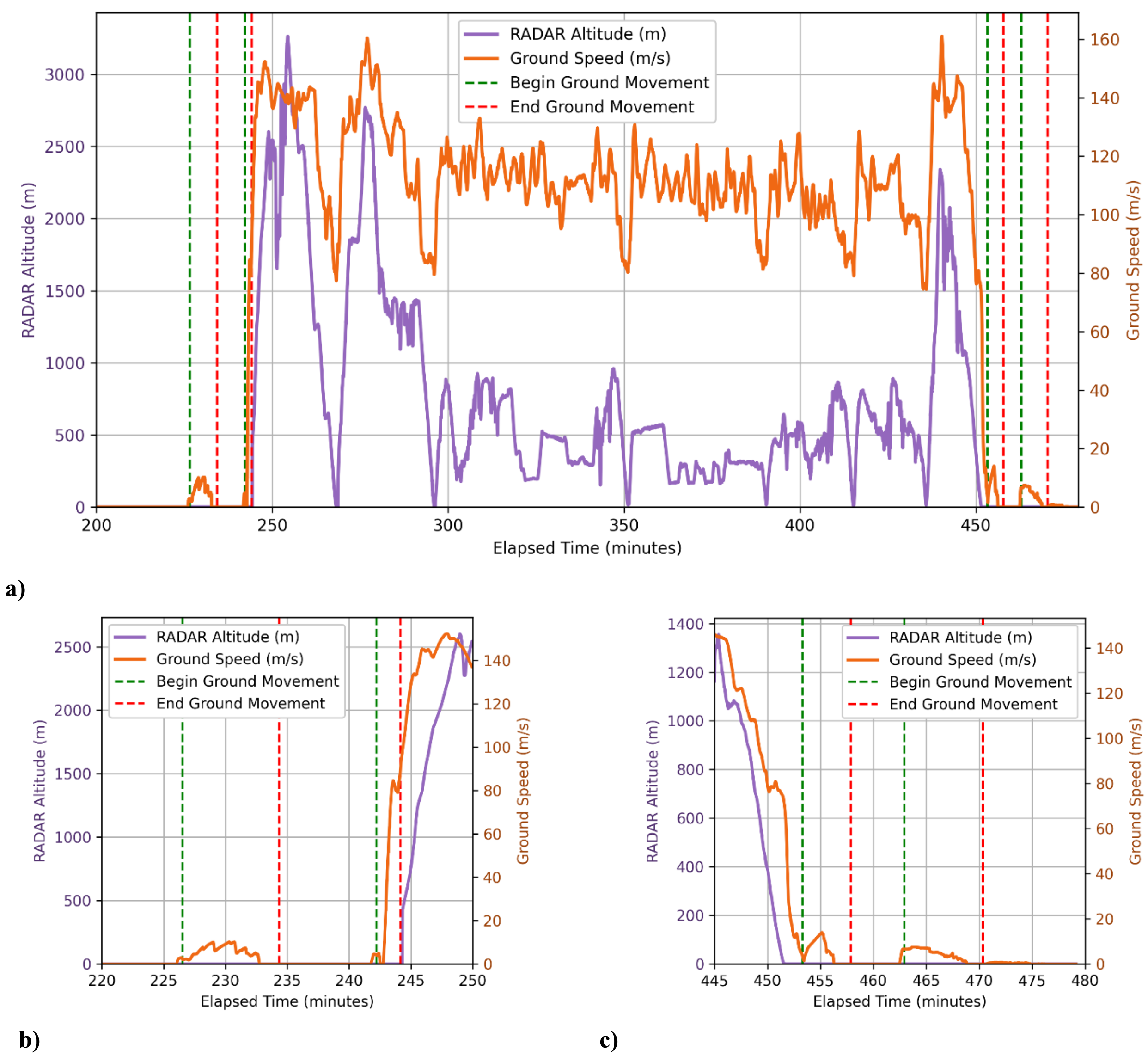

Fig. 2 Radar altitude and ground speed versus elapsed time.

Figure 2 shows the radar altitude and ground speed of an aircraft throughout the duration of its flight. Red and blue vertical lines indicate the start and end of ground movement. Figures 2(b) and 2(c) are shown, magnified, on the start and end of the flight. The four taxi events identified correspond to the following (in chronological order): taxiing to the runway, takeoff roll, landing roll, and taxiing to the ramp.

Once the ground movement events in every available flight log have been identified, the latitude and longitude of the aircraft during the ground movement is compared to a database of International Civil Aviation Organization (ICAO) airport coordinates to determine the specific airport where the aircraft is located.

A metadata spreadsheet is provided with the dataset that contains information about each ground event. This information includes the aircraft tail number, ASP campaign name, ground movement date and time, starting Coordinated Universal Time (UTC) timestamps for each video, and airport ICAO code. Additionally, the dataset includes a manually labeled column indicating the type of ground movement for each taxi event, such as taxi, landing roll, or takeoff roll. Another column contains manual, qualitative labels for the phase of the day, categorizing events as daytime, evening, or nighttime. Finally, a column is provided that includes labels regarding video quality, with notes on conditions such as rain, smudges on the camera, or video compression artifacts.

The video file names include information about the aircraft tail number, campaign, and date. There are thirty-six campaigns with video on the ASP archive and twenty-one of them indicate the video start time (in UTC) as part of the video file names. To temporally correlate flight data from the IWG1 logs with video frames, this dataset consists of ground movement from the twenty-one timestamped campaigns.

By comparing the timestamps and aircraft tail numbers, the ground movement flight logs are associated with their respective video. The ASP archive divides flight videos into fifteen-minute segments. Some taxi events span two or more of these fifteen-minute videos. Due to slightly mismatched starting and ending keyframes, manually concatenating these videos can lead to a desynchronization between the video timestamp and the true UTC time. For this reason, multiple videos for a single-taxi event are stored individually with their own starting UTC timestamp. To determine the UTC timestamp of any given frame in the video, the UTC timestamp must be adjusted by the frame offset time, that is, the time at which the frame occurs within the video. For a fixed-frame rate video, this frame offset time could be calculated by multiplying the frame number by the video frame rate. The videos in the ASP archive, however, were recorded at a variable frame rate; therefore, for the convenience of the user, a separate CSV file is provided for each video, listing the frame offset time for each frame in the video.

Finally, CSV and rosbag2 files are created to store flight data that has been extracted from the relevant sections of the IWG1 files. Both file formats contain the same flight data, allowing the user to choose the format that best suits their workflow. To minimize file sizes, videos are saved as .mp4 files rather than being stored as individual frames within the rosbag2 files.

III. Results

The data set includes over 33 hours of total video across three-hundred ten taxi events at twenty-six airports. The middle 50 percent of taxi events have a duration between 235 and 509 seconds. Figure 3 shows the distribution of taxi events by airport.

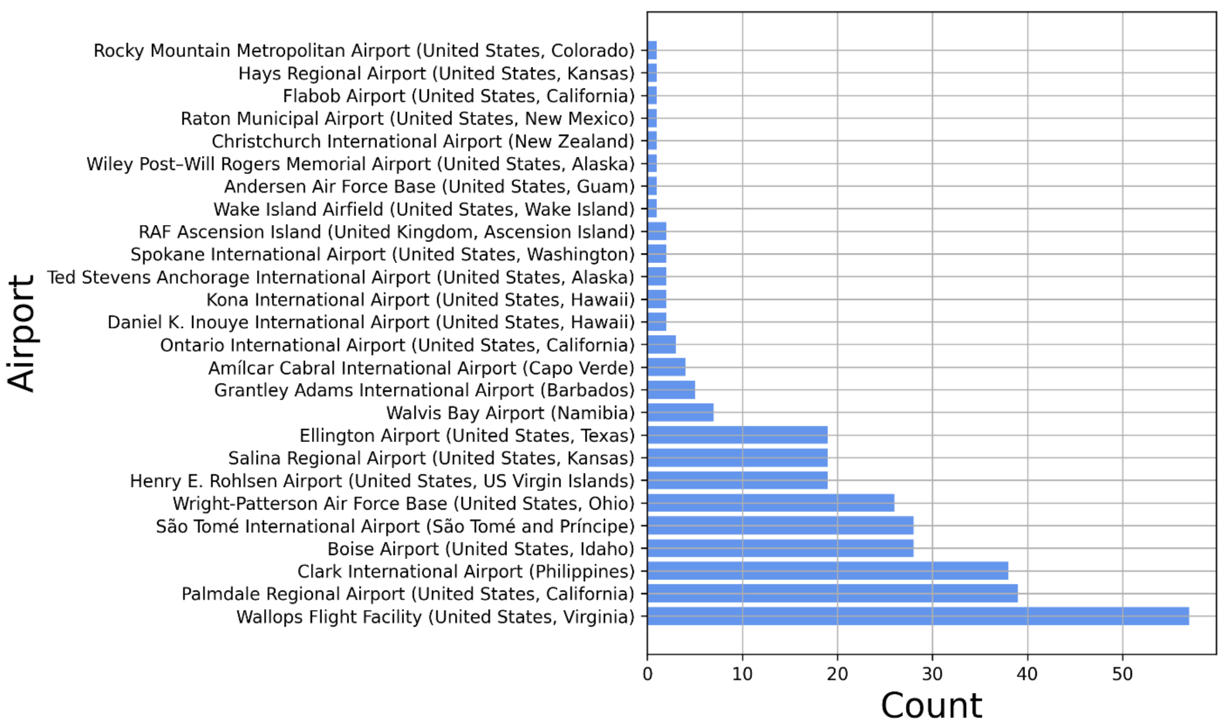

Fig. 3 The number of ground movement events in this dataset at each airport.

Out of three hundred ten total ground movement events recorded in the dataset, two hundred fifty events take place during daytime and evening hours (captured in color video), whereas sixty events happen during nighttime hours (captured in black and white video). Figure 4 illustrates examples of both daytime and nighttime ground movement events.

Daytime	 Palmdale Regional Airport (Palmdale, California, USA)
Evening	 Clark International Airport (Pampanga, Philippines)
Nighttime	 Palmdale Regional Airport (Palmdale, California, USA)

Fig. 4 Example frames of video from taxi events during a variety of lighting conditions.

IV. Conclusion

In this paper, we have introduced a high-quality, temporally synchronized database of flight logs and forward-facing videos of aircraft ground movement events. The dataset includes over 33 hours of 4k, 1080p, and 720p footage, recorded in a variety of geographical locations and lighting conditions. We believe that this dataset will be useful in furthering the development of computer vision tools in the aerospace domain, specifically in areas such as autonomous navigation and guidance.

V. Acknowledgements

The authors express their appreciation to the NASA Airborne Science Program for having the foresight to record and publish the flight data and video from their campaigns.

References

[1] Fei-Fei, L., Marco, A., M'A, Ranzato, Pietro, P., Caltech Data by Caltech Library, "Caltech 101," Published April 6, 2022, Version 1.0, URL: https://data.caltech.edu/records/mzrjq-6wc02 [retrieved 3 December 2024].

[2] LeCun, Y., Cortes, C., J.C. Burges, C, "THE MNIST DATABASE of handwritten digits," URL: https://yann.lecun.com/exdb/mnist/ [retrieved 3 December 2024].

[3] Stanford Vision Lab, Stanford University, Princeton University, "ImageNet," URL: https://image-net.org [retrieved 3 December 2024].

[4] Armato, S. G., *et al.*, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a Completed Reference Database of Lung Nodules on CT scans," *Medical Physics*, vol. 38, no. 2, pp. 915–31, 2011, doi: 10.1118/1.3528204.

[5] Geiger, A., Lenz, P., Stiller, C., and Urtasun, R., "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, Aug. 2013, doi:10.1177/0278364913491297.

[6] "ASP Archive," *NASA Airborne Science Program*, 2004. https://asp-archive.arc.nasa.gov [retrieved 3 December 2024].

[7] "IWG1 ASCII Packet Definition," *Earth Observing Laboratory*, April 22, 2009. URL: https://archive.eol.ucar.edu/raf/Software/iwgadts/IWG1_Def.html [retrieved 3 December 2024].

[8] The Commercial Aviation Safety Team (CAST) and the International Civil Aviation Organization (ICAO) (CAST/ICAO Common Taxonomy Team (CICTT), "Phase of Flight Definitions and Usage Notes," April 2013, (1.3), URL: https://www.ntsb.gov/safety/data/Documents/datafiles/PhaseofFlightDefinitions.pdf [retrieved 3 December 2024].

[9] Richardson, L., "Beautiful soup documentation," URL: https://beautiful-soup-4.readthedocs.io [retrieved 3 December 2024].