Missing Data? How do we know?

In-Situ Image Data Analysis

National Institute of Standards and Technology

Problem Statement

Real-time monitoring is a critical process control technique in additive manufacturing (AM), particularly for powder bed fusion. Modern AM machines are equipped with multiple in-situ sensors that generate massive amounts of data per build, characterized by high speed and large volume. At the National Institute of Standards and Technology (NIST), our testbed's co-axial camera is capable of capturing up to 20,000 frames per second. However, due to data transfer bottlenecks, occasional image loss can occur. Techniques such as image similarity analysis and feature extraction are employed to identify missing images, yet managing this missing data remains a significant challenge.

We have observed this issue when producing numerous parts simultaneously on our testbed. Upon reviewing the collected in-situ data, discrepancies have been found between the number of images recorded by the camera and those actually collected. For instance, while the camera log may indicate 200,000 images were taken, the actual number of images could be short by 1 to 50 frames, indicating a clear loss of data.

The challenge is to develop a method to locate these missing frames within a series of melt pool monitoring (MPM) images. You are free to use any approach to solve this problem. Your solution will be evaluated after seven days, with results subsequently released. Your final score will be based on accuracy, creativity, and clarity of presentation.

The problem is divided into two parts:

1. Part 1: Assume the missing images can be recovered but have lost their spatial and temporal information, representing an ideal scenario.

2. Part 2: Reflects a more realistic scenario where the missing frames are permanently lost.

1. Part 1

In this scenario, you will be provided with the missing images and need to determine the location of each missing frame within the sequence.

- The first 200 frames are complete with no missing images.

- There are no consecutive missing frames.

You are required to guess the position of all missing frames. You may choose to predict the exact location using the specified criteria, known as Criteria 1a. This criterion offers high rewards for accurate predictions but carries a higher risk of losing points for larger errors.

Table 1. Criteria 1

Prediction Error (frames)	**Points**
0	10
[-10, 10]	2
[-25, -10), (10, 25]	1

Or you choose the Criteria 1b, then points will be given based on the following table.

Table 2. Criteria 2

Prediction Error (frames)	**Points**
[-10, 10]	3
[-25, -10), (10, 25]	2
[-50, -25), (25, 50]	1

2. Part 2

This problem simulates a real-world scenario where users are unaware of the specific details regarding missing images, such as their location, timing, and identity. Therefore, the missing images will not be provided. You are required to infer where the images are likely missing based on another experiment with similar image properties. Points will be awarded based on Criteria 2.

Table 3. Criteria 2.

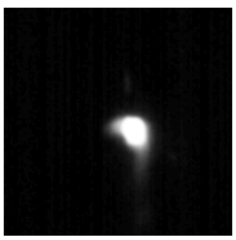
Prediction Error (frames)	**Points**
0	10
[-10, 10]	5
[-25, -10), (10, 25]	3
[-50, -25), (25, 50]	1

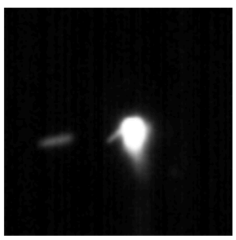
About the Data

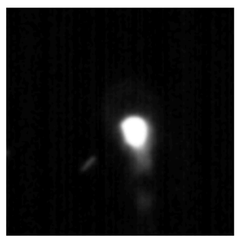
- Sample Dataset: Released two weeks before the hackathon, this dataset includes 400 original frames plus 4 missing frames, with their actual locations provided.
- Part 1 Dataset: Released at the start of the hackathon, this dataset includes approximately 5000 original frames from an experiment and will also provide the missing frames.
- Part 2 Dataset: Also released at the start of the hackathon, this dataset includes the same amount of original frames from another experiment, but the missing frames are permanently lost.

Example

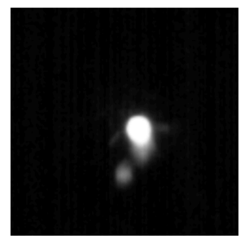
To illustrate, consider a sequence of 5 original frames with one missing frame. If the missing frame is located between Frame 4 and Frame 5, the correct answer is 4. Your task is to estimate the correct location of the missing frames, and your score will be based on the accuracy of these estimations.

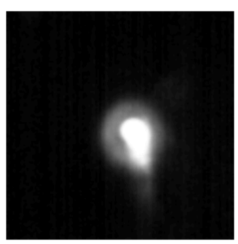

Frame 1


Frame 2


Frame 3


Frame 4


Frame 5


Missing frame 1

This table illustrates the points received using different criteria in Part 1 based on your predicted positions relative to the actual position of the missing frame.

Table 4. Examples of point allocation based on different criteria.

<table>
<tr><th>Example</th><th>Actual Position</th><th>Predicted Position</th><th>Points (Criteria 1a)</th><th>Points (Criteria 1b)</th></tr>
<tr><td>1</td><td rowspan="3">Between Frame 4 and 5 (4)</td><td>4</td><td>10</td><td>3</td></tr>
<tr><td>2</td><td>2</td><td>2</td><td>3</td></tr>
<tr><td>3</td><td>35</td><td>0</td><td>1</td></tr>
</table>

Submission

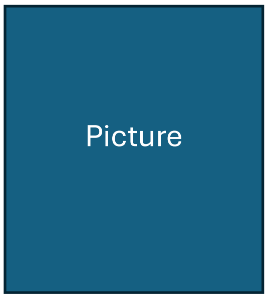
For Part 1, you need to edit the file 'submission_p1_name_c1a.csv' or 'submission_p1_name_c1b.csv' based on the criteria you choose (c1a for Criteria 1a, c1b for Criteria 1b). Replace 'name' with your team name. For Part 2, you need to edit the 'file submission_p2_name.csv', replacing 'name' with your team name. Each file should have two columns: the first column lists all missing frame numbers, and the second column should contain your estimated positions for the missing frames.

Table 5. The Judging Criteria

Category	Criteria	Scoring
Technical Approach (35%)	• Problem formulation • Literature review and exploration of ideas • Development and design of the idea • Scientific soundness of the approach • Creativity of the approach • Readiness of the idea and approach	Excellent (9-10 pts) Very good (7-8 pts) Good (5-6 pts) Limited (3-4 pts) Poor (1-2 pts)
Results (45%)	• Accuracy • 50% - 50%	Excellent (9-10 pts) Very good (7-8 pts) Good (5-6 pts) Limited (3-4 pts) Poor (1-2 pts)
Overall Presentation (20%)	• Title, headings, labels: appropriate size, location, spelling, and content • The demonstration of teamwork • Structure and clarity	Excellent (9-10 pts) Very good (7-8 pts) Good (5-6 pts) Limited (3-4 pts) Poor (1-2 pts))

Domain Experts and Support

Zhuo Yang, NIST



Lisha White, NIST

Yan Lu, NIST