ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-WEPS50

PROGRESS ON THE AUTONOMOUS EVENT DETECTION SYSTEM FOR THE LASER PARTICULATE COUNTER*

H. Zhang†, A. Sy, J. Gubeli, Jefferson Lab, Newport News, VA 23606, USA
S. Zhong, OmniSensing Photonics LLC, Columbia, MD 21046, USA

Abstract

Field emission is one of the most important issues that limits the performance of the superconducting radio frequency (SRF) systems and leads to SRF cavity trips at the Continuous Electron Beam Accelerator Facility at Jefferson Lab. Studies have confirmed that particulates are the dominant source of field emitters and the particulates can be transported into a cavity from other parts of the accelerator. To monitor the transportation of the particulates, a prototype of a novel, non-invasive laser particulate counter (LPC) has been developed and tested. Experiments have been done to validate the capability of the LPC. We are developing autonomous event detection system to continuously monitor the readout from the LPC and to recognize real events generated by particulates from noises using machine learning model. In this report, we will present how the data are prepared and how the model is trained. We will also discuss the performance of the model.

INTRODUCTION

Electron field emission (FE) is a phenomenon that electrons emit from conducting surface under exposure to intense electron field, which exists in both normal conducting RF cavities and SRF cavities while SRF cavities are more susceptible to it [1]. The FE leads to SRF cavity trips and could cause radiation damage to the accelerator [1]. In the several decades operation of CEBAF, the Continuous Electron Beam Accelerator Facility, at Jefferson Lab (JLab), it is known that after running the beam for some time the SRF cavities could suffer from the "gradient loss" – cavities have to run in lower gradient for stable operation because the FE becomes more severe, which increases the trip rate [2]. Since 2014, studies have been carried out to investigate the root cause of the FE in CEBAF SRF cavities [1-4]. It has been confirmed that small foreign particulates are the dominating cause of the FE and evidence has been found that the particulates can be transported into the cavities from other components of CEBAF. Experimental study is needed to further understand this process in CEBAF. After all, FE is not a specific problem for CEBAF, but a key challenge to the whole SRF community [5].

To investigate the transportation of particulates, a laser particulate counter (LPC) is under development a JLab. A prototype has been build and tested using rotating acrylic disks with a defect to mimic the motion of a particulate. It has been confirmed that such a “particulate” can generate identifiable signal and the minimum size of the particulate

* Work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177
† hezhang@jlab.org

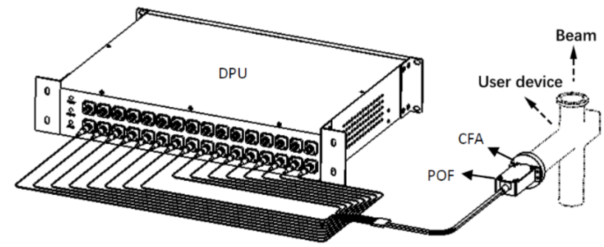

Figure 1: Schematic diagram of the LPC.

continuously generate data from the detection channels but most are noises. We are now testing whether a machine learning model can be developed to distinguish a real signal from the noises, based on which we expect to build an autonomous event detection system that can monitor the readouts from the LPC, detect the signals generated by particulates, and achieve the respective data.

LASER PARTICLE COUNTER

The LPC is a non-invasive particle detector concept built by OmniSensing Photonics, LLC that intends to direct laser beams through an AR-coated window into a CEBAF pump drop chamber, using the internal perpendicular surface of the pump drop as the laser focal plane [6]. The purpose of the LPC is to detect the particulate when it passes the monitored area and measure the moving direction and speed of the particulate. The LPC system includes three key components: a detection processing unit (DPU), a passive optical front-end (POF), and a customized flange adapter (CFA). Figure 1 shows a schematic diagram of the counter system, in which one end of the POF is connected to the DPU and the other end is installed through the CFA to the user device. The principle of particulate detection is based on the optical interference of two coherent laser beams, which are created by splitting a laser beam. A phase modulator is applied to one of the beams. The other beam goes through the monitored area, reflected back by a reflective surface on the opposite side, and is interfered with its phase modulated split in the DPUs detector. The total intensity of the two beams depends on the path length difference between the two beams. When a particulate passes through a laser beam, the intensity and phase of the interference are both disturbed, from which one can detect the existence of the particulate. The optical system contains the laser diode, the optical lens set, an array of laser detectors integrated into a photonic integrated circuit chip and the POF. The CFA works as an installation adapter for POF. The DPU pre-processes the signals from the POF and provides more user-friendly data. In the current prototype of the LPC, two groups of lasers and respective sensors are

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-WEPS50

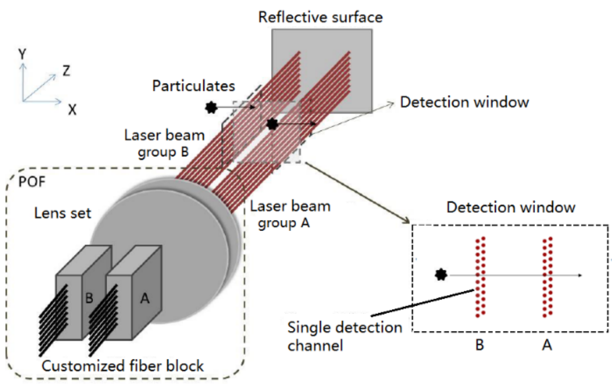

Figure 2: Layout of the laser channels and schematic of 48 laser beams as seen in the laser focal plane. A particulate traversing the path traced by the black arrow intercepts two laser beams.

implemented in the POF, as illustrated in Fig. 2. The 48 laser channels (laser beams and sensors) are arranged in two pairs of arrays, with 12 channels per array. The two arrays in each pair are slightly offset to create overlapping for better coverage of the active area, and the pairs of arrays are separated by approximately 1 cm. Each channel connects to the DPU with an individual fiber. The distance from the optical front end to the reflective surface is adjustable per the request of measurement. Assuming the particulates travel along X direction, the detection window can be defined as the rectangular area in X/Y plane in which the counter can detect the particulates. When a particulate enters the detection window from outside, it interacts first with detection channels in array B and then with those in array A. Signals will be generated from those channels. Knowing the positions of the channels in the X/Y plane and the time difference between the peak signals from array B and array A, the velocity of the particulate in X/Y plane can be calculated.

EXPERIMENTS

To test the capability of the prototype LPC, experiments were carried out from 2022 to early 2023. The experimental test requires a method of deterministically passing a particulate of a known size across a specified point in the laser array. After several iterations, we determined the best method is to use a large rotating disk that has high transmission at the laser wavelength. We mimic particulates by creating defects in the disk that changes the index of refraction. The defects were created by a 350 nm UV laser. The size of the defect was controlled by adjusting the disk position relative to a focusing lens, the laser power, and the duration of the exposure. The experimental setup consisted of an eight-inch diameter acrylic disk attached to a stepper motor shaft, shown in Fig. 3. The mounting allowed for translation transverse to the laser array. The rotational velocity of the disk could also be controlled. In the experiments, 24 channels of lasers in the LPC were used, and output from all the channels was achieved for future analysis.

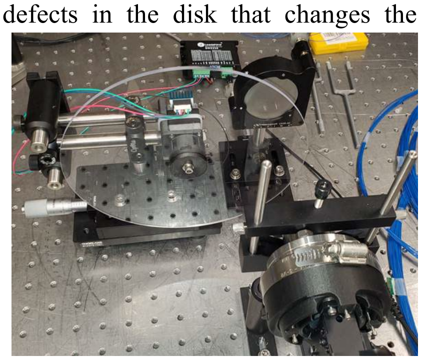
Figure 3: Use defects on a rotating disk to approximate moving particulates.

AUTONOMOUS EVENT DETECTION

No matter whether a particulate is detected or not, the LPC continuously generates data through the DPU. It would be inefficient, even impossible practically, to store all the data. An autonomous event detection system is proposed to monitor the output from the LPC in real-time, automatically distinguish the signals caused by particulates from the background noises, and respond to the events properly, such as saving the data and triggering a warning to the operators. Machine learning could be an appropriate technique to classify the LPC readout into signals or noises in a moving time window. We are attempting to verify this idea using the data obtained in the test experiments.

For data preparation, we developed a program to process all the data files automatically. Each file saves the raw data from an individual channel over a period of time. Using the code provided by the manufacturer, we can convert the raw data into a more user-friendly format. An example of the raw data in the user-friendly format is shown in Fig. 4. We can clearly see periodicity from the data, which is believed to be caused by the circular motion of the disk. After smoothing with Gaussian filters and normalizing the data, together with other numerical treatment, eventually we present the data as the green lines in Fig. 5. In the figure, we can see a group of taller peaks and another group of shorter peaks overlapped on the red lines. The taller peaks are generated by a straight line located 180 degrees from the defect on the circulating disk. This line cuts through all the laser beams and the taller peaks are referred to as "ticks". Between two consecutive ticks, a shorter peak is generated by the defect. These shorter peaks are referred to as "signals". We see ticks from all the channels but see signals from only those channels that are disturbed by the defect. For each data file, which saves the readout from an individual channel over a short period of time, to check whether this channel is disturbed by the defect, we first locate all the ticks and therefore calculate the potential signal

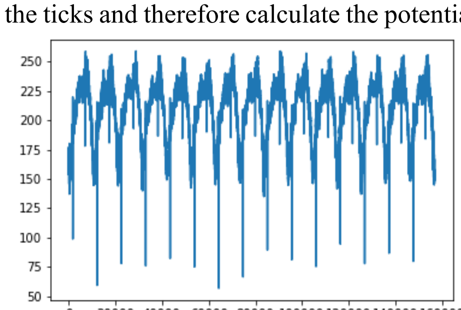

Figure 4: Raw data in the user-friendly format (a sample).

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-WEPS50

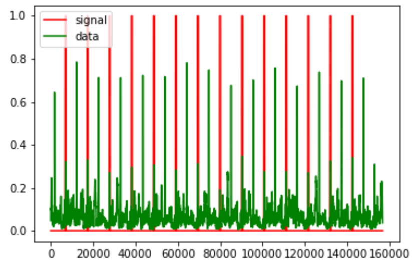

Figure 5: A sample of processed data.

positions in the middle of two consecutive ticks. Finally, the following algorithm is used to check if there are valid signals.

(1) Assume D_i is the 1000 data points centered at the i-th tick; D_{ci} is the 200 data points centered at the i-th tick; S_{ci} is the 200 data points centered at the i-th signal position; and m is the mean of all the data in this file.

(2) Calculate $D_{pi} = \max D_{ci}$, $S_{pi} = \max S_{ci}$, $\overline{D_i} = \overline{D_i - D_{ci}}$;

(3) Calculate $\overline{D_p} = \overline{D_p - \max D_p - \min D_p}$, $\overline{\overline{D}} = \overline{\overline{D} - \max \overline{D} - \min \overline{D}}$, $\overline{S_p} = \overline{D_p - \max S_p - \min S_p}$;

(4) Signals are validated if all three conditions are met: $\left|\overline{D_p} - \overline{\overline{D}}\right| > 0.2$, $|\max \overline{S_p} - \min \overline{S_p}\,| < 0.4$, $m < 0.25$.

We want to note here that the signals in Fig. 5 have significantly larger amplitude than the background noises but this is not always the case. Many valid signals that satisfied the aforementioned criteria have similar or even smaller amplitudes than the background noises and are hard to distinguish in a visual check. We still treat them as valid signals when training the machine learning model.

After going through all the data files, we found the system ran more stably in the last few experiments than in the earlier ones. So the data obtained on Dec. 09, Dec. 12, and Dec. 21 in 2022, and on Jan. 12 in 2023 are selected for training the model. After finding all the data files that contain the signals using the program, we manually checked all the plots generated and the data will be discarded in the following two scenarios: (1) Multiple ticks with similar amplitude show up next to each other. It is hard to tell which one is the real tick although the program always picks one, and hence it is difficult to find the exact position of the signal. (2) Signals found by the program have similar or higher amplitude compared to the ticks. This rarely happens. But in such a case it is hard to tell which are the ticks and which are the signals. With all the selected data files, we cut 200 data points centered at each valid signal position as signal samples. We cut five noise samples before each signal sample and five after it. Each noise sample also contains 200 data points. In total, we get 8402 signal samples and 84020 noise samples. In the noise samples, we randomly choose 8402 to use in training. Eighty percent of the signal samples and same ratio of the selected noise samples are randomly mixed together to form the training set, and all the remaining data are used as the testing set. Within the training set, 20 percent is used for validation. In this process, the processed data are used to locate the signals while the raw data are used as training and testing data.

Table 1: Three-Layer Neural Network

Layer	Type	Neuron #	Activation	Param #
input	dense	128	relu	25728
hiden	dense	64	relu	8256
output	dense	1	sigmoid	65

A simple three-layer neural network with 34049 parameters has been trained using TensorFlow [7]. The properties of the model are listed in Table 1. An accuracy of 0.9167 and a precision of 0.9673 can be easily achieved. Using four-layer or five-layer model did not lead to better results. We also tried to extend the noise sample range from ± 5 pieces around the signal to ± 15 pieces and still selecting the same amount of the noise samples randomly for training. Did not see significantly changes on the accuracy and the precision. To remove the periodicity in the noise samples, we shuffled the data points in each noise sample before putting them into the training. In this case we observed a slight improvement in the precision with an accuracy of 0.9184 and a precision of 0.9822. This result is far from perfect but still encouraging. The following approaches can be applied to build a better model. First, improve the data quality. The circulating disk inevitably disturbs the lasers, which results in the periodicity in the data. In many cases, the amplitude of the signals is similar to or smaller than the background noises, which makes the signals visually indistinguishable. This may prevent the model from achieving higher accuracy. If we could conduct the experiments in a more realistic condition, such as replacing the disk with real particulates inside a vacuum chamber, much better data can be expected. Second, we did not consider the correlation between different channels. As shown in Fig. 2, a particulate tends to disturb several channels close to each other rather than just a single channel. If we include this information in the model, *e.g.* using a decision tree, the model may work better. Third, we can consider developing different machine learning model for different particulate size and combining them together. Fourth, more sophisticated model can be implemented.

SUMMARY

A prototype of the LPC has been developed at JLab to study the transportation of particulates inside beam lines. Experiments have been carried out to test the capability of the LPC. Using the experimental data, we are testing the idea of an autonomous event detection system based on machine learning. Although the result is not perfect, it still suggests that machine learning is a feasible tool for the system. It is reasonable to believe that a better machine learning model can be developed when better quality data is available and more intricate approaches are implemented.

ACKNOWLEDGMENT

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177.

ISBN: 978-3-95450-247-9 ISSN: 2673-5490 doi: 10.18429/JACoW-IPAC2024-WEPS50

REFERENCES

[1] R. L. Geng, A. Freyberger, R. A. Legg, R. Suleiman, and A. S. Fisher, "Field Emission in Superconducting Accelerators: Instrumented Measurements for Its Understanding and Mitigation", in *Proc. IBIC'17*, Grand Rapids, MI, USA, Aug. 2017, pp. 470-477. doi:10.18429/JACoW-IBIC2017-TH1AB1

[2] R. L. Geng, J. F. Fischer, E. A. McEwen, and O. Trofimova, "Nature and Implication of Found Actual Particulates on the Inner Surface of Cavities in a Full-Scale Cryomodule Previously Operated With Beams", in *Proc. SRF'15*, Whistler, Canada, Sep. 2015, paper MOPB035, pp. 164-168.

[3] R. Geng, "Root causes of field emitters in SRF cavities placed in CEBAF tunnel", in *Proc. 7th Int. Particle Accelerator Conf. (IPAC'16)*, Busan, Korea, Canada, May 2016, pp. 3198–3201. doi:10.18429/JACoW-IPAC2016- THOBB03

[4] R. L. Geng, "Root Causes of Field Emitters in SRF Cavities Placed in CEBAF Tunnel", in *Proc. IPAC'16*, Busan, Korea, May 2016, pp. 3198-3201. doi:10.18429/JACoW-IPAC2016-THOBB03

[5] R.L. Geng *et al.*, "Field Emission Suppression in High-Gradient SRF Cavity Systems", Snowmass 2021 LOI, 2021.

[6] J. Zhu and S. Zhong, "Vacuum particle remote detection system specification", vo1.4, OmniSensing Photonics, LLC, Tech. Rep., Aug. 2019.

[7] TensorFlow Developers, TensorFlow. Zenodo, 2024. doi:10.5281/ZENODO.4724125