Journal of Electrical Engineering and Automation (ISSN: 2582-3051)
www.irojournals.com/iroeea/

Real-Time Outlier Detection and Removal in Data Collection with LabVIEW using Peirce's Criteria

Vishnu V[1], Manoj S G[2], C M Maheshan[3], H Prasanna Kumar[4]

[1]PG Student, Department of EEE, UVCE, Bangalore University, Bengaluru, India 560001

[2]Engineer, RTDG Group, Ace Designers Limited, Peenya, Bengaluru, India 560058

[3]Associate Professor, Department of EEE, UVCE, Bangalore University, Bengaluru, India 560001

[4]Professor, Department of EEE, UVCE, Bangalore University, Bengaluru, India 560001

Email: [1]vishnurv.srv@gmail.com, [2]manojgowda76@gmail.com, [3]mcm_eeuvce@bub.ernet.in, [4]uvcehpk@gmail.com

Abstract

Testing and experimental processes in industries and research institutes play important roles in understanding systems and developing accurate models. Experimental uncertainties in variables introduce deviations (errors) in measured data, which can stem from various factors, one of which is the presence of outliers. Outliers impact the accuracy of measurements significantly during analysis and lead to inaccurate conclusions. Outliers can be due to environmental changes, drift in measurements, etc., during assessment. Identifying outliers and eliminating them is a crucial step during data analysis. This research aims to investigate real-time outlier detection and removal using LabVIEW. Peirce Criteria are assessed for detecting and eliminating outliers from displacement data obtained through LVDT sensors. The research demonstrates that Peirce's criteria are particularly well suited for small datasets. By employing LabVIEW and Peirce criteria, this study presents a practical approach to accurately detect and remove outliers in real time to enhance the reliability of data analysis.

Keywords: Data collection, Outliers CNC, LabVIEW, LVDT

DOI: https://doi.org/10.36548/jeea.2023.4.005
Received: 08.11.2023, received in revised form: 04.12.2023, accepted: 21.12.2023, published: 02.01.2024

1. Introduction

In industrial sectors, accurate and precise measurements [10] are important for ensuring optimal machine performance. Turning CNC machines are employed for a wide range of manufacturing processes. These machines experience thermal growth during operation. Thermal growth is a change in materials (expansion and contraction of materials) due to a change in temperature. To understand this thermal growth, experiments are conducted using LVDTs (linear variable differential transformers). The LVDT transducer has high sensitivity and ability to measure displacement (linear) with high precision [12]. In the context of Turning machines, LVDTs are used to track minute variations in the machine's dimensions (expansion and contraction) caused by changes in temperature [11]. Data collection from these experiments also has challenges; multiple readings are obtained after each operation cycle to accurately capture effect caused by thermal growth. One challenge faced in these experiments is the occurrence of outliers. Outliers are points significantly away from the rest of the measurements, which compromises the overall reliability and accuracy of the results. [2].

Outliers in LVDT measurements can originate from voltage fluctuations during the experimental process; magnetic interference (possibly from nearby equipment) can distort measurement readings; the loading effect (effect on the source by load impedance) may impact LVDT accuracy. Drift error and other reasons can contribute to outlier occurrences [4]. To address this problem, Peirce's criteria are utilized, and a custom LabVIEW code has been developed to process the data acquired from LVDT and implemented an outlier detection approach in real time. LabVIEW environment is a powerful and flexible platform, allowing efficient data analysis and quick response to presence of outliers.

A structured approach was followed in this research to understand and improve data analysis. Section 2 explains the principles behind LVDT. Section 3 highlights the importance of DAQ (data acquisition). Section 4 explains about the LabVIEW software. Section 6 explains about the Peirce's Criteria, rule to detect outliers. Section 6 explains about the experimental setup, outlining how the test was conducted and how data was collected using sensors and DAQ, and Section 7 focuses on analysis and results.

2. LVDT

LVDT is a transducer (converts one form of energy into another form) used for linear position or linear displacement measurement [5]. LVDT consists of primary and secondary coils in a cylindrical core. When primary coil is supplied with alternating current, a magnetic field is generated around the core [6].

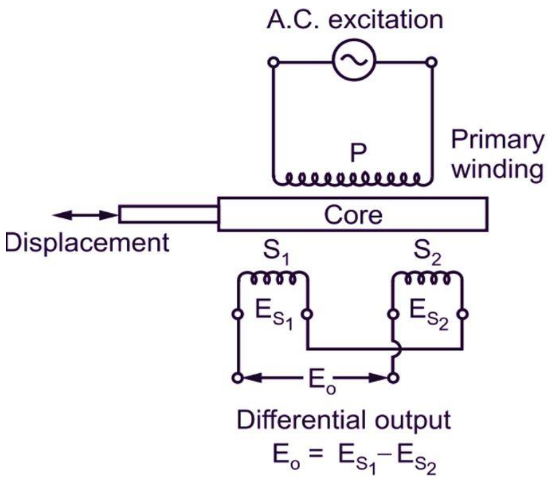

Figure 1. Layout of LVDT [13]

The voltage induced is equal and opposite across two secondary coils when the core is centered within the coil assembly, resulting in a net output voltage of zero. When code is linearly displaced, differential voltage is developed across the secondary coil, and the output voltage is proportional to the core's displacement. The phase relationship between the two secondary voltages is determined by the direction of displacement [9]. The Layout of LVDT is shown in Figure. 1. Displacement can be calculated using the following formulae:

Displacement = $(V_{out}*L)/V_{max}$

V_{max} = Maximum voltage output of LVDT

L = Stroke length of LVDT

V_{out} = Output voltage of LVDT

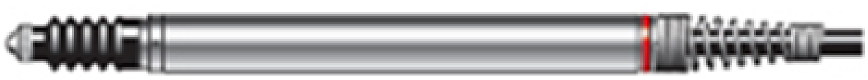

Figure 2. LVDT Pencil Probe [14]

The LVDT shown in Figure. 2 was used in the experiment to measure displacement growth between the tool and workpiece position due to thermal effects in Turning the CNC machine. Specifications of LVDT are available in Table 1.

Table 1. LVDT Specifications

Parameters	Details
Material	Stainless Steel (SS)
Maximum stroke	2mm
Measuring range	+1mm to -1mm
Operating voltage	24VDC
Output analog signal	±10VDC

3. DAQ

Data acquisition (DAQ) is one of the fundamental processes in experimental research. The DAQ system includes sensors and devices that help in collecting data. It converts physical signals (pressure, temperature, etc.) into digital signals (current, voltage, etc.) that can be analyzed and stored on the PC. The DAQ system helps in analyzing real-world phenomena. The DAQ system plays an important role in monitoring and controlling complex systems in industries.

The NI 6003 USB DAQ, shown in Figure 3, is employed in this experiment, with its specifications provided in Table 2 for reference.

Figure 3. NI 6003 USB DAQ [15]

Table 2. NI 6003 USB DAQ Specifications

Parameters		**Details**
Analog Input	Differential	4
	Single-ended	8
Analog Input Range		±10V
ADC resolution		16-bits
Input Impedance		>1G Ω
Analog Outputs		2
Digital I/O		13
Counter		1
USB Bus speed		12Mb/s
Sampling Rate		100kS/s

4. LabVIEW

LabVIEW (Laboratory Virtual Instrument Engineering Workbench) is a graphic-based programming language. LabVIEW contains two screens [8]:

1. Block diagram: In this screen, the actual logic of the program is constructed by dropping essential icons and connecting wires to define the flow of the program [7].

2. Front panel: On this screen, controls and indicators of the program, constructed in a block diagram, are represented, allowing users to interact and take necessary actions.

5. Peirce's Criteria

Peirce's criteria are a powerful technique used to identify outliers in the dataset [3]. Outliers are points that are significantly away from the rest of the measurements. Peirce's criteria utilize Gaussian curve analysis and provide a systematic approach to detecting and eliminating outliers, thereby improving the reliability of statistical analysis [1]. Peirce's criteria help in identifying outliers within smaller datasets.

Steps to apply Peirce's criteria:

1. Calculate the mean (x_m) and standard deviation (σ) of the data set.
2. Obtain the value of R (R is the value used to establish the threshold for identifying outliers) from Peirce's criteria table corresponding to the number of values in the dataset and the number of doubtful observations in the dataset.
3. Calculate the maximum deviation that can be allowed using: $x_{max} = \sigma * R$
4. Eliminate the value if $|\text{ value} - x_{max} | > \sigma * R$

Where R, refers to the maximum allowable deviation of the measured value from the data mean to the standard deviation [1].

6. Experimental Setup

Thermal growth in the CNC is captured by the measuring displacement that is generated between the tool and the workpiece, as well as the measuring temperatures at certain locations (critical heat zones) of CNC Turning machine.

Temperature is measured using RTDs. These RTDs are connected to the ADAM 4015 DAQ module (supports 3 Wire RTD configuration), which is connected to the ADAM 4520 DAQ module which, converts the RS 485 signals to RS 232 signals, allowing easy communication with the PC. For displacement measurement, the LVDTs (Linear Variable Differential Transformers) are used and connected to signal conditioner. The output of signal conditioner is linked to analog input of the NI 6003 USB Device. This DAQ device serves as an interface between the LVDT and the PC and is responsible for capturing, analyzing, and processing data. The hardware connection is represented in Figure. 4. LVDT is set to a specific reference point on the workpiece and called to that point after each operation cycle; this process is shown in Figure. 4.1.

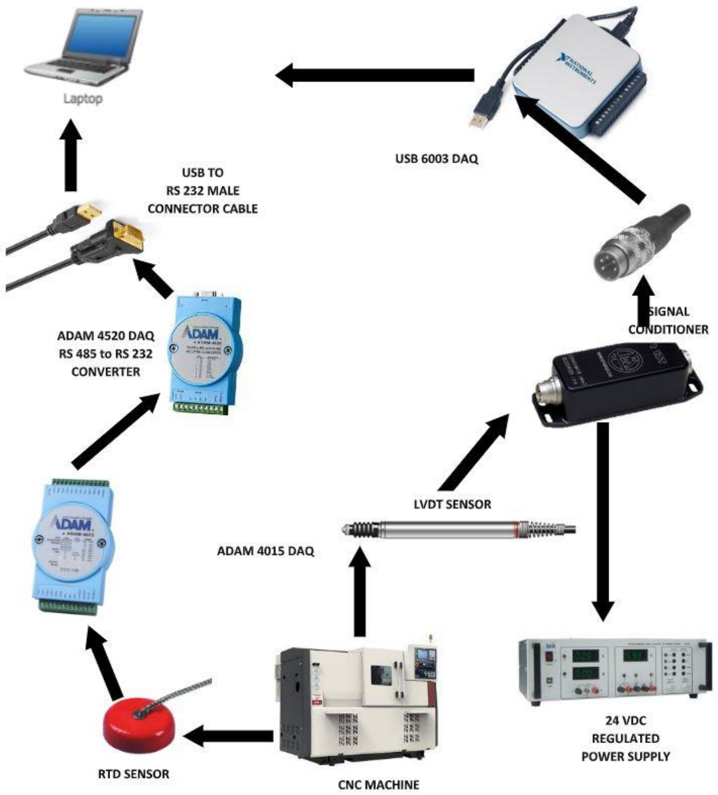

Figure 4. Pictorial Representation of Hardware Connection.

As the machine undergoes multiple cycles, the thermal growth in certain parts of the machine causes changes in the LVDT readings at the reference point. In this research, only displacement data was used for checking the presence of outliers in the data set.

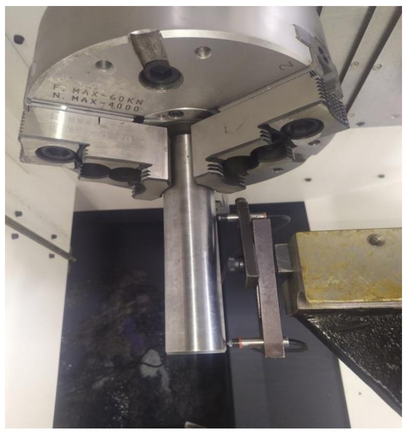

Figure 4.1. LVDTs Set to Reference Point on Workpiece

In LabVIEW, the sensor readings were captured using a single variable, and a time delay was implemented such that the sensor could read 10 samples per second (sample rate =

10 samples per second). The measurements are recorded by the machine after each operation cycle. Immediately after each operation cycle, the machine might account for vibration for a few seconds, so a time delay known as dwell time is given to the machine to reach a stable condition. During dwell time, LVDT readings are not considered. After dwell time is over, a digital trigger is sent to PLC (CNC machine) to DAQ. This trigger indicates that dwell time is over and values can be logged. Block diagram of DAQ system for displacement measurement with a digital trigger is shown in Figure. 4.2.

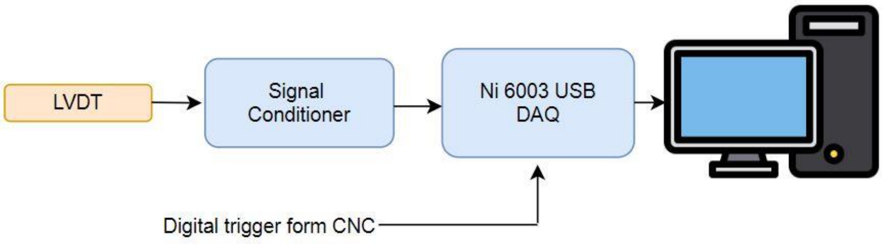

Figure 4.2. DAQ System for Displacement Measurement with Digital Trigger.

When a trigger event occurs, data is logged for a few seconds, and the last 20 points are stored by "*mean vi*" (icon in LabVIEW) in the form of an array, and their average value is logged (old method). This data set in the array may contain potential outliers, which can lead to inaccurate results.

In the proposed method, real-time outlier detection and removal are done using two steps

Step1: Storing Sensor Data into an Array

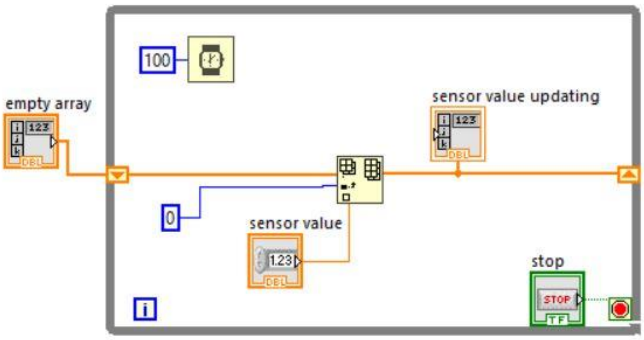

Figure 5. LabVIEW Code for Sensor Data Acquisition

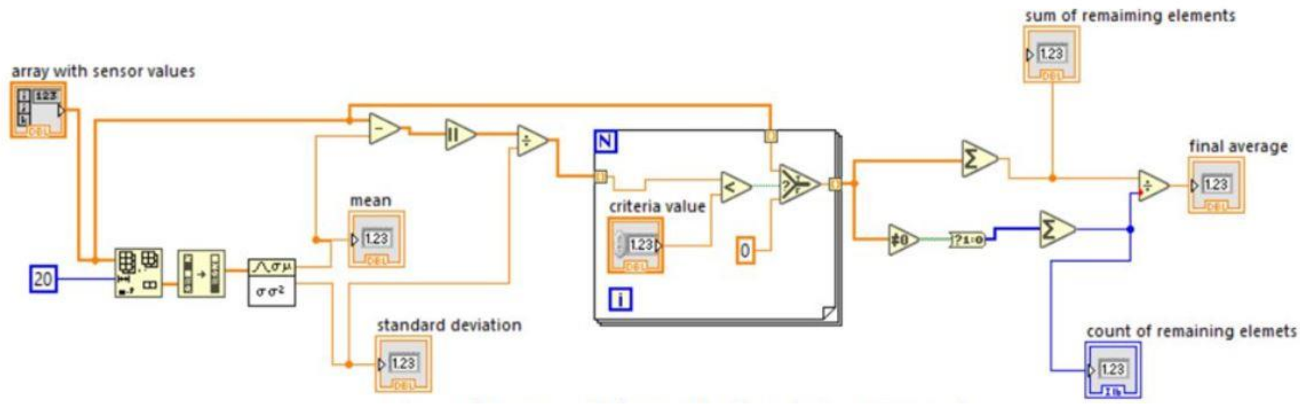

Figure 6. LabView Code for Outlier Detection and Removal

The purpose of this LabVIEW code is to store continuous sensor data into an array so further statistical analysis can be done to remove outliers. The following LabVIEW code is shown in Figure 5. An empty array is used to create and hold data. Sensor data reading code is enclosed within a while loop; this allows repeated iterations to ensure new values are obtained.

Step 2: Data Segmentation and Outlier Removal Code

The purpose of this LabVIEW code is to isolate responsible readings, remove outliers from the dataset, and process the remaining data. The following LabVIEW code is shown in Figure. 6. The code is written in a case structure that is connected to a digital trigger (digital trigger from the PLC is given to the digital input pin in DAQ) upon completion of the cycle, and when the digital trigger is activated, the case structure is evaluated and the true and subsequent code is executed. An array containing sensor values is connected to *"Delete from Array vi"* with a sample size of 20; this makes sure only 20 values enter into this array.

The values present in the array are passed through "*Standard deviation and Variance vi"* for calculation of mean and standard deviation. To find the deviation, each value in the array is subtracted from mean values. To ensure all the deviations are positive the values are passed through "*Absolute vi*". Following Peirce's criteria, deviation to standard deviation ratio is obtained by dividing each deviation by the standard deviation. While employed to compare each ratio value with a predefined criteria value. If the ratio value is below the criteria value, its corresponding measurement is included in a new array, If any ratio value is about the criteria value, its corresponding measurement is considered an outlier and replaced with zero in the output. In the final stage, the aggregate (sum) of all values in the new array is calculated. The count of all non-zero values in the new array are determined. These values are used to compute the average value by dividing the aggregate by the count.

By following this method, the outlets can be detected and removed from the data set, leading to more accurate results when calculating the average. The process of comparing values with criteria values and removing outlets ensures improved average accuracy. Code is converted into "*Sub vi*," so it takes up less space and can be managed easily. *Sub vi* of the outlier code is shown in Figure. 7.

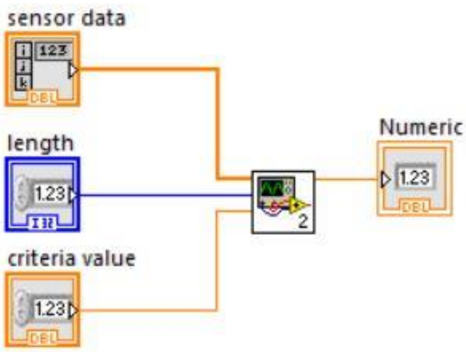

Figure 7. Sub VI Outlier icon

7. Analysis and Results

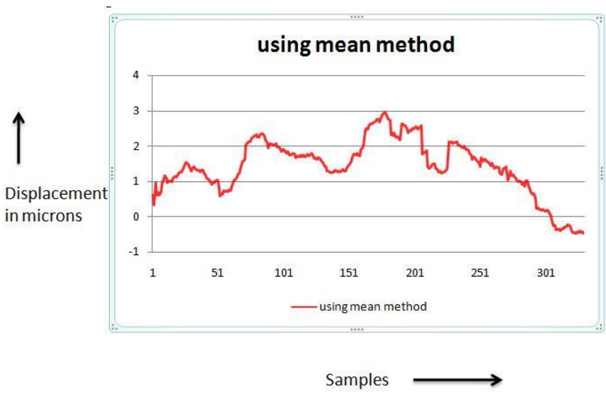

Figure 8. Values obtained using Mean Method (Old Technique)

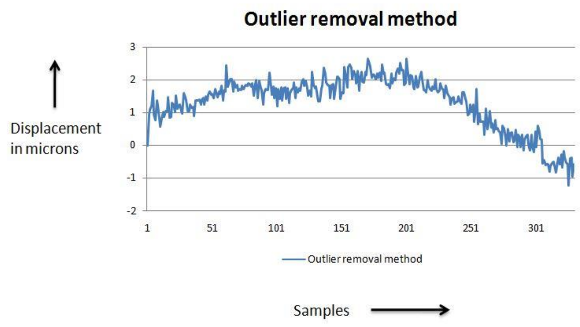

Figure 9. Values obtained using Outlier Removal Method

LVDT-measured data was split into two streams. The first stream followed the previous technique, which did not involve outlier removal. The average of this data stream was calculated after reaching the machine cycle and is represented by the red line in Figure. 8. The second stream was passed through the outlier removal code, and the resulting data was indicated by the blue line in Figure. 9.

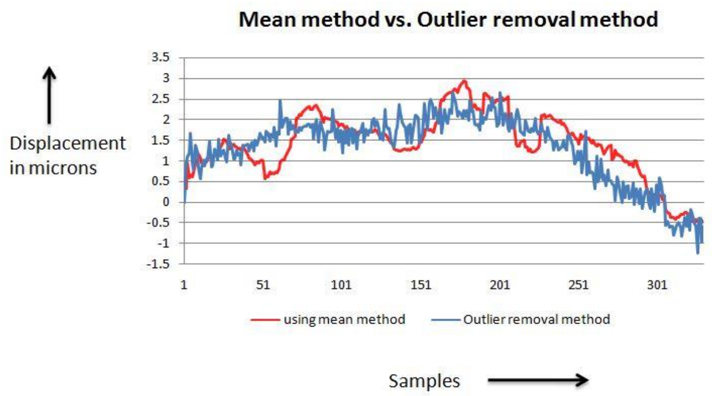

Figure 10. Comparison of Mean method and Outlier Removal Method

By comparing the two lines, it was noticed that red line exhibited significant fluctuations, while blue line appeared relatively smooth and most of the readings in the red line were off by approximately 0.5 microns. However, in one particular instance the deviation was as high as 1.47 microns. The proposed outlier detection method using Peirce's criteria significantly improved the accuracy of measurement by systematically identifying and eliminating outliers in real time. Comparison between Mean method and Outlier removal revealed consistent data trend, confirming the method's success in producing reliable results. Peirce's criteria effectiveness was supported by results shown in Figure 10, exhibiting its impact on measurement accuracy. This method highlighted its role in eliminating outliers, which helps in enhancing the reliability of data and guarantees more precise and accurate measurements in industrial data analysis.

8. Conclusion

The implementation of real-time outlier detection and removal using LabVIEW proved to be effective in improving the accuracy of measured data. By storing continuous sensor data into an array and applying data segmentation techniques, outliers were identified and eliminated from the dataset. LabVIEW code successfully captured sensor data and stored it in an array for additional analysis. Outliers were detected and removed using Peirce's criteria. The average was calculated after outlier removal, which demonstrated improved accuracy compared to the initial average. The comparison between the red line (representing data without outlier removal) and the blue line (representing data with outlier removal) clearly shows the benefits of the outlier removal process. The blue line exhibited smoother variation, indicating a more reliable and consistent measurement. Most of the values represented in the red line are approximately 0.5 microns in most cases and up to 1.47 microns in one instance, further emphasizing the necessity of outline removal for accurate results. Real-time outlier detection and removal implemented in LabVIEW proved to be a valuable technique for improving the accuracy and reliability of data analysis. This provides an automated response to anomalous data points, reducing the potential for incorrect analysis and enabling more precise measurements.

References

[1] Ross, Stephen M. "Peirce's criterion for the elimination of suspect experimental data." ,Journal of engineering technology, 2003, pp. 38-41.

[2] Lin, Lily, and Paul D. Sherman. "Cleaning data the Chauvenet way." The Proceedings of the Southeast SAS users Group, SESUG Proceedings, 2007, pp. 1-11.

[3] Rochim, Adian Fatchur Rochim Adian Fatchur. "Chauvenet's Criterion, Peirce's Criterion, and Thompson's Criterion (Literatures Review)." ,2021.

[4] Limb, Braden J., Dalon G. Work, Joshua Hodson, and Barton L. Smith. "The Inefficacy of Chauvenet's Criterion for Elimination of Data Points." Journal of Fluids Engineering 139, no. 5 (2017): 054501.

[5] Masi, Alessandro, Alessandro Danisi, Roberto Losito, Michele Martino, and Giovanni Spiezia. "Study of magnetic interference on an LVDT: FEM modeling and experimental measurements." Journal of Sensors 2011 (2011).

[6] Masi, Alessandro, Alessandro Danisi, Roberto Losito, Michele Martino, and Giovanni Spiezia. "Study of magnetic interference on a LVDT prototype." In 2010 IEEE Instrumentation & Measurement Technology Conference Proceedings, pp. 219-223. IEEE, 2010.

[7] Ayuba, Yusuf. "Temperature control and data acquisition method for factory using LabVIEW." International Journal of Computer Engineering and Technology 7, no. 2 (2016): 01-14.

[8] Kalkman, Cor J. "LabVIEW: A software system for data acquisition, data analysis, and instrument control." Journal of clinical monitoring 11 (1995): 51-58.

[9] George, Ashline, Ashly Sunny, Jerin Cyriac, and Manu Francis. "Signal Processing in LVDT for automatic calibration." In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 1114-1118. IEEE, 2016.

[10] Niculescu, Titu, Florin Gabriel Popescu, Marius Marcu, and Ioan Răzvan Slusariuc. "Improving the quality of measurements in electrical processes using NI USB data acquisition systems." Calitatea 20, no. S1 (2019): 293.

[11] Hinduja, S., D. Mladenov, and M. Burdekin. "Assessment of force-induced errors in CNC turning." CIRP Annals 52, no. 1 (2003): 329-332.

[12] Joshi, Sarthak, and Shrikant Madhav Harle. "Linear variable differential transducer (LVDT) & its applications in civil engineering." Int. J. Transp. Eng. Technol 3, no. 4 (2017): 62.

[13] ElectricalWorkbook, https://electricalworkbook.com/lvdt-linear-variable-differential-transformer

[14] Displacementsensors, https://www.marposs.com/eng/product/lvdt-hbt-usb-and-digitized-probes

[15] National Instruments, https://www.ni.com/en-in/shop/model/usb-6003.html

Author's biography

Vishnu V received his BE from Presidency University in 2021, pursuing M. Tech in Control and Instrumentation at University Visvesvaraya College Engineering (UVCE), Bangalore University, K.R. Circle, Bengaluru.

Manoj S G received his BE from Visvesvaraya Technological University, Belagavi in 2017, currently working as Engineer in Research Technology and Development Group at Ace Designers Limited, Peenya, Bengaluru.

C. M. Maheshan received his BE and ME from Bangalore University in 1999 and 2003 respectively. He is pursuing his Ph.D. in Electrical Engineering from VTU, Belagavi. He is currently working as an Associate Professor at the Department of Electrical Engineering, UVCE, K.R.Circle, Bangalore University, Bengaluru. His research areas include Power systems engineering, Image Processing, and Control systems.

H. Prasanna Kumar received his BE from Karnataka University in 1997, M. Tech from Visvesvaraya Technological University (VTU) in 2005 and received his Ph.D. during 2015 at Anna University, Chennai. He is currently working as Professor in the Department of Electrical Engineering, University Visvesvaraya College Engineering (UVCE), Bangalore University, K.R. Circle, Bengaluru. His research interests cover Biomedical Image processing, Instrumentation and Process Control.