IMEKO TC8, TC11 and TC24 Joint Conference
October 11-13, 2023, Madeira, Portugal

DEVELOPMENT OF OPEN-SOURCE TOOLS FOR THE DIGITAL AND MACHINE-READABLE CALIBRATION OF FLOWMETERS WITH NUMERICAL DISPLAYS

G. Esteves Coelho/Presenter[1], A. Pinheiro[2],A. Silva Ribeiro[3], C. Simões[4], L. Martins[5]

[1] Laboratório Nacional de Engenharia Civil, Lisboa, Portugal, gfcoelho@lnec.pt
[2] Laboratório Nacional de Engenharia Civil, Lisboa, Portugal, apinheiro@lnec.pt
[3] Laboratório Nacional de Engenharia Civil, Lisboa, Portugal, asribeiro@lnec.pt
[4] Laboratório Nacional de Engenharia Civil, Lisboa, Portugal, csimoes@lnec.pt
[5] Laboratório Nacional de Engenharia Civil, Lisboa, Portugal, lfmartins@lnec.pt

Abstract:

This paper presents a methodology for obtaining digital machine-readable measurements from numerical displays images. The proposed method provides means to automate and digitalize a previously manual and labour-intensive laboratory procedure for flowmeters calibration.

The proposed method allows to obtain machine-readable readings from remote numerical displays with available-off-the-shelf hardware and open-source software. By using smartphones for remote image capture and streaming and the Tesseract open-source OCR engine, is possible to leverage the infrastructure's digital transition, improve procedures efficiency and effectiveness while promoting sustainable actions with cost reductions.

Keywords: Metrology; Automation; Process efficiency; Machine-readable data; Open-source programming.

1 INTRODUCTION

The Industrial 4.0 era and particularly the Industrial Internet of Things (IIoT) lead to an unprecedented ability to use different types of sensors and generate enormous amount of measurement data, requiring traceability. The heterogeneous nature of the equipment and the laboratory with inadequate Information Technology (IT) penetration, is often the bottleneck for a fully digitalized calibration process.

The development and establishment of digital processes and digital infrastructures offers enormous potential for overall calibration process efficiency. However, this process is often doomed to manual interaction and is usually labour-intensive. Nevertheless, it is possible to reduce human interaction and increase process efficiency and effectiveness by increasing the infrastructure's IT penetration, towards a fully digital and machine-readable calibration process.

The current accessibility to technology combined with the maturity of open-source software, including Artificial Intelligence (AI), allows the introduction and interoperability of digital processes with available off-the-shelf hardware and custom software, with reduced financial investments. The method proposed for the calibration of flowmeters with numerical displays, aims to aggregate and automate in a single computer all the data collection and storage of a typical decentralized and manual calibration process, gathering images remotely with cameras (image synchronization is also discussed), followed by pre-processing the images and obtaining the machine-readable readings from the numerical displays by means of Optical Character Recognition (OCR).

This method is scalable to virtually any number of cameras and can be applied to both laboratory calibration and *in-situ* measurement processes. The software is developed with open-software solutions implemented in Python, with OpenCV for image pre-processing and Tesseract for OCR engine implementation.

2 INFRASTRUCTURE DESCRIPTION

The Unit of Hydraulic Metrology (UHM) is a R&DI infrastructure jointly coordinated by the Department of Hydraulics and Environment (DHA) and the Scientific Instrumentation Center (CIC) of LNEC, with competence and capabilities to develop research in Hydrology and Hydraulics and to provide traceability to instrumentation and systems applied in a wide range of measurement quantities, namely, flow rate (mass and volumetric), flow speed, volume, level, and precipitation.

The laboratory infrastructure (figure 1) has several hydraulic test benches allowing to establish different conditions to obtain flow rate by the primary gravimetric measurement using two weighing platforms (reaching 3 ton and 30 ton of mass) and the measurement of time using universal time counters, all traceable to the primary standards of IPQ (Portuguese Institute for Quality, the Portuguese National Metrology Institute). The main experimental facility as the following operational capabilities:

- volumetric flow rate ≤ 0.500 m^3/s;
- mass flow rate ≤ 400 kg/s;
- nominal diameter ≤ DN 400;
- maximum operating pressure ≤ 1.0 MPa;
- power ≤ 250 kW of electric power groups;
- power ≤ 75 kW for electric pumps not coupled to drive motors.

This Unit supports the skills that allow LNEC-UHM to be a Designated Institute for the flow rate and flow speed for liquids, according to the international recognition accepted by the BIPM in 2021 and confirmed by EURAMET in 2022. The management of UHM is developed according to the LNEC Quality Management System complying with the requirements of the ISO/IEC 17025 standard [1].

The R&DI develops methods and apply processes to provide traceability and to perform metrological characterization related with several types of measuring instruments, namely, ultrasonic flowmeters, turbine meters, positive displacement flowmeters, differential pressure flowmeters, rotameters, mass flowmeters, Parshall flumes, among others [2-5].

This infrastructure has human resources and technologies capable of promoting hydraulic metrology services and metrological information management, in a variety of areas of water resources management (water supply, undue inflow, agricultural uses and wastewater treatment), and in different frameworks (management entities, industry, manufacturers, and customers).

Figure 13: UHM infrastructure overview

3 THE DIGITAL TRANSITION CHALLENGE AND IMPLEMENTATION

The previous described infrastructure has been subject to successive equipment modernization to foster and leverage the digital technologies and transition. More recently, the infrastructure's automation legacy system was upgrade in the perspective of the Industry 4.0 and towards the Industrial Internet of Things (IIoT) [6]. The integration of standard industrial communication protocols and modern technologies leverages the infrastructure's automation capabilities and big data analysis, aiming to improve efficiency and effectiveness of the calibration procedures, the auditing process and the infrastructure's maintenance actions.

In the same perspective, the work presented here aims to contribute to the modernization of this infrastructure and contribute to the ongoing digital transition.

The main goal with this work is to acquire the measurements from the flowmeter under calibration in a machine-readable format to able to collect the measurements directly in a digital media or support (e.g. spreadsheet). However, given the wide range of flowmeters devices that are currently calibrated at this infrastructure, with diverse interfaces and different numerical displays, a one-size-fits-all solution approach is difficult to reach. Additionally, the remote location of the equipment in the plant usually restricts the use of cables and wired devices to obtain the measurements of the flowmeters under calibration. Nevertheless, the majority of the flowmeter's devices are at least equipped with one numerical display that enables the measurement by inspection.

To tackle these issues, the proposed approach general principle employs imaging devices to capture the remote images from the numerical displays in real-time and obtain the measurements in a numerical machine-readable format by means of Optical Character Recognition (OCR).

In figure 2 the implemented system architecture is presented. From left to right: Near the flowmeter under calibration an imaging device (such as smartphone) is installed. The imaging device is positioned in order to obtain an image of the flowmeter's numerical display; The image is stream over Wi-Fi to a local PC; The PC implements the OCR engine and image processing and also provides Network Time Protocol (NTP) for clock synchronization between the capturing devices and the PC; The OCR engine converts the numerical display from the image to machine-readable numbers format that can be stored in digital support such as spreadsheets files.

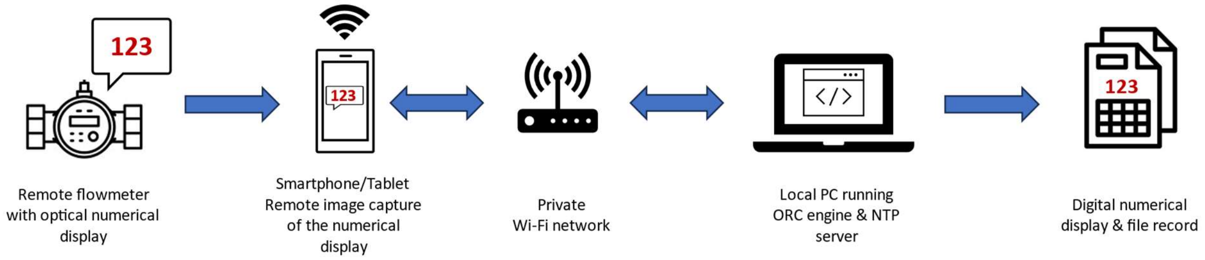

Figure 14: System implementation architecture

a. Experimental system implementation

The remote flowmeter's numerical display can be capture with available off-the-shelf smartphones, considering that these devices have good image quality with low optical distortion and are Wi-Fi capable. The smartphone is mounted in a tripod with the camera properly facing the flowmeter's numerical display, as depicted in figures 3 and 4.

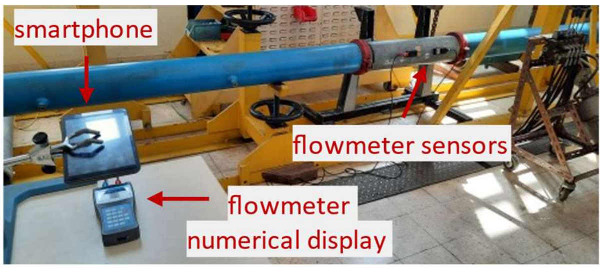

Figure 15: Experimental setup example for ultrasonic flowmeter calibration at the UHM plant

There are several implementations and software available to stream the video from a smartphone to a PC. In this case we used the android IP Webcam app (available in Google Play), that provides unicast video streaming with session validation capabilities. Although, private Wi-Fi network are available at UHM´s infrastructure, the later option is used *in-situ* calibrations environments where private networks are not available, and communications security are mandatory.

Figure 16: Close up of the flowmeter's numerical display and the smartphone for image streaming

Although the image processing is conducted in an asynchronous model by design, the NTP synchronization of the devices provides means of measurement synchronization between the flowmeter and the standard measurement and provides traceability and auditability of the overall process. For this purpose, the image streaming is time stamped locally in the smartphone device and the PC provides another time stamp for the measurements during the calibration process. In order to provide a temporal reference of the entire process, the capturing devices (smartphones) and the image processing unit (PC) are time synchronized with the same NTP server. This NTP server is provided from the PC to have the same time reference.

With the IP Webcam app installed in the smartphone, the video stream is accessed through the device's IP. This configuration provides a real-time video stream of the remote flowmeter's numerical display to a PC located in our monitoring and command room, see figure 17.

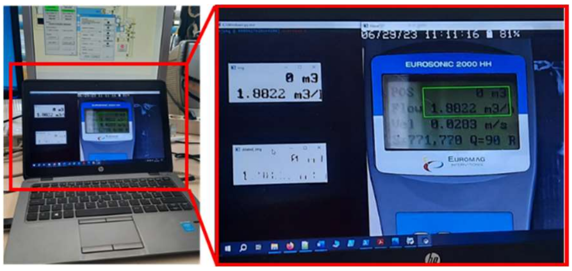

Figure 17: Remote image of the flowmeter display in the PC in the monitoring and command room

The image processing application and OCR engine was developed in Python with fully open-source software: using OpenCV for the image pre-processing and using Tesseract for the OCR engine.

The first step is to is to create a bounding box around the area of numbers of interest in the image, namely in the area corresponding to the measurement of the flowmeter. Since the OCR engine is sensitive to the image quality, e.g., light exposure, display contrast and especially the presence of light reflections in the display surface, the next step the numerical display image is converted in binary (black and white) video format. This step is needed to achieve adequate confidence levels from the OCR results. With the binary image output (see figure 18), the operator can evaluate the appropriate image quality to send to OCR engine.

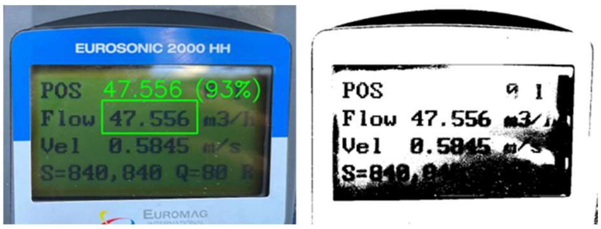

Figure 18: (left) Original image of the flowmeter numerical display with OCR results overlayed and (right) corresponding binary image

With a single command in the PC, the binary image is sent to Tesseract OCR engine, where the machine-readable of the numerical display from the OCR engine is overlayed in the image along with the confidence value (see figure 18). The digital information of the OCR's numerical display and the corresponding confidence value is stored in a spreadsheet file, along with the PC time stamp and the original processed picture stored path, for auditing purposes.

4 RESULTS

To assess the proposed methodology, several pictures of the flowmeter's numerical displays were acquired in different angles and light conditions and evaluated with the OCR engine. In general, the results show that the implemented OCR engine is adequate to convert the numerical measurements from the numerical display to machine-readable format, to the majority of the flowmeters under calibration.

In figure 19 an example of the numerical display of an ultrasonic flowmeter (model eurosonic 2000 HH) is presented. The OCR results are overlayed in the picture showing the obtained numerical conversion (left) and a confidence level of the OCR (right). Figure 8 shows the binary image output of the pre-processed image in the bounding box of figure 7 before sending to the OCR engine to obtain the measurements in machine-readable format.

Figure 19: OCR result and confidence overlayed in the numerical display image

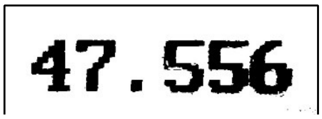

Figure 20: Binary image output of the pre-processed image of interest in figure 7

With proper display's contrast and light conditions, the OCR results in a correct numerical conversion. However, in the presence of low brightness and/or light reflections, the conversions results are inadequate as expected. This is an inevitable limitation of the OCR engine, nonetheless it can be mitigated by assuring a fair quality of the images before applying it to the OCR engine. To implement this, the image stream is preprocessed to obtain the binary counter part of the images. Image binarization enables to obtain the best contrast between the background and the numbers to be converted and provide useful information to the operator about the quality of the image stream in the OCR's perspective.

Figure 21 shows an example of OCR output results in presence of adverse light conditions. In this case, although the image appears to be in fair conditions for human interpretation, the OCR engine outputs the incorrect numerical measurements when the binary image quality is inferior (see figure 22).

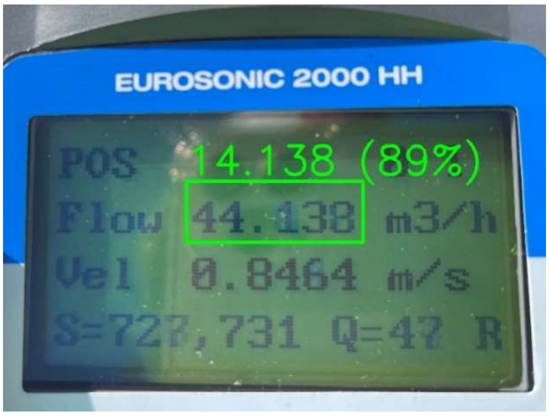

Figure 21: Influence of light reflections in the OCR output

Figure 22: Binary image output in the presence of adverse light effects

Consequently, image pre-processing is also used in the early stage of the smartphone positioning in front of the numerical display of the flowmeter under calibration, to achieve the best possible quality of the image stream to the OCR engine.

Nevertheless, the OCR can also output incorrect values due to limitation in the OCR model. The case presented in figure 23 is an example of the incorrect identification of the zero value, even in good light conditions (see figure 24). This is a limitation of the current OCR engine for the slashed zero "Ø" that is incorrectly converted to number "6". This problem only affects displays using slashed zero and is further discussed in section 5. However, the confidence low values give some insight of this limitation, and for this reason the OCR output is always evaluated and stored along with the confidence values to detect and potentially manage these situations.

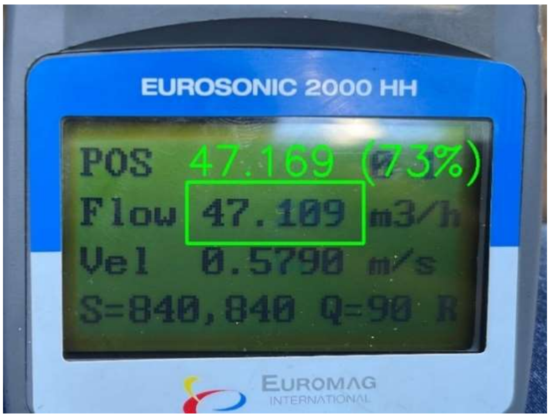

Figure 23: Slashed zero OCR engine problem

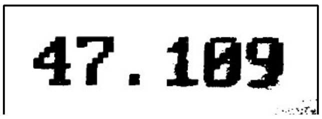

Figure 24: Corresponding binary image

5 DISCUSSION AND FUTURE WORK

A solution to automate and digitalize a manual and labour-intensive laboratory procedure was presented. With available off-the-shelf hardware (smartphone and laptop) and open-source software was possible to develop and implement a new tool that enables to read remote numerical displays of a flowmeter and obtain its measurements in a machine-readable format ready to store directly in a computer with minimal human interaction.

The results showed that its possible to improve laboratory procedure efficiency and effectiveness with minimal financial investment. The proposed solution also promotes sustainable practices, important in the digital transformation, by reducing intermediate processes that are usually supported by pen-and-paper methods.

Since the image streaming is implemented over standard ethernet protocols, this method is scalable to any number of smartphones and feasible for calibration of multiple flowmeters simultaneously.

Although the laboratory procedure for flowmeter's calibration usually requires human interaction, there are some procedures that is desirable to automate to further increment the process efficiency. Since the process for the calibration of a flowmeter device involves gathering simultaneous and synchronous measurements (both the device under calibration and the standard measurement), future work addresses further reduction in human interaction towards a semi-automated procedure. In order assure the OCR best performance, one potential improvement is to implement a threshold approach to the confidence level of the OCR output by implementing a while-loop with a time-out mechanism: The flowmeter measurements are only accepted based on the OCR's confidence level threshold that in turn triggers the standard measurement (standard measurements are obtained in machine-readable format and in real-time) and both measurements are obtained simultaneously.

To reduce the influence of adverse light conditions resulting in the poor OCR performance, on solution is to use artificial and controlled light environment, combined with low reflective backgrounds.

The problem regarding the slashed zero mismatch is a well-know and documented issue in the Tesseract open-source community. To solve this issue adequate OCR models are need. These models can be obtained directly from the open-source community or by training new OCR models with adequate training data. Nevertheless, as already stated this limitation only affects the measurements with devices that employs this specific type of numerical font.

6 REFERENCES

A. S. Ribeiro, L. Dias, A. C. Martins, H. Dimas, R. Mendes, "Flow measurement traceability used to improve water management", Test & Measurement 2018 International Conference. NLA – South Africa, Western Cape (South Africa), 2018

M. A. Silva, C. Amado, A. S. Ribeiro, D. Loureiro, "Uncertainty evaluation in time-dependent measurements", Measurement 196, 2022. : DOI:10.1016/j.measurement.2022.111196

A. Ribeiro, D. Loureiro, M. C. Almeida, M. Cox, J. A. Sousa, M. Silva, L. Martins, R. Brito, A. C. Soares, "Uncertainty evaluation of totalization of flow and volume measurements in drinking water supply networks", FLOMEKO 2019 – 18th International Flow Measurement Conference, 26-28 June, LNEC, Lisboa (Portugal), 2019.

C. Simões, A. S. Ribeiro, M. C. Almeida, D. Covas, "Uncertainty evaluation related with the fitting of probability distributions to rainfall experimental data", IMEKO XXIII World Congress (IMEKO 2021) - Measurement: sparking tomorrow's smart revolution. Yokohama (Japan). 30Aug.–3Sep, 2021.

R. S. Brito, M. C. Almeida, A. Ribeiro, "Enhancing hydraulic data reliability in sewers", Water Practice & Technology Vol 17 No 1, 431, 2022 DOI: 10.2166/wpt.2021.093.

G. E. Coelho; A. S. Ribeiro;J. dos Reis, J.; C. Simões, – "Industry 4.0 Legacy Systems Integration Case Study", IMEKO TC11 & TC24 Joint Hybrid Conference, Dubrovnik, Croatia, October 17-19, 2022.