Adaptive Process Control and Sensor Fusion for Process Analytical Technology

Niall O’Mahony, Trevor Murphy, Krishna Panduru, Daniel Riordan, Joseph Walsh
IMAR Technology Gateway, Institute of Technology Tralee, Tralee, Ireland
E-mail: niall.omahony@research.ittralee.ie

Abstract—Increased globalisation and competition are drivers for process analytical technologies (PAT) that enable seamless process control, greater flexibility and cost efficiency in the process industries. This research aims to introduce an integrated process control approach, embedding novel sensors for monitoring in real time the critical control parameters of key processes in the minerals, ceramics, non-ferrous metals, and chemical process industries. The paper will discuss smart sensors, data fusion and process modelling and control in industrial applications with an emphasis on solutions enabling the real-time data analytics of sensor measurements that PAT demands.

Keywords—Machine learning; predictive control; data fusion; process analytical technology.

I. INTRODUCTION

This paper will give a snapshot of industrial applications of data fusion and process modelling and control encountered in research to determine the key elements of a global control platform for use in industrial processes as part of their implementation of PAT [1] shown in Figure I-1. The FDA have defined PAT to be “a system for designing, analysing, and controlling, manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality.” [2].

A concept closely aligned with PAT is Quality by Design (QbD) by which key strategic process specific attributes are identified in order to devise a robust control strategy which is monitored and constantly updated for continuous process improvement [3]. Section II will mention some chemometric techniques used to develop process models for such a strategy.

The key word in the definition of PAT is ‘timely’ as PAT is concerned with observing the progression of reactions/processes (in batch or semi-continuous processes) as well as controlling process stability (in continuous processes), which requires real-time, in-line or on-line process measurements and the application of real-time data analytics. This paper will focus on the latter topic with Section III reviewing some approaches involving machine learning in adaptive process control.

Section IV will overview the benefits of smart sensors to a PAT system, particularly in terms of cost, ease of integration and maintenance. Section V will investigate the application of a wider control platform, with respect to data fusion, to tackle the pre-existing barriers to the adoption of PAT- the lack of a flexible platform for sensor integration and handling the large volumes of data generated.

The research project receives funding from the European Community’s Framework programme for Research and Innovation Horizon 2020 (2014-2020) under grant agreement no. 637232.
III. ADAPTIVE CONTROL

Adaptive/Predictive Control involves using process understanding of the effects of adjusting critical control parameters at any given time on future product quality in order to optimise process development. There is an inherent link between adaptive control and process modelling in the previous section, as this is where the models developed by chemometrics are applied to improve process efficiency. Adaptive Control involves monitoring the process in order to become familiar with its behaviour. This procedure can be done by process experts (through Model Predictive Control (MPC) [7], [8] or other algorithms [9]) or be automated with the use of machine learning techniques such as Kalman Filters, Neural Networks or Adaptive Neuro-Fuzzy Inference Systems.

A. Kalman Filters

The Kalman Filter and its extensions are remarkably simple and effective algorithms [10] that have been used in applications in trajectory estimation, robot motion planning and control, sensor fusion [11] and sensor calibration. It is an “optimal recursive estimator of the state of an uncertain dynamic system” contaminated by white Gaussian noise [12]. Its main advantage is its ability to operate in real-time [12] and overcome noise [13] although it has limitations when outliers are present [10].

Implementing Kalman filters in a microcontroller digitally is challenging however because of the high computational complexity and effect of quantisation error on stability. Technological advancements such as FPGA based System On Chip (SoC) platforms have made it possible and mean that the application of Kalman filters is likely to spread in the near future [12]. A smart sensor implementation of Kalman Filters was encountered in this research in a radio wave moisture sensor where the sensor applies Kalman Filters for situations with non-continuous material flow to achieve more intelligent continual averaging and produce an analog output for process control in concrete mixer and fluid bed dryer applications [14].

B. Artificial Neural Networks

Artificial Neural Networks are a computing paradigm inspired by the functioning of the human brain. Like the human brain, it is composed of many computing cells or ‘neurons’ that each perform a simple operation and interact with each other to make a decision.

The simplest form of a neuron is a perceptron (where inputs and outputs are binary), with more complicated neurons simply having floating point inputs and outputs and a transfer function applied to give better sensitivity at the threshold of decision making. Each neuron multiplies each of its inputs by a weight dependent on the importance of the input to the decision the neuron is making [15]. To give an example in industry, if we wanted to determine if our process was at a particular stage of its development cycle, we may treat one of the inputs, say viscosity, with a greater weight than another input, e.g. moisture content, based on the fact that viscosity is a better indicator of process development. The sum of products of the inputs and their weights is then compared with a bias value to determine the output of the neuron. Adjusting the proportions of the weights with respect to each other or with respect to the bias greatly influences the decision-making process. When many neurons are combined in a network as in Figure III-1 [16], quite complicated decisions can be made.

Neural networks learn using a set of input for which the correct outputs are given, called a training set. The network adjusts the weights and biases of the neurons so that the correct output is produced by the neuron without influencing the decision making of the rest of the network. Machine learning can be performed on a pre-collected training set or can occur online as a process is running (which can be supervised or unsupervised). A cost function, otherwise known as a loss or objective function is used to determine how well the network is learning by adjusting the weights and biases to find the minimum mean squared error from the correct output for example. Since finding the minimum of a function with many variables is not computationally easy using differentiation, a gradient descent technique (among other techniques) runs simulations to estimate the minimum cost function [17]. The area of Deep Learning is all about learning or ‘credit assignment’ across many layers of a neural network accurately, efficiently and without supervision and is of recent interest due to enabling advancements in processing hardware [18].

The success of machine learning relies on ‘Big Data’ (very large data sets) and on the fact that more data beats wiser algorithms. The algorithms produced are tailored to their specific process and focus is taken away from developing the perfect process model meaning less time and labour is spent on manually fine-tuning a model which may have to change anyway. The acquisition and storage of vast amounts of data, the vanishing gradient problem [19] and overfitting (developing over complicated models that fit the training data precisely) are challenges associated with neural networks. Ortega-Zamorano et al. highlight the difficult task of selecting the right neural network architecture for any particular application and proposes Constructive Neural Networks (CoNNs) that generate networks that grow as input information is received. The proposed solution also has a short training period and employs competition between neurons and filtering to prevent overfitting [20].

Such is the versatility of Neural Networks that they have been implemented for many sensor applications including metrological performance enhancement [21] (calibration, nonlinearity correction, offset, identification), actuation control, sensor fault detection and classification [22], [23], process control [24], large sensor arrays [25] and sensor fusion [26], [27].
C. Adaptive Neuro-Fuzzy Inference Systems

Fuzzy logic is a method by which a human’s linguistic interpretation can be translated to and from measurements readable by a microprocessor (in numerical format). For example, we recognise colour as belonging to a subset (a membership set) or even a number of subsets under which a range of colours can lie, e.g. red, amber, yellow. Fuzzy logic can be used to quantify colour sensor measurements in the same way and use them for industrial control applications [28]. A fuzzy control unit performs three basic processes; fuzzification (translating the sensor readings to the degrees of ownership to each of the controller’s membership sets), rule evaluation (the Fuzzy Inference Unit (FIU)), and defuzzification (translating the fuzzy outputs to system outputs for process control) [29]. At the core of the controller is the FIU which uses a set of fuzzy rules defined by the user to map fuzzy inputs to fuzzy outputs [30], e.g. in a polymerisation example dealt with in this research, if the colour shade of the resin gets too dark, then stop the process.

Fuzzy control offers the advantage of being able to easily translate operator insight of a process into controller nonlinearities. Thus, it offers a better user interface that enables internal expertise to be exploited [31]. The technology has been applied to improving control in conventional PID controllers in industrial processes, e.g. polymerisation, distillation [31] and manufacturing process control [32]. Fuzzy logic alone does not provide adaptive control however. ANFIS [33] combine Fuzzy logic and neural networks to provide the advantages of both technologies [34], [35] and mostly start with a fuzzy system to which neural network learning is applied [29]. Combined Fuzzy Logic and Neural Network systems that start with neural networks and apply fuzzy logic are in more recent development and adapt in a more sophisticated manner [29].

D. Genetic Algorithms

Genetic algorithms, or more broadly, Evolutionary algorithms are global search and optimisation techniques inspired by the principles of natural evolution and genetics. It is a heuristic approach that works well on complicated real-world problems where traditional optimization methods frequently fail or perform poorly [36]. Frank describes evolutive learning as the ‘ultimate technology’ for automatically applying machine learning to any application [29]. It is still an emerging technology that has not yet been widely applied in industry [34].

IV. SMART SENSORS

According to Malar and Kamaraj “20% of the total cost of a data acquisition system belongs to the hardware configuration and calibration of sensors” [37].

A smart sensor is a sensor with some additional functionality it owes to the addition of a microprocessor [38]. An increase in cost is introduced by the addition of a microcontroller, however, the advantages introduced offset this cost. Wei et al. detail how making a sensor IEEE 1451 conformant (a smart sensor family of standards aimed at enabling plug and play capability) provides signal conditioning and processing, self-recognition and self-documentation easing system integration. As a result, the time taken to setup a sensor network is drastically reduced along with the related costs [39].

Costs can also be incurred from incorrect calibration arising from the manual entry of important parameters. The resultant inaccuracy leads to a deviation from optimum process control. Smart sensors enable remote/automatic calibration so that the sensor’s parameters can be corrected promptly without the need for manual intervention and process downtime. Thus, maintenance costs are reduced as are costs associated with poor product quality [40].

Ongoing development in MicroElectroMechanical Systems has brought many improvements to the smart sensor industry. MEMS sensors are serially manufactured by means of mature technologies derived from the semiconductor industry. With such mass production, comes reduced cost. MEMS is now a mature industry where a lot of the research has been done in the area of the engineering process [41], cofabricating sensors with electronics (CMOS-MEMS) [42], reducing production time [43], improving yields [44] and developing new manufacturing technologies [45].

According to Gaura and Newman, smart sensors process the signal from the sensor element, to eliminate non-linearity due to sensor imperfections, and communicate the information digitally. Gaura and Newman also define two further degrees of intelligence; Intelligent and cogent sensors. Intelligent sensors is a term, although sometimes interchangeable with ‘smart sensors’, to define sensors that go a processing step further and provide a measurement relevant to the sensors application [46].

Data processing features may include multivariate analysis to produce correlated measurements (e.g. temperature/pressure compensation). The reduction of data to ‘real world’ measurements means that the central control platform needs only to be concerned with the process information and not the specific protocols and conversion ratios of each sensor in its network. Cogent sensors reduce the data down to information the application requires using data fusion techniques, trend recognition and decision-making [47]. Deviations from normal operation or natural stages in product development can be identified and transmitted. Unnecessary data is not passed on which reduces network traffic and makes the sensor very suitable for use in very large distributed control systems.

V. DATA FUSION

Data fusion, or Sensor fusion, refers to the analysis of multiple sensor outputs to achieve better process understanding than that which would have been possible when considering the sensor outputs independently [48].

Sensor fusion is implemented on a basic level in sensors with temperature compensation (or pressure compensation, etc.). Sensor fusion can be used to create virtual or ‘soft sensors’. Soft sensors or virtual sensors can be used to boost process understanding achievable from available hardware. For one reason or another, the sensors used in a process control approach have limitations such as poor accuracy, poor field of measurement, or perhaps the parameter we want to measure cannot be measured due to limited budget, extreme process conditions or nonexistence of sensor technology. To give some examples of data fusion encountered in this research application; mass flow can be better calculated using outputs from both a microwave sensor and a Particle Size Scatterometer. Also, an
NIR sensor can be calibrated in real-time against the supporting sensors. Larger scale data fusion is often implemented with neural networks [46] and other machine learning approaches discussed previously. A number of fusion algorithms have been developed to perform tasks such as deciding on what sensors to fuse based on their reliability [50] and performing distributed sensor fusion [51], [52].

Data fusion can be categorised in a number of ways, e.g. based on inputs and outputs or on configuration, as in Table V-1.

A. Architectures

Data fusion architectures describe the distribution of sensors and the treatment to be applied to the data from each sensor. Working with a diversity of sensors with distinct output formats and periodicity requires a robust protocol to communicate between all components of the system. A fusion architecture can apply different levels of data representation to the inputs dependent on how influential they are to the decision-making process. Architectures should also have protocols dictating how to deal with issues such as sensor failure, corrupt or incompatible data and any other difficulties that could arise [53].

Multi-sensor fusion architectures are classified according to how the data/workload is distributed and can be centralised, decentralised, publish/subscribe or mobile agent based [53]. Architectures usually have 3-4 layers:

- Physical layer (communication protocols, data alignment)
- Fusion layer (performs processing on data)
- Data presentation layers (output data or perform decisions)

There is a trade-off to be considered when deciding to implement data processing at the sensor or at the fusion centre (or ‘on the cloud’) as outlined in Table V-2. Sensor processing can be limited by constraints such as computational ability, storage space, battery life and communication ability. However, doing all of the processing at the fusion centre means the communication of raw data and heavy network traffic. A balance is therefore required where the data processing duties are divided between the sensor and the fusion centre [54].

Table V-1 Categorisation based on Sensor Configuration [48].

<table>
<thead>
<tr>
<th>Category</th>
<th>Competitive</th>
<th>Complementary</th>
<th>Cooperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fused Sensors measure:</td>
<td>the same phenomenon</td>
<td>different phenomena</td>
<td>multiple phenomena</td>
</tr>
<tr>
<td>Fusion achieves:</td>
<td>Better reliability, accuracy &amp; robustness</td>
<td>A more complete view of the process</td>
<td>Data unavailable from individual sensors</td>
</tr>
<tr>
<td>Aspects</td>
<td>Redundant configuration</td>
<td>Sensors are not interdependent-offers a more complete view</td>
<td>Difficult - accuracy &amp; reliability are reduced</td>
</tr>
<tr>
<td>Example</td>
<td>Sensor arrays</td>
<td>NIRS calibration</td>
<td>E-nose [49]</td>
</tr>
</tbody>
</table>

1 Categories are not exclusive, sensor fusion can be a blend of the categories.

Benoit and Fouloy present a novel method of leaving further processing to higher levels while avoiding complexity by simply communicating the function definitions to be implemented that are coherent and specific to each node, i.e. lower nodes transmit source code for higher nodes to execute. The required processing capability of the microcontroller at each node is reduced, reducing cost [55]. Publish/subscribe and mobile agent-based architectures reduce the required bandwidth, tolerate faults and offer better network scalability and adaptability [53].

Distributed Control becomes increasingly advantageous, particularly as the average number of sensors in a typical network has grown in recent years. Distributed Control allows greater freedom - a constant communication link is not required and reduced network traffic enables less bulky (fewer wires) or even wireless communication protocols to be used.

B. Frameworks

A number of frameworks describing the underlying structure of a data fusion system have been proposed in an effort to organise the knowledge about the systems environment, e.g. Mitchell and Esteban frameworks [53]. The best-known fusion framework is the JDL framework which defines 4 levels of data fusion [56]:

1. Object Refinement: data alignment and data association.
2. Situation Assessment: describes the situation in terms of indications of warnings and plans of action.
3. Impact Assessment: awareness of the future consequences of decisions to process development.
4. Process Assessment: an ongoing assessment of the other fusion stages to ensure optimal results.

Smith and Singh give an overview of contemporary techniques in distributed data fusion respective to the 4 levels of the JDL model as well as the remaining challenges in data fusion, for instance, uncertainty management, out of sequence measurements and data correlation [56].
The commercial sensors and a point-to-point communication between a multipoint communication network using RS485 for sensors in a network can communicate. RS485 has been identified as a communications interface suited to harsh industrial environments and was determined to be most appropriate for use in this project as all the selected sensors utilise this interface (or use RS232 serial interface which can be easily used on an RS485 bus with a converter module).

The acquisition system to be developed will be a hybrid solution between a multipoint communication network using RS485 for the commercial sensors and a point-to-point communication system such as RS232 for complex sensors (NIR and particle size sensors) being developed in the project.

Since a number of the sensors use different communication protocols (Modbus, Profibus, IMPbus, etc.), the development of a custom universal communications protocol is required. The sensors send frames that include a field with its identifier on the network enabling multiple devices to be daisy-chained onto the RS485 bus. There is a need to be able to assign IDs to each of the devices on the bus in the event that sensors have the same ID ex-factory. The time stamp of every acquisition should also be included in the packets in order to easily associate data at higher levels of data fusion.

VI. FUTURE TRENDS

With the proliferation of MEMS technology and more affordable sensors comes the revolution of the Industrial Internet of Things (IIoT) where large sensor networks can be implemented to continually monitor industrial environments with unprecedented levels of detail [57]. Such pervasive sensing applications have been described as a dream by Gaura & Newman which is getting closer to being realised due to distributed control, plug and play capability, and improving MEMS manufacturing techniques [58].

This paper has identified some recent research in machine learning for industrial applications to provide an autonomous versatile control platform. Autonomous sensors reduce cost as less expert analysis is required at the central control unit. A counter-argument to such unsupervised, decentralised control is that removing human interaction is a major loss. There is no processor more powerful than the human brain (work is still ongoing in understanding our learning and memory mechanisms with a view to using this understanding in computer science [59]) and the input from experts on the particular process and algorithm developers is often very valuable. That being said, the cost of such expert analysis is very high, unfeasible and unnecessary in the long term. Therefore autonomous control is favourable for most applications and is an active area of research as increasing processing capabilities of microcontroller hardware empowers more sophisticated algorithms.

Machine learning is an exciting technology that mimics human intelligence. Gaura and Newman give an interesting suggestion of how a sensor network could be inspired by a human’s central nervous system, where transmitting raw data from all sensors (eyes, ears, nose, skin, etc.) is not feasible due to the sheer amount of data, and the sensors have a neural network memory element that learns what is normal so that sensors only transmit when something abnormal happens [46]. A similar system architecture model found in nature as inspiration for data fusion networks is localised decision-making where cells talk to their neighbours to work out how to position themselves [60]. Such mimicry of nature is common in research into technological advancement striving to achieve the efficiency and simplicity present in biological systems.

VII. CONCLUSION

This paper has reviewed industrial applications of predictive control and data fusion relevant to the implementation of PAT. Sensor fusion will be utilised to improve performance in the sensor networks of PAT applications. The first challenge to overcome in order to achieve this was discussed in Section V-C—the collection and alignment of sensor outputs. An overview was given of the benefits smart sensor technology provides to a system-wide control approach. The paper also investigated machine learning methods that can be utilised in higher levels of data fusion to achieve better process understanding and control and has identified upcoming trends in the area. ANFIS has been identified as a method with wide applications that provides adaptive control while also taking advantage of internal process knowledge. Taking into account the background knowledge detailed in this review, it is now possible to go about developing a versatile global control platform for PAT.

ACKNOWLEDGMENT

Funding is received from the European Community’s Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under grant agreement no. 637232.

REFERENCES


