Gapless Power-Quality Disturbance Recorder

L. R. M. Silva, E. B. Kapisch, C. H. N. Martins, L. M. A. Filho, A. S. Cerqueira, C. A. Duque, *Senior Member, IEEE*, and P. F. Ribeiro, *Fellow, IEEE*

***Abstract*—This paper presents and discusses the concepts of a gapless-power-quality disturbance recorder (G-PQDR) and a novelty detector. In the G-PQDR, power signals are divided into continuous frames and the energy difference between an actual frame and a reference frame is used for comparison. If the current frame is different from the reference frame, it will be recorded in the equipment and it becomes the new reference frame; otherwise, the frame is discarded and the frequency of the fundamental component is stored. The G-PQDR is able to detect and record the waveforms of the previously known PQ disturbances as well as those of the PQ disturbances not yet observed in power systems, thanks to the novelty detection concept. This characteristic is important in the context of smart grids where hidden disturbances can be detected by this methodology. For efficient data storage by minimizing the volume of the data stored, the G-PQDR uses lossy and lossless compressors. This paper describes the entire process of detection and compression, and compares the performance of the proposed method with that of a commercial-power-quality data recorder. In addition, some of the prototype implementation details are presented.**

***Index Terms*—Gapless recording, novelty detection, power quality (PQ), signal compression.**

I. Introduction

NOWADAYS there has been a growing concern regarding Power Quality (PQ) problems. These problems are due to the increasing use of non-linear loads and electronic-based equipment in residences, commercial centers, and in industrial plants. In addition, the proliferation of distributed generation in power systems can cause several PQ problems such as voltage variations and flicker, harmonic distortion, voltage dips, etc. as discussed in [1]. Therefore, the monitoring of electric power systems (EPS) in real-time, along with an offline analysis using both centralized and decentralized schemes, has grown in importance [2].

In several EPS applications such as the PQ, operation and protection, continuous data acquisition, and data storage are necessary. This is sustained since the postprocessing of this data can reveal information not observed previously, enabling system enhancement, troubleshooting, algorithm optimization, etc. For example, high frequency disturbances like transient oscillations or switching processes are only visible in the full waveform. The signature "hidden" in the raw data can be used to predict the breaking of cables. Although aggregation is useful for data reduction and comparison, deep data analysis should be enabled considering the possibilities of a full observation [3]. Continuous raw data recordings of an electrical signal are not simple because of the large amounts of data to be recorded and later transferred. Besides, few commercial instruments are currently available for recording continuous raw data at a high sampling rate [4]. Most of the conventional recorders are application oriented and are used only for acquiring short intervals of abnormal signals [5], [6].

Power system signal compression has been widely studied in the previous decades. Signal compression can be divided into lossless and lossy compression. In the former, no information is lost, while in the later, some information is discarded in order to increase the compression ratio. Lossless compression techniques such as the Lempel-Ziv and the Huffman coding are employed in some of the works, such as [7], [8]. The Power Quality Data Interchange Format (PQDIF) also provides a lossless compression [9]. Other techniques are found in the literature such as the utilization of high order delta modulation that is computationally simple and produces a high sample rate is proposed in [10] .

For achieving higher compression ratios, lossy compression is used. In this approach, the compression ratio increases with the volume of the discarded information. However, it is important to note that the reconstructed signal must preserve the essential characteristics of the signal for posterior analysis [11]. Lossy techniques generally use an orthogonal transform such discrete Fourier transform (DFT), discrete cosine transform (DCT), discrete wavelet transform (DWT), and wavelet packet transform (WPT) followed by a thresholding process in the transform domain [12]–[16].

The wavelet transform (WT) is widely used to deal with non-stationary signals. PQ disturbances are time-varying in nature, hence, the WT is an appropriate tool for compression [12]. The methodologies that use the WT for compressing the PQ disturbances can be divided into the following steps: (i) signal decomposition (ii) thresholding of the coefficients (iii) application of a coding algorithm [14], [15].

In this paper, a Gapless Power Quality Disturbance Recorder (G-PQDR) is described and evaluated. It is a system capable of recovering the entire power system signal at a high sampling rate without employing continuous raw data recording. The signal is segmented and only the segments comprising the detected

Manuscript received December 3, 2015; revised March 4, 2016; accepted April 4, 2016. Date of publication April 21, 2016; date of current version March 22, 2017. Paper no. TPWRD-01685-2015.

L. R. M. Silva, E. B. Kapisch, C. H. N. Martins, L. M. A. Filho, A. S. Cerqueira, and C. A. Duque are with the Electrical Engineering Department of Federal University of Juiz de Fora, Juiz de Fora MG 36036-330, Brazil (e-mail: leandro.manso@ufjf.edu.br; eder.kapisch@engenharia.ufjf.br; chnmartins@gmail.com; luciano.andrade@engenharia.ufjf.br; augusto.santiago@ufjf.edu.br; carlos.duque@ufjf.edu.br).

P. F. Ribeiro is with the Electrical Engineering Department, Federal University of Itajubá, Itajubá MG 35903-087, Brazil (e-mail: pfribeiro@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRD.2016.2557280

electrical disturbances are packed, compressed, and recorded for an offline analysis of the entire signal. Additionally, information regarding the fundamental frequency estimation of each signal cycle is also recorded. This information is used to reconstruct the signal between the two detected segments, recovering the entire electrical signal at any time.

The proposed power quality recorder is based on the novelty detection concept. This concept addresses the detection of any abnormal variations in the signal waveform; it detects not only the known disturbances, being suitable for application in the smart grid scenario, where the type of the disturbances is not completely known.

The traditional novelty detection concept, widely used in the biomedical area [17], [18], in industrial failure detection [19], video surveillance [20], [21], sensor networks [22] and in other applications, is slightly different from the one used in this paper. Traditional novelty detection is used when a very large number of examples of the "normal" condition are available and where there is insufficient data to describe the "abnormalities"; therefore, it is usually based on a training stage where the system is trained with the "normal" condition set; hence, it is capable of recognizing the "abnormalities" condition during the test [23].

In this paper, the novelty detection concept is used to determine if a signal frame is different from a reference frame and its application is as follows: (i) the signal is segmented in fixed length frames (ii) the current frame is compared with a reference frame (iii) if they are different, a novelty is detected. It differs from the classical novelty detection because a training set is not used.

The main contributions of the G-PQDR are: (i) the proposed novelty detection algorithm (ii) handling of the fundamental frequency variations by the detector (iii) a signal reconstruction method that avoids the occurrence of the signal discontinuities. Signal discontinuities may appear at the reconstructed signal step owing to an inaccuracy in the frequency estimation. This paper demonstrates how the discontinuity can be avoided.

Regarding the proposed handling of the fundamental frequency variations by the detector, the contribution of this paper is to build a detector with a low sensitivity to the variations in the fundamental frequency. The main idea is that a slow frequency variation should not be considered as a novelty. This is accomplished by filtering the fundamental component and the lower order harmonic components before the computation of the frame energy. Consequently, minor frequency variations seldom trigger the novelty detector, increasing the compression ratio. Further, the detector becomes more sensitive to the disturbances because the higher frequency components become less disguised and are therefore more visible and easier to detect; this is a desired characteristic for a gapless PQ disturbance recorder.

It is worth mentioning that the methodology was implemented with a holistic view. This implies that the G-PQDR functions range from novelty detection to the reconstruction of the electrical signals through the loss and lossless compression stages. The paper also discusses the choice of the mother wavelet and the number of levels in the WT, the number of bits of stored data, and its real-time implementation in Field Programmable Gate Array (FPGA) platform.

A performance comparison was carried out with a high quality (class A) Commercial Power Quality Data Recorder (C-PQDR), based on the FFT [24].

The paper is organized as follows. Section II provides the definition of a G-PQDR device and shows a review of the gapless recording equipments. Section III presents the detection and compression system for the G-PQDR. Section IV provides some of the constructive aspects of the G-PQDR hardware implementation. Section V shows the performance of the proposed G-PQDR and a comparison with the C-PQDR. Finally, the conclusions are presented in Section VI.

II. A Gapless PQ Disturbance Recorder

The term gapless recording has been used in technical literature with a dual meaning. Sometimes, and more frequently, it is related to the gapless PQ parameters, i.e., none of the PQ parameters is lost. In other cases, as in [24], it is used to denote a continuous waveform recorder. Nevertheless, in the new context of PQ in smart grids, it is possible that new events that have not been observed before, can occur. For this reason, it is important to avoid losing valuable information from the power signal. However, it is illogical to save data that has the same information, for example, the steady state of the voltage or current signal, even if they have harmonic and interharmonic content. A method to overcome this limitation is the novelty concept. The equipment that does not lose valuable information (disturbance) is named in this paper as a G-PQDR. Of course, a continuous waveform recorder is a G-PQDR but it saves the redundant information also. On other hand, the novelty method does not have this drawback and the loss of information can be controlled by an appropriate choice of the internal parameters.

A thorough survey of the leading manufacturers of power quality analyzers and data loggers was made. Nine of the top brands were examined, totaling 27 devices. From the investigation of their manuals and data sheets, it was observed that all of them are able to record the PQ parameters for a considerable time period. Depending on the aggregation time, some of the equipment is able to record the PQ parameters for over a year. Nevertheless, only two of them are able to save the waveform recordings for a long period. The one described in [24] was the only one used in this paper for comparison because it is able to record the gapless waveform for over a year, using the compression technique.

III. Detection and Compression System for a Gapless Recording

The proposed system works using the *novelty concept* to detect disturbances [25]. The signal is divided into frames and the aim is to detect whether a frame presents differences compared to a reference frame. These frames are named novelty frames. By this method, not only the common PQ disturbances (sag, swell, oscillatory transients, etc.) but any variation in the signal that may be useful for an offline analysis, can be detected. Therefore, a gapless reconstruction of the signal is possible.

Aiming to reduce the volume of data to be recorded, the proposed methodology has three compression levels: (i) novelty detection that determines the frames that will be stored; (ii)

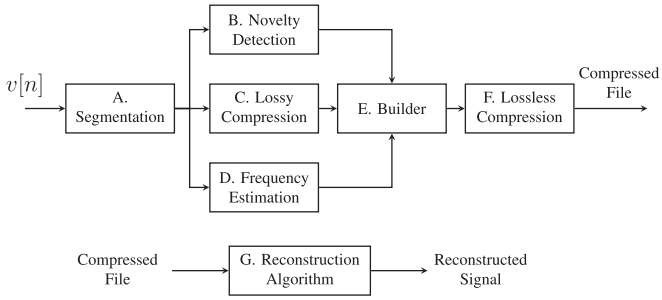

Fig. 1. Block diagram of the proposed G-PQDR.

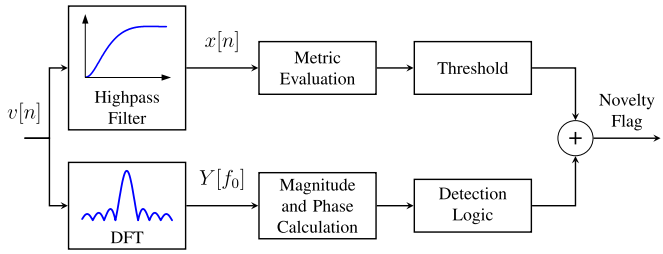

Fig. 2. Block diagram of the proposed novelty detector.

lossy compression applied in the novelty frames (iii) lossless compression performed on the packed data.

A block diagram of the proposed methodology is shown in Fig. 1. The first step involves dividing the power signal into fixed length windows. For each window, the novelty detection is performed, the fundamental frequency is estimated, and the lossy compression technique is applied. The information from the novelty detection, frequency estimation, and lossy compression blocks are combined in the builder block before the application of the lossless compression technique that is the final step in the compression stage of the G-PQDR.

A. Segmentation

Partitioning the signal into disjointed segments or frames (segmentation) is a necessary preprocessing step before effective analysis methods can be applied. The segmentation length can be fixed or variable (adaptive), depending upon the signal behavior and the processing technique.

Adaptive segmentation has been extensively used in biomedical and speech signals [26] . The main goal is to find the segment length where the signal can be represented by the same model or parameters. A similar concept has been introduced in [27] to analyze power system disturbances. Adaptive segmentation requires extra processing to identify the length for each segment before the application of a signal processing algorithm to extract information.

In this paper, fixed length segmentation is used owing to its simplicity. The frame length was chosen empirically as four cycles of the fundamental component. If the frame length is too long, the detection sensibility is reduced; if it is too short the compression ratio can deteriorate. Further studies are required to determine the optimum frame length.

B. The Novelty Frame Detection

The first step to guarantee a good compression performance is the design of a detection system capable of only detecting frames that are different from the reference frame according to certain useful metrics and a threshold value.

The proposed Novelty Detector is composed of two detection branches, as shown in Fig. 2. The upper branch is used for detecting disturbances related to a wide frequency band range (referred to here as wideband novelty detection), while the lower branch is related to the detection of low frequency disturbances (referred to here as low frequency novelty detection).

1) Wideband Novelty Detection: Wideband novelty detection is responsible for effectively detecting novel frames with a wide frequency range content. The first processing block is a high-pass filter that attenuates the fundamental component and the low frequency harmonics, aiming to improve the immunity against smooth variations in the system frequency. A 10^{th} order Finite Impulse Response (FIR) filter with a cutoff frequency of 720 Hz was used. Several filters with different cutoff frequencies were tested and the performances of the detector using each one was compared to select the best one.

The filtered version of the segmented window $(x\,[n])$ feeds the metric evaluation block, where a comparison with the reference frame (last detected novelty window) is performed. The choice of the energy parameter was based on the detection theory because it maximizes the Signal to Noise Ratio (SNR) for the detection of unknown signals corrupted by noise [28]; therefore, the Difference of Frame Energies (DFE) is used in this work. The DFE is calculated as the absolute value of the difference of the current frame energy and the energy of the reference frame, as shown in (1). If the DFE is higher than a given threshold, a novelty frame is detected, and this frame is labeled as the new reference frame.

$$\mathrm{DFE} = |\mathrm{EF}_c - \mathrm{EF}_r| \tag{1}$$

The energy of each frame is calculated according to

$$\mathrm{EF} = \sum_{n=0}^{M-1} |x\,[n]\,t|^2, \tag{2}$$

where $x\,[n]$ is the n-th sample of the frame and M is the number of samples per frame. It is important to note that $x\,[n]$ is the output of the highpass filter shown in Fig. 2.

Owing to the use of the high-pass filter, the DFE metric is not capable of detecting low frequency disturbances such the flicker and the sub-harmonics. Therefore, an auxiliary detection scheme is needed.

2) Low Frequency Novelty Detection: Different from the wideband novelty detection, where the energy of the segmented window is used for detection, a DFT centered at the fundamental frequency, performed for each cycle of the segmented window is used for low frequency novelty detection. The magnitude and phase of the DFT is computed. The main reason for using the DFT is that the low frequency disturbances (frequencies near the fundamental) cause a spreading spectrum around the fundamental component that can be detected.

The detection logic was built to detect low frequency disturbances, while being immune to smooth system frequency variations. It was observed that low frequency disturbances such as the flicker or the sub-harmonics cause significant variations in the magnitude of $Y[f_0]$ while its phase varies only slightly. On the other hand, for frequency variation scenarios, the phase of $Y[f_0]$ varies as well as its magnitude.

The value of the magnitude of the current cycle is subtracted from the magnitude of the reference cycle, generating an error signal and the same is done for the phase. The detection is performed based on these two error signals, and an event is detected when the magnitude error is greater than a specified threshold and the phase error is smaller than a specified threshold. The reference frame is updated when an event is detected.

C. Lossy Compression

Once a novelty frame is detected, the current frame becomes the reference and must be saved in the mass memory of the G-PQDR. Before being saved, the frame is processed in order to decrease the amount of information to be stored, generating a sparse frame (with several coefficients equal to zero). The sparser the frame, the better is the compression ratio.

The original signal is normally corrupted by noise, therefore, the first processing step should be the application of a denoising technique, i.e., the sample carrying no information should be zeroed. As discussed earlier, the method commonly used in this step is the DWT. To select the best decomposition filter bank, several families of the mother wavelet were tested.

To obtain the number of the decomposition levels, the entropy criterion [29] was applied for several mother wavelets. The entropy is calculated based on (3).

$$H(\mathbf{x}) = -\sum_{n=1}^{N-1} |x_n|^2 \cdot \log_2\left(|x_n|^2\right), \tag{3}$$

where x_n is the n-th sample of the vector x and N is the number of samples of this vector. According to this criterion, the optimal number of decomposition levels to be used for compressing the signal is five for all the tested mother wavelets.

The choice of the mother wavelet was done based on the Minimum Description Length (MDL) proposed in [30]. This criterion aims to minimize the relationship between the number of retained coefficients and the reconstruction error. The MDL is described by

$$\begin{aligned}\mathrm{MDL}(k,m) = \ & \min\left\{\frac{3}{2}k\log_2(N) + \frac{N}{2}\log_2\right. \\ & \left.\times\left(\left\|\left(\mathbf{I}-\Theta^{(k)}\right)\mathbf{W}_m^T\mathbf{d}\right\|^2\right)\right\} \\ & 0\le k<N; 1\le m\le M,\end{aligned} \tag{4}$$

where k is the number of retained coefficients, m is the index of the mother wavelet, N is the number of samples of the signal $\mathbf{d}$, $\mathbf{W}_m$ is the wavelet basis, and $\Theta^{(k)}$ is a thresholding operator.

Based on these criteria, a DWT of five levels with a biorthogonal (bior3.1) mother wavelet was chosen to implement the G-PQDR and its decomposition filter bank is shown in Fig. 3.

Two important additional important parameters for the lossy compression efficiency are the threshold and the wavelet coefficients thresholding and quantization. A hard threshold was adopted [31] and the threshold value was calculated using the methodology described in [32], that uses the following equation:

$$\lambda_u = \sigma\sqrt{2\log(N)} \tag{5}$$

where N is the length of the vector and σ is its standard deviation. The standard deviation σ can be estimated as described in [32] using the Median Absolute Deviation (MAD). The equation that describes this estimator is

$$\sigma = \frac{MAD(\mathbf{C}_{D1})}{q}, \tag{6}$$

where C_{D1} is the vector of level 1 detail wavelet coefficients, and q is a scaling factor.

In real time applications, the median calculation is not a simple task because the vector must be ordered. If the noise has a Gaussian distribution, it is possible to use the mean instead of the median [33]. Then, the equation that describes the *MAD* $(\mathbf{C}_{D1})$ is

$$MAD(\mathbf{C}_{D1}) = \overline{|\mathbf{C}_{D1} - \bar{\mathbf{C}}_{D1}|} = \frac{1}{N}\sum_{i=1}^{N}|C_{D1i} - \overline{C_{D1}}|. \tag{7}$$

where the overline operator indicates the averaging operation.

The retained coefficients must be quantized before being transmitted and/or stored. This is necessary for reducing the number of bits and for achieving a better compression ratio. The coefficients are quantized using

$$c_Q = \left\lfloor \frac{c}{\Delta}\cdot 2^{N_b}\right\rfloor, \tag{8}$$

where c is the original value of the coefficient, Δ is a normalization factor that depends on the signal magnitude (in this paper, considering that the input signal is always in the interval $[-1,1]$ and the frequency response of the decomposition filter bank, $\Delta = 2^4$) and N_b is the number of bits used for the quantization.

In order to find the number of necessary bits, several signals in each disturbance class were generated, decomposed, thresholded, and quantized with various number of bits. Then, they were reconstructed and compared with the reconstructed signal using the coefficients before quantization. The metric used to measure the quality of the reconstruction was the Mean Squared Error (MSE). It was noted that the error is significantly small for a number of bits equal to or greater than, for all disturbance classes. Thus, it can be concluded that bits is a suitable value for the quantization of the coefficients.

D. Frequency Estimation

One of the premises of the G-PQDR is that the smooth fundamental frequency variations in the input signal $(v[n])$ should not trigger the novelty detector. Small frequency variations are always present in the power system signal (voltage and current) because of the unbalance between the generation and the load [34]. Although small, if the detector is not insensitive to these variations, the novelties will be triggered frequently. This situation worsens in the new scenario of the smart grid where the

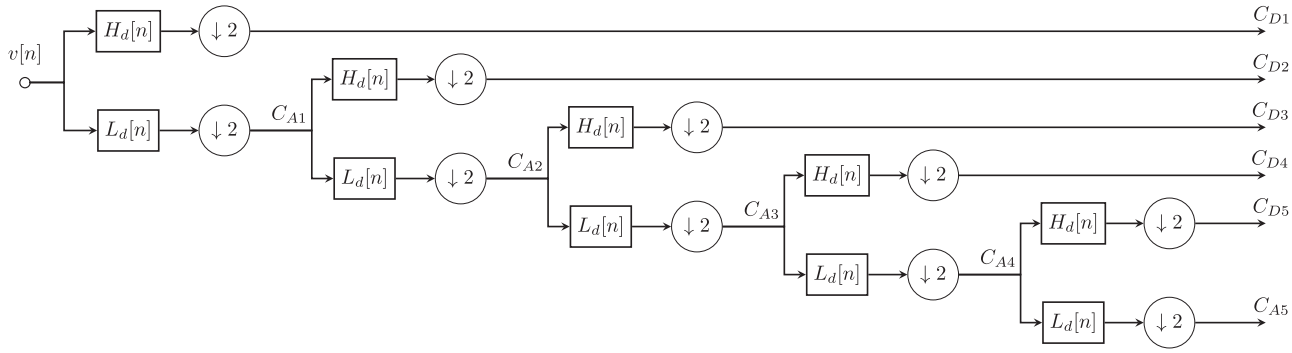

Fig. 3. Wavelet decomposition tree used in the G-PQDR with five levels.

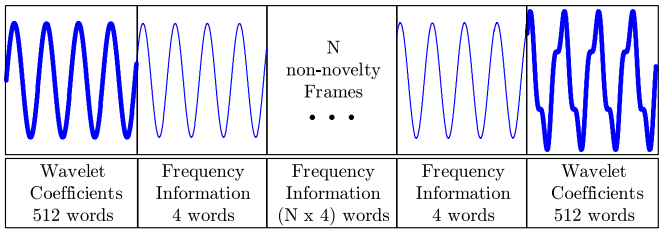

Fig. 4. Packed information according with the frame status.

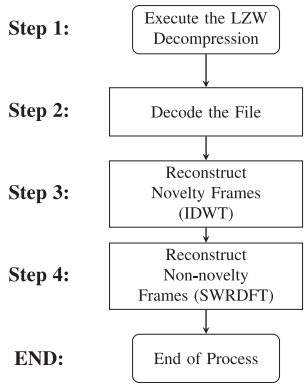

Fig. 5. Block diagram of the reconstruction algorithm.

frequency variations are expected to be higher than those in the conventional grid.

The fact that the small frequency variations do not trigger the novelty detector imposes a challenge during the signal reconstruction step that is an offline process. When the signal between two novelty frames, with a long time gap between them, needs to be reconstructed, the fundamental frequency information between them needs to be known. Hence, a frequency estimator must be included in the G-PQDR. The frequency estimator used is based on the phase locked loop (PLL) presented in [35], [36]. Instead of saving the instantaneous frequency generated by the PLL, only the average frequency of every cycle is recorded. This procedure allows the reconstruction of the entire signal between the two novelty frames, as illustrated in Section V.

It is known that frequency estimation algorithms are highly sensitive to disturbances in the input signal, generating an erroneous estimation. As the value of the estimated frequency is very important to the reconstruction of the signal, this value must be reliable. To overcome this limitation, in situations when the estimated frequencies present abrupt variations, a novelty is triggered. Therefore, when the estimated fundamental frequency presents a considerable deviation from the correct value, a novelty is triggered, avoiding the use of the wrong estimated fundamental frequency in the signal reconstruction.

E. The Builder

The builder block is responsible for packing the information of the other blocks. The information that will be packed depends upon the status of the current frame. An example is shown in Fig. 4.

The upper part of Fig. 4 depicts the signal that needs to be stored. The frames plotted with thick lines are the novelty frames and the ones plotted with thin lines are the non-novelty frames. The information that needs to be saved for each frame is shown in the lower part of Fig. 4. For the novelty frames, the 512 words related to the wavelet coefficients are packed and for the non-novelty frames, only four words of the frequency information are packed. Note that there is an interval of $(\mathrm{N}+2)$ non-novelty frames between the two novelty frames. For this interval, only the frequency information for each frame is packed, significantly reducing the amount of data that will be stored.

F. The Lossless Compression

In the last step, the packed information is compressed by a lossless algorithm. The lossless compression must be done online. Hence, the Lempel-Ziv Welch (LZW) [37] algorithm was chosen because of its capacity to be adapted to run in real time. The adaptation that was made is the limitation of the maximum size of the dictionary that is constructed during the execution. The data compressed by the LZW is saved in the mass memory of the G-PQDR.

G. The Reconstruction Algorithm

The proposed decompression system runs offline and is based on the steps shown in Fig. 5.

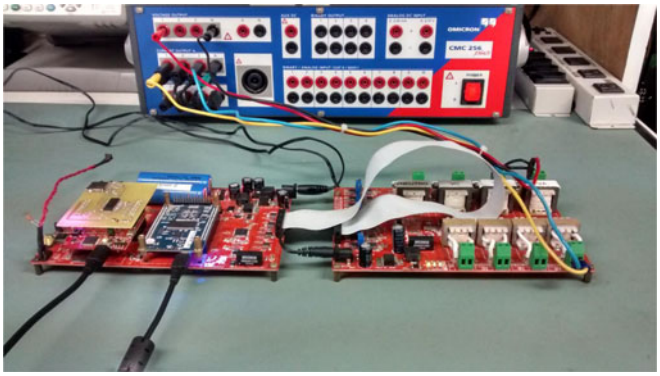

Fig. 6. G-PQDR prototype.

In Step 1, the compressed file is decompressed using the LZW algorithm generating a coded file with all the necessary information to reconstruct the signal. In Step 2, the file is decoded and analyzed in order to identify the novelty frames, collect the wavelet coefficients to reconstruct them and the frequency information to reconstruct the non-novelty frames.

The reconstruction of the novelty frames is done first. In Step 3, they are reconstructed through an inverse discrete wavelet transform (IDWT) using the wavelet coefficients that were stored in the file. Taking the case shown in Fig. 1 as an example, only the frames plotted with the thick lines are reconstructed after the execution of Step 3.

The first frame of Fig. 4 will be used as a reference for the reconstruction of the frames in the interval between the two novelty frames. In Step 4, this frame is decomposed into harmonic components using the sliding window recursive discrete Fourier transform (SWRDFT) algorithm [38]. These harmonic components will be modulated according to the averaged frequencies in each cycle of the non-novelty frames and combined using the SWRDFT algorithm to reconstruct the signal in the non-novelty interval. Therefore, the entire signal is reconstructed, as can be seen in Fig. 4. The necessity of modulating the harmonic components according to the estimated frequencies is to avoid discontinuities in the reconstructed signal. This fact will be explored in more detail in Section V.

IV. Hardware Description

A hardware prototype was built based on the proposed methodology. The prototype consists of two modules: the analog module that is responsible for the voltage and the current signal conditioning, and the processing module that implements all the described algorithms. A 16 bit analog to digital converter (ADC) with eight simultaneous channels is used. A picture of the prototype is shown in Fig. 6.

In Fig. 6, the Omicron source (used for the generation of test signals during the evaluation tests) can be seen on the top of the picture. Analog signal conditioning for the three-phase voltages and currents is performed on the circuit board seen on the right side of the picture. Signal processing, compression, and data management are performed on the circuit board on the left side of the picture, where the FPGA and the ARM Processor are located.

The processing module is assembled using two complementary platforms: a FPGA [39], and an ARM Cortex M4 [40] . The former is responsible for the entire signal processing as the Novelty Detector, the Frequency Estimation, the Wavelet and the Builder blocks; the latter uses a software framework to implement the Coding block, the SD card control, management of the system, and communication with the user. The signals from the channels are compressed and saved in the SD Card for posterior reconstruction using a PC-based application.

In the FPGA, the architecture is based on custom processors, specially developed for this application. Those processors execute tasks for the 8 channels sequentially. The processor topology chain is divided into a three-stage pipeline processing. The first stage comprises three processors that compute in parallel the Novelty Detection, the Frequency Estimation and the Wavelet Decomposition. Then, based on the Detection decision, a second pipeline stage composed of a fourth processor is responsible for packaging the information. If a novelty is detected, the information from the wavelet processor is used, otherwise, only the frequency estimation is taken into account. The external ARM processor evaluates the last pipeline stage, where the lossless compression algorithm is executed on the packed data.

The custom FPGA processors are based on a RISC architecture with separate memories for the instruction and data words (Harward Architecture) [41]. Only the internal FPGA resources are employed such as the on-chip memories and the DSPs blocks. Moreover, the Arithmetic-Logic Unit (ALU) contains only the necessary resources to evaluate the corresponding algorithm for each processor. The ALU implements a floating-point arithmetic, simultaneously providing fast software development and accurate output results. A set of only 18 instructions was developed, composing a very light and fast processor.

Regarding the software development framework, a compiler was designed using lexical and parser facilities from the GNU: the Flex [42] and Bison [43] frameworks, respectively. The compiler implements a sub-class of the C language, where only the necessary commands for the flow control and array handling were included. The resulting binary instructions are copied to the instruction memory during the final hardware synthesis and cannot be modified during nominal execution. This philosophy permits only the minimal resources necessary for the targeted algorithm evaluation.

V. Results

In this section, the performance of the proposed G-PQDR is evaluated. The Omicron CMC-256-6 plus source was used to generate the synthetic signals for testing the prototype. The prototype was tested with various disturbance signals such as the sags, swells, interruptions, transients, etc. The same signals were applied to a C-PQDR and the compression results were compared.

To evaluate the quality of the reconstructed signal the following quantities are used: the Normalized Mean Square Error (NMSE) defined by (9); the cross-correlation (COR) between

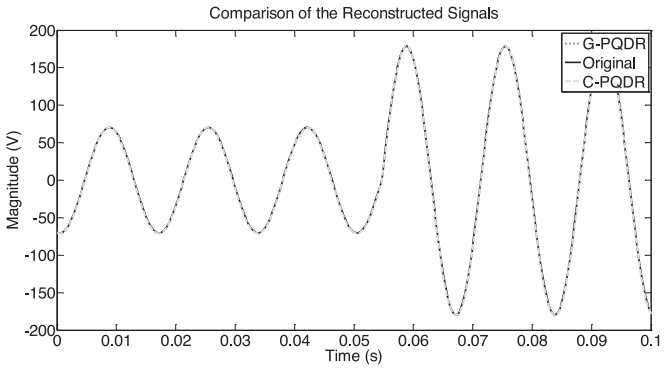

Fig. 7. Reconstructed signal during the last frames of a Sag event. The signal reconstructed by the G-PQDR (black dotted line), the signal reconstructed by the C-PQDR (gray dashed line), and the original signal (black solid line) are depicted.

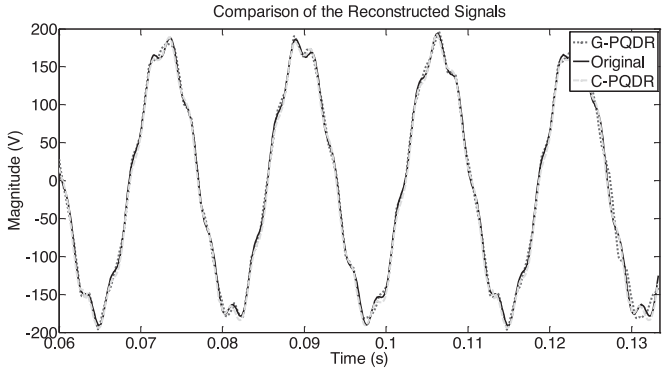

Fig. 8. Reconstructed signal during the interharmonics. The signal reconstructed by the G-PQDR (black dotted line), the signal reconstructed by the C-PQDR (gray dashed line,) and the original signal (black solid line) are shown.

the reconstructed signal and the original signal defined in (10); and the percentage of retained energy (RTE(%)), defined in (11).

$$\mathrm{NMSE} = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|^2}{\|\mathbf{x}\|^2} \tag{9}$$

$$\mathrm{COR} = \frac{\mathbf{x^T} \cdot \hat{\mathbf{x}}}{\mathbf{x^T} \cdot \mathbf{x}} \tag{10}$$

$$\mathrm{RTE}(\%) = \frac{\sum_{n=0}^{N} x[n]^2}{\sum_{n=0}^{N} \hat{x}[n]^2} \tag{11}$$

where, $\mathbf{x}$ is the vector notation for the original signal $x\,[n]\,, \widehat{\mathbf{x}}$ is the vector notation of the reconstructed signal $\hat{x}\,[n]$ and N is the length of the signal.

A. Case 1

In this case, the test signal is composed of a sequence of different disturbances interleaved with a pure sinusoidal signal. The disturbances that were used are: Sags, Swells, Even and Odd Harmonics, Interharmonics, Notches, and ramp magnitude variations. A signal of approximately 13 min was generated.

To store the entire signal with a sampling rate of 7680 Hz (128 samples per cycle) with 16-bit resolution, a 12.15 MB file would be necessary. The size of the file compressed by the G-PQDR is is 102.29 kB, constituting a compression ratio of approximately 119: 1. The size of the compressed file generated by the C-PQDR for the same test signal is 254.34 kB, more than two times larger than the compressed file of the G-PQDR. Additionally, it is worth to mention that only 12% of the frames were detected as novelty frames, which contributed to the high level of compression ratio achieved.

The obtained $\mathrm{NMSE} = 0.0343, \mathrm{COR} = 0.991$ and $\mathrm{RTE}(\%) = 98.37\%$ demonstrate the considerable quality of the reconstructed signal.

The signals reconstructed by the G-PQDR and the C-PQDR are shown in Fig. 7, along with the original signal. This part corresponds to the moment where the sag disturbance ends. It can be noted that both the PQDRs performed considerably well, reconstructing the signal almost perfectly because it is difficult to distinguish the reconstructed signals from the original.

Fig. 8 shows the reconstruction of a signal segment with the interharmonics. In this case, it can be noted that there exists a reconstruction error for both the PQDRs; however, the C-PQDR performs better than the G-PQDR. This fact is evident in the edges of the signal, where the shape of the signal reconstructed by the G-PQDR does not match with the original signal. This is due to the fact that the current reconstruction method used in the G-PQDR has a 15 Hz frequency resolution because the frame length is four cycles. High frequency resolution methods could be used in an offline reconstruction algorithm to overcome this problem, without modifying the online detection methodology. This is currently under investigation.

B. Case 2

The test signal for Case 2 is a composition of 2 s of harmonic distortion (by even harmonics or odd harmonics) interleaved with a pure sinusoidal signal for 1 s. A 10 min signal was acquired and the total size needed to store the signal without compression would be 9.25 MB. The signal compressed by the G-PQDR resulted in a 16.22 kB file, constituting a compression ratio of approximately 570:1. The size of the file compressed by the C-PQDR was 36.00 kB.

In this case, the high compression ratio achieved is owing to the fact that the signal has only harmonic components with an almost stationary behavior. In this case, novelty frames are detected only in 1.15% of the frames. The frames, detected as novelties, are the frames near to the transition of a pure sinusoidal signal to the signal with harmonics.

The frames detected as novelties, are the frames close to the transition stage of the pure sinusoidal signal into the signal with harmonics. The quality of the reconstruction was slightly better than Case 1 because the $\mathrm{NMSE} = 0.0270$ is smaller and the $\mathrm{COR} = 0.992$ and $\mathrm{RTE}\,(\%) = 98.54\%$ are greater than the ones achieved for Case 1.

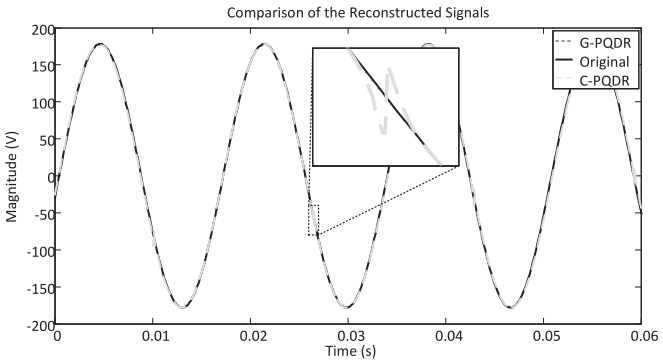

Fig. 9. Reconstructed signal with frequency variations. The signal reconstructed by the G-PQDR (black dotted line), the signal reconstructed by the C-PQDR (gray dashed line,) and the original signal (black solid line) are displayed. The zoom box highlights a reconstruction error related to the C-PQDR.

C. Case 3

The test signal in Case 3 is a pure sinusoidal signal with a ramp magnitude variation. The magnitude varies between 127 V and 180 V. To store the entire signal (10 min duration) without compression would require a 9.1 MB file. The file compressed by the G-PQDR has only 20.22 kB, achieving a compression ratio of 448: 1. The size of the file compressed by the C-PQDR is 22.67 kB, that is considerably close to the size of the file compressed by the G-PQDR.

The reconstruction quality metrics are: $\mathrm{NMSE} = 0.0277, \mathrm{COR} = 0.989$ and $\mathrm{RTE}\,(\%) = 97.80\%$, which is similar to Case 1 and slightly worse than Case 2.

D. Case 4

The test signal in Case 4 is a pure sinusoidal signal with a fundamental frequency variation. The fundamental frequency varies between 59 Hz and 61 Hz. The duration of the stored signal is 86 s and the file size without compression would be 1.32 MB. The G-PQDR compressed file has 28.34 kB, constituting a compression ratio of 47: 1. The size of the compressed file by the C-PQDR is 122.2 kB, that is significantly larger than the size of the file compressed by the G-PQDR. This illustrates an important advantage of the G-PQDR over the C-PQDR. Signals with minor frequency variations are highly compressed with a good quality by the proposed G-PQDR as a result of the G-PQDR methodology, based on the frequency estimation of each cycle and the SWRDFT for the signal reconstruction.

The reconstruction quality metrics are: $\mathrm{NMSE} = 0.07, \mathrm{COR} = 0.992$ and $\mathrm{RTE}\,(\%) = 98.85\%$, that are similar to Case 2. The NMSE is the worst among all the cases owing to the short duration of the test signal that introduces a higher uncertainty level.

The quality of the reconstruction can be seen in Fig. 9. Additionally, another advantage of the G-PQDR over the C-PQDR is highlighted in the zoom box of Fig. 9, where it is possible to see the discontinuities in the C-PQDR reconstructed signal that are not present in the original one.

TABLE I
SUMMARY OF THE COMPRESSION RESULTS

Signal	Case 1	Case 2	Case 3	Case 4
Original File Size	12.15 MB	9.25 MB	9.06 MB	1.32 MB
Compressed File Size G-PQDR	102.29 kB	16.22 kB	20.22 kB	28.34 kB
Compressed File Size C-PQDR	254.34 kB	36.00 kB	22.67 kB	122.2 kB
NMSE G-PQDR	0.0343	0.0270	0.0277	0.07
COR G-PQDR	0.991	0.992	0.989	0.992
RTE(%) G-PQDR	98.37	98.54	97.80	98.85

TABLE II
RESULTS FOR CASE 5

	Voltage	Current	Total
CR - G-PQDR	749:1	372:1	533:1
CR - C-PQDR	-	-	651:1
File Size - G-PQDR	9.8 MB	17.2 MB	27 MB
File Size - C-PQDR	-	-	22 MB
NMSE - G-PQDR	0.0093	0.0174	
COR - G-PQDR	0.996	0.990	
ENE(%) - G-PQDR	99.14	97.97	

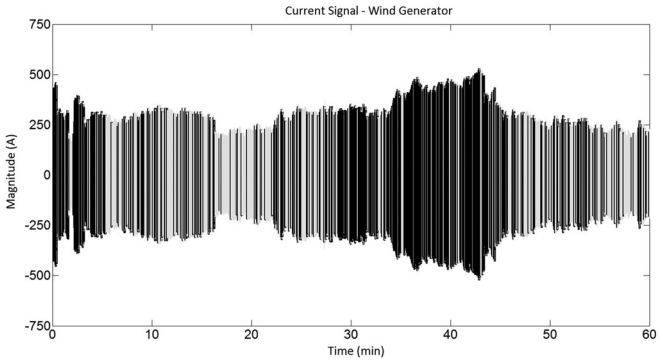

Fig. 10. Current signal durinf a period of one hour.

The four cases comprised the most common disturbances in electrical systems and Table I summarizes the achieved results.

E. Case 5: Real Signal

The signal used in Case 5 was acquired from a wind generator. There are eight acquisition channels, four for the voltages and four for the currents. The signal duration is 34 h. Table II shows the results of the compression and the quality obtained by the G-PQDR for the voltage and current channels, separately, and the total compression obtained by the C-PQDR.

Observing Table II, it can be noted from the quality metrics that the G-PQDR could reconstruct the signal with a good quality. The compression ratio obtained for the voltage is higher than the compression ratio for the current. This is owing to the fact that the current signals present more variations, i.e., more novelties are detected for the current signals. The current signal for a period of 1 h is shown in Fig. 10.

In Fig. 10 the black frames are the ones that were detected as novelty frames and the gray frames are the non-novelty frames.

It can be observed that several frames were detected as novelty frames during this interval, contributing to the reduction in the compression ratio, compared to the voltage signal.

On comparing the compression ratios, it can be observed that in this case the performance of the G-PQDR is quite close that of the CPQDR.

VI. Conclusion

This paper presented a G-PQDR based on the novelty detection concept. The methodology has three levels of data compression including two lossy compression levels (the novelty detection and the wavelet processing) and one lossless compression. Further, signals with slow variations in the fundamental frequency are not detected, thereby, the compression ratio increases. The proposed methodology is particularly suitable for the complex scenarios of smart grids because of its ability to detect any novelty in the power system signal. Moreover, as the method provides high compression ratios, PQ monitors based on this technique can be spread out along the network after the amount of data to be stored and transmitted are significantly reduced.

The developed prototype has been tested and its performance was compared with a C-PQDR. It was possible to verify that the G-PQDR outperformed the C-PQDR in all cases in terms of the compressed file size. Regarding the quality of the reconstruction, both presented similar performances, except for the case of the fundamental frequency variation, where the proposed method presented a better quality and the case of a signal with interharmonics, where the proposed method presented a poorer quality.

The next step in this work concerns the study, development, and application of a high frequency resolution reconstruction algorithm to increase the resolution of the signal reconstruction of the G-PQDR.

Acknowledgment

This work was supported in part by CNPq, CAPES and FAPEMIG (Brazilian research agencies).

References

[1] Y. Yang and M. H. Bollen, *Power Quality and Reliability in Distribution Networks With Increased Levels of Distributed Generation*, Stockholm, Sweden: Elforsk, 2008.

[2] M. Tcheou *et al.*, "The compression of electric signal waveforms for smart grids: State of the art and future trends," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 291–302, Jan. 2014.

[3] H. Maaß *et al.*, "Data processing of high-rate low-voltage distribution grid recordings for smart grid monitoring and analysis," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, pp. 1–21, 2015.

[4] H. Maass *et al.*, "First evaluation results using the new electrical data recorder for power grid analysis," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2384–2390, Sep. 2013.

[5] M. Moreto and J. G. Rolim, "Using phasor data records and sequence of events to automate the classification of disturbances of power generating units," *Elect. Power Syst. Res.*, vol. 81, no. 7, pp. 1266–1273, 2011.

[6] C. Qiong and W. Zhao-Hui, "Research and design of portable fault recorder based on fpga," *Energy Proc.*, vol. 12, pp. 686–692, 2011.

[7] K. Sayood, *Introduction to Data Compression*, Pittsburgh, PA, USA: Newnes, 2012.

[8] F. Lorio and F. Magnago, "Analysis of data compression methods for power quality events," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, 2004, pp. 504–509.

[9] *IEEE1159, IEEE Recommended Practice for the Transfer of Power Quality Data*IEEE Standard 1159.3-2003, 2004.

[10] D. Zhang, Y. Bi, and J. Zhao, "A new data compression algorithm for power quality online monitoring," presented at the *Int. Conf. Sustain. Power Gen. Supply*, Nanjing, China, 2009.

[11] M. P. Tcheou, A. L. Miranda, L. Lovisolo, E. A. da Silva, M. A. Rodrigues, and P. S. Diniz, "How far can one compress digital fault records? analysis of a matching pursuit-based algorithm," *Digital Signal Process.*, 2012.

[12] P. Pillay and A. Bhattacharjee, "Application of wavelets to model short-term power system disturbances," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 2031–2037, Nov. 1996.

[13] S. Santoso, E. Powers, and W. Grady, "Power quality disturbance data compression using wavelet transform methods," *IEEE Trans. Power Del.*, vol. 12, no. 3, pp. 1250–1257, Jul. 1997.

[14] C. Hsieh and S. Huang, "Disturbance data compression of a power system using the huffman coding approach with wavelet transform enhancement," in *Proc. Inst. Elect. Eng, Gen., Transm. Distrib.*, pp. 7–14, 2003.

[15] S.-J. Huang and M.-J. Jou, "Application of arithmetic coding for electric power disturbance data compression with wavelet packet enhancement," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1334–1341, Aug. 2004.

[16] G. Panda, P. Dash, A. Pradhan, and S. Meher, "Data compression of power quality events using the slantlet transform," *IEEE Trans. Power Del.*, vol. 17, no. 2, pp. 662–667, Apr. 2002.

[17] J. A. Quinn and C. K. Williams, "Known unknowns: Novelty detection in condition monitoring," *Pattern Recogn. Image Anal.*, New York, USA: Springer, 2007, pp. 1–6.

[18] L. Clifton *et al.*, "Identification of patient deterioration in vital-sign data using one-class support vector machines," in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, 2011, pp. 125–131.

[19] C. Surace and K. Worden, "Novelty detection in a changing environment: A negative selection approach," *Mechan. Syst. Signal Process.*, vol. 24, no. 4, pp. 1114–1128, 2010.

[20] C. P. Dieh *et al.*, "Real-time object classification and novelty detection for collaborative video surveillance," in *Proc. Int. Joint Conf. Neural Netw.*, 2002, vol. 3, pp. 2620–2625.

[21] M. Markou and S. Singh, "A neural network-based novelty detector for image sequence analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1664–1677, Oct. 2006.

[22] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Commun. Surveys Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[23] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Process.*, vol. 99, pp. 215–249, 2014.

[24] P. Nisenblat, A. M. Broshi, and O. Efrati, *Method of compressing values of a monitored electrical power signal* Patent 7,415,370 U.S.

[25] C. A. Duque, M. V. Ribeiro, F. R. Ramos, and J. Szczupak, "Power quality event detection based on the divide and conquer principle and innovation concept," *IEEE Trans. Power Del.*, vol. 20, no. 4, pp. 2361–2369, Oct. 2005.

[26] H. Azami, H. Hassanpour, J. Escudero, and S. Sanei, "An intelligent approach for variable size segmentation of non-stationary signals," *J. Adv. Res.*, 2014.

[27] M. H. Bollen and I. Gu, *Signal Processing of Power Quality Disturbances*, Hoboken, NJ, USA: Wiley, 2006, vol. 30.

[28] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[29] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

[30] E. Y. Hamid and Z.-I. Kawasaki, "Wavelet-based data compression of power system disturbances using the minimum description length criterion," *IEEE Trans. Power Del.*, vol. 17, no. 2, pp. 460–466, Apr. 2002.

[31] G. Jie, "Wavelet threshold de-noising of power quality signals," in *Proc. 5th Int. Conf. Natural Comput.*, Aug. 2009, vol. 6, pp. 591–597.

[32] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[33] P. Z. Peebles, J. Read, and P. Read, *Probability, Random Variables, and Random Signal Principles*, New York, USA: McGraw-Hill, 2001, vol. 3.

[34] M. H. Bollen, *Understanding Power Quality Problems*, Piscataway, NJ, USA: IEEE, 2000, vol. 3.

[35] S. da Silva, R. Novochadlo, and R. Modesto, "Single-phase pll structure using modified p-q theory for utility connected systems," in *Proc. IEEE Power Electron. Specialists Conf.*, Jun. 2008, pp. 4706–4711.
[36] M. Ciobotaru, R. Teodorescu, and F. Blaabjerg, "A new single-phase pll structure based on second order generalized integrator," in *Proc. 37th IEEE Power Electron. Special. Conf.*, Jun. 2006, pp. 1–6.
[37] U. Nandi and J. Mandal, "A compression technique based on optimality of lzw code (olzw)," in *Proc. 3rd Int. Conf. Comput. Commun. Technol.*, 2012, pp. 166–170.
[38] D. Fabri, C. Martins, L. Silva, C. Duque, P. Ribeiro, and A. Cerqueira, "Time-varying harmonic analyzer prototype," in *Proc. 14th Int. Conf. Harmon. Quality Power*, Sep. 2010, pp. 1–7.
[39] ALTERA, *Cyclone IV Device Handbook*, 2010.
[40] Texas Instruments, *Tiva TM4C123GXL microcontroller datasheet*, 2013.
[41] R. Kamal, *Microcontrollers: Architecture, Programming, Interfacing and System Design*, Princeton, NJ, USA: Pearson, 2011.
[42] Flex, *Version 2.5: A fast scanner generator.*https://www.gnu.org/software/flex/
[43] Bison, *The Yacc-compatible parser generator.*http://www.gnu.org/software/bison/

L. R. M. Silva, photograph and biography not available at the time of publication.

E. B. Kapisch, photograph and biography not available at the time of publication.

C. H. N. Martins, photograph and biography not available at the time of publication.

L. M. A. Filho, photograph and biography not available at the time of publication.

A. S. Cerqueira, photograph and biography not available at the time of publication.

C. A. Duque (SM'14), photograph and biography not available at the time of publication.

P. F. Ribeiro (F'03), photograph and biography not available at the time of publication.