A Compression Method for Power Quality Data

R. E. Dapper[3], C. D. P. Crovato[2], A. A. Susin[3], S. Bampi[1]

[1] Informatics Institute – Microelectronics Program (PGMicro)
Federal University of Rio Grande do Sul (UFRGS)
Campus do Vale – Av. Bento Gonçalves, 9500, Porto Alegre (Brazil)
Phone/Fax number:+55 (51) 3308-6156 / +55 (51) 3308-7308, e-mail: **{lbsoares, bampi}@inf.ufrgs.br,**

[2] Feevale University
ERS239, 2755, Novo Hamburgo – CEP: 93352-000 (Brazil)
Phone number:+55 (51) 3358-4016, e-mail: **cesarc@feevale.br**

[3] Electrical Engineering department
Federal University of Rio Grande do Sul (UFRGS)
Av. Osvaldo Aranha, 103, Porto Alegre (Brazil)
Phone number:+55 (51) 3308-3515, e-mail: **{roque.dapper, altamiro}@ufrgs.br**

Abstract. This paper presents a compression method for power quality (PQ) signals. This method consists of polynomial approximation algorithm and a lossless compression algorithm The proposed method uses the Deflate algorithm as lossless compression algorithm. The polynomial approximation is intended to decrease the entropy of the signal, thus increasing the compression ratio of the Deflate algorithm. The proposed method is characterized by high compression ratio and low computational cost, making feasible its implementation in embedded systems.

Key words

Power Quality, Power Systems, Deflate, LZ77, Huffman Coding, Signal Compression.

1. Introduction

Due to evolution of the electronic technology, electronic devices with switching power supplies, rectifiers and inverters, and non-linear loads are becoming more present in industries and residences [1]. These devices have contributed to the increase of power quality disturbances, for example, increase of harmonic distortion present in the power network[1]. Beside affecting the PQ parameters these electronic loads are very sensitive to disturbances in the power network [2, 3].

Power quality problems can cause malfunctions in sensitive equipments, production stoppages or poor operation of protection systems [1]. In order to monitor these disturbances more and more power quality analyzers are being distributed in different points of the power network to monitor and analyze the parameters of power quality [1,4].

The power quality parameters are defined by international standards, such as IEC 61000-4-30. However regional standards are also edited to complement and add new parameters and characteristics necessary for certain region [4].

The IEC 61000-4-30 standard defines a list of power quality parameters to be evaluated, as well as their methods of calculation. These parameters are mostly evaluated every period of approximately 200mS. However other aggregation times are defined, for example, 3 seconds and 10 minutes [5]. The measurement campaigns can last a few hours in small industries up to weeks or months in large distribution or transmission systems.

Many power quality analyzers do not have processing capacity to analyze and store all power quality parameters at the same time. However, with the increasing processing capacity, more and more devices have the ability to simultaneously process and store all power quality parameters required by the standards. However, to achieve long periods of measuring, even with the increasing size of memory available, it is necessary to compress the information saved in memory.

Currently, most researches and papers are focused on the compression of short sampling periods of high frequency transient signals or even in the compression of the power network oscillography. In the research weren't found any academic reference about the compression of power quality parameters, only some incomplete methods and/or patent protected methods implemented in totally different hardware targets, sampling rates, resolutions, etc, making any comparison extremely hard or in some cases no sense.

This paper aims to present a compression method of power quality parameters calculated as defined by IEC 61000-4-30. The proposed method is characterized by high compression ratio and low computational cost, making feasible its implementation in embedded systems.

2. Signal Characteristics

Signals as rms voltage or current, among other measures defined by IEC 61000-4-30, are typically stochastic signals. These signals can be represented as mean value added to a low amplitude non stationary signal. The characteristics of this non-stationary signal are very diverse. The frequency band of the signal may be flat, thus approaching a white noise, may have the behavior of a band-limited signal and may exhibit high or low dispersion. These characteristics are dependent on the accuracy of the measuring equipment that generated the measures, and depends on the characteristics of the electric system. The Fig. 1, illustrates an rms voltage measure.

According to Fig. 1, once removed the mean value, remains a non-stationary stochastic signal. Compress signals in PQ not necessarily requires a lossless compression. However, the intrinsic error of the equipment added to the error inserted by the compression process may not exceed the limits set for each power quality parameter, as defined in [5]. So the method to be defined should be able to lossless compress the signals or with insignificant losses when such signals admit such losses. The point of admitting losses in a measurement is typically associated with the issue of decreasing the entropy of the signal, once low entropy signal is well compressed by lossless compression methods such as Huffman, LZ77/78, among others.

The signal compression will be evaluated by three parameters, compression rate, absolute maximum error and rms error. The compression rate is defined as
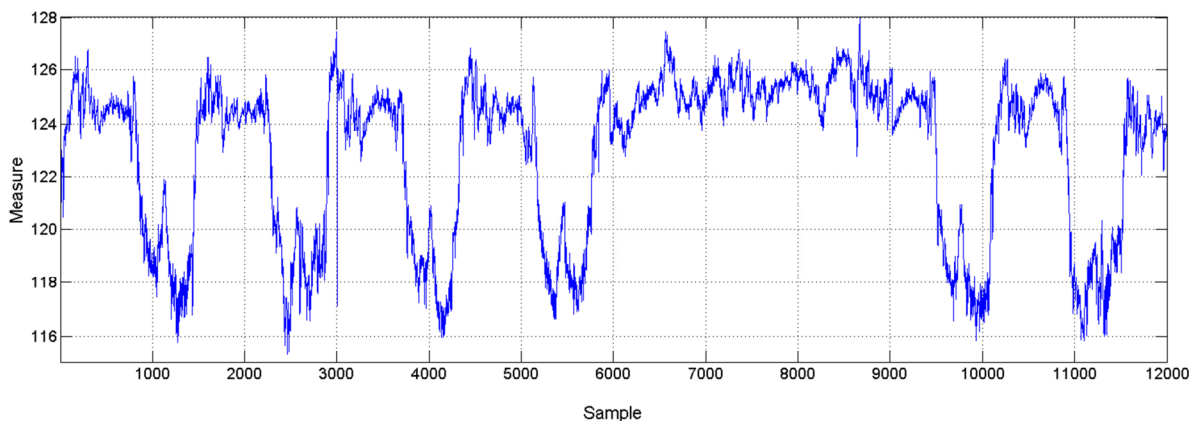
$$Cr = \left(\frac{Sc}{S}\right) \cdot 100\%..$$ (1)

where S_c represents the size in bytes of the compressed signal and S represents the size in bytes of the original signal. The rms error is defined as (2).

$$E_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N}\left(S_{uncompress}(i) - S(i)\right)^2}$$ (2)

3. Proposed Algorithm

The proposed algorithm is based on a polynomial approximation followed by a lossless compression method. The algorithm was originally developed for use with data represented in floating point, but can be used in systems represented in fixed point.

The first step of the algorithm is given by grouping N consecutive measurements of a given magnitude, for example, rms voltage from channel A calculated every 200mS. This set of N points is approximated by a straight line by linear regression. Through this line is calculated a vector of differences, which consists of the difference between each point in the original vector and the corresponding point on the approximate straight line. This process is illustrated in Fig. 2.The linear regression is very exhaustive from the computational point of view, thus, the smaller is the vector size lower is the computation time. Simulations showed that vectors of 64 positions offered a low computational cost, as well as produced good approximations.

According to Fig. 2a, a block of 64 points is approximated by a line, as illustrated by the red line. A vector of differences, as illustrated in Fig. 2b, is calculated by the difference between each point from the vector and the line. This vector of differences is then quantized.
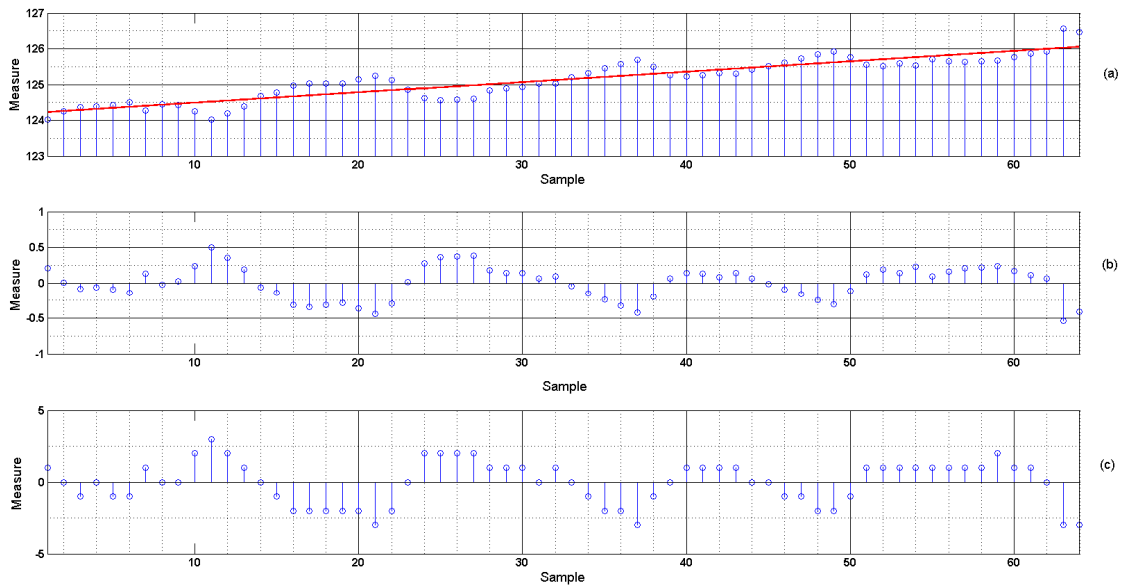

Fig. 1. Rms voltage measure.

Fig. 2. Polynomial approximation.

The process of quantization generates loss of signal information, therefore this introduced error should be less than the precision required by the signal. The quantization is an important step in the compression process. If this process uses many bits to represent a signal, the compression rate will be very small. However, using few bits to represent the signal, the compression ratio will be too high, but the information will be degraded, not reaching the accuracy required. As parameters for the quantization process were used the reference value of the measure and accuracy of reference defined by IEC 61000-4-30 for the given measure. The quantization is defined in (3).

$$d_q = \text{round}\left(\frac{d[i]}{V_{ref}} \cdot \frac{P}{100\%}\right) \qquad (3)$$

The variable d represents the difference between each point measured and their approximation, while V_{ref} is defined as the reference value for the measure that is being approximated. The precision required for the measurement variable is represented by P in (3).

The Fig. 2c illustrates the quantization of the vector of differences for the voltage vector shown in Fig. 2b. In this sample was used as reference value, as defined in [5], U_{din}, which in this example has a reference value of 127V and required accuracy of 0.1%. This precision is required for a device classified as Class A. The quantized data will be stored into a fixed point variable of 16bits, while the coefficients of approximation of the line will be represented in single-precision floating point.

The next step of the compression process is given by the compression of vector of differences and their coefficients of approximation. As lossless compression method, was chosen deflate. This method was chosen because it is simple to implement and because provides good compression ratios in different situations. Deflate is a lossless data compression algorithm that combines the LZ77 algorithm and Huffman coding [6,7].

The LZ77 algorithm is based on the use of blocks that have been already read as a dictionary, replacing subsequent occurrences of the same blocks by the position (absolute or relative) of its last occurrence. To limit the search distance, previous occurrences are limited by a "sliding window" that have fixed size and "slides" over the vector, delimiting the beginning and end area where the repetitions will be searched. The window size is the primary factor to adjust the performance of this algorithm. Due to memory limitations of embedded devices, it was chosen compress small memory buffers of 16kbytes, thus there is no need for the use of the sliding window. To mount a buffer of 16kbytes, 128 vectors of differences are grouped, each containing 64 points represented in fixed point of 16 bits.

Once the vector of differences are represented with fixed size, 16 bits, in many approximations are needed few bits to represent the differences, thus the most significant bits of most points will be zero.

The compression using deflate is better as lower is the entropy. With this purpose it is proposed a transformation in this signal in order to decrease the entropy. Table I represents the vector of differences as a matrix where the columns represent the points in the vector of differences and the lines represent the bits of each point represented in fixed point.

Table I – Vector of differences represented as a matrix

P$_0$	P$_1$	...	...	P$_{n-1}$	P$_n$
b15	b15	b15	b15	b15	b15
b14	b14	b14	b14	b14	b14
b13	b13	b13	b13	b13	b13
...	...	...	...	...	...
b3	b3	b3	b3	b3	b3
b2	b2	b2	b2	b2	b2
b1	b1	b1	b1	b1	b1
b0	b0	b0	b0	b0	b0

Regardless of the original data representation, the deflate treats data as 8 bit words. Consequently, data must be transferred in words of 8 bits. However, the bytes are formed grouping bits of same weight, as shown by arrows in Table I. Starting by the most significant bits, long strings of 0 will be formed in the most significant bits, as well as long strings of 1 will be formed in least significant bits, thus decreasing the entropy of the signal.

To be true this behavior the fixed point representation of the points cannot be coded in 2's complement, because in 2's complement negative number are like number with large amplitude, which would be mixed with small amplitude points, thus breaking the long strings of 0 in the most significant bits. Therefore, the best way to represent this is through bit encoding signal, where the most significant bit represents the sign and the remaining bits represent the amplitude of the signal.

4. Results

All power quality parameters defined in IEC 61000-4-30 can be compressed with this algorithm, but each parameter has its required accuracy defined by the standard. The rms voltage is the parameter which requires the most accurate measure. For a Class A equipment the accepted error is just 0.1% of Udin for the rms voltage[5]. The results are very dependent on the signal characteristics and the required accuracy. The compression ratios are better when the matrix of differences has low entropy. This lower entropy is achieved when the signal has low dispersion and the quantization process requires few bits to represent the signal.

Although this method depends on the nature of the signal, the compression ratio of rms voltage for real measurements achieved ratios between 7 and 15% if considered accuracy of 0.1%. But for a required accuracy of 0.5%, as required for a Class S equipment, the achieved compression ratios are around 4% an 10%.

4. Conclusion

This algorithm has as main characteristics the low computational cost and high compression rate. Despite the simplicity of implementation of the algorithm, compression rates are quite significant. Although this algorithm is a lossy compression algorithm, the losses can be controlled, which makes this algorithm a very interesting alternative for compression of power quality data. The losses can be as small as necessary.

The algorithm requires a floating point MAC for the linear approximation module and 16kbytes of memory for the Deflate. This implementation in embedded systems or in FPGA is perfectly viable, even the smallest FPGAs available in the market.

Acknowledgement

References

[1] Lin Lin et al, "Review of Power Quality Signal Compression Based on Wavelet Theory", International Conference on Test and Measurement, 2009, vol. 1, pp. 235–238, Dez. 2009.

[2] BOLLEN, Math H. J.; GU, Irene Y. H. *Signal Processing of Power Quality Disturbances*. New York: Wiley-IEEE Press, 2006, 861p. ISBN 0471731684.

[3] M. V. Ribeiro, C. A. Duque, and J. M. T. Romano, "An improved method for signal processing and compression in power quality evaluation," *IEEE Trans. Power Del.*, vol. 19, no. 2, pp. 464–471, Apr. 2004.

[4] Bubla, Viktor; Kraus, Jan; Tobiska,Tomas. "Loosless Encoding and Compression Algorithms aplied on power quality datasets", International Conference and Exhibition on Electricity Distribution, 2009, pp. 1 - 4 , Jun. 2009.

[5] IEC 61000-4-30 – International Standard – Electromagnetic compatibility (EMC) – Part 4-30: Power Quality Measurement Methods

[6] Huan, Wang; Huijiang, Guo; Jingbao, Fan; Wu Weimin; Yi, Hu. "Improvable Deflate Algorithm, IEEE Conference on Industrial Electronics and Applications, 2008, vol. 1, pp. 1572 - 1574 , Jun. 2008.

[7] Jalaleddine, S.M.S.; Hutchens, C.G. ; Strattan, R.D. ; Coberly, W.A. "ECG data compression techniques-a unified approach. Biomedical Engineering", IEEE Transactions, vol. 37, pp 329-343, 1990.