Digital Signal Processing ••• (••••) •••–•••

Contents lists available at SciVerse ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

How far can one compress digital fault records? Analysis of a matching pursuit-based algorithm

Michel P. Tcheou [a,b,*], André L.L. Miranda [b], Lisandro Lovisolo [c], Eduardo A.B. da Silva [a], Marco A.M. Rodrigues [b], Paulo S.R. Diniz [a]

[a] LPS/PEE/COPPE/Poli – UFRJ, Cx.P. 68504, Rio de Janeiro, RJ, 21941-972, Brazil
[b] CEPEL, Cx.P. 68007, Rio de Janeiro, RJ, 21941-911, Brazil
[c] PROSAICO/PEL/DETEL/FEN – UERJ, Rio de Janeiro, RJ, 20550-900, Brazil

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Signal compression
Electric disturbance signals
Fault analysis
Matching pursuit
Damped sinusoids

ABSTRACT

In this paper, we verify how far electric disturbance signals can be compressed without compromising the analysis of encoded fault records. A recently proposed compression algorithm, referred to as Damped Sinusoidal Matching Pursuit (DSMP) has the remarkable feature of obtaining both compact and physically interpretable representations. However, for fault analysis applications, one is primarily interested in how accurate can be the analysis performed on compressed signals, instead of evaluating mean-squared error figures. Unlike previous works in digital fault records compression, the performance of the DSMP compression method is evaluated using a protocol based on fault analysis procedures commonly performed by expert engineers. This protocol is applied for comparing the results obtained in the analysis of both uncompressed records and their compressed versions at different compression ratios. The results show that the DSMP is a reliable compression system since it achieves high compression ratios (6.4:1) without causing fault analysis misinterpretation.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Digital Fault Recorders (DFRs) are being widely used in modern power systems in order to better understand features related to faults and power quality analysis. DFRs provide sampled versions of current and voltage signals, referred to as oscillograms, usually ranging from hundreds to few thousands of samples per second, usually stored as 16-bit wide numbers. Most DFRs are installed in substations, usually far from central offices of utilities. The transport of DFR data to a fault analyst computer can be performed in a variety of ways: using digital media as tapes, CDs or DVDs, a modem over a telephone line (wired or microwave), a local or a wide area network. Even with high capacity communications technologies (like optical fibers or satellite links), the capacities of the channels are shared with other functions of major importance, such as protection, supervision and control, and also with other services of economical relevance, such as band allocation for third parties. Thus, data compression can be useful to reduce transmission time and bandwidth requirements.

A compression technique suitable for DFR files is highly desirable as long as it has two important features: providing high compression ratios in order to make its use worthwhile and also introducing low distortion in the acquired waveforms in order to avoid misinterpretation of the recorded phenomena. Many compression frameworks for power system signals can be found in literature; most of them use wavelet-based techniques and variants [1–5]. One important reason for this is that wavelet analysis techniques have properties that are particularly suited for the analysis of transients. Nonetheless, wavelets provide poor representation of components of specific frequency that are not well-localized in time such as the fundamental and harmonic components. In these cases, it is necessary to adopt other techniques for estimating their related parameters such as amplitude, frequency and phase [3]. Other compression approaches based on compression techniques originally developed for other signals have been also tried out for compressing electric signals. For example, in [6] electric signals are converted into 2D-representations (images) and compressed using standard image compression techniques.

In [7–9], the Damped Sinusoidal Matching Pursuit (DSMP) compression method was introduced. It is a lossy compression algorithm that uses an analysis-by-synthesis approach. That is, the DSMP first analyzes the signal into elementary waveforms that are relevant to the electric disturbance analysis process; then the parameters of these waveforms are encoded. The decoder simply synthesizes the signal from the encoded parameters. The

* Corresponding author at: LPS/PEE/COPPE/Poli – UFRJ, Cx.P. 68504, Rio de Janeiro, RJ, 21941-972, Brazil.
E-mail addresses: pompeu@lps.ufrj.br (M.P. Tcheou), miranda@cepel.br (A.L.L. Miranda), lisandro@uerj.br (L. Lovisolo), eduardo@lps.ufrj.br (E.A.B. da Silva), mamr@cepel.br (M.A.M. Rodrigues), diniz@lps.ufrj.br (P.S.R. Diniz).

1051-2004/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.dsp.2011.11.003

analysis-by-synthesis procedure is carried out by using an adaptive decomposition algorithm based on the Matching Pursuit (MP) algorithm [10,11]. This compression method presents rate-distortion performance comparable to the state-of-the-art wavelet-based methods such as the one in [3]. For instance, in [3] it is reported that for 1 bit/sample the achieved MSE lies around −30 dB. In [8,9] it is verified that DSMP accomplishes nearby distortion levels, around −45 to −35 dB at the same rate.

Due to the adaptability of DSMP (as described in Section 2), it enables for some flexibility in the signal representation, in contrast to other decomposition approaches such as Fourier and wavelets transforms. DSMP provides the representation of voltage or current signals as a linear combination of damped sinusoidal components which may be well localized. Such components include dumped or un-dumped fundamental and harmonics that are well fitted for modeling stationary signals as well as transients and discontinuities. In essence, these components are highly correlated with power systems phenomena. A remarkable feature of DSMP consists of its ability to obtain electric disturbance signals representations that are both compact and physically interpretable. This provides efficient ways to de-noise and filter disturbance signals [8]. Another important feature that makes DSMP to be different from other compression methods is the use of a rate-distortion scheme, proposed in [9], wherein the parameters of the damped sinusoidal components are encoded regarding the best signal reproduction given a desired compression level.

The main use of DFRs is for fault and power quality analysis [23]. Therefore, when it comes to compressing disturbance waveforms, it is of paramount importance to verify the loss of information introduced by compression. In this work it is investigated how much compression is possible before leading to fault misinterpretation by an expert engineer or an expert system. It is important to point out that, taking this into consideration, the requirements for the compression of disturbance waveforms present characteristics quite different from the ones used in compression systems of various signals, as for example image and sound. When evaluating the performance of sound and image compression schemes, one is interested in the amount of information that can be eliminated from the original signal such that the ear/brain or eye/brain processing does not perceive the missing information [12,13]. Oscillograms require a different approach since they are often processed using specialized algorithms in order to extract parameters that are used to interpret and analyze oscillograms. Therefore, one must evaluate how much information has to be retained.

Considering this scenario, the main purpose of this paper is to evaluate the performance of the DSMP compression method by using traditional and proven tools employed by expert engineers or specialists for fault analysis. We have used a commercial automated fault analysis tool that was developed at CEPEL (Centro de Pesquisas de Energia Elétrica, a power systems research center in Brazil) [14,15]. It was designed for the analysis of faults in high voltage transmission lines, and it is based on the experience on disturbance analysis of protection relay engineers. This tool comprises traditional algorithms already proposed in literature and widely used by power systems experts [16–19]. These algorithms are used here to verify the maximum compression level DSMP can achieve without compromising the analysis results of compressed fault records.

Usually the analysis of fault records encompasses fault detection, fault classification, fault location, event analysis, and operation analysis of the protection system and circuit breakers [23]. The strategy for evaluating the DSMP compression method employed in this work employs only a few metrics or figures of merit: the instant of fault inception, the phasor values measured from the waveforms, the fault type and the fault location. These are reviewed in Section 3. One should note the role of the “expert” or “specialist” which is responsible for evaluating the fault records employing these figures of merit. He employs the automated system for computing the figures of merit or metrics, which are in turn used in the analysis of the fault records for extracting relevant information about the fault.

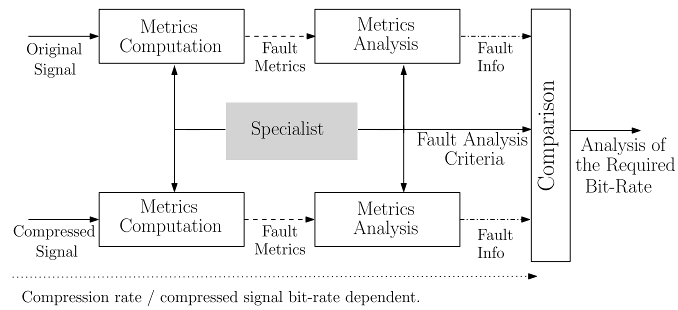


Fig. 1. Methodology employed for the assessment of the performance of the compression method.

An important result of fault analysis procedures for transmission line is the dispatch of a maintenance team to the correct location. Therefore, the accuracy of the information extracted from the signal analysis known as “fault location” is one of the most relevant factors for analyzing the impact of the compression of electric signals. Nonetheless, for locating the fault it is necessary to first estimate the fault inception and then, from this metric, the phasors during the fault and previous to it are computed. From these phasors one classifies the fault, i.e., guesses the system phases involved in the fault. Then, using the fault classification information and the computed phasors values the fault location is estimated in terms of percentage of the length of the transmission line.

In this work, these fault analysis procedures comprise the evaluation protocol for the DSMP compression method. Fig. 1 summarizes this protocol – the methodology employed for the assessment of the DSMP “rate-usability” performance. The objective is to assess at which extent electric signals can be compressed without compromising the fault analysis procedures. In other words, one wants to determine what is the minimum rate required for compressing electric signals using the DSMP so that their analysis is not jeopardized.

The experiments for evaluating the DSMP compression method are held using a blind protocol. The specialist is not aware of which fault records are compressed or uncompressed. The specialist employs the automated system for analyzing the original signals and some compressed versions of them, computing the relevant metrics and extracting the relevant information on fault classification and location. The analysis results for the signals are then compared using error criteria that are acceptable from the specialist’s viewpoint. These criteria are applied to compare the analyses of the fault record for different compression levels and, therefore, different distortions. These comparisons aim at verifying the maximum acceptable compression ratio or equivalently the minimum acceptable transmission/storage rate (in bits per sample – bps) that does not compromise the fault records analyses when the DSMP is used for compressing them.

The analysis results of original files of two different power systems and their respective compressed versions at different compression ratios are presented in Section 4. The original files are a set of 18 oscillograms acquired by digital fault recorders of the Brazilian power system grid, each one containing 6 to 8 analog channels (voltage and current in 3 phase circuits and grounds). The waveforms are originally stored in IEEE COMTRADE format [25]. The utility in which the data is collected has 138 kV transmission lines and operates with multiple parallel lines and with a large number of generation plants and loads. The set of oscillograms is

Please cite this article in press as: M.P. Tcheou et al., How far can one compress digital fault records? Analysis of a matching pursuit-based algorithm, Digital Signal Process. (2011), doi:10.1016/j.dsp.2011.11.003

composed by disturbances in distinct situations such as voltage sags, current swells, harmonics, transients and circuit switching. The sampling frequencies in this set of files are 1200, 2000, 2040 and 3840 Hz, and the length of the waveforms varies from 378 to 1991 samples.

The analysis results presented in Section 4 are compared in order for determining from which compression ratios the use of the automated system starts to fail in providing the same results for both the uncompressed and compressed oscillograms. These results allow adjusting the parameters of the DSMP compression technique in order to obtain the largest possible compression ratio that does not cause any degradation in the performance of the fault analyzers. In addition, in Section 5 we present a comparison of the DSMP to other compression methods for electric signals.

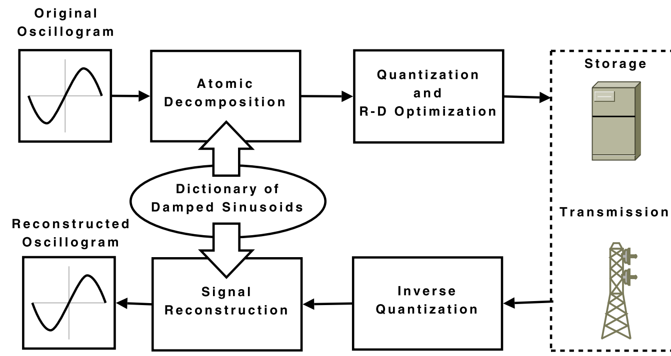


Fig. 2. General overview of the compression process of electric disturbance signals.

2. The DSMP compression method

This section describes the Damped Sinusoidal Matching Pursuit (DSMP) [7,9,8], the compression method for voltage and current signals in oscillograms, which is analyzed in this work.

2.1. Faults and power quality phenomena

DFR files contain signals corresponding to snapshots of the evolution in time of power-system disturbances [14]. This evolution is generally characterized by sinusoidal oscillations of varying amplitudes. Sometimes these oscillations may have sustained waveform distortions which are highly influenced by circuit switching, as well as by nonlinear behavior of some equipments [20].

In fault analysis tasks, the knowledge of phase quantities prior to the fault may reveal conditions leading to power system degradation. Generally speaking, these Power Quality (PQ) issues may impact transmission and distribution, jeopardizing consumers' equipments and devices. From a power distribution point of view, the most common phenomena affecting PQ are harmonics, transients, sags and swells [20]. In transmission line fault analysis, the PQ state of the system is important because it may explain the behavior of the protection system, which may require further adjustments to avoid unwanted triggering of protection actions induced by these phenomena.

2.1.1. A model for electric disturbance signals using damped sinusoids

In order to achieve good compression performance for electric disturbance signals, it is crucial to use a model capable of precisely representing the intrinsic phenomena existing in those signals [8]. In this work, an electric disturbance signal $x(t)$ is represented by means of a sum of damped sinusoids which may be well localized in time, that is [7]

$$x(t) = \sum_{m=1}^{M} \alpha_m e^{-\rho_m (t-t_m^s)} \cos(2\pi k_m F t + \phi_m) \times \left[u\left(t - t_m^s\right) - u\left(t - t_m^e\right)\right], \quad (1)$$

where M is the number of elements employed, F is the fundamental frequency (50/60 Hz), $u(\cdot)$ corresponds to the unit step function; each component is represented by a 6-tuple $(\alpha_m, k_m, \rho_m, \phi_m, t_m^s, t_m^e)$. In this 6-tuple, α_m is the amplitude, k_m is an integer multiple of the fundamental frequency, ρ_m is the decaying factor, ϕ_m is the phase, and t_m^s and t_m^e are respectively the starting and ending times of the mth signal component.

2.1.2. The DSMP compression system

Fig. 2 illustrates the compression process. First, each signal of the oscillogram is decomposed into a linear combination of damped sinusoidal components using an adaptive decomposition algorithm. This algorithm selects from a dictionary of damped sinusoids a subset of elements, also called atoms or structures, which are highly correlated with the signal patterns. Defining $\mathbf{x}$ as the signal, one obtains the M-term approximation of the signal:

$$\mathbf{x} \approx \hat{\mathbf{x}} = \sum_{m=1}^{M} \alpha_m \mathbf{g}_{\gamma(m)}, \quad (2)$$

where each component is characterized by its corresponding coefficient α_m and a set of parameters $\gamma(m)$ defining the damped sinusoid $\mathbf{g}_{\gamma(m)}$. The sequence of pairs $(\alpha_m, \gamma(m))$, $m = 1, 2, \ldots, M$, forms the structure book. For compression, the coefficients and the parameters of the structure book are quantized generating a bit-stream that is transmitted or stored. Bit-allocation is performed among the parameters through a rate-distortion optimization scheme [9]. After decoding, since there is loss of information due to quantization, one recovers an approximation of the encoded structure book by de-quantizing the bit-stream. This way, the quantization is performed over the parameters and not directly on the waveforms samples. Finally, the signal is reconstructed through the weighted sum of the elements of the recovered structure book, as in Eq. (2). Note that larger quantization errors in the structure book lead to larger distortion in the reconstructed signal.

2.2. Decomposition algorithm

The adaptive decomposition algorithm inside DSMP used to obtain signal representations according to Eq. (1) is based on the Matching Pursuit (MP) algorithm [10,11]. The MP successively approximates a signal by selecting the dictionary atom leading to the best possible approximation in each of its iterations. A set of heuristics is introduced in the decomposition loop of the MP core algorithm. These have the purpose of improving the matching between the chosen atom in the iteration and the intrinsic patterns in the residual signal. Details can be found in [8,7]. Other decomposition methods for electric signals do exist using a similar signal model for electric signals to the one proposed in [7]. One example is the one presented in [26] which employs a quantum-inspired evolutionary algorithm for speeding up the decomposition process.

2.3. Compression

For compression, the coefficients and atoms parameters are quantized after signal decomposition by using uniform scalar quantizers as described in [21]. The quantizers employed for each parameter is defined by its dynamic range and number of levels [9]. Rate-Distortion (R-D) optimization is carried out to achieve the best signal reproduction for a desired compression rate target [22]. In the presented compression framework, one uses an R-D algorithm that trades-off the number of atoms in the signal representation and the amount of bits allocated among the coefficients α_m and atom parameters $\gamma(m)$. It basically consists of minimizing the mean square error (MSE) between the original and the reconstructed signal for a desired bit-rate – the number of bits per sample; for details see [9].

3. Automated fault analysis system

In this work, we employ an automated analysis system to extract fault characteristics, which are in turn used to evaluate the compression capabilities of the DSMP. The algorithms employed for that analysis are adaptations of traditional tools broadly used by power systems experts [16–19]. In addition, it must be emphasized that this work focuses on data acquired from the monitoring of transmission lines. Thus, the fault characteristics considered are: (i) the instant of fault inception, (ii) the phasors values measured some cycles before and after the fault, (iii) the type of the fault, and (iv) the distance of the fault along the transmission line.

The procedures of the automated system used to obtain these fault characteristics are described in the subsequent sections. There is an interconnection among them. The fault inception time is first detected. Next, phasors values are computed some cycles before and after the fault inception time. Then, based on the computed phasors, fault classification is performed. Finally, the fault distance is calculated with respect to the phasors and the type of fault. It is essential to remark that the analysis system adopted here has its parameters empirically tuned considering a large set of real oscillograms acquired by monitoring different transmission lines with voltages ranging from 138 kV to 375 kV in the Brazilian power system. These signals form a different database than the one used for evaluating DSMP in the presented work. Furthermore, the analysis system uses heuristics when it is needed to make the analysis results less sensitive to the line conditions or to power system variations. Therefore, this fault analysis system was designed to present good performance (although not necessarily optimal) for very distinct line conditions.

3.1. Computation of the phasors of voltage and current signals

The steady state voltage or current waveforms in AC power systems are usually represented by phasors. A phasor represents a sinusoidal waveform of frequency ω_0 and possibly time varying magnitude $|v(t)|$ and angle $\phi_v(t)$; in the particular case of

$$v(t) = A\cos(\omega_0 t + \varphi_0), \tag{3}$$

one has $|v(t)| = A$ and $\phi_v(t) = \varphi_0$. Phasors are usually measured at the fundamental frequency, and are fundamental for the analysis of power systems protection and control behavior [24].

Power system disturbances may introduce system dynamic oscillations, harmonics, transients, this means that voltage and current waveforms cannot in this case be modeled by Eq. (3), that is, the amplitude A and the phase φ_0 may be varying in time. There are several techniques targeted at obtaining, in such cases, an acceptable representation of these voltage and current waveforms in phasor form [16]. The most widely used is based on the Fourier filter, which filters the fundamental component before phasor computation [19]. It uses rectangular cosine and sine sliding windows to calculate the phasor values at each signal sample k, as described below for a one-cycle Fourier filter implementation:

$$Y_C(k) = \frac{2}{N_{sc}} \sum_{l=0}^{N_{sc}-1} v(k + l - N_{sc} + 1) \cdot \cos\left(\frac{2\pi}{N_{sc}} l\right),$$

$$Y_S(k) = \frac{2}{N_{sc}} \sum_{l=0}^{N_{sc}-1} v(k + l - N_{sc} + 1) \cdot \sin\left(\frac{2\pi}{N_{sc}} l\right),$$

$$\left|v(k)\right| = \sqrt{Y_C^2(k) + Y_S^2(k)}, \qquad \phi_v(k) = -\arctan\frac{Y_S(k)}{Y_C(k)}, \tag{4}$$

where N_{sc} denotes the number of samples per cycle of the fundamental frequency. N_{sc} can be written as the ratio $N_{sc} = \frac{F_s}{F}$, where F_s and F are the sampling and fundamental frequencies respectively. Details about the half-cycle and two-cycle implementations of this filter can be found in [19].

Other phenomena can also appear in voltage or current waveforms such as superimposed decaying exponentials and sub-synchronous components, switching noise, power system fundamental frequency deviation and TC saturation. These phenomena lead to imprecise phasor computation even using the Fourier filter. The signal decomposition technique discussed earlier was successfully used to filter out exponential components reducing the lack of precision in phasor computation [7,8]. Nevertheless, the aim here is to evaluate the DSMP compression of voltage and current waveforms. Due to the vast use of the Fourier filter scheme it is considered suitable for evaluating the performance of the DSMP described in this work.

3.2. Fault inception time

The fault inception time algorithm presented here consists of detecting phasor magnitude variation from the pre-fault to the fault period. Sometimes impulsive noise can produce small magnitude variations leading to premature fault inception identification. In order to minimize this situation, one considers a variability index which is calculated by using two concatenated sliding windows: the reference and the data window. In this way, impulsive noise and other fast phenomena are filtered out. The transition into a faulty state is considered only if the phasor magnitude values are sufficiently distinct in these windows. The variability index is defined as

$$I_{var} = \frac{|\bar{w}_p - \bar{w}|}{\bar{w}}, \tag{5}$$

where $\bar{w}_p$ and $\bar{w}$ are the mean magnitude calculated using the data window and reference window, respectively. The fault inception time is detected when the variability index is greater than a defined threshold. The threshold value used is 0.04. The lengths of the windows are input parameters of the algorithm and can be defined according to the number of samples corresponding to the fundamental frequency.

3.3. Fault classification

Fault classification is an important issue in the fault understanding process. It is also a key ingredient when fault location in transmission lines is performed based on information from one terminal, as explained in the next section.

For fault classification the phasor values in the pre-fault period are compared to the ones in the fault period in order to determine whether a phase is or is not involved in the fault. The pre-fault values are obtained from the pre-fault segment of the oscillogram using a one-cycle Fourier filter, ending half cycle prior to the fault inception point. The fault values are calculated in the same way, but using the fault segment information instead, positioned usually about 1.75 cycles after the fault inception point.

The algorithm for fault classification is:

1. If any current phasor has magnitude greater than its steady state value scaled by a constant (> 1) then the phase is considered to be in fault. This constant is empirically determined and the value 1.4 yields good results.
2. If more than one phase is considered to be in fault by the above criterion, the ratio between the current magnitudes of

Please cite this article in press as: M.P. Tcheou et al., How far can one compress digital fault records? Analysis of a matching pursuit-based algorithm, Digital Signal Process. (2011), doi:10.1016/j.dsp.2011.11.003

distinct phases is analyzed. If this ratio is greater than a certain value then only the largest phase is considered faulty. A value equal to 2.5 leads to good results.

3. Otherwise, if no phase is considered faulty by the previous tests, the voltage phasor magnitudes are analyzed. If the magnitude of any voltage phase is lower than 90% of its steady state magnitude, then it is considered a faulty phase.
4. To determine if the fault involves the ground, the symmetrical components of currents are calculated. If the ratio between the zero sequence and the positive sequence current of the fault is greater than the ratio between the zero sequence and the positive sequence current of the pre-fault, then the ground is considered to be involved in the fault.

These empirical values have been defined through massive tests of this fault classification algorithm over a large set of real oscillograms acquired from the Brazilian power system.

3.4. Fault location

For reliability of power system operation, it is of major importance to correctly estimate the fault location. In case of a permanent fault, it allows the dispatcher to quickly request a maintenance crew to the line section where the fault occurred for restoring it back. For nonpermanent faults, fault location can help finding the line section where an intermittent problem occurs due, for example, to trees or fire below the line.

The fault location algorithm used here is based on [17] and [18]. Phasor values for the pre-fault and fault segments are determined in the fault classification step, Section 3.3. From these, the fault distance is computed as described below:

- phase–ground faults

$$d = \frac{\mathrm{Im}(V_F \times I_F^*)}{\mathrm{Im}[\sum_{k=A,B,C}(Z_{Fk} \times I_k^*) \times I_F^*]} \tag{6}$$

  where $F \in \{A, B, C\}$ is the faulty phase and $\mathrm{Im}(d)$ is the imaginary part of d;
- phase–phase faults and three-phase faults

$$d = \frac{\mathrm{Im}\{(V_{F1} - V_{F2}) \times (I_{F1} - I_{F2})^*\}}{\mathrm{Im}[\sum_{k=A,B,C}(Z_{F1k} - Z_{F2k}) \times I_k \times (I_{F1} - I_{F2})^*]} \tag{7}$$

  where for phase–phase faults $F1, F2 \in \{A, B, C\}$, and $F1 \neq F2$ are the faulty phases; and for three-phase faults, $F1$ and $F2$ can be any phase given that $F1 \neq F2$.

In these equations, Z_{Fi} corresponds to each element of the impedance matrix of the transmission line, where $i \in \{1, 2\}$; and I_k denotes the pre-fault current of the phase $k \in \{A, B, C\}$. One should note that this algorithm is highly sensible to phasor calculation and fault type classification errors.

Imprecision in phasor computation, line parameter measurement, nonuniform transposition of the phases and weather and soil conditions can lead to an imprecise location of the fault by this algorithm. However, the purpose here is not to assess the performance of this location method. The goal is to use this traditional tool for evaluating the compression technique applied to electric disturbance signals described in this work.

4. Compression system evaluation

In this section, the performance of the DSMP compression method is assessed using the results of an automated system for fault analysis. As earlier stated the main purpose is to evaluate how much can electric signals be compressed using the DSMP. This evaluation considers the use of the analysis procedures described in Section 3. In the evaluation process, a set of 18 oscillograms acquired by digital fault recorders of the Brazilian power system grid is considered, each one containing 6 to 8 analog channels (voltage and current in 3 phase circuits plus V0 and I0 channels). The waveforms are stored in IEEE COMTRADE format [25]. The utility, which provided these real DFR data, has 138 kV transmission lines and operates with multiple parallel lines and with a large number of generation plants and loads connected with short length lines. The observed fault response is strongly impacted by a large number of neighbor circuits. The set of oscillograms is composed by disturbances in distinct situations such as voltage sags, current swells, harmonics, transients and circuit switching. The waveforms samples are represented using 16 bits, whereas the fundamental frequency of the Brazilian power system is 60 Hz. The sampling frequencies in this set of files are 1200, 2000, 2040 and 3840 Hz, and the length of the waveforms varies from 378 to 1991 samples.

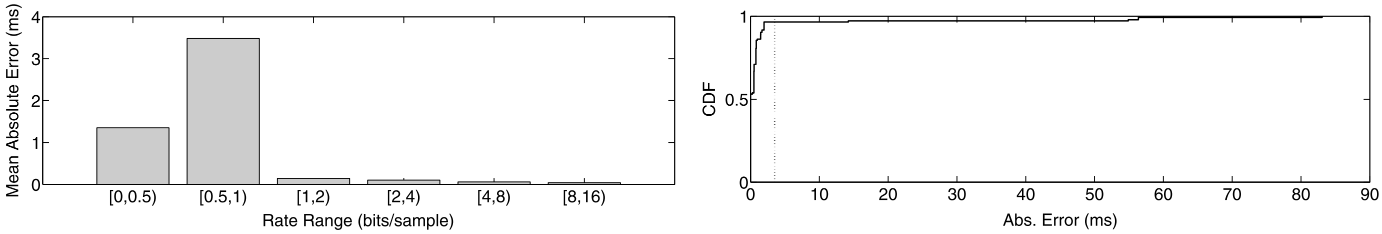
The automated system software is applied to these DFR files as well as to their compressed versions at different bit-rates. By gathering all possible encoded versions of the 18 files, one obtains a complete set of 321 files. The results obtained for the analyses of these compressed files using the automated system are compared to the results achieved by the analysis of their uncompressed versions.

The analysis procedures described in the previous section depend on the value of the fault-inception time threshold. It is important to remark that the threshold used for the analysis of the compressed files is the same used for the uncompressed ones. That is, for the analyses of the compressed DFR files the value of this threshold equals the one commonly used in practice by expert engineers which is also applied for the uncompressed (original) DFR. In addition, the rate × distortion optimization employed for DFR compression does not take into account the DFR data analysis. Therefore, by comparing the analysis results for compressed and uncompressed DFRs one can evaluate the required bit-rate for compressed DFR data. That is, one evaluates the level of compression that can be applied to DFR data without impacting their analyses by standard methodologies of fault-analysis that are used by expert engineers. Thus, the results that are presented following provide a bridge between the compression of DFR data and its practical usage.

4.1. Influence on fault inception time

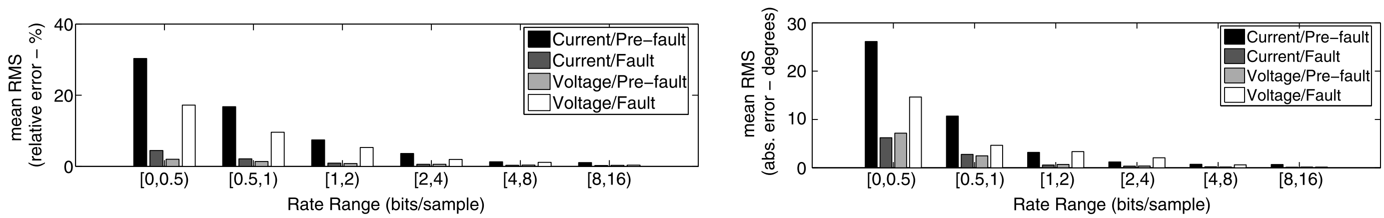
Initially, it is evaluated how much compression affects the detection of the fault inception time, the starting point of the fault analysis process. A correct fault inception time detection is important for fault classification and therefore for fault distance computation. For observing the information on the fault inception time that is retained in the compressed signal one employs the absolute error. The absolute error is the absolute value of the difference between the fault inception time instants obtained for the compressed and uncompressed oscillograms. It is represented in milliseconds. The results obtained are shown in Fig. 3. Fig. 3(a) exhibits the mean across all oscillograms of the absolute error for different bit-rate intervals. This graph is computed by separating the compressed oscillograms according to the rate employed in the compression process (it should be highlighted that a given oscillogram is compressed in this work using different rates, generating different compressed versions). For each interval the mean of the absolute error is computed. It is possible to observe that the mean absolute error does not exceed 3.5 ms, which is reasonably small when compared with the time duration of other fault events such as protection relay operation or switch breaker opening.

To obtain a more detailed analysis of the bit-rate region with the largest acceptable absolute error values, the Cumulative Distri-



(a) (b)

Fig. 3. Detection of the fault time instant: (a) mean absolute error versus bit-rate ranges, (b) CDF of the absolute errors of the fault time instant detection for bit-rates up to 1 bit/sample, the dashed line indicates 3.5 ms.



(a) Magnitude (b) Angle

Fig. 4. Magnitude and angle computation of voltage and current in the pre-fault and fault periods: (a) mean RMS of relative errors of magnitude versus bit-rate ranges, (b) mean RMS of absolute errors of angle versus bit-rate ranges.

bution Function (CDF: $F_X(x) = P(X \leqslant x)$) of the absolute errors for oscillograms compressed with 0 to 1 bit/sample in Fig. 3(a) is plotted in Fig. 3(b). One observes that although the maximum absolute error can be as high as 80 ms, approximately 97% of the absolute errors are smaller than 3.5 ms (indicated by the dashed line). This high percentage of small errors demonstrates that our compression system (the DSMP) even when operating at low bit-rates preserves the fault inception time detection. A more detailed analysis shows that for bit-rate values above 0.7 bits per sample (bps) the absolute error in the fault inception time does not exceed 3.5 ms; therefore, in order to avoid significant fault inception time errors, the oscillograms should be compressed using the DSMP at bit-rates above 0.7 bps.

4.2. *Influence on phasors computation*

As previously discussed a correct computation of the magnitude and the angle of voltage and current phasors are of utmost importance. Therefore, the effect of DSMP encoding distortion on the computation of these values has to be evaluated. The magnitude and angle are calculated some cycles before and also some cycles after the fault inception time. These phasor parameters are used as input parameters for fault classification and computation of fault distance. Since these quantities rely on the detection of the fault inception instant, the compressed oscillograms with unacceptable absolute errors (greater than 3.5 ms) are not considered.

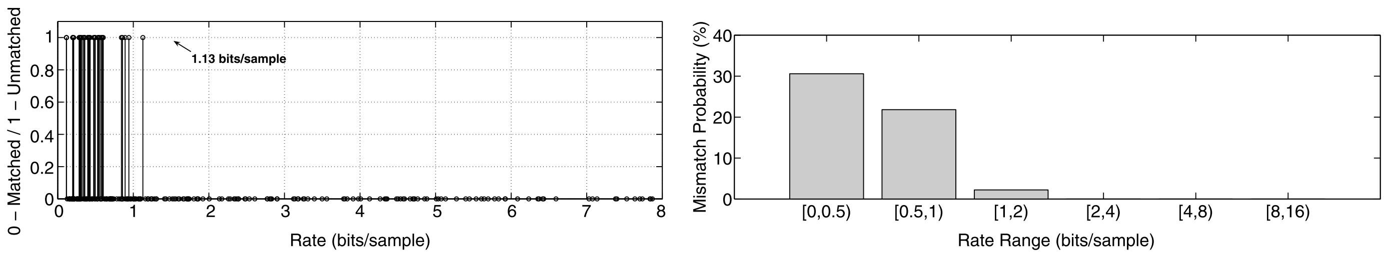
Fig. 4(a) depicts the mean, across different oscillograms, of the root mean square between compressed and uncompressed files for voltage and current waveforms, employing distinct bit-rate ranges. Fig. 4(b) depicts the mean RMS value, for the distinct oscillograms, of the absolute angle errors between compressed and uncompressed voltage and current signals with respect to the same bit-rate ranges. These graphs are obtained in a manner similar to the one described in Section 4.1 for the graph in Fig. 3(a). However, instead of using the mean of the absolute error, one employs the RMS value. Both figures show separately the RMS values of the errors for current and voltage signals and for each signal type, the pre-fault and fault periods are considered. From these results, one notes that for current signals the distortion introduced in the pre-fault is larger than in the fault period, and that the opposite occurs for voltage signals, particularly at low bit-rate. This happens because, in general, faults produce current swells and voltage sags. Therefore, for current signals the signal energy is highly concentrated in the fault period whereas, for voltage signals, the signal energy is concentrated in the pre-fault. As the amount of compression increases, for each oscillogram being compressed, the components of lower energy are discarded before the ones of higher energy. Hence, the DSMP tends to affect more severely current signals in the pre-fault period and voltage signals in the fault period.

The analysis of phasor distortion is not enough for determining the tolerable amount of compression, since for this purpose one must know how much distortion on the phasor magnitude and angle are acceptable. In order to do so, other measurements [14, 15] must be performed. For example, the phasors computed some cycles before and after the fault inception are used for fault classification and fault location. Therefore, the compression method is evaluated investigating at how extent it is able of not impairing these fault analysis procedures. The procedure employed is as follows: first, the fault is classified (Section 3.3); once the fault is classified, one computes the fault distance as in Section 3.4.

4.3. *Influence on fault classification*

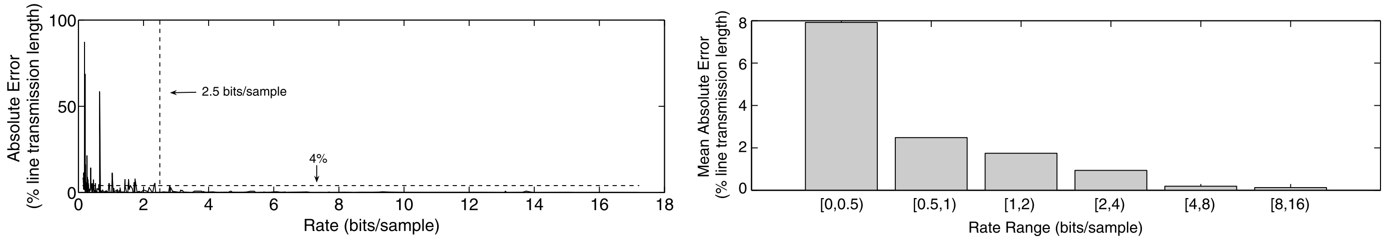
One way to evaluate the influence of DSMP compression for fault analysis procedures is to compare the fault type classification match for original and compressed oscillograms. Fig. 5(a) presents such an analysis; it indicates which bit-rate values are associated to erroneous classification (0 – correct; 1 – incorrect). Fig. 5(b) shows the probability of fault classification mismatch in different bit-rate ranges. For bit-rates greater than 1.13 bits per sample there is no mismatch in fault classification. Therefore, DFR files should be compressed over this bit-rate using the DSMP. In Section 4.1, it was verified that a rate greater than 0.7 bits per sample is required to obtain very low fault inception time errors. Therefore, for the set of signals being used to evaluate the DSMP, a coding rate greater than 1.13 bits per sample guarantees no fault classification mismatch and negligible inception time errors.

Please cite this article in press as: M.P. Tcheou et al., How far can one compress digital fault records? Analysis of a matching pursuit-based algorithm, Digital Signal Process. (2011), doi:10.1016/j.dsp.2011.11.003



(a) (b)

Fig. 5. Fault classification: (a) indicates which bit-rates values lead to classification mismatch and (b) shows the probability of fault classification mismatch for different bit-rate ranges.



(a) (b)

Fig. 6. Computation of fault distance: (a) absolute error versus bit-rate, (b) mean absolute error versus bit-rate ranges.

4.4. Influence on fault location

As previously discussed an important function of oscillograms is their use for fault location computation. This section establishes how much compression can be applied to oscillograms using the DSMP without largely impairing fault location.

It is important to note that the accuracy of fault classification is crucial for the computation of fault distance. Thus, when the fault classification of the compressed signal does not match with the one of the uncompressed one, it becomes useless to compare the results of fault location between them. Thereby, the performance of the DSMP is assessed regarding the fault location, given that the fault type identified for the compressed file matches the one identified for the original file. Fig. 6(a) shows the absolute errors between the computed fault distances for the compressed and the uncompressed signals. Fig. 6(b) exhibits the mean absolute error in fault distance computation for different bit-rate intervals. The distances are calculated in percentage of the transmission line length. From the results obtained one can see, for example, that if one considers absolute errors that are lower than 4% to be acceptable, the signals can be compressed using a rate larger than 2.5 bps as indicated by the dashed lines in Fig. 6(a).

4.5. Overall performance

A protocol to be employed to evaluate and validate a compression system for power system signals must start by studying the influence of signal compression on the computation of fault inception time, since this is the first step on fault analysis. Therefore, compressed signals leading to large errors in the fault inception time are not acceptable (we considered 3.5 ms as the maximum acceptable error). Thus, quantizers (conversely, bit-rates) leading to such unacceptable errors must be discarded. Following, the second step of the protocol shall evaluate the influence of compression on the computation of current and voltage phasors on the post-fault and pre-fault (note that the transition between pre-fault and post-fault is given by the fault inception time). The relevance of computing these values lies on the fact that they are input to the fault classification process. However, no conclusions can be taken solely from phasor distortion; one shall rather consider the influence of phasor distortion on fault classification and location. Thus, the next step of the protocol shall evaluate the match between the fault classification for original and compressed signals. Quantizers leading to compressed signals that generate classification mismatches shall thus be discarded. The last step of the fault analysis protocol is fault location. It shall be applied only on those signals for which the classification was correct.

Concerning the overall results, one observes that the amount of compression that can be employed when using the DSMP technique is determined by the fault location procedure. In other words, the bit-rate needed to achieve a tolerable error in the fault location is more than sufficient to ensure both an acceptable estimate of the fault inception time and an error-free fault classification. Nonetheless, one should have in mind that the amount of tolerable error in fault location depends on the transmission line length. In conclusion, for the oscillogram test set used in this work and considering as acceptable a 4% fault location error, a bit-rate as low as 2.5 bit/sample can be employed for compressing DFR data using the DSMP. Thus, this indicates that the DSMP algorithm is able to provide high compression ratio (6.4:1) with reasonable reliability and robustness, without compromising the fault analysis process of compressed DFRs. It is important to remark that we also verified the DSMP compression performance when applied to different line conditions and obtained similar results.

4.6. Results for different transmission line conditions

Here, we verify how much variation occurs on the results of the compression evaluation with respect to line condition modifications. For this purpose the evaluation procedures previously described are applied on a set of 4 oscillograms acquired by monitoring 375 kV transmission lines and their respective compressed versions. The utility that provided theses files operates with large line lengths with few (typically 1) parallel circuits. In this case, the observed system response usually shows the fault behavior more clearly, as compared to the 138 kV system presented before. Considering this set of oscillograms, it is verified that in order to obtain absolute errors of fault inception time smaller than 3.5 ms, the fault records should be encoded at bit-rates larger than 0.16 bit/sample. In addition, no fault classification error was found.

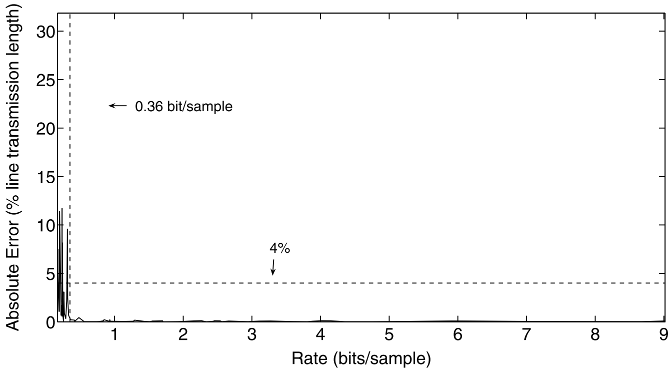


Fig. 7. Computation of fault distance modifying line conditions: absolute error versus bit-rate.

Fig. 7 shows the absolute errors in fault distance computation versus bit-rate considering the whole set of 4 DFR files. It can be seen that considering as acceptable a fault location error of 4%, bit-rates above 0.36 bit/sample can be employed. Hence, for the two different line conditions presented above the fault location is still the main limiting factor for determining the required bit-rate that is necessary for DFR compression using the DSMP. Note that for the second line condition the bit-rate that is necessary to compress its DFR data without compromising the analysis of the compressed files is much smaller (0.36 bits/sample) than for the first line condition investigated (2.5 bits/sample).

5. Discussion

In this section, we compare quantitatively the DSMP to other compression methods for electric signals. In [6], one presents compression results for power system signals when using lossless and lossy off-the-shelf CODECs developed for audio and image signals. As can be verified in [6], the SPIHT (Set Partitioning In Hierarchical Trees) image encoder [27] provided a better compression performance for electric signals when compared, for instance, to JPEG and MP3 encoders [21]. Therefore, we decided to use SPIHT to encode the same set of fault records compressed through DSMP. Previously, the signals of each fault record are concatenated, scaled to be in the 0–255 range and then converted to matrix representation.

Table 1 provides the compression performance of DSMP and SPIHT lossy encoders and also of the Lempel–Ziv lossless encoder [21]. In order to accomplish a fair comparison between DSMP and SPIHT, one compresses each fault record with the minimum possible bit-rate without compromising any of the fault analysis procedures. The same bit-rate is employed when using SPIHT, so that both compression methods can be quantitatively compared through the signal-to-noise ratio distortion metric ($\mathrm{SNR} = 10\log_{10}(\frac{\|\mathbf{x}\|^2}{\|\mathbf{x}-\tilde{\mathbf{x}}\|^2})$ dB). One may observe that both DSMP and SPIHT methods provide a considerably higher Compression Ratio (CR) than the Lempel–Ziv lossless encoder. Hence, it is very advantageous to use lossy compression approaches. On the other hand, when comparing DSMP and SPIHT, one can note that for high compression ratio (> 10) DSMP achieves a larger SNR than SPIHT whereas for low compression ratio the inverse occurs.

As the results show, at low bit-rate the performance of the DSMP largely surpasses the one of the SPIHT, while for higher rates their performances are roughly equivalent, with SPIHT being slightly better. Therefore, one can see that the DSMP is a competitive compression method for electric signals specially at high compression ratio levels when compared to off-the-shelf encoders. A very important practical aspect is that, at the same time, the DSMP does not compromise the fault analysis.

Table 1
Compression performance of DSMP, SPIHT and Lempel–Ziv (LZ) encoders. When there is an SNR difference larger than 3 dB, the method with the highest SNR is marked in boldface.

Fault record	DSMP			SPIHT			LZ
	CR	Rate (bps)	SNR (dB)	CR	Rate (bps)	SNR (dB)	CR
1	14.71	1.09	**21.47**	16.00	1.00	17.82	1.29
2	9.55	1.67	22.80	9.52	1.68	23.99	1.35
3	20.43	0.78	**26.09**	20.51	0.78	12.50	1.23
4	83.80	0.19	**14.00**	84.21	0.19	4.89	1.61
5	101.41	0.16	**26.73**	100.00	0.16	7.05	1.38
6	70.48	0.23	**37.16**	69.57	0.23	7.42	1.72
7	98.61	0.16	**29.95**	100.00	0.16	3.03	1.26
8	32.19	0.50	**28.20**	32.00	0.50	14.48	1.25
9	27.76	0.58	**33.02**	27.59	0.58	20.91	1.19
10	55.02	0.29	**26.35**	55.17	0.29	15.33	1.20
11	18.74	0.85	19.73	18.60	0.86	**22.77**	1.39
12	75.58	0.21	**30.14**	76.19	0.21	7.41	1.46
13	6.81	2.35	24.35	6.81	2.35	**32.94**	1.37
14	9.20	1.74	21.14	9.25	1.73	**30.15**	1.36
15	16.54	0.97	**46.91**	16.49	0.97	28.40	1.46
16	62.24	0.26	**35.12**	61.54	0.26	9.26	1.47
17	6.56	2.44	23.41	6.56	2.44	**33.95**	1.42
18	11.19	1.43	35.46	11.19	1.43	36.09	1.26

6. Conclusions

In this paper, the performance of the DSMP compression method for voltage and current waveforms representing faults in transmission lines was investigated. Although not usual in the literature, estimates of how much one can compress digital fault records were obtained by considering fault analysis procedures. With this approach we could achieve more reliable compression performance validation results than using mean-squared error versus bit rate graphs and tables. The fault analysis system employed in this work was capable of performing basic fault analysis procedures such as computing phasors, identifying the fault inception time, classifying the fault type and calculating the fault distance.

It has been observed that the required bit-rate may vary a lot depending on the line being monitored. However, for the set of oscillograms explored, having different line conditions, one can employ bit rates as low as 2.5 bits/sample without compromising any of the fault analysis procedures considered. This corresponds to a compression ratio of 6.4 to 1. Considering the overall results obtained using the performance assessment here presented, one can infer that the DSMP is a reliable and state-of-the-art method for the compression of electric disturbance signals.

References

[1] S. Santoso, E. Powers, W. Grady, Power quality disturbance data compression using wavelet transform methods, IEEE Trans. Power Deliv. 12 (1997) 1250–1257.
[2] T.B. Littler, D.J. Morrow, Wavelets for the analysis and compression of power system disturbances, IEEE Trans. Power Deliv. 14 (2) (Apr. 1999) 358–364.
[3] M.V. Ribeiro, H. Park, J. Marcos, T. Romano, S.J. Mitra, A novel MDL-based compression method for power quality applications, IEEE Trans. Power Deliv. 22 (1) (Jan. 2007) 27–36.
[4] P. Dash, B. Panigrahi, D. Sahoo, G. Panda, Power quality disturbance data compression, detection, and classification using integrated spline wavelet and s-transform, IEEE Trans. Power Deliv. 18 (2) (Apr. 2003) 595–600.
[5] S.-J. Huang, M.-J. Jou, Application of arithmetic coding for electric power disturbance data compression with wavelet packet enhancement, IEEE Trans. Power Syst. 19 (3) (Aug. 2004) 1334–1341.
[6] O.N. Gerek, D.G. Ece, Compression of power quality event data using 2D representation, Electric Power Syst. Res. 78 (6) (Jun. 2008) 1047–1052, doi:10.1016/j.epsr.2007.08.006.
[7] L. Lovisolo, E.A.B. da Silva, M.A.M. Rodrigues, P.S.R. Diniz, Efficient coherent adaptive representations of monitored electric signals in power systems using damped sinusoids, IEEE Trans. Signal Process. 53 (10) (Oct. 2005) 3831–3846.
[8] L. Lovisolo, M.P. Tcheou, E.A.B. da Silva, M.A.M. Rodrigues, P.S.R. Diniz, Modeling of electric disturbance signals using damped sinusoids via atomic de-

compositions and its applications, EURASIP J. Adv. Signal Process. 2007 (2007), doi:10.1155/2007/29507, Article ID 29507, 15 pp.
[9] M.P. Tcheou, L. Lovisolo, E.A.B. da Silva, M.A.M. Rodrigues, P.S.R. Diniz, Optimum rate-distortion dictionary selection for compression of atomic decompositions of electric disturbance signals, IEEE Signal Process. Lett. 14 (2) (Feb. 2007) 81–84.
[10] S. Mallat, Z. Zhang, Matching pursuits with time–frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (Dec. 1993) 3397–3415.
[11] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1998.
[12] ITU-R Rec. BS.1387, Method for objective measurements of perceived audio quality, International Telecommunications Union, Geneva, Switzerland, 1998.
[13] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (Apr. 2004) 600–612.
[14] M.A.M. Rodrigues, M.V.F. de Figueiredo, A.L.L. Miranda, S.S. Diniz, Oscillography for power system operational planning, in: VII SEPOPE, Curitiba, Brazil, May 2000, pp. 23–28.
[15] M.A.M. Rodrigues, J.C.C. Oliveira, A.L.L. Miranda, M.V.F. Figuereido, R.B. Sollero, L.A.C.D. Castillo, F.D. Campos, N.F. Costa, Evaluation of protection system performance using DFR and relay comtrade files, in: Cigre Session 42, 2008.
[16] A.G. Phadke, J.S. Thorp, Computer Relaying of Power Systems, Research Studies Press Ltd, England, 1988.
[17] T. Takagi, Y. Yamakoshi, J. Baba, K. Uemura, T. Sakaguchi, A new algorithm of an accurate fault location for EHV/UHV transmission lines: Part I – Fourier transformation method, IEEE Trans. Power Apparatus Syst. PAS-100 (3) (Mar. 1981) 1316–1323.
[18] T. Takagi, Y. Yamakoshi, J. Baba, K. Uemura, T. Sakaguchi, A new algorithm of an accurate fault location for EHV/UHV transmission lines: Part II – Laplace transformation method, IEEE Trans. Power Apparatus Syst. PAS-101 (3) (Mar. 1982) 564–573.
[19] E.O. Schweitzer III, D. Hou, Filtering for protective relays, in: WESCANEX 93, 'Communications, Computers and Power in the Modern Environment', Conference Proceedings, IEEE, 1993, pp. 15–23.
[20] W. Xu, Component modeling issues for power quality assessment, IEEE Power Eng. Rev. 21 (11) (Nov. 2001) 12–15, 17.
[21] K. Sayood, Introduction to Data Compression, 2nd ed., Morgan Kaufman Publishers, San Francisco, USA, 2000.
[22] A. Ortega, K. Ramchandran, Rate-distortion methods for image and video compression, IEEE Signal Process. Mag. 15 (6) (Nov. 1998) 23–50.
[23] M. Kezunovic, T. Popovic, D.R. Sevcik, A. Chitambar, Requirements for automated fault and disturbance data analysis, in: CIGRE Colloquium, Sydney, Australia, Sep. 2003.
[24] J.J. Grainger, W.D. Stevenson, Power System Analysis, McGraw–Hill, 1994.
[25] IEEE Std C37.111-1999, IEEE standard common format for transient data exchange (COMTRADE) for power systems, Mar. 1999.
[26] G.-X. Zhang, Time–frequency atom decomposition with quantum-inspired evolutionary algorithms, Circuits Syst. Signal Process. 29 (3) (2010) 209–233.
[27] J. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans. Signal Process. 41 (12) (Dec. 1993) 3445–3462.

Michel P. Tcheou was born in Rio de Janeiro, Brazil. He received the Engineering degree in Electronics from Federal University of Rio de Janeiro (UFRJ) in 2003, the M.S. and D.S. degrees in Electrical Engineering from COPPE/UFRJ in 2005 and 2011, respectively. He is currently a researcher at the Electric Power Research Center (CEPEL) in Rio de Janeiro, Brazil. His research interests are in signal processing and operation planning of hydrothermal power systems.

André L.L. Miranda was born in Rio de Janeiro, Brazil, in 1971. He received the Engineering degree in Electronics from the State University of Rio de Janeiro in 1994, and the M.S. in Signal Processing from Federal University of Rio de Janeiro in 2005. He has been working at the Electric Power Research Center (CEPEL) in Rio de Janeiro, since 1997, in algorithms for data analysis, and signal processing applications related to protection for power systems. He is also a professor in an undergraduate course at FABES, a private college in Rio de Janeiro, teaching classes of basic electronic and electrical circuits, since 2004. His current research interests are signal processing, oscillographic fault and disturbance analysis automation, and power system protection. He is a member of Cigre (Conseil International des Grands Reseaux Electriques).

Lisandro Lovisolo was born in Neuquen, Argentina. He lives in Rio de Janeiro, Brazil, since 1976. He concluded his doctorate studies in Electrical Engineering at COPPE, Universidade Federal do Rio de Janeiro (UFRJ) in 2006. He is currently an Associate Professor at the Department of Electronics and Communications at Universidade do Estado do Rio de Janeiro (UERJ) and at the post-graduate program in Electronics Engineering in the same institution. At UERJ, he has the founded the PROSAICO (Laboratório de PROcessamento de Sinais, Aplicações Inteligentes e COmunicações – http://www.prosaico.uerj.br). His interests and activities are mainly focused in signal processing and communications.

Eduardo A.B. da Silva was born in Rio de Janeiro, Brazil. He received the Electronics Engineering degree from Instituto Militar de Engenharia (IME), Brazil, in 1984, the M.Sc. degree in Electrical Engineering from Universidade Federal do Rio de Janeiro (COPPE/UFRJ) in 1990, and the Ph.D. degree in Electronics from the University of Essex, England, in 1995.

In 1987 and 1988 he was with the Department of Electrical Engineering at Instituto Militar de Engenharia, Rio de Janeiro, Brazil. Since 1989 he has been with the Department of Electronics Engineering (the undergraduate dept.), UFRJ. He has also been with the Department of Electrical Engineering (the graduate studies dept.), COPPE/UFRJ, since 1996. He has been head of the Department of Electrical Engineering, COPPE/UFRJ, Brazil, for the year 2002. In 2007 he has been a Visiting Professor at the University of Nice Sophia-Antipolis. His teaching and research interests lie in the fields of digital signal, image and video processing. In these fields, he has published over 160 referred papers. He won the British Telecom Postgraduate Publication Prize in 1995, for his paper on aliasing cancellation in subband coding. He is also co-author of the book "Digital Signal Processing – System Analysis and Design", published by Cambridge University Press, in 2002, that has also been translated to the Portuguese and Chinese languages, whose second edition has been published in 2010.

He has served as an associate editor of the IEEE Transactions on Circuits and Systems – Part I, in 2002, 2003, 2008 and 2009, of the IEEE Transactions on Circuits and Systems – Part II in 2006 and 2007, and of Multidimensional, Systems and Signal Processing, since 2006. He has been a Distinguished Lecturer of the IEEE Circuits and Systems Society in 2003 and 2004. He is Technical Program Co-Chair of ISCAS2011.

He has given training and consultancy for several Brazilian cable and satellite television companies on digital television. He was part of the team that worked in the development of the Brazilian Digital Television System. His research interests lie in the fields of digital signal and image processing, especially signal compression, digital television, wavelet transforms, mathematical morphology and applications to telecommunications. He is a senior member of the IEEE.

Marco A.M. Rodrigues was born in Rio de Janeiro, Brazil, in 1964. He received the Engineering degree in Electronics, the M.S. and the D.S. degrees in Electrical Engineering from the University of Rio de Janeiro in 1986, 1991, and 1999, respectively. He has been working at the Electric Power Research Center (CEPEL) in Rio de Janeiro, Brazil, since 1987, in data acquisition systems design, software design, algorithms for data analysis, and control and signal processing applications related to power systems. He is also an Invited Professor in a postgraduate course in protection for power systems, held at the University of Rio de Janeiro. His current research interests are signal processing, power system measurements, oscillographic analysis automation, and power system protection systems. Dr. Rodrigues is a Senior Member of IEEE and Member of Cigre (Conseil International des Grands Reseaux Electriques) and of SBrT (the Brazilian Telecommunications Society). He volunteers in the IEEE Rio de Janeiro section as vice-president (2010–2011). He participates in the technical committee Brazilian Protection and Control Technical Seminary (STPC) since 2003.

Paulo S.R. Diniz was born in Niterói, Brazil. He received the Electronics Eng. degree (Cum Laude) from the Federal University of Rio de Janeiro (UFRJ) in 1978, the M.Sc. degree from COPPE/UFRJ in 1981, and the Ph.D. from Concordia University, Montreal, P.Q., Canada, in 1984, all in electrical engineering.

Since 1979 he has been with the Department of Electronic Engineering (the undergraduate dept.) UFRJ. He has also been with the Program of Electrical Engineering (the graduate studies dept.), COPPE/UFRJ, since 1984, where he is presently a Professor. He served as Undergraduate Course Coordinator and as Chairman of the Graduate Department. He is one of the three senior researchers and coordinators of the National Excellence Center in Signal Processing. He has also received the Rio de Janeiro State Scientist award, from the Governor of Rio de Janeiro state.

From January 1991 to July 1992, he was a visiting Research Associate in the Department of Electrical and Computer Engineering of University of Victoria, Victoria, B.C., Canada. He also holds a Docent position at Helsinki University of Technology. From January 2002 to June 2002, he was a Melchor Chair Professor in the Department of Electrical Engineering of University of Notre Dame, Notre Dame, IN, USA. His teaching and research interests are in analog and digital signal processing, adaptive signal processing, digital communications, wireless communications, multirate systems, stochastic processes, and electronic circuits. He has published several refereed papers in some of these areas and wrote the books *ADAPTIVE FILTERING*: *Algorithms and Practical Implementation*, Springer, NY, third edition 2008, and *DIGITAL SIGNAL PROCESSING*: *System Analysis and Design*, Cambridge University Press, Cambridge, UK, second edition 2010 (with E.A.B. da Silva and S.L. Netto).

He has served as General co-Chair of ISCAS2011 and Technical Program Chair of the 1995 MWSCAS both held in Rio de Janeiro, Brazil. He was also the Technical Program co-Chair of SPAWC2008. He has been on the technical committee of several international conferences including ISCAS, ICECS, EUSIPCO and MWSCAS. He has served Vice President for region 9 of the IEEE Circuits and Systems Society and as Chairman of the DSP technical committee of the same society. He is also a Fellow of IEEE (for fundamental contributions to the design and implementation of fixed and adaptive filters and electrical engineering education). He has served as associate editor for the following Journals: IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing from 1996 to 1999, IEEE Transactions on Signal Processing from 1999 to 2002, and the Circuits, Systems and Signal Processing Journal from 1998 to 2002. He was a distinguished lecturer of the IEEE Circuits and Systems Society for the year 2000 to 2001. In 2004 he served as distinguished lecturer of the IEEE Signal Processing Society and received the 2004 Education Award of the IEEE Circuits and Systems Society.

Please cite this article in press as: M.P. Tcheou et al., How far can one compress digital fault records? Analysis of a matching pursuit-based algorithm, Digital Signal Process. (2011), doi:10.1016/j.dsp.2011.11.003