

Sistema de tempo real para processamento de sinais de descargas parciais focado em implementação por hardware

Daniel Luiz de Souza¹

Jéssica Luiza Cabral Guimarães¹

Jordy Braga Xisto¹

Welington Carlos Barbosa¹

Rubem Geraldo Vasconcelos Machado²

RESUMO

Este artigo descreve melhorias em um agente desenvolvido para processar sinais provenientes de monitoramento visando identificar ocorrências de descargas parciais e viabilizando o processo de eliminação de ruídos inerentes ao processo de amostragem. O agente original utiliza conceitos de Multiple Instance Learning, Self-Organizing Maps e é baseado em Manhattan Distance como métrica de distância. Os aprimoramentos propostos, visando aprimorar a seleção de referências e seleção de coeficientes relevantes, baseia-se em Support Vector Machine e Differential Evolution. Com estas adaptações consegue-se um agente classificador simples, auto ajustável e com pouca demanda por recursos, viabilizando sua implementação em hardware, permitindo uso de paralelização e pipelining de forma a se obter melhores taxas de amostragem.

Palavras-chave: Descargas Parciais. Máquinas de Vetores de Suporte. Evolução Diferencial. Mapas Auto Organizáveis. Processamento de Sinais.

ABSTRACT

This article describes improvements in an agent developed to process signals from monitoring to identify occurrences of partial discharges and making possible the elimination of noise inherent in the sampling process. The original agent uses Multiple Instance Learning concepts, Self-Organizing Maps, and is based on Manhattan Distance as a distance metric. The proposed improvements, aiming to improve the selection of references and selection of relevant coefficients, is based on Support Vector Machine and Differential Evolution. With these adaptations, a simple, self-adjusting classifier agent is achieved, with little demand for resources, enabling

¹ Estudantes de Engenharia Elétrica. Faculdade Doctum de João Monlevade

² Engenheiro Eletricista. Professor Mestre Orientador.

its implementation in hardware, allowing the use of parallelization and pipelining in order to obtain better sampling rates.

Keywords: Partial Discharges. Support Vector Machine. Differential Evolution. Self-Organizing Maps. Signal Processing.

1 INTRODUÇÃO

Os sistemas elétricos de potência tornaram-se essenciais às atividades humanas e, assim, eventuais falhas nestes podem acarretar inúmeros prejuízos, inclusive financeiros, às várias camadas desse sistema (geração, distribuição e consumo). Por isso, procedimentos de manutenção são fundamentais para que se garanta a sua confiabilidade, tornando-se necessárias estratégias mais eficientes para viabilização destes, com uso, sempre que possível, de rotinas e/ou ações preditivas. Para que manutenções preditivas se tornem efetivas, torna-se fundamental uma eficiente coleta de dados, permitindo que a identificação de falhas seja efetivada em tempo e de forma correta.

Conjunto representativo de falhas que ocorrem em sistemas elétricos de potência está associado a problemas no isolamento elétrico e estes podem ser identificadas através da detecção, tratamento e posterior análise de sinais de descargas parciais (DP). Em Bartnikas (2002) discorre-se sobre o mecanismo de ocorrência destes eventos, bem como procedimentos de detecção e medida. O grande problema é que as ocorrências de descargas parciais são aleatórias, tanto no tempo em que ocorrem, quanto na frequência e na forma de onda associadas. Com isso, o ideal é que a taxa de amostragem seja a mais alta possível, evitando-se a perda de ocorrências dos eventos. A identificação de falhas pode ser baseada no número de ocorrências por unidade de tempo, na alteração da forma de onda ou em ambos.

Outro fator relevante é que, quando da detecção e medição de sinais de DP, dadas suas características (baixa amplitude), ruídos, associados a fatores ambientais e aos próprios procedimentos de amostragem, acabam poluindo os sinais coletados. A viabilização de monitoramento on-line com processamento local destes sinais de DP pode permitir economia na largura de banda dos canais de

comunicação e na eliminação de gargalos existentes no processamento centralizado destes sinais. Assim, permite-se, além da eliminação de ruídos e identificação de padrões de interesse, a associação das ocorrências de DP com condições de ambiente e/ou de operação dos equipamentos.

Em Mota et. al. (2011), Khamseh et. al. (2014), Abdel-Galil et. al. (2008) demonstra-se que a utilização de transformadas de wavelets é uma ferramenta efetiva para a eliminação de ruídos existentes nos sinais provenientes do monitoramento visando identificação de ocorrências de descargas parciais, bem como a identificação de padrões indicativos de falhas. Em Hussein et. al. (2016), utiliza-se transformada rápida de Fourier na operação de *denoising*. Em Raymond et. al. (2017), utilizam-se técnicas de reconhecimento de padrões (PCA, *Principal Component Analysis*) e um classificador, baseado em SVM (*Support Vector Machine*) com o mesmo intuito. Em Nimmo et. al. (2017) compara-se o desempenho de agentes baseados em PCA e t-SNE (*t-Distributed Stochastic Neighbor Embedding*) na efetiva detecção de padrões para eliminação de ruídos em meio a sinais provenientes de monitoramento visando identificação de DP. Nas referências citadas, embora os problemas foco sejam diferentes, em todas procura-se explorar assinaturas bastante características associadas aos sinais de DP (frente de onda abrupta, de curta duração, e decaimento lento após este impulso inicial). O grande problema, em todos estes trabalhos, é que as operações são demasiadamente complexas para que exista a possibilidade de viabilização de agente local para processar, em tempo real, as operações necessárias ao monitoramento de ocorrências de DP e que tenha um custo mais acessível.

Em Machado (2015) propõe-se uma forma de se simplificar o processamento de sinais provenientes de monitoramento visando identificar ocorrências de DP visando uma implementação em hardware deste procedimento. A proposta se baseia em conceitos associados a MIL (Multiple Instance Learning), descrito em Settles (2008), e SOM (*Self-Organizing Maps*), descrito em Kohonen (1990), utilizando como métrica a distância Manhattan. O mérito da proposta é conseguir simplificar bastante o procedimento de identificação de DP e operação de *denoising*. Esta simplificação viabiliza a implementação em hardware (FPGA) dos procedimentos necessários à operação de eliminação de ruídos. Em Machado

(2016) discute-se sobre a efetiva implementação do agente proposto em hardware e, além disso, assumindo o uso de transformada de *wavelets*, efetua-se a proposta de utilização de coeficientes de aproximações ao invés de coeficientes de detalhes (módulo de máximos locais) conforme proposto em Machado (2015).

O presente trabalho propõe melhorias nos procedimentos adotados em Machado (2015), incluindo rotinas para definição da melhor camada de decomposição, quando da aplicação de transformada de *wavelets*, e para identificação de componentes relevantes, de forma a reduzir, ainda mais, a quantidade de procedimentos necessários à operação de eliminação de ruídos. Para estas novas funcionalidades, utilizou-se um agente baseado em SVM e de um algoritmo de otimização baseado em DE (*Differential Evolution*) que, no presente caso, apresentou melhor desempenho que outros algoritmos utilizados para reconhecimento de padrões, como LDA (*Linear discriminant analysis*), PCA e t-SNE.

2 SSGC-2 (SELF-SCALABLE GRID CLASSIFIER – UPGRADE)

Neste trabalho foram mantidas as operações básicas propostas em Machado (2015) para o SSGC (*Self-Scalable Grid Classifier*), pois, conforme destacado, tem-se um procedimento simples, pouco custoso computacionalmente e que permite minimizar a demanda por recursos quando da implementação em hardware. Desta maneira, continuam presentes as operações de montagem de bags, seguindo o paradigma de MIL, que permite dividir o espaço do problema em regiões nas quais torna-se mais fácil, utilizando transformadas de *wavelets* para geração dos vetores de entrada, separar sinais e ruídos (além da identificação de padrões de interesse). Basicamente, ao invés de um procedimento global mais complexo, opta-se por dividir o espaço do problema em *clusters (bags)*, nos quais é mais simples identificar os padrões locais de entrada.

No método original, a escolha da camada de decomposição era um parâmetro do algoritmo, ou seja, algum conhecimento prévio deveria existir a respeito do sinal para a definição da camada de decomposição a ser utilizada. As referências utilizadas para classificação (utilizando o SSGC) são os coeficientes de detalhes

(módulos máximos locais) em Machado (2015) e aproximações associadas à parcela do sinal sendo analisada em Machado (2016). Em Machado (2015), dependendo do problema e das características dos sinais sendo tratados, sempre se operou na sexta ou na décima camada de decomposição. Em Machado (2016), com a utilização dos coeficientes de aproximações, pode-se optar por trabalhar na sexta camada de decomposição, na qual estavam disponíveis oito coeficientes de aproximações e com isso, viabilizou-se a mesma operação de classificação porém, com uma menor demanda por decomposições utilizando transformadas de *wavelets*.

Neste trabalho, operou-se na tentativa de reduzir o número de componentes nos vetores de referência para se viabilizar procedimentos mais simples, que demandam menos recursos e que sejam mais facilmente paralelizáveis em uma implementação em hardware. Associando-se isto à utilização de *pipelining* abre-se a possibilidade de que o agente opere com maiores taxas de amostragem. Para a minimização das referências necessárias, utilizou-se um agente baseado em SVM e DE para seleção das componentes relevantes e também para escolha da melhor camada de decomposição. Como função objetivo a ser maximizada, definiu-se a acurácia do classificador e foram penalizadas soluções que utilizassem mais componentes dos vetores de entrada utilizados no processo de treinamento do agente. Em algumas simulações, acrescentou-se também penalizações para conjuntos de referências cujas distâncias entre classes opostas estavam muito próximas no espaço do problema de dimensão reduzida. Deste modo, o que no método de referência deveria ser fornecido como parâmetro, passa a ser definido pelo algoritmo e pelas características dos dados informados no conjunto utilizado para treinamento. A seleção da camada de decomposição e dos coeficientes relevantes permite que sejam entregues ao classificador um conjunto de referências otimizado, tanto na dimensão quanto na capacidade de representar os dados do problema e portanto, na manutenção de desempenho do agente classificador.

3 BASES DE DADOS UTILIZADAS E CONSIDERAÇÕES

Visando permitir uma avaliação mais precisa das modificações propostas, optou-se por utilizar algumas das bases de dados que serviram de referência em

Machado (2015) e em Machado (2016). Para avaliar o desempenho dos classificadores, utilizou-se como métrica apenas a acurácia. Na Tabela I apresenta-se a composição dos conjuntos de dados utilizados neste trabalho, em que se tem tanto bases de dados contendo coeficientes de módulos de máximos locais (coluna “Origem” definida como *det*) quanto bases de dados compostas por coeficientes de aproximações (coluna “Origem” definida como *aprox*). São utilizados conjuntos de dados tanto gerados por medição, quanto por simulação. Essas bases de dados, para efeito de teste, possuem sinais de amplitudes e tempos de ocorrência aleatórios e estão associados a ruídos Gaussianos de pequena amplitude para simular interferências de fundo.

Além disso, em cada conjunto apresentam-se ruídos normalmente possíveis de serem encontrados em processos de medição visando identificação de sinais associados a ocorrências de descargas parciais (como ruídos AM, síncronos, transitórios e aleatórios, pulsativos) provenientes de ocorrências em ar ou em sólido.

Outro conjunto de dados utilizados, não apresentados na Tabela I, é composto por alguns sinais contendo séries temporais provenientes de medição ou simulação de sinais contendo descargas parciais e ruídos. Neste caso, cada conjunto é composto por amostras somente de dados (temporais) de ruído, por amostras somente de sinais contendo eventos de descargas parciais, ambos utilizados para treinamento do agente, e de amostra para teste, composta por sinais medidos ou simulados de descargas e ruídos típicos diversos, como descrito no parágrafo anterior. A utilização destes conjuntos permite avaliar a operação de separação de sinal e ruído, utilizando procedimento similar ao que deverá ser executado em hardware, FPGA no caso de Machado (2016). Em todos os conjuntos de dados utilizados, para os dados medidos, os conjuntos de treinamento foram gerados por medições independentes de sinais e ruídos ou por partes de um mesmo sinal em que sinais e ruídos não se sobreponham.

Base de dados	Tipo	Origem	Número de componentes	Amostras de PD	Amostras de ruídos	Amostras totais
1	Simulado	Det	6	605	6293	6898

2	Simulado	Det	6	40	242	282
3	Simulado	Det	6	320	3260	3580
4	Medido	Det	10	1172	5610	6782
5	Medido	Det	6	653	411	1064
6	Medido	Det	6	660	111	771
7	Medido	Aprox.	8	16	13709	13725
8	Medido	Aprox.	8	19	7937	7956
9	Medido	Aprox.	8	21	15994	16015
10	Medido	Aprox.	8	16	16129	16145
11	Simulado	Aprox.	8	64	993	1057
12	Simulado	Aprox.	8	64	993	1057

Tabela I. Características dos conjuntos de dados utilizados no treinamento

A performance do agente com as alterações propostas foi avaliada utilizando como referência um agente baseado em SVM e outro baseado no padrão SSGC original, proposto em Machado (2015) e utilizou-se a mesma família de wavelets (Daubechies 6). Os valores publicados se referem à média de acertos na identificação de sinais de DP e de ruídos ao longo de cinquenta execuções sucessivas. Em cada iteração, aleatoriamente foram montados conjuntos de treino e teste e os mesmos foram fornecidos aos três agentes utilizados (SSGC, SVM e SSGC2). Estes conjuntos foram assim definidos: aleatoriamente, foram montados um conjunto de treino e um de teste com, respectivamente, 85% e 15% das amostras disponíveis de cada uma das bases de dados explicitadas na Tabela I. Além da verificação da acurácia obtida pelo agente, avaliou-se o número de referências definidas e o número de componentes destas referências definido pelo SSGC2. Como na prática a definição destas referências não será implementada em hardware, em nenhum momento avaliou-se o tempo gasto por cada agente, para definição das mesmas. O agente SVM, da mesma forma que em Machado (2015) foi configurado utilizando como função *kernel*, a função *rbf* (*radial basis function*).

4 RESULTADOS

Na Figura 1, são destacados os resultados médios para percentual de avaliações corretas e na Figura 2 o número de componentes relevantes selecionadas (valores arredondados para inteiros mais próximos), após cinquenta execuções para cada base de dados. Pode-se observar que a escolha da melhor camada de decomposição (normalmente, a quarta ou quinta) e a redução de dimensão dos vetores com características dos padrões de entrada, para alguns dos conjuntos de dados utilizados, permitiu, inclusive, melhorar a distinção entre as classes do problema e, conseqüentemente, observamos maiores percentuais de acertos nas classificações realizadas. No geral, a proposta SSGC2 manteve a acurácia próxima à das referências utilizadas (SVM e SSGC). Na Figura 3 são apresentados os valores médios (e com arredondamento para inteiro mais próximo) para o número de referências (neurônios ou vetores de suporte) definidas pelo agente e que, efetivamente, deverão ser utilizados na implementação em hardware. A Figura 3.(a) apresenta os resultados para as bases de dados cujos vetores de características são compostos por detalhes (módulo de máximos locais) e a Figura 3.(b) apresenta os resultados para as bases de dados cujas componentes são formadas por coeficientes de aproximações. Pode ser observado que, para os conjuntos de dados numerados de 7 a 12, nos quais foram utilizados os coeficientes de aproximações, os resultados foram muito mais consistentes, tanto com relação à acurácia, quanto no número de referências obtidas, para todos os agentes classificadores utilizados. Este resultado já havia sido destacado em Machado (2016). Para estes conjuntos que utilizam coeficientes de aproximações, deve-se destacar que o valor médio (arredondado para inteiro mais próximo) foi de uma componente relevante selecionada.

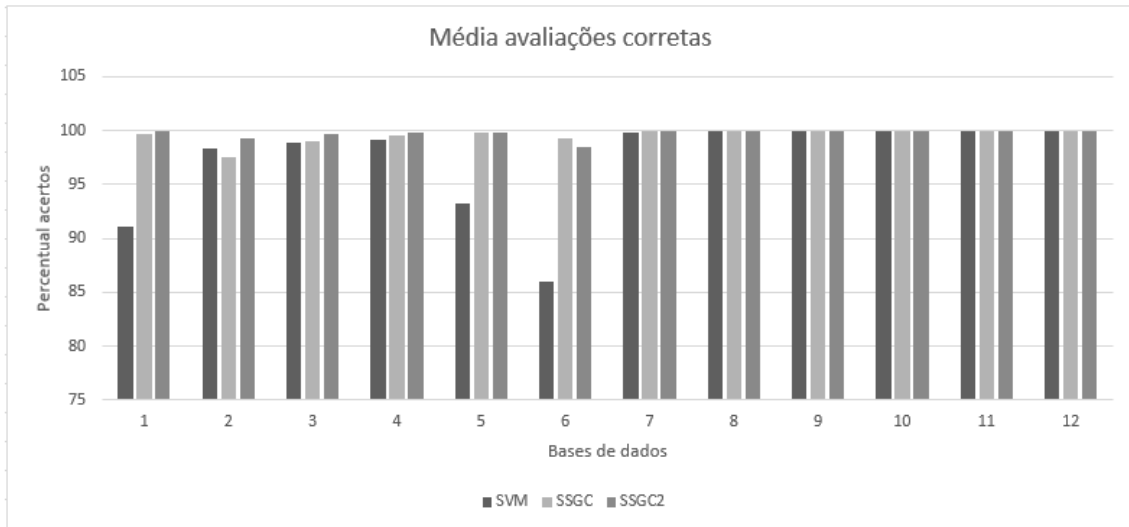


Figura 1. Resultados obtidos pelos classificadores – percentual acurácia.

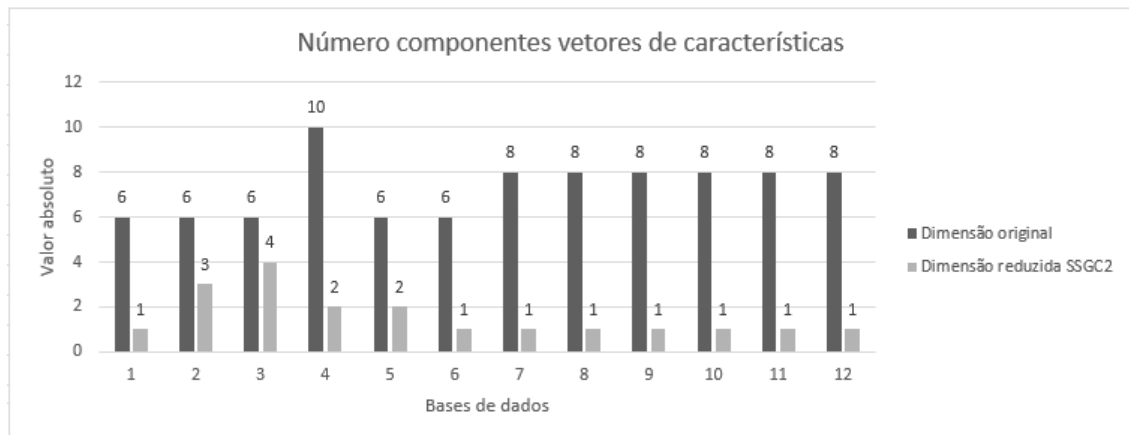


Figura 2. Resultados redução de dimensão obtidos pelo SSGC2.

Pela análise da Figura 3 pode ser observado que a redução de dimensionalidade implicou no aumento no número de referências necessárias para a correta identificação dos padrões de entrada do problema. Assim, se por um lado temos ganhos na redução das dimensões necessárias para caracterizar os vetores de características dos padrões de entrada, por outro, temos perdas com o aumento na quantidade de referências necessárias para correta classificação destes padrões. Na prática, as características do dispositivo de hardware a ser utilizado é que determinarão se o conjunto de referências com dimensão reduzida efetivamente pode ser ou não mais vantajoso para uso.

Na Figura 4 é apresentado um dos resultados práticos obtidos, com relação à eliminação de ruídos, utilizando-se o algoritmo SSGC2. Nesta Figura 4, que se refere à base de dados com rótulo 8 na Tabela I, podemos observar duas posições destacadas por setas. A primeira seta, identificada por “a” na Figura 4, tem-se uma parcela associada a uma ocorrência de descarga parcial e na Figura 5 temos um melhor detalhamento desta parcela do sinal. Na Figura 6 apresenta-se um melhor detalhamento da parcela indicada por “b” na Figura 4, em que temos uma posição associada com ruído impulsivo. Da mesma forma que em Machado (2016), uma vez identificada as ocorrências de DP, na reconstrução operou-se de forma a preservar as componentes de alta frequência nas posições associadas à frente de onda do sinal detectado. Estas posições são facilmente identificadas após a operação realizada pelo classificador. Nas demais posições, incluindo faixas associadas a ruídos, temos a eliminação de todos as componentes de detalhes.

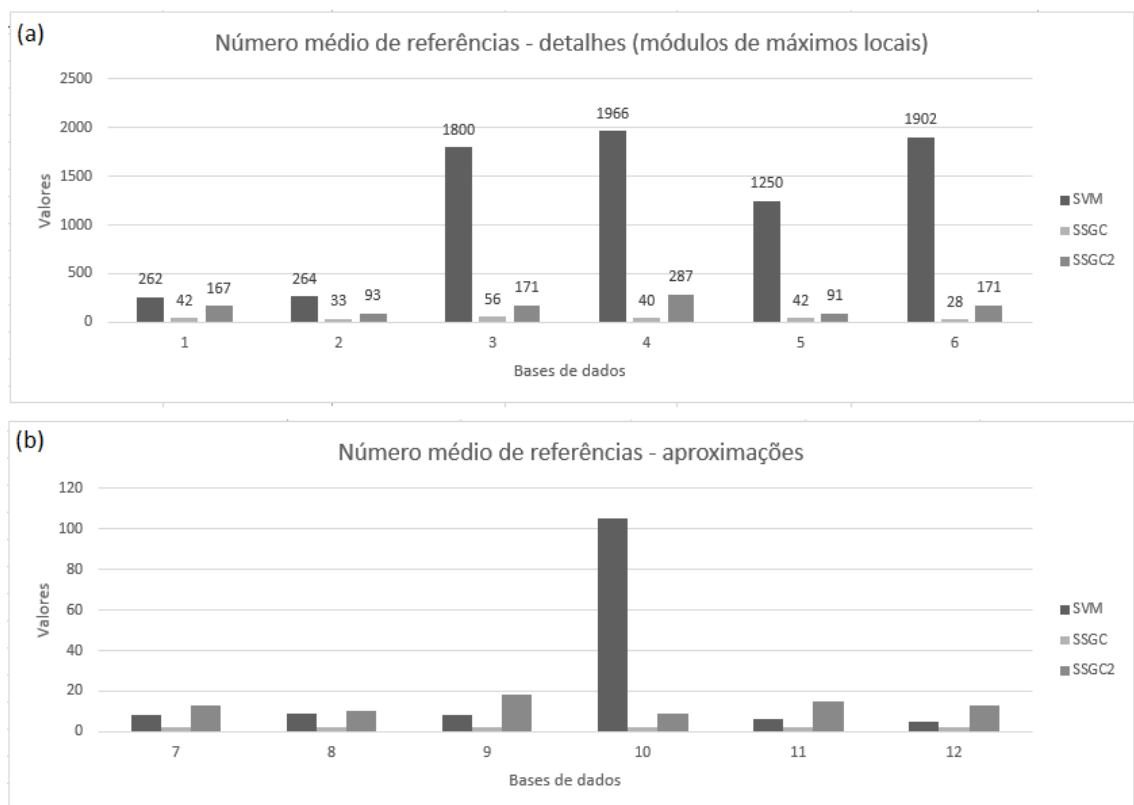


Figura 3. Número médio de referências definidas.

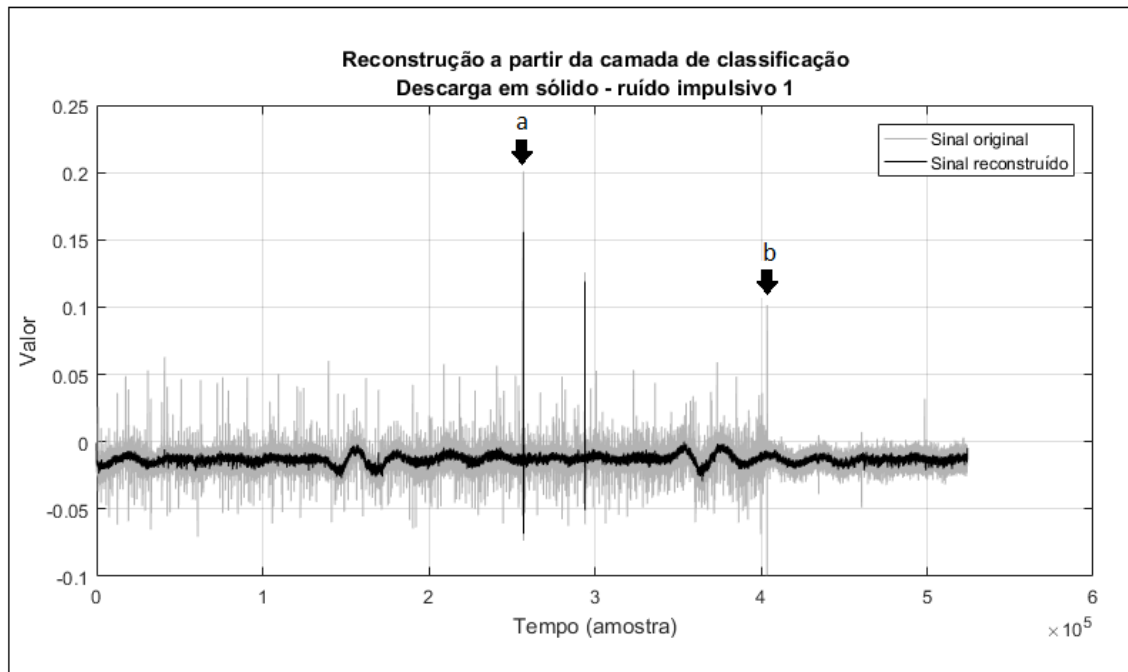


Figura 4. Exemplo de sinal reconstruído.

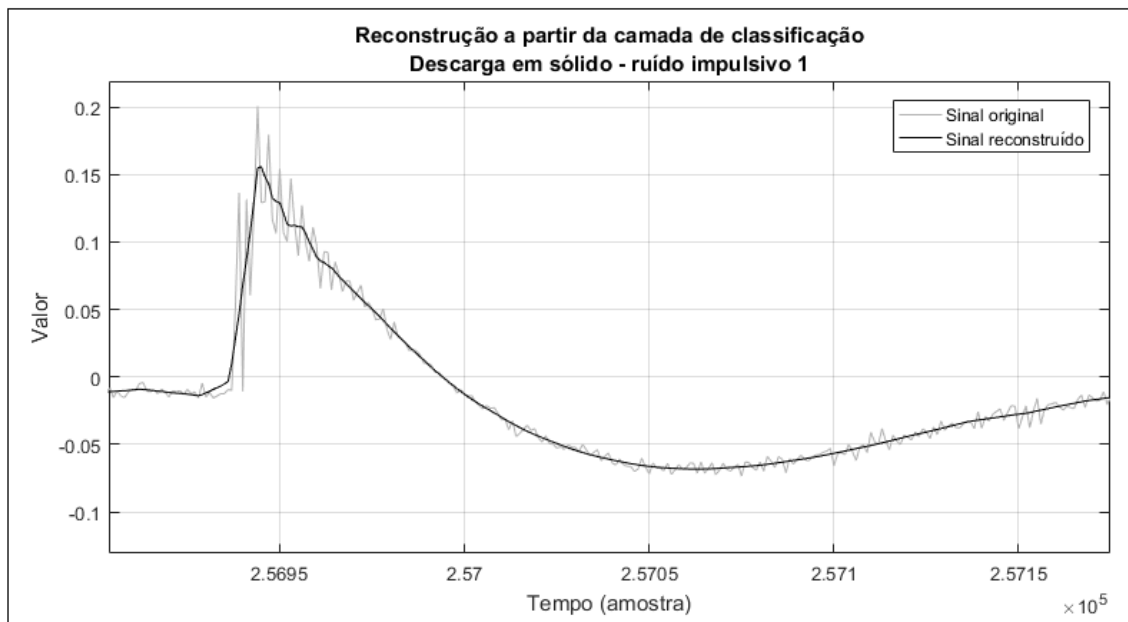


Figura 5. Exemplo de sinal reconstruído – detalhe posição “a” Figura 4.

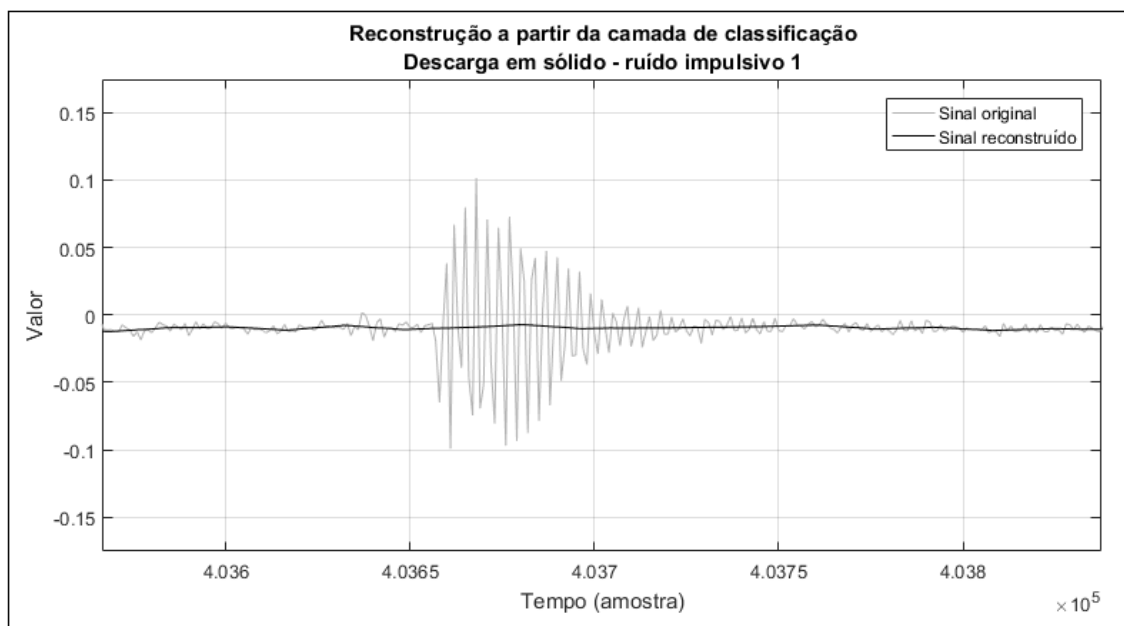


Figura 6. Exemplo de sinal reconstruído – detalhe posição “b” Figura 4.

5 CONCLUSÃO

Neste trabalho tomou-se como ponto de partida a proposta de algoritmo SSGC apresentada em Machado (2015), focado em separação de ruídos em meio a sinais amostrados para identificação de ocorrências de descargas parciais e, utilizando-se de algoritmo baseado em SVM e *Differential Evolution*, obteve-se procedimento para, de maneira autônoma, selecionar a melhor camada de decomposição, quando da execução da transformada de *wavelets*. Simultaneamente, também, pode-se selecionar quais componentes efetivamente são relevantes para identificação dos padrões de interesse. Os resultados obtidos mostraram que as camadas que permitem melhor separação entre os padrões de entrada, para o problema foco, são a quarta ou quinta, ou seja, os procedimentos de execução da transformada de *wavelets* não necessitam de ser tão profundos quanto os anteriormente implementados no método de referência. Isso, por si só, já representa algum ganho quando da implementação em hardware.

Verificou-se que a redução de dimensionalidade está associada, para o problema foco, a um aumento no número de referências a serem implantadas na implementação em hardware. A rigor, na prática, a utilização das referências e

SOUZA, D. L. S.; GUIMARÃES, J. L. C. G.; XISTO, J. B.; BARBOSA, W.C.; MACHADO, R. G. V. Sistema de tempo real para processamento de sinais de descargas parciais focado em implementação por hardware dimensões obtidas pelos métodos SSGC e SSGC2 deve ser selecionada de acordo com as características, recursos e/ou demandas do hardware a ser utilizado para implementação. A depender do hardware, um ou outro padrão dos vetores de entrada pode permitir melhores resultados para maximização de operações paralelas, viabilização ótima de *pipelining* e minimização de ciclos de *clock* para processamento dos dados, proporcionando uma maior taxa de amostragem. Continuou-se com uma estrutura que demanda operações simples (baseada em cálculo de distância Manhattan) para identificação dos padrões de entrada.

No presente trabalho, utilizou-se somente a acurácia como métrica de avaliação de desempenho da abordagem proposta e utilizou-se como referência resultados obtidos por agentes baseados em SVM e na versão original do algoritmo (SSGC). Os resultados obtidos, com a redução de dimensionalidade e seleção de camada de decomposição, mantiveram o mesmo desempenho de acurácia.

Por último, deve-se destacar que os procedimentos utilizados para eliminação de ruídos podem ser replicados para identificação de alterações nas formas de onda dos sinais de DP, o que permitiria identificação de condições de operação indicativa de defeito nos equipamentos sendo monitorados.

Embora não tenha sido implementado neste trabalho, outra melhoria que se apresenta é incluir no procedimento a seleção da melhor família de wavelets de acordo com as características dos conjuntos de dados utilizados para treinamento.

REFERÊNCIAS

ABDEL-GALIL, T., EL-HAG, A., VAOU DA, A., SALAMA, M., e BARTNIKAS, R., *De-noising of partial discharge signal using eigen-decomposition technique*, IEEE Trans. Dielectr. Electr. Insul., vol. 15, no. 6. p. 1657– 1662. Dec. 2008.

BARTNIKAS, R., *Partial discharges. Their mechanism, detection and measurement*, IEEE Trans. Dielectr. Electr. Insul., vol. 9, no. 5. p.763–808. 2002.

SOUZA, D. L. S.; GUIMARÃES, J. L. C. G.; XISTO, J. B.; BARBOSA, W.C.; MACHADO, R. G. V. Sistema de tempo real para processamento de sinais de descargas parciais focado em implementação por hardware

HUSSEIN, R., BASHIRSHABAN, K. e EL-HAG, A. H., *Denoising of acoustic partial discharge signals corrupted with random noise*. IEEE Transactions on Dielectrics and Electrical Insulation, v. 23, n. 3. p. 1453-1459. 2016.

KHAMSEH, H. B., RUELA, V. S. P., VASCONCELOS, F. H. e MOTA, H. de O., *Mining undecimated Wavelet Transform maxima lines!: An effective way to denoise Partial Discharge signals*, In: The 32nd IEEE Electrical Insulation Conference, Philadelphia. 2014.

KOHONEN, T., *The self-organizing map*, Proceedings of the IEEE International Conference on Neural Networks, San Francisco, p. 1147-1156. 1990.

MACHADO, R. G. V. e MOTA, H. D. O. *Simple self-scalable grid classifier for signal denoising in digital processing systems*. In: Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on. IEEE. p. 1-6. 2015.

MACHADO, R. G. V., Desenvolvimento de sistema de tempo real baseado em FPGAs para processamento de sinais de descargas parciais. Minas Gerais. 2016. 161f. Dissertação (Mestrado em Engenharia Elétrica) – PPGEE, Universidade Federal de Minas Gerais. Belo Horizonte.

MOTA, H. D. O., DA ROCHA, L. C. D., SALLES, T. C. D. M. e VASCONCELOS, F. H., *Partial discharge signal denoising with spatially adaptive wavelet thresholding and support vector machines*, Electr. Power Syst. Res., vol. 81, no. 2. p. 644–659. Feb. 2011.

NIMMO, R. D., CALLENDER, G. e LWIN, P. L., *Methods for wavelet-based autonomous discrimination of multiple partial discharge sources*. IEEE Transactions on Dielectrics and Electrical Insulation, v. 24, n. 2, p. 1131-1140. 2017.

RAYMOND, W. J. K. et al., *High noise tolerance feature extraction for partial discharge classification in XLPE cable joints*. IEEE Transactions on Dielectrics and Electrical Insulation, v. 24, n. 1. p. 66-74. 2017.

SETTLES, B., CRAVEN, M. e RAY, S. *Multiple instance active learning*, *Advances in Neural Information Processing Systems*, no. 20, MIT Press. p. 1289-1296. 2008.