

ANÁLISES DE TÉCNICAS DE SELEÇÃO DE PARÂMETROS APLICADAS À QUALIDADE DE ENERGIA

Fernando Miranda Vieira Xavier

Universidade Federal do Rio de Janeiro

fernandomvx@gmail.com

Diego Fagundes Fabri, Augusto Santiago Cerqueira, Carlos Augusto Duque

Universidade Federal de Juiz de Fora

diego.fabri@engenharia.ufjf.br, augusto.santiago@ufjf.edu.br, carlos.duque@ufjf.edu.br

Resumo – Este trabalho apresenta estudos de 4 técnicas de seleção de parâmetros relacionados a distúrbios de qualidade de energia elétrica (QEE): Teste de Relevância das Características (TRC), Discriminante de Fisher (DF), a Análise das Componentes Principais (PCA) e as Componentes Principais de Discriminação (PCD). Para a extração dos parâmetros representativos dos distúrbios é utilizada Estatística de Ordem Superior (HOS). A PCA e PCD são utilizadas para representar os parâmetros extraídos num outro espaço, com reduzida dimensão. O TRC e o DF são empregados para selecionar os parâmetros de maior caráter discriminatório. Por fim, os parâmetros selecionados são usados como entrada de classificadores especialistas baseado em redes neurais, de forma a verificar a eficiência de classificação.

Palavras-chave – Qualidade de Energia, Seleção de Parâmetros, Análise de Componentes Principais, Componentes Principais de Discriminação, Discriminante de Fisher

Abstract – This paper presents four studies of techniques for feature selection related to disturbances of power quality (QEE): Relevance of Characteristics Test (TRC), the Fisher Discriminant (DF), the Principal Component Analysis (PCA) and Components Principal of Discrimination (PCD). It is used to Higher Order Statistics (HOS) for the extraction of representative parameters of the disturbances. The PCA and PCD are used to represent the parameters extracted in another space with reduced dimension. The TRC and DF are used to select the parameters of greater discriminatory character. Finally, the selected parameters are used as input of experts classifiers based on neural networks in order to verify the classification efficiency.

Keywords – Power Quality, Feature Selection, Principal Components Analysis, Principal Component of Discrimination, Fisher Discriminant

1. Introdução

O questão da qualidade da energia elétrica (QEE) vem tendo crescente importância visto que os distúrbios no sistema elétrico podem causar problemas significantes tanto para o usuário quanto para as concessionárias. Estes distúrbios são causados principalmente por dispositivos chaveados, cargas não-lineares, sistemas desbalanceados, dispositivos microprocessados, retificadores e inversores [1]. A crescente utilização destes dispositivos têm causado diversos distúrbios nas correntes e tensões, como harmônicos, transitórios, elevações (*swell*), afundamentos (*sag*) e interrupções. Estes distúrbios podem causar mau funcionamento e redução da durabilidade dos equipamentos elétricos, o que pode levar a grandes perdas econômicas para a indústria e para os consumidores residenciais.

Desta maneira, o monitoramento e controle dos sistemas de energia têm se tornado indispensáveis. Diversas ferramentas vêm sendo pesquisadas e desenvolvidas e várias técnicas baseadas em processamento digital de sinais vêm sendo utilizadas em aplicações de detecção e classificação de distúrbios. Os problemas de classificação de distúrbios encontram-se num ramo conhecido como reconhecimento de padrões. Este ramo pode ser dividido em três estágios distintos: a extração dos parâmetros, a seleção dos parâmetros e a classificação [2].

A extração de parâmetros pode ser entendida como uma operação de pré-processamento que transforma um padrão de sua forma original para uma nova forma adequada ao processo [3]. Existem diversas técnicas de extração de parâmetros. Em qualidade de energia, a estatística de ordem superior (HOS) tem sido fortemente empregada [4–6]. Outras técnicas bem utilizadas são a Transformada de Fourier [7], a Transformada S [8] e a Transformada Wavelet [9].

Os parâmetros extraídos devem ser selecionados e fornecidos como entradas de sistemas que têm como objetivo a classificação do sinal de entrada. Dentre as técnicas utilizadas para o estágio de classificação, pode-se destacar a Lógica Fuzzy [10] e as Redes Neurais Artificiais [3, 9, 11], largamente utilizadas em problemas de reconhecimento de padrões.

A seleção de parâmetros é um estágio de vital importância no processo de classificação, mas muitas vezes tem sido negligenciada. O conjunto de parâmetros extraídos pode ter dimensões muito elevadas, e muitas vezes estes parâmetros podem

fornecer informações redundantes ou irrelevantes [2]. Sendo assim, há a necessidade de técnicas que permitam selecionar adequadamente os parâmetros a serem fornecidos como entradas do estágio de classificação. Apesar de encontrar-se na literatura diversas técnicas de seleção de parâmetros, poucos estudos têm sido apresentados na área de qualidade de energia e classificação de distúrbios. Dentre algumas técnicas de seleção pode-se destacar a função discriminante de Fisher (DF) [12, 13], os algoritmos genéticos [1], as redes neurais artificiais [2, 14], a análise de componentes principais (PCA) e as componentes principais de discriminação (PCD) [15].

Este trabalho apresenta um estudo comparativo de técnicas de seleção aplicadas a um conjunto de parâmetros extraído através das estatísticas de ordem superior. Quanto às técnicas de seleção, serão utilizadas a PCA, a PCD, o discriminante de Fisher e o teste de relevância baseado no classificador perceptron [16]. Para ilustrar e comparar a eficiência dos métodos de seleção de parâmetros, é realizada uma aplicação de classificação de distúrbios.

O artigo é dividido da seguinte forma: na Seção 2 é apresentado o estágio de extração dos parâmetros representativos dos distúrbios; a Seção 3 apresenta o estágio de seleção de parâmetros e as técnicas implementadas são descritas; na Seção 4 os parâmetros selecionados pelos diferentes métodos são usados como entrada de classificadores especialistas baseado em redes neurais artificiais; por fim, na Seção 5 são apresentadas as conclusões.

2 Extração de Parâmetros para Análise de Distúrbios de QEE

Os parâmetros extraídos para a análise de distúrbios utilizados neste trabalho são baseados na HOS. Mais precisamente, os parâmetros são os cumulantes de 2ª ordem, relativos ao sinal sem a componente fundamental [4]. A equação para o cálculo do cumulante de 2ª ordem é dada por:

$$\hat{c}_{2,z}(i) := \frac{1}{N} \sum_{n=0}^{N-1} z(n) z[n \bmod (n+i, N)], \quad (1)$$

onde i é o i -ésimo atraso, $z(n)$ é o n -ésimo elemento do vetor \mathbf{z} e $\bmod(n+i, N)$ é o resto inteiro da divisão de $n+i$ por N .

Serão considerados os seguintes eventos de QEE: elevação (*swell*) e afundamento (*sag*), harmônicos, interrupções, transitórios e flutuações (ver Figura 1). Devido à similaridade do sinal do distúrbio sem a componente fundamental, elevação e afundamento serão tratados como uma única classe.

Os eventos utilizados aqui foram gerados sinteticamente seguindo o padrão *IEEE Recommended Practice for Monitoring Electric Power Quality* (1995), com 1024 amostras, sendo 256 amostras por ciclo.

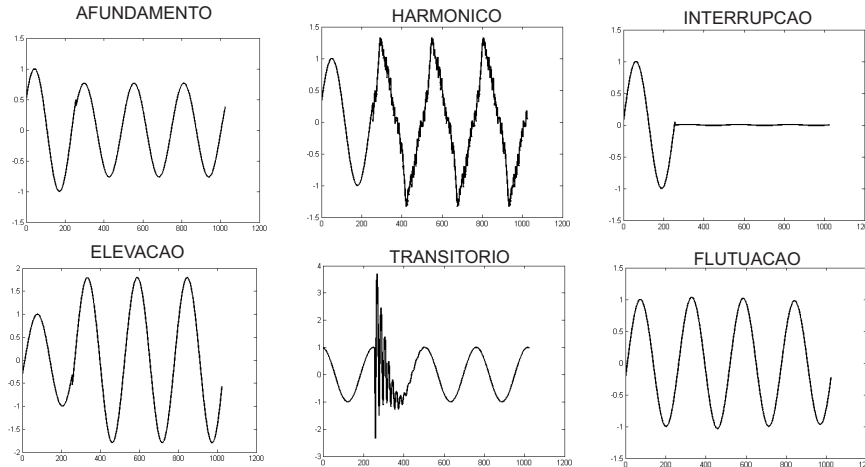


Figura 1: Exemplos de eventos de distúrbios.

3 Seleção de Parâmetros

A seleção de parâmetros é um procedimento que visa representar os dados originais em um número reduzido de dimensões. Além disso, pretende-se obter uma melhor representação dos dados, remover informações redundantes (compactar informação), ou até mesmo otimizar o conjunto de parâmetros para uma determinada aplicação.

Neste trabalho, duas técnicas são usadas para selecionar diretamente as características de maior caráter discriminatório dentro do espaço original dos dados: uma baseada num teste de relevância (TRC) e a outra baseada no discriminante de Fisher (DR). Além destas, outras duas técnicas são também empregadas, tendo como característica a representação dos dados num novo espaço. São elas, as componentes de discriminação (PCD) e as componentes principais (PCA).

3.1 Relevância das Características - TRC

A técnica da análise de relevância tem por objetivo medir a importância de cada uma das variáveis de entrada para um classificador. Nesta técnica, avalia-se a contribuição da variável à composição da saída substituindo-se seu valor, a cada evento, pelo valor médio da variável com respeito a todos os eventos disponíveis na entrada do sistema de discriminação. Quanto maior a variação (valor absoluto da medida de relevância), mais relevante é considerada a variável. Assim sendo, observando-se a variação da saída é possível estimar a contribuição daquela variável ao processo discriminatório.

A Figura 2 mostra a variação do desempenho para o classificador perceptron usado para o teste de relevância, tendo como entrada os 1024 cumulantes de 2ª ordem.

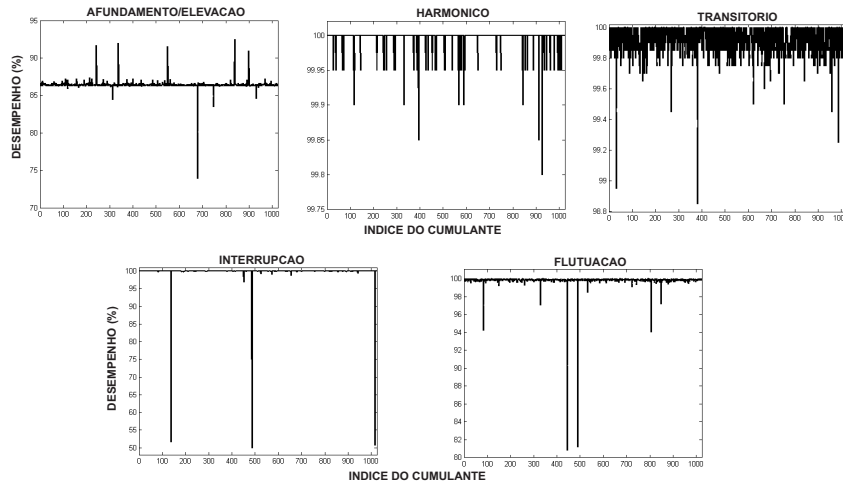


Figura 2: Teste de relevância para os 1024 cumulantes.

3.2 Discriminante de Fisher - DF

O discriminante de Fisher é uma técnica largamente empregada em reconhecimentos de padrões e utilizada para quantificar a discriminação entre classes, parâmetro a parâmetro [17]. O objetivo do uso do DF é encontrar os parâmetros que melhor separam duas classes, através da maximização da função custo J_c , dada pela Equação 2,

$$J_c = (\mathbf{m}_1 - \mathbf{m}_2)^2 \odot \frac{1}{\mathbf{D}_1^2 + \mathbf{D}_2^2} \quad (2)$$

onde $\mathbf{J}_c = [J_1 \dots J_{Ll}]^T$, Ll é o número total de parâmetros, \mathbf{m}_1 e \mathbf{m}_2 , e \mathbf{D}_1^2 e \mathbf{D}_2^2 são, respectivamente, os vetores de média e variância dos vetores de parâmetros referentes às classes 1 e 2. O símbolo \odot refere-se ao produto de Hadamard [17].

O i -ésimo elemento do vetor de parâmetros \mathbf{J}_c de maior valor é considerado de maior caráter discriminatório. A Figura 3 ilustra o vetor \mathbf{J}_c calculado para os 1024 cumulantes.

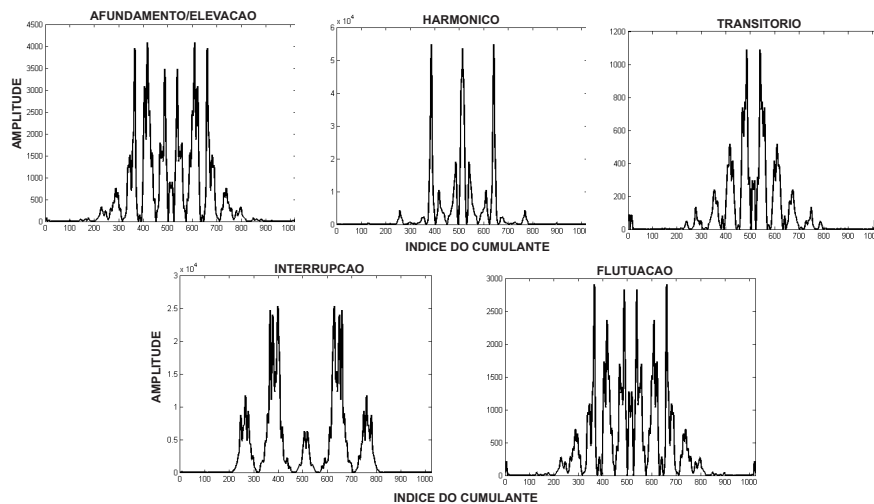


Figura 3: J_c para os 1024 cumulantes.

3.3 Análise de Componentes Principais - PCA

A análise de componentes principais foi desenvolvida por Karl Pearson, em 1901. Seu objetivo é encontrar a melhor representação para os dados no espaço de parâmetros. Ela pode ser geometricamente compreendida como uma rotação dos eixos do sistema de coordenadas original para um novo conjunto de eixos ortogonais ordenados em termos da variância dos dados originais [18]. A Figura 4 mostra as 2 PCAs para um conjunto de dados formado pela região de elipse.

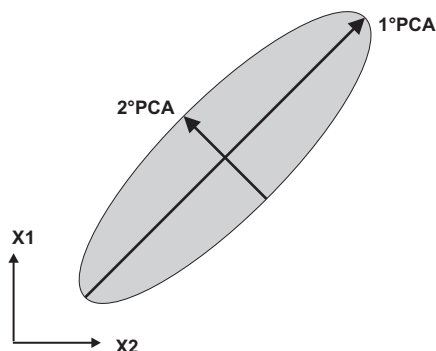


Figura 4: Extração das PCAs.

Seja então uma matriz \mathbf{X} $m \times n$, onde m é o número de parâmetros e n é o número de eventos. Cada linha da matriz deve ter média zero, ou seja, a média de cada parâmetro deve ser nula. Através de uma transformação linear, a matriz \mathbf{X} é transformada em uma nova matriz \mathbf{Y}

$$\mathbf{Y} = \mathbf{P}\mathbf{X}. \quad (3)$$

Através desta nova matriz \mathbf{Y} pretende-se minimizar a redundância entre os parâmetros e maximizar o sinal. O processo de redução de redundância é feito através da minimização da covariância entre os parâmetros. Ou seja, deseja-se que a matriz de covariância de \mathbf{Y} (supondo média nula) seja diagonal, onde

$$\mathbf{C}_Y \equiv \frac{1}{n-1} \mathbf{Y}\mathbf{Y}^T. \quad (4)$$

O processo de diagonalização baseia-se na propriedade de decomposição em autovetores. Deve-se então obter uma matriz \mathbf{P} ortonormal, de forma que a transformação linear dada pela Equação (3) gere uma matriz de covariância diagonal, conforme a Equação (4). Para isso, as linhas da matriz \mathbf{P} devem ser os autovetores de $\mathbf{X}\mathbf{X}^T$, ordenados de forma crescente pelos seus autovalores. As linhas de \mathbf{P} são as chamadas componentes principais de \mathbf{X} e a variância da projeção de \mathbf{X} em \mathbf{P} (energia) pode ser usada para verificar o nível da representação dos dados originais \mathbf{X} .

3.4 Componentes Principais de Discriminação - PCD

Diferentemente da PCA, a PCD busca por direções que possam maximizar a diferença entre as classes, ou seja, direções que maximizem a eficiência de discriminação. Essa busca visa identificar sutis diferenças entre estas classes.

Neste trabalho, as componentes principais de discriminação são extraídas utilizando redes neurais artificiais. Uma rede com um único neurônio na camada escondida e um neurônio na saída é treinada para maximizar a discriminação das classes, onde os pesos sinápticos que ligam a entrada à camada intermediária correspondem a primeira componente de discriminação.

A partir da 1ª PCD \mathbf{w} , é então realizada a subtração da projeção dessa componente nos dados originais \mathbf{X} , processo aqui denominado de deflação [19]. Agora, com os dados sem a projeção da 1ª PCD (\mathbf{X}_N), aplica-se novamente os métodos de extração das PCDs, obtendo-se, assim, a segunda componente de discriminação. Este processo de deflação é repetido até a extração de N PCDs, em que N é igual ou menor ao número de parâmetros dos dados e o processo é representado pelo diagrama de blocos da Figura 5.

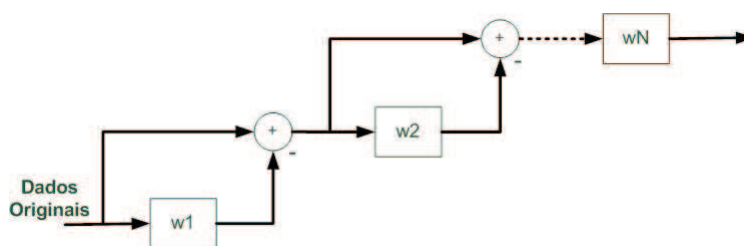


Figura 5: Processo de Deflação.

4 Aplicação - Classificação de Distúrbios

Nesta seção, as técnicas de seleção de parâmetros são aplicadas ao problema de classificação de distúrbios. É importante salientar que o objetivo deste trabalho não é obter o melhor classificador e sim ilustrar aplicações das técnicas apresentadas para seleção de parâmetros.

A técnicas de seleção de parâmetros através da PCA ou PCD pode ser aplicada em problemas que utilizam extração de parâmetros através HOS, de wavelets, etc, ou até mesmo podem ser diretamente aplicadas como ferramentas de extração.

O sistema de classificação de distúrbio usado neste trabalho pode ser dividido em 4 estágios:

1. Estágio de extração de parâmetros: cálculo dos cumulantes de 2ª ordem.
2. Estágio da seleção de parâmetros: aplicação das técnicas de PCA, PCD, DF e TRC.
3. Estágio de classificação: utilização de redes neurais artificiais para a classificação.
4. Estágio de decisão: decisão final do sistema de classificação.

O estágio de classificação consiste de 5 redes neurais que discriminam entre uma classe e todas as outras, ou seja, cada rede compara um tipo de distúrbio com todos os outros 4. O sistema de decisão dá a resposta final do classificador. O banco de dados usado para a verificação dos desempenhos dos classificadores é formado por conjuntos de treino e teste, onde cada conjunto consiste de 1000 eventos de uma classe e 1000 eventos das outras classes restantes, divididos igualmente.

Todos os classificadores empregados (baseado em redes neurais), apresentam o mesmo número de neurônios na camada intermediária (4) e um único neurônio da camada de saída, com a tangente hiperbólica como função de ativação.

A primeira aplicação foi extrair os 1024 cumulantes de 2ª ordem e extrair as 2 primeiras PCAs e PCDs. Os resultados dos classificadores podem ser vistos nas tabelas 1 e 2.

Tabela 1: Desempenho dos classificadores referente à extração das PCAs dos 1024 cumulantes.

Nº DE COMPONENTES BANCO	1			2		
	TREINO	TESTE	ENERGIA (%)	TREINO	TESTE	ENERGIA (%)
<i>SWELL/SAG</i>	74,8%	74,8%	55,6	97,6%	96,9%	82,3
<i>HARMONICO</i>	90,7%	89,3%	71,3	96,8%	95,8%	85,5
<i>INTERRUPCAO</i>	93,4%	93,2%	55,1	99,8%	99,8%	85,2
<i>TRANSITORIO</i>	94,0%	94,4%	16,0	96,2%	96,4%	31,2
<i>FLUTUACAO</i>	90,4%	91,5%	86,3	99,9%	99,9%	96,7

Tabela 2: Desempenho dos classificadores referente à extração das PCDs dos 1024 cumulantes.

Nº DE COMPONENTES BANCO	1		2	
	TREINO	TESTE	TREINO	TESTE
<i>SWELL/SAG</i>	85,8%	85,6%	99,1%	98,3%
<i>HARMONICO</i>	100,0%	100,0%	100,0%	100,0%
<i>INTERRUPCAO</i>	99,1%	99,3%	99,3%	99,2%
<i>TRANSITORIO</i>	94,4%	94,4%	99,7%	99,5%
<i>FLUTUACAO</i>	93,7%	93,6%	98,8%	98,2%

Tabela 3: Desempenho dos classificadores usando os 7 mais relevantes cumulantes como entrada.

METRICA BANCO	TRC		DF	
	TREINO	TESTE	TREINO	TESTE
<i>SWELL/SAG</i>	99,7%	97,9%	95,7%	93,4%
<i>HARMONICO</i>	98,9%	96,6%	100,0%	100,0%
<i>INTERRUPCAO</i>	99,9%	99,7%	99,2%	99,0%
<i>TRANSITORIO</i>	98,8%	96,8%	87,5%	85,9%
<i>FLUTUACAO</i>	99,8%	99,0%	98,4%	97,4%

Assim, a partir dos resultados das tabelas 1 e 2 conclui-se que duas componentes são suficientes para uma ótima eficiência de classificação. Porém, é necessária a extração dos 1024 cumulantes, que, do ponto de vista de uma possível classificação *on-line* e por razões de esforço computacional, este método torna-se praticamente inviável.

Para sanar este problema, foram utilizados o teste de relevância (ver Figura 2) e DF (ver Figura 3) para a seleção dos cumulantes de maior caráter discriminatório. Usando agora os 7 mais relevantes cumulantes como entrada dos classificadores, é gerada a Tabela 3.

Logo, baseado nos resultados da Tabela 3, nota-se também uma ótima eficiência de classificação para ambos os métodos de seleção dos cumulantes.

5 Conclusões

Neste trabalho são apresentadas técnicas de seleção de parâmetros aplicadas à classificação de distúrbios de QEE. A seleção é feita a partir dos parâmetros extraídos usando HOS, mediante o cálculo dos cumulantes de 2ª ordem. Duas destas técnicas, teste de relevância das características (TRC) e discriminante de Fisher (DF), são empregadas para selecionar os parâmetros de maior caráter discriminatório. Já a análise das componentes principais (PCA) e as componentes principais de discriminação (PCD) são usadas com intuito de uma melhor representação dos dados, reduzindo o conjunto de parâmetros a serem analisados.

Para verificar o efeito dos parâmetros selecionados na classificação de distúrbios, foram utilizadas redes neurais artificiais como classificadores especialistas para cada tipo de distúrbio. Os resultados mostraram que tanto a seleção de parâmetros por PCA e PCD quanto pelo TRC e DF ofereceram bons resultados, reduzindo o número de parâmetros e obtendo-se um bom desempenho de classificação.

Assim, como propostas futuras, pretende-se trabalhar na aplicação dos métodos de seleção em novos parâmetros extraídos dos distúrbios, e também explorar com mais detalhes as propriedades da PCA e PCD em outras aplicações.

REFERÊNCIAS

- [1] B. K. Panigrahi and V. R. Pandi. “Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm”. *IET Generation, Transmission and Distribution*, vol. 3, pp. 296–306, 2009.
- [2] A. Verikas, M. Bacauskiene and K. Malmqvist. “Selecting Features for Neural Network Committees”. In *Proceedings of International Joint Conference on Neural Networks*, volume 25, pp. 1879–1891, 2002.
- [3] A. M. Gaouda, S. H. Kanoun, M. M. A. Salama and A. Y. Chikhani. “Pattern Recognition Applications for Power System Disturbance Classification”. *IEEE Transactions on Power Delivery*, vol. 17, pp. 677–683, July 2002.
- [4] D. D. Ferreira, C. A. G. Marques, A. S. Cerqueira, C. A. Duque and M. V. Ribeiro. “Sistema Automático de Detecção e Classificação de Distúrbios Elétricos em Qualidade de Energia Elétrica”. *Sba: Controle e Automação Sociedade Brasileira de Automática*, vol. 20, pp. 53 – 62, 03 2009.
- [5] D. D. Ferreira and A. S. Cerqueira. “HOS-based method for classification of power quality disturbances”. *Electronics Letters*, vol. 45, no. 3, pp. 183 –185, 29 2009.
- [6] M. V. Ribeiro, C. A. Marques, C. A. Duque, A. S. Cerqueira and J. L. R. Pereira. “Power quality disturbances detection using HOS”. In *Power Engineering Society General Meeting, 2006. IEEE*, p. 6 pp., 0-0 2006.
- [7] G. T. Heydt, P. S. Fjeld, C. C. Liu, D. Pierce, L. Tu and G. Hensley. “Applications of the Windowed FFT to Electric Power Quality Assessment”. *IEEE Transactions on Power Delivery*, vol. 14, no. 4, pp. 1411–1416, October 1999.
- [8] M. Zhang, K. Li and W. Hu. “Automated Classification of Power Quality Disturbances Using the S-Transform”. In *Proceedings of the 2008 International Conference on Wavelets Analysis and Pattern Recognition, Hong Kong, August 2008*.
- [9] B. Perunicic, M. Mallini, Z. Wang and Y. Liu. “Power Quality Disturbance Detection , Classification Using Wavelets , Artificial Neural Networks”. In *Proceeding of 8th International Conference on Harmonics and Quality of Power*, pp. 77–82, October 1998.
- [10] B. Bizjak and P. Planinsic. “Classification of Power Disturbances using Fuzzy Logic”. In *Proceedings of 12th International Power Electronics and Motion Control Conference, EPE-PEMC, 2006*.
- [11] B. Perunicic, M. Mallini, Z. Wang and Y. Liu. “Power Quality Disturbance Detection and Classification Using Wavelets and Artificial Neural Networks”. In *Proceeding of 8th International Conference on Harmonics and Quality of Power*, pp. 77–82, October 1998.
- [12] M. Wang, G. I. Rowe and A. V. Mamishev. “Classification of Power Quality Events Using Optimal Time-Frequency Representations - Part 2: Application”. *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1496–1503, July 2004.

- [13] P. R. Manke and S. B. Tembhurne. “Artificial Neural Network Classification of Power Quality Disturbances using Time-Frequency Plane in Industries”. In *Proceedings of First International Conference on Emerging Trends in Engineering and Technology*, pp. 564–568, 2008.
- [14] V. Onnia, M. Tico and J. Saarinen. “Feature Selection Method Using Neural Network”. In *Proceedings of International Conference on Image Processing*, pp. 513–516, 2001.
- [15] D. F. Fabri, F. M. V. Xavier, A. S. Cerqueira and C. A. Duque. “Análise de Componentes Principais e de Discriminação aplicadas à Qualidade de Energia Elétrica”. In *XVIII Congresso Brasileiro de Automática (CBA 2010)*, pp. 2099–2105, 2010.
- [16] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1998.
- [17] R. O. Duda, P. E. Hart and D. G. Stork. *Pattern Classification*. Wiley - Interscience, second edition, 2000.
- [18] A. Webb. *Statistical Pattern Recognition*. Wiley, 2002.
- [19] L. P. Calôba, J. M. Seixas and F. S. Pereira. “Neural Discriminating Analysis For a Second-Level Trigger System”. In *Proceedings of the International Conference on Computing in High Energy Physics*, September 1995.