

ANÁLISE DO EFEITO DA PRECISÃO FINITA NO ALGORITMO ADAPTATIVO SIGMOIDAL

José de R. S. Fonseca

Ewaldo E. C. Santana

André C. B. Cavalcante

Allan K. Barros

ribamar.slz@hotmail.com

UFMA - Universidade Federal do Maranhão.

Av. dos Portugueses, 1966, CEP 65080-805, São Luís-MA, Brasil

Raimundo C. S. Freire

UFCG - Universidade Federal de Campina Grande

Rua Aprigio Veloso, 882 - B. Univ., CEP 58429-900, Campina Grande-PB, Brasil

Resumo. A filtragem adaptativa constitui atualmente uma ferramenta importante no processamento estatístico de sinais, especialmente quando é necessário processar sinais de ambientes com estatísticas desconhecidas que variam com o tempo. O estudo de filtragem adaptativa foi impulsionado com o desenvolvimento do algoritmo Least Mean Square (LMS) em 1960. Novos algoritmos tem surgido, entre eles, o algoritmo Sigmoidal (SA) que se apresentou superior ao LMS, em relação à taxa de convergência e o desajuste em implementação em precisão infinita. Nos dispositivos de hardwares, tais como DSPs, Microcontroladores e FPGAs, os algoritmos adaptativos são implementados em precisão finita, em geral, com aritmética de ponto fixo. Alguns efeitos nessa implementação podem afetar o seu desempenho. Em última análise, levar à divergência devido aos erros de quantização, especificados na aproximação dos valores das variáveis de seus valores originais. Este artigo propõe analisar o desempenho do algoritmo adaptativo Sigmoidal (SA) em precisão finita, quando implementado utilizando aritmética de ponto fixo. Em particular, analisar sua curva de desempenho e o desajuste, comparando-os em diferentes comprimentos de palavras (número de bits). Os resultados apresentados neste artigo propõe uma aproximação em série de Taylor do gradiente da função de custo $\ln(\cosh e)$ do algoritmo SA para implementação em precisão finita. Analisamos a sua curva de desempenho para diferentes comprimentos de palavras. Mostra-se que o algoritmo apresenta estabilidade em seu desempenho em relação à convergência, para diferentes comprimentos de palavras, e que o aumento no nível do desajuste em estado estacionário é sensível ou influenciado pela quantização dos valores das variáveis envolvidas nos cálculos desse algoritmo.

Palavras-chave: Filtros Adaptativos, Aritmética de ponto-fixo, Sigmoidal

1 INTRODUÇÃO

Erros de quantização podem degradar o desempenho de um filtro adaptativo de várias maneiras. Por exemplo, eles podem afetar a estabilidade do filtro e, em última análise, conduzir à sua divergência. Também podem degradar o desempenho em estado estacionário do filtro fazendo-o atingir um erro médio quadrático maior do que o esperado a partir de uma análise de precisão infinita (Gupta et al, 2000). A degradação do desempenho tende a ser mais grave para o algoritmo *Recursive Least Square* (RLS) convencional podendo ser mais sensível a erros de precisão finita e divergir como resultado (Sayed, 2013) em oposição ao algoritmo *Least Mean Square* (LMS) (Gupta et al, 2000). Isto porque para filtros do tipo LMS, erros gradiente de ruído são mais significativos do que os erros de precisão finita durante a fase transitória. No entanto, em estado estacionário, um significativo excesso de erro quadrático pode ser formado (Verhaegen, 1989). Trabalhos recentes foram publicados em que analisam filtros adaptativos em precisão finita utilizando aritmética de ponto fixo, como em (Maluenda, 2005) que apresenta nova modelagem para o algoritmo LMS, (Meher, 2014) traz uma arquitetura eficiente para o algoritmo LMS e deriva expressões de erros de estados. Algoritmos de filtragem adaptativa com base em erro não-médio-quadrático tem sido sugeridos por vários autores, quer como uma forma de melhorar o desempenho de adaptação destes algoritmos para ambientes estatísticos particulares, ou como forma de simplificar o cálculo do algoritmo em hardware (Douglas Meng, 1994). Desta forma propomos uma análise de um algoritmo baseado em erro não-médio-quadrático dado em precisão finita, denominado Sigmoidal (SA) (Santana, 2006).

2 FILTRAGEM ADAPTATIVA E OS EFEITOS DA PRECISÃO FINITA

Considere o diagrama em blocos mostrado na Figura 1, referente ao problema de identificação de sistemas em filtragem adaptativa.

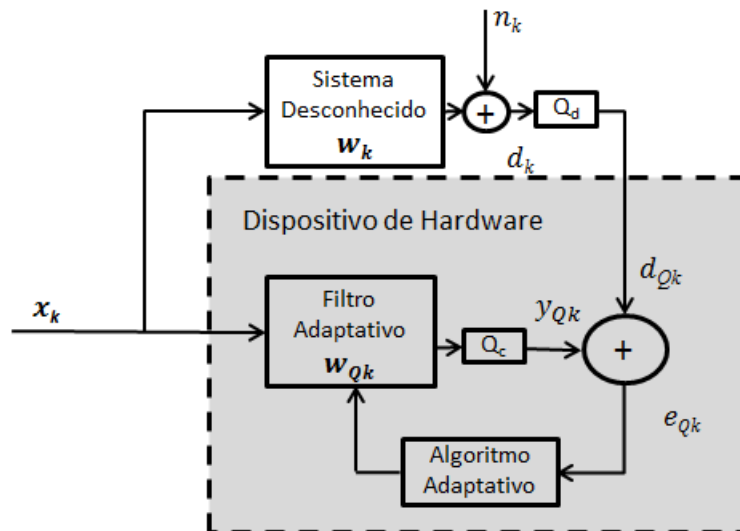


Figura 1: Identificação de sistema implementado em precisão finita. Onde x_k designa o sinal de entrada. Os blocos Q_d e Q_c são quantizadores de b bits. Assim, o quantizador, Q_d , usa b_d bits para os dados de entrada e Q_c quantiza os coeficientes em b_c bits. O sinal de saída y_{Qk} indica a forma quantizada. O sinal desejado definido por d_k do sistema desconhecido e d_{Qk} sua forma quantizada. O ruído de medição n_k e o sinal de erro quantizado e_{Qk} . Os coeficientes w_k do sistema desconhecido e w_{Qk} do filtro adaptativo. (Gupta, 2000)

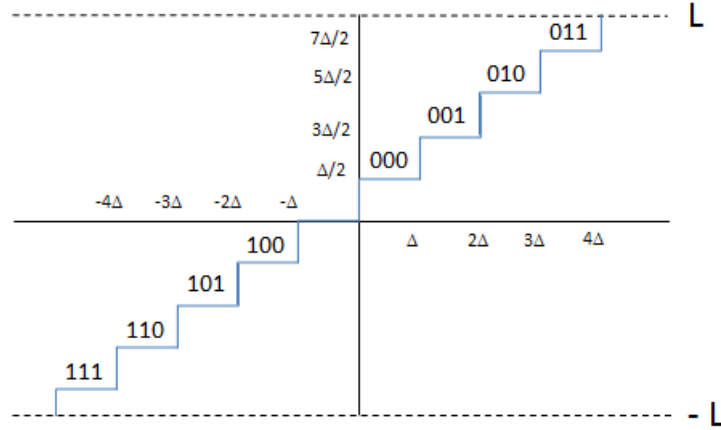


Figura 2: Representação esquemática de níveis de quantização para o caso de $b = 2$ bits e $\Delta = L/4$. Consideramos um valor L limiar. Usando b bits, o intervalo de $[0, L]$ pode ser dividido em 2^b níveis de largura $\Delta = L/2^b$ cada. Se, além disso, usamos 1 (um) bit de sinal, os resultantes $(b + 1)$ bits divide o intervalo $[-L, L]$ em 2^{2b} níveis. Aqui assumimos operação de arredondamento. (Sayde, 2013)

O filtro adaptativo Sigmoidal (SA) tem como função de custo $F_k(e) = \cosh(\alpha e)$ aplicada sobre o erro. Essa função é uma não linearidade par, contínua, simétrica e não tem mínimo local, apenas o mínimo global (Santana, 2006).

A partir desta função, pode-se gerar uma família de funções, $\cosh(\alpha e)$, multiplicando o argumento e por um inteiro positivo α , ou seja, quando aumenta α , aumenta-se a inclinação da superfície de desempenho, onde as curvas destas funções tem inclinações maiores do que a curva da função quadrática do algoritmo Least Mean Square (LMS) (Santana, 2006). Para desenvolver o algoritmo SA, estimamos o gradiente de $F_k(e)$. Onde ficamos com a Eq. (1):

$$\hat{\nabla} F_k(e) = -\alpha \tanh(\alpha e) \mathbf{x}_k \quad (1)$$

Logo o algoritmo adaptativo será dado pela Eq. (2):

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu \alpha \tanh(\alpha e) \mathbf{x}_k \quad (2)$$

A Equação (2) caracteriza o algoritmo adaptativo Sigmoidal (SA). Onde o μ é uma constante que regula a velocidade e a estabilidade da adaptação (Santana, 2006).

Considerando a Figura 1, e a estrutura do algoritmo SA, as equações que definem a atualização dos coeficientes do filtro adaptativo em precisão finita são:

$$d_{Qk} = (d_k + n_k)_Q \quad (3)$$

$$y_{Qk} = (\mathbf{w}_{Qk} \mathbf{x}_k)_Q \quad (4)$$

$$e_{Qk} = (d_{Qk} - y_{Qk})_Q \quad (5)$$

$$\mathbf{w}_{Q(k+1)} = (\mathbf{w}_{Qk} + \mu \alpha \tanh(\alpha e_{Qk}) \mathbf{x}(k))_Q \quad (6)$$

Note que Q é a operação de quantização. Todos os escalares e elementos do vetor do algoritmo SA irá desviar os seus valores originais, devido aos efeitos da quantização, Figura

2. Podemos definir os erros ou ruídos de quantização nas variáveis e quantidades relacionadas com algoritmo SA como segue:

$$n_\theta = \theta_k - \theta_{Qk} \quad (7)$$

Onde θ representa a variável ou quantidade relacionada no algoritmo e o subscrito Q indica a sua forma quantizada.

3 ANÁLISE DE DESEMPENHO

Invocamos uma série de pressupostos sobre a natureza estatística dos dados de entrada e ruído característico do modelo, comumente utilizados na literatura. Nesta análise, nós linearizamos uma função não linear arbitrária $f(e_k)$ sobre um ponto de funcionamento adequado (amostra de ruído n_k) e, em seguida, prosseguimos com a análise de covariâncias em estado estacionário.

Pressupostos:

A1: Assume-se que \mathbf{x}_k é um sinal aleatório discreto com $-1 < \mathbf{x}_k \leq 1$ não sofre quantização e cada \mathbf{x}_k vetor de dados de entrada é estatisticamente independente de todos os vetores de dados anteriores \mathbf{x}_j , $j < k$.

A2: A sequência de ruído amostrado por medição ou erros de quantização gerado em qualquer etapa no algoritmo é considerado como uma variável aleatória com média zero, que é estatisticamente independente de quaisquer outros erros e/ou quantidades relacionadas.

A3: O parâmetro de tamanho do passo μ deve ser escolhido suficientemente pequeno de tal modo que o excesso de erro médio quadrático na iteração k seja muito menor do que o mínimo erro médio quadrático perto de convergência.

B1: As variações dos erros dependem do tipo da quantização e aritmética que vai ser utilizada na implementação do algoritmo. Em nosso caso utilizamos aritmética de ponto fixo. Utilizando aritmética de ponto fixo o erro pode ser modelado em processo estocástico com média zero, com variância dada por Eq. (8) (Diniz, 2013):

$$\sigma^2 = C \frac{2^{-2b}}{12} \quad (8)$$

Onde C é uma constante.

A saída do sistema desconhecido pode ser expressa pela Eq. (9):

$$d_k = \mathbf{w}^{*t} \mathbf{x} + n_k \quad (9)$$

Onde \mathbf{w}^* é o coeficiente ótimo.. Supõe-se que o sinal de entrada e o sinal desejado não sofre quantização, de modo que apenas a quantização interna de computação são tidas em conta.

Podemos expressar o erro em sua forma quantizada pela Eq. (10)

$$e_{Qk} = n_k - n_e - \mathbf{v}_k^t \mathbf{x}_k \quad (10)$$

Onde definimos n_e como o ruído devido a quantização e \mathbf{v}_k^t é o vetor desvio do peso dado por: $\mathbf{v}_k = \mathbf{w}_k - \mathbf{w}^*$.

Expressando a não linearidade $f(e_k)$ em uma expansão em série de Taylor sobre o valor do ruído, obtemos: Eq.(11).

$$\begin{aligned} f(e_k) &= \sum_{i=0}^{\infty} \frac{f^i}{i!} (-n_{ek} - \mathbf{v}_k^t \mathbf{x}_k)^i \\ &= f(n_k) - f'(n_k)(n_e + \mathbf{v}_k^t \mathbf{x}_k) + \frac{f''(n_k)}{2}(n_e + \mathbf{v}_k^t \mathbf{x}_k)^2 - \frac{f'''(\delta)}{6}(n_e + \mathbf{v}_k^t \mathbf{x}_k)^3 \end{aligned} \quad (11)$$

Onde f^i denota a derivada (i) da função f , temos utilizado a forma derivada para representar o termo restante de Taylor tal que δ encontra-se no intervalo $[0, n_{ek} + \mathbf{v}_k^t \mathbf{x}_k]$. Tecnicamente esta representação de série de Taylor limita a não linearidade $f(\cdot)$ para uma classe restrita de funções que são diferenciáveis até a terceira ordem ao longo de toda extensão do erro e_k .

Agora vamos examinar o comportamento da Eq. (11) no caso em que o vetor de erro peso é pequeno, de tal modo que o termo $\frac{f'''(\delta)}{6}(n_e + \mathbf{v}_k^t \mathbf{x}_k)^3$ pode ser desprezado em relação à menor ordem do termo. Em seguida a algoritmo se torna: Eq. (12)

$$\mathbf{v}_{Q(k+1)} = \mathbf{v}_k + \mu [f(n_k) \mathbf{x}_k - f'(n_k)(n_e + \mathbf{v}_k^t \mathbf{x}_k) \mathbf{x}_k + \frac{f''(n_k)}{2}(n_e + \mathbf{v}_k^t \mathbf{x}_k)^2 \mathbf{x}_k] - n_v \quad (12)$$

Em que n_v é um vetor de ruído devido à quantização do vetor peso \mathbf{v}_k .

A fim de determinar equações para desajustes, utilizamos a Eq. (12) e multiplicando ambos os lados por sua transposta. Tomando as expectativas e invocando os pressupostos A1, A2. Assumindo o vetor erro de peso pequeno, podemos ignorar os termos de ordem O^3 .

$$\begin{aligned} E[\mathbf{v}_{Q(k+1)} \mathbf{v}_{Q(k+1)}^t] &= E[\mathbf{v}_{Qk} \mathbf{v}_{Qk}^t] - 2\mu [f'(n_k)] \mathbf{R} E[\mathbf{v}_{Qk} \mathbf{v}_{Qk}^t] + \mu^2 E[f^2(n_k)] \mathbf{R} \\ &+ \mu^2 E[f^2(n_k)] \mathbf{R} \sigma_e^2 + \mu^2 E[f'^2(n_k)] E[\mathbf{x}_k \mathbf{x}_k^t \mathbf{v}_k \mathbf{v}_k^t \mathbf{x}_k \mathbf{x}_k^t] \\ &+ \mu^2 E[f(n_k) f''(n_k)] E[\mathbf{x}_k \mathbf{x}_k^t \mathbf{v}_k \mathbf{v}_k^t \mathbf{x}_k \mathbf{x}_k^t] + \sigma_v^2 I \end{aligned} \quad (13)$$

Para $R = E[\mathbf{x}_k \mathbf{x}_k^t]$, $E[\mathbf{n}_v \mathbf{n}_v^t] = \sigma_v^2 I$ e $E[n_e^2] = \sigma_e^2$.

Podemos desprezar o quinto e sexto termo na Eq. (13), porque se μ e \mathbf{v}_{Qk} são pequenos, pressuposto A3, esses termos são muito menores do que os outros termos. Podemos utilizar a transformação de similaridade para expressar a Eq.(13) como um conjunto de equações iterativas dissociadas. Definindo a matriz unitária \mathbf{L} de autovetores e a matriz diagonal Λ de autovalores podemos expressar a matriz de valor real e covariância simétrica como $\mathbf{R} = \mathbf{L} \Lambda \mathbf{L}^t$. Assim, definindo o vetor de erro de peso no domínio de coordenadas transformadas como $\tilde{\mathbf{v}}_{Qk} = \mathbf{L}^t \mathbf{v}_{Qk}$, e para $f(n_k) \gg f'(n_k)$. Obtemos a Eq. (14).

$$E[\tilde{\mathbf{v}}_{Qk} \tilde{\mathbf{v}}_{Qk}^t] = \frac{1}{2E[f'(n_k)]} (\mu E[f^2(n_k)] + \frac{\sigma_v^2}{\mu}) I \quad (14)$$

Que possui a solução de $E[\tilde{\mathbf{v}}_{Qk}\tilde{\mathbf{v}}_{Qk}^t] = E[\mathbf{v}_{Qk}\mathbf{v}_{Qk}^t]$. Sabe-se que o excesso de erro médio quadrático na saída do filtro é causado por flutuações aleatórias (i) nos coeficientes de peso durante a adaptação. No qual assumimos a independência de \mathbf{x}_k e \mathbf{v}_k , bem como a natureza não correlacionada do vetor de erro de peso. Eq. (15) e (16).

$$E[(\mathbf{v}_{Qk}^t \mathbf{x}_{Qk})^2] = E[(\sum_{i=1}^N v_{i,k} x_{i,k})^2] \quad (15)$$

$$E[(\mathbf{v}_{Qk}^t \mathbf{x}_{Qk})^2] = \frac{1}{2E[f'(n_k)]} (\mu E[f^2(n_k)] + \frac{\sigma_v^2}{\mu}) E[x_{i,k}^2] \quad (16)$$

O desajuste é definido como excesso de erro médio quadrático sobre o erro mínimo médio quadrático e dado por: Eq. (17).

$$M = \frac{E[(\mathbf{v}_{Qk}^t \mathbf{x}_{Qk})^2]}{E[n_k^2]} \quad (17)$$

$$M = \frac{tr[\mathbf{R}]}{2E[f'(n_k)]\sigma_v^2} (\mu E[f^2(n_k)] + \frac{\sigma_v^2}{\mu}) \quad (18)$$

Onde na Eq. (18) o $tr[\cdot]$ denota a operação de traço $E[x_{i,k}^2] = tr[\mathbf{R}]$ e $E[n_k^2] = \sigma_v^2$ variância do ruído de medição.

Para $\frac{\sigma_v^2}{\mu} = 0$ não tem o ruído devido a quantização.

Assim, estas equações podem representar uma generalização de muitos resultados específicos baseados em uma família de algoritmos adaptativos baseados em uma função de custo de erros não-médio-quadrático agindo sobre o erro utilizando a técnica do gradiente estocástico desenvolvida em precisão finita. Em que podemos analisar o desajuste em estado estacionário em sua forma quantizada.

4 RESULTADOS E SIMULAÇÕES

Para utilizar a função $\tanh(e)$ em precisão finita utilizando aritmética de ponto fixo, foi necessário realizarmos uma aproximação desta função em série de Taylor até o quinto grau.

Aplicamos o algoritmo adaptativo em um problema de identificação de sistemas desconhecido, Figura (1) onde temos uma planta representada pela função de transferência polinomial $P(z) = 0.2037z^{-1} + 0.5926z^{-2} + 0.2037z^{-3}$, cuja saída é corrompida por um ruído, n_k . Nosso objetivo é encontrar, de modo adaptativo, um modelo de planta, $\hat{P}(z)$ em precisão finita utilizando aritmética de ponto fixo. Com este objetivo, utilizamos o algoritmo SA em diferentes comprimentos de palavras. O sinal de entrada foi simulado como um sinal aleatório uniformemente distribuído, limitado no intervalo $[-1, 1]$. Como ruído utilizou um sinal de ordem 10^{-3} com distribuição uniforme de probabilidades. O sinal desejado foi posto como a soma do sinal de entrada mais ruído. O parâmetro tamanho do passo para o algoritmo SA foi $\mu_{SA} = 0.01$ e utilizamos para o parâmetro α com o valor 3. Além do bit das variáveis utilizamos aqui um bit de sinal, resultante $(b + 1)$. Foram realizadas 100 simulações de Monte Carlo.

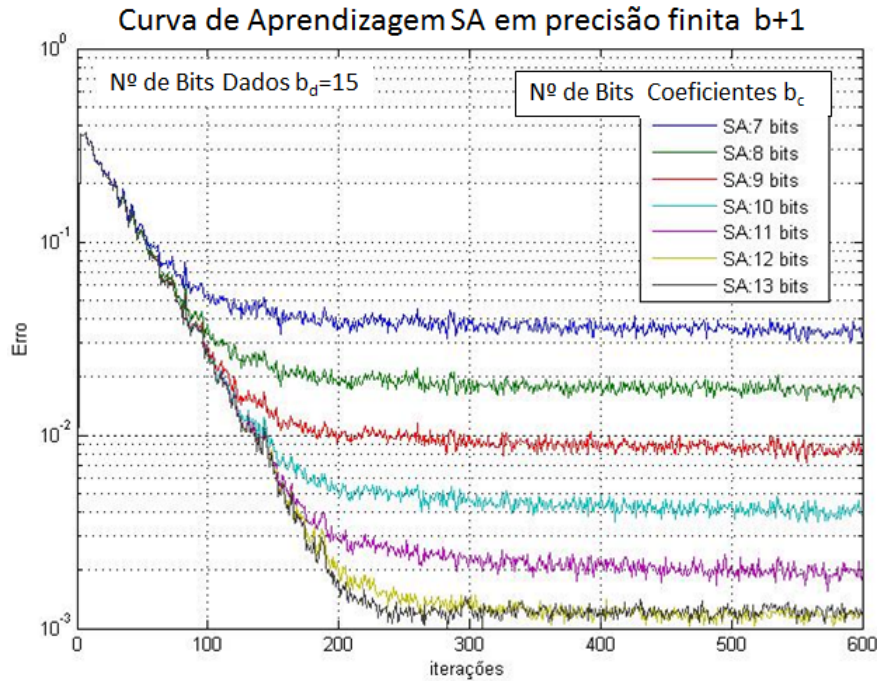


Figura 3: Curva de aprendizagem SA em precisão finita para $b_d = 15\text{bits}$ e $C = 0,0590$ em diferentes comprimentos de palavras

Tabela 1: Desajuste e MSE do SA em Precisão Finita para $b_d = 15\text{bits}$ e $C = 0,0590$

N° de Bits b_c	Experimental	Teorico	MSE
7	4,4654	1,9999	$4,4654 \cdot 10^6$
8	1,1144	0,5000	$1,1145 \cdot 10^6$
9	0,2570	0,2500	$2,5703 \cdot 10^7$
10	0,0705	0,0625	$7,0529 \cdot 10^8$
11	0,0187	0,0156	$1,8703 \cdot 10^8$
12	0,0064	0,0078	$6,4395 \cdot 10^9$
13	0,0068	0,0078	$6,7681 \cdot 10^9$

5 DISCUSSÃO

Estes resultados sugerem um aumento do nível de erro ou ruído adicional devido à quantização, para alguns comprimentos de palavras (7, 8, 9, 10 e 11 bits). Que dependem do problema envolvido, tal característica também pode ser encontrada em outros algoritmos adaptativos como o LMS. Em nosso modelo para uma previsão adequada do desajuste, pode ser necessário escolher um comprimento de palavra, de forma que o excesso de erro médio quadrático (MSE) seja muito menor que o erro mínimo médio quadrático, conforme tabela (1). Pode-se ainda especificar uma relação entre o comprimento de palavra mínimo para os coeficientes de modo a garantir a convergência com um ruído de quantização desprezível, no caso (12 e 13 bits), para o algoritmo SA.

6 CONCLUSÃO

Analizamos o algoritmo SA em precisão finita utilizando aritmética de ponto fixo. Para isto derivamos expressões analíticas para o desajuste em precisão finita e realizamos uma aproximação em série de Taylor, o que tornou possível realizar essa implementação. Os resultados preliminares encontrados sugerem a possível implementação em dispositivos de hardware desse algoritmo. Isso implica em novas alternativas para resolução de problemas de processamento de sinais encontrados em diversas áreas, como por exemplo, na medicina, militar e outros. Os próximos rumos dessa pesquisa será continuarmos avaliando estes resultados analíticos para outros algoritmos adaptativos em precisão finita, baseados em funções de custo de erro não-médio-quadrático, para diferentes sinais de entrada e ruídos. Como pesquisa futura sugerimos a implementação desse algoritmo em dispositivo de hardware.

AGRADECIMENTOS

Ao PPGEE-UFMA (Programa de Pós Graduação em Engenharia Elétrica-Universidade Federal do Maranhão), ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), a FAPEMA (Fundação de Amparo à Pesquisa e Desenvolvimento Científico do Maranhão) e ao PIB (Laboratório da Informação Biológica).

REFERÊNCIAS

- Bershad, N. J., & Bermudez, J., C., M., 1996. A Nonlinear analytical model for the quantized LMS algorithm - The power-of-two step size case. *IEEE Transactions on Sig. Proc.*, vol. SP-44, no. 11, pp. 2895-2900.
- Diniz, P., 2013. *Adaptive filtering*, Rio de Janeiro.
- Douglas, S., C., & Meng, T., H., 1994. Stochastic gradient adaptation under general error criteria. *IEEE Transactions on Sig. Proc.*, Vol. 42, no. 6.
- Gupta, R., & Hero, A., O., 2000. Transient Behavior of fixed point LMS adaptation,. *University of Michigan, Ann Arbor*, MI 48109-2122.
- Haykin, S., 2014. *Adaptive filter theory and Applications*, USA
- Maluenda, Y., R., M., & et al, 2005. Propriedades do algoritmo LMS operando em precisão finita. *XXII Simpósio Brasileiro de Tele.*, SBrT'05.
- Meher, P., K., & Park, S., Y., 2014. Area-delay-power efficient fixed-point LMS adaptive filter with low adaptation-delay. *IEEE Transactions on Very large scale int.(VLSI) Systems*, Vol. 22, no. 2.
- Oppenheim, A. V., & et al, 1996. *Signals and Systems*, Prentice Hall, New Jersey.
- Sayed, A. H., 2013. *Fundamentals of Adaptive Filtering*, IEEE Press, University of California.
- Santana, E., 2006. Estudo e desenvolvimento de uma família de algoritmos não lineares para filtragem adaptativa. In Olhoff, N. & Rozvany, G. I. N., eds, *UFMA*.
- Widrow, B., & Stearns, S., D., 1985. Adaptive signal processing, Prentice Hall, Englewood Cliffs. *Prentice Hall, Englewood Cliffs*.