

Comparação de Desempenho Computacional e Qualidade de Voz dos Algoritmos LMS e NLMS Aplicados ao Cancelamento de Eco em Voz sobre IP

Guilherme Rauter Corsetti, Ricardo Becker

{gcorsetti, rbecker}@gparc.org

Grupo de Pesquisas Avançadas em Redes de Comunicação e Tecnologia da Informação (GPARC&TI)

Faculdade de Engenharia Elétrica – Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS

CEP: 90.619-900 – Porto Alegre – RS – Brasil

Abstract- The main objective of this paper is to make a comparison between implementations of Least Mean Square (LMS) and Normalized Least Mean Square (NLMS) algorithms in voice over IP echo cancellation. There are two main points to make this comparison. The first is the computational performance and the second is the voice quality after echo cancellation is done. To perform the algorithms will be simulated using different delays and attenuations of the echo.

I. INTRODUÇÃO

O eco é um problema comum que ocorre quando há uma reflexão do sinal de voz que ocorre, por exemplo, nas redes telefônicas. Existem dois tipos de eco na rede de telefonia convencional: o eco acústico e o eco híbrido. O primeiro é causado por ruídos existentes no ambiente, como reflexões do som nas paredes onde a pessoa está falando e pela realimentação do alto-falante para o microfone. O segundo é gerado em transformadores híbridos, usado na rede pública de telefonia comutada (RPTC) para a transmissão de dados, onde um link de 4 fios tem que ser transformado em um link de 2 fios [5].

Neste artigo é mostrada a implementação de um filtro adaptativo LMS [1] e a implementação de um filtro adaptativo NLMS, cuja função será cancelar o eco de uma rede IP ligada a uma rede de telefonia convencional. Após os algoritmos serem explicados é feita uma comparação entre os mesmos. Primeiramente é comparado o desempenho computacional de cada algoritmo, posteriormente sendo feita uma comparação da qualidade da voz após o cancelamento de eco.

II. FILTROS ADAPTATIVOS E CANCELADOR DE ECO IP

Os filtros adaptativos LMS e NLMS têm como função, no cancelador de eco, gerar um sinal na saída do filtro (y) que seja capaz de eliminar o eco de um sinal de voz ($s + r$). A eliminação do eco ocorre através da subtração do sinal com eco pelo sinal de saída do filtro. O processo adaptativo dos filtros ocorre através da utilização de pesos (W). O que diferencia os filtros LMS e NLMS é a forma com que o passo de adaptação (μ) dos pesos é calculado. O μ é a constante de ganho que regula a velocidade e estabilidade do processo de adaptação. A Equação 1 mostra como os pesos são calculados.

$$W_{k+1} = W_k + 2\mu\epsilon_k X_k \quad (1)$$

onde, X_k representa é o vetor de amostras de entrada e ϵ_k é o erro da saída do filtro. Para o algoritmo LMS o μ um valor fixo que fica compreendido entre

$$0 < \mu < 2/\text{tr}\{R\} \quad (2)$$

onde, $\text{tr}\{R\}$ é o somatório da diagonal da matriz de autocorrelação dos vetores do sinal de entrada [1]. O valor de

μ para o algoritmo NLMS varia a cada iteração do sistema e é recalculado pela Equação 3.

$$\mu = 2/\text{tr}\{R\} \quad (3)$$

Um cancelador de eco tem como função eliminar o eco que está sendo gerado em uma chamada. O eco gerado pelo transformador híbrido pode ser eliminado já no gateway de entrada da Rede IP [4]. O eco acústico também existe na rede IP, ele é gerado da mesma forma que na rede de telefonia convencional [5]. As implementações de canceladores de eco que são mostradas neste artigo objetivam cancelar apenas o eco acústico. Na Figura 1 é mostrado um modelo geral de um cancelador de eco de uma ligação de um telefone IP para o outro, utilizando apenas a rede IP:

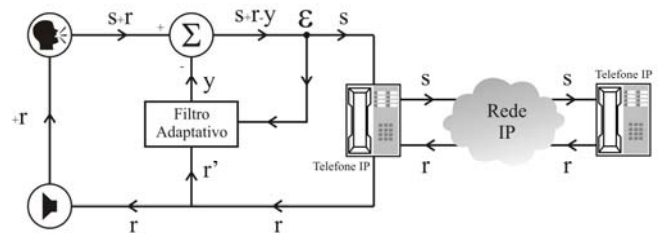


Figura 1 – Modelo de cancelador de eco de telefone IP para telefone IP

III. DESENVOLVIMENTO DOS FILTROS

Para a implementação dos algoritmos de cancelamento de eco, foi escolhido o ambiente Delphi. Foram gravadas amostras de voz com codificação PCM (padrão da telefonia), a 8 bits e 8 kHz. Essas amostras foram utilizadas para que os algoritmos pudessem ser testados. Sobre a amostra de voz utilizada foi acrescentado eco, primeiramente com um atraso fixo e depois foi feito o atraso do eco variar. Como o eco na realidade sofre uma atenuação em relação ao sinal original, essa característica também foi simulada. O objetivo disso é simular uma fala com um sinal mais atrasado e atenuado com relação ao original (eco acústico). Sobre o arquivo gerado foi aplicado o algoritmo de cancelamento de eco que foi desenvolvido. Foi utilizado um vetor de 128 pesos na implementação dos algoritmos. Esse número de pesos foi utilizado pelos dois algoritmos, pois com esse número ambos os canceladores de eco obtiveram uma boa performance no processo de eliminação do eco. O μ utilizado pelo LMS foi igual a 0,00007.

IV. LMS X NLMS

Para comparar o desempenho computacional de cada algoritmo foi utilizada uma função chamada de "QueryPerformanceCounter", desenvolvida em [7], essa

função é responsável por checar o tempo de processamento do algoritmo. O computador utilizado para verificar o tempo de processamento dos algoritmos é um Pentium IV, 2,4 GHz e 512 MB de memória RAM. O sistema operacional utilizado foi o Microsoft Windows XP Professional.

O tempo de processamento calculado para fazer a comparação entre os algoritmos representa o tempo necessário para eliminar o eco de um pacote de 20 ms de voz. Isso se deve ao fato de voz sobre IP ser transmitido por pacotes de voz [5]. O Gráfico 1 mostra o tempo de processamento de cada algoritmo, pelo atraso do eco. Quanto maior o atraso do eco menor é o tempo de processamento dos algoritmos, isso ocorre porque no início da análise dos arquivos (arquivo com eco e arquivo sem eco), os algoritmos não precisam convergir, pois os arquivos são iguais.

Para comparar a qualidade de voz de cada algoritmo foi feita uma adaptação do método ACR da recomendação P.800 (MOS) [6] do ITU-T para a realização do teste de opinião por escuta. Para a realização dos testes de escuta foram gravadas quatro sentenças, ou quatro amostras de voz em uma sala silenciosa fechada, não acusticamente isolada e com a utilização de um microfone dinâmico de alta performance.

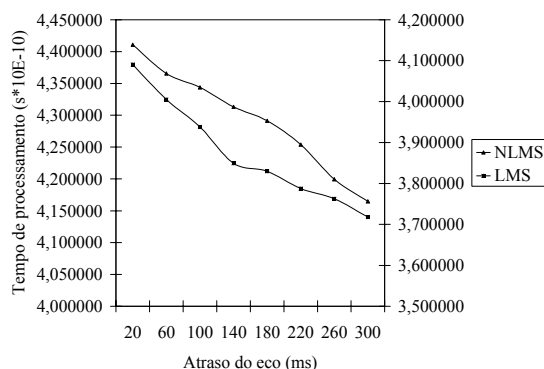


Gráfico 1 – Comparação do tempo de processamento dos algoritmos

Estas amostras foram submetidas ao algoritmo que insere eco nas mesmas. Foi criado um arquivo de cada amostra com eco. O eco inserido em todas as amostras foi de 120 ms de atraso e 2 dB de atenuação. Após isso foi removido o eco de cada arquivo utilizando os algoritmos de cancelamento de eco apresentados nesse artigo, gerando assim os arquivos que foram submetidos para a avaliação do público em geral.

Para a realização dos testes, 50 pessoas foram convidadas aleatoriamente a participar. Todos os participantes, no momento da avaliação, encontravam-se sentados e usando *headphones*. Os mesmos ouviam duas sentenças e depois faziam uma marcação na pontuação referente ao esforço necessário para entender o significado das sentenças. Cada participante do público foi submetido às sentenças que estavam apenas sob o efeito de uma técnica de cancelamento de eco. Ao final do experimento estavam contabilizadas 25 avaliações por técnica. No condizente ao nível de esforço necessário para o entendimento das quatro sentenças, o resultando é apresentado no Gráfico 2. O resultado é uma média das 25 notas, entre um e cinco, dadas para cada técnica avaliada pelo público entrevistado.

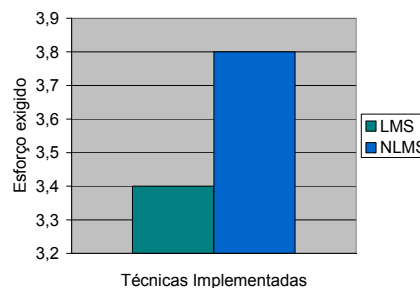


Gráfico 2 – Nível de esforço para o entendimento das amostras de voz

V. CONCLUSÕES

Após a implementação dos algoritmos pode-se concluir que ambos os algoritmos possuem baixa complexidade computacional, pois necessitam de poucos cálculos para realizar o processo adaptativo e esses cálculos são operações simples de se realizar.

Os dois algoritmos são lentos para iniciar o processo adaptativo. Isso porque eles iniciam a sua adaptação apenas quando o número de amostras analisadas pelo filtro é igual ao tamanho do vetor de pesos, no caso dos algoritmos que foram mostrados nesse artigo, eles começam a se adaptar apenas depois de 128 amostras, ou seja, após 16 ms de fala. Como foi mostrado no artigo o algoritmo LMS é mais rápido que o NLMS, pois executa um cálculo a menos (não calcula o passo de adaptação).

Com relação à qualidade dos algoritmos, o NLMS obteve um melhor resultado nos testes realizados e isto se deve ao seu processo adaptativo que é menos ruidoso, pois calcula o passo de adaptação a cada amostra analisada.

REFERENCIAS

- [1] Widrow, B. and Stearns, S. D. (1985) *Adaptive signal processing*, Prentice-Hall, Englewood Cliffs, pp. 1-364.
- [2] Becker, R., *Técnicas de Detecção e Supressão de Silêncio para Sistemas de Voz sobre IP*, Trabalho de conclusão de curso, Engenharia Elétrica, PUCRS, Porto Alegre, 2004.
- [3] Corsetti, G. R.; Coster M. et all, *Implementação de um Filtro Adaptativo LMS Aplicado ao Cancelamento de Eco em Voz sobre IP*, II Escola Regional de Redes de Computadores, Canoas, 2004.
- [4] Douskalis, B. *IP Telephony the Integration of Robust VoIP Services*, Prentice-Hall, 1999, pp. 253-258.
- [5] Hersent, O.; Guide, D. and Petit, J. *Telefonia IP – Comunicação multimídia baseada em pacotes*, Addison Wesley, São Paulo, 2002.
- [6] ITU-T. *Methods for Subjective Determination of Transmission Quality*. (ITU-T Recommendation P.800), 1996.
- [7] JEDI. <http://www.delphi-jedi.org/>.