

# Análise da relação sinal/ruído a priori como parâmetro básico em técnicas de redução de ruído em sinais de voz

Fabrcio Sato e Jozue Vieira Filho

**Resumo** — A reduo de ruído em sinais de voz é uma área importante em telecomunicações, com aplicaões em reconhecimento automático de fala, telefonia, codificao, etc. As técnicas que exploram o uso da relao sinal/ruído (SNR) a priori como parâmetro básico são as que têm apresentado resultados mais satisfatórios. Isto tem despertado o interesse de muitos pesquisadores, que propõem novas metodologias para a estimao da SNR a priori em sistemas reais. Porém, neste trabalho é apresentada uma análise de algumas técnicas importantes, onde mostra-se que o uso da SNR a priori real pouco altera o resultado qualitativo do processamento, quando comparado com o uso de uma SNR a priori estimada. As comparaões foram feitas usando-se a PESQ (*Perceptual Evaluation of Speech Quality*) e as conclusões obtidas são importantes para o direcionamento de pesquisas futuras.

**Palavras-Chave** – Reduo de ruído, Relao sinal/ruído a priori, Qualidade de voz, Processamento de voz.

**Abstract** — Speech Enhancement is important in different context of telecommunication systems, such as telephony, speech coding, automatic speech recognition, etc. Noise reduction techniques based on a priori signal-to-noise ratio (SNR) have present good results so far. As a consequence, there are many researchers investigating new methodologies to estimate the a priori SNR. However, in this work it is demonstrated that some techniques based on a prior SNR do not increase significantly their results when the real a prior SNR is used. The comparisons were carried out using the PESQ (*Perceptual Evaluation of Speech Quality*) and the results are important to head new researches on this subject.

**Keywords** – Speech enhancement, A priori signal-to-noise ratio, Speech quality, Speech processing.

## I - INTRODUO

As técnicas de reduo de ruído baseadas na SNR a priori foram impulsionadas pelo brilhante trabalho desenvolvido por Ephraim e Malah [1]. Estudos iniciais [2] indicavam que a mesma era mais eficiente do que outras técnicas tradicionais, como a subtrao espectral [3], pois permite a obteno de um sinal ausente de ruído musical [4]. Estudos posteriores comprovaram a eficiência da técnica proposta por Ephraim e Malah (EM). Porém, outros trabalhos tentaram explorar e desvendar os detalhes dessa técnica.

Fabrcio Sato e Jozué Vieira Filho, Departamento de Engenharia Elétrica, Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista (UNESP), Ilha Solteira, SP, Brasil. E-mails: [fsato\\_12@yahoo.com.br](mailto:fsato_12@yahoo.com.br) e [jozue@dee.feis.unesp.br](mailto:jozue@dee.feis.unesp.br). Trabalho financiado parcialmente pelo CNPq.

Em 1996, Scalart e Vieira-Filho [5] verificaram que o uso da SNR a priori era um fator de destaque da técnica de EM, apesar da excelente explorao do modelo gaussiano para um sinal de voz, e que a elevada carga computacional poderia limitar muitas aplicaões práticas. Estudando as técnicas clássicas, como a subtrao espectral (SE) e a filtragem de Wiener (FW), Scalart e Vieira-Filho mostraram que o ruído musical presente nos sinais processados pelas mesmas era fruto do uso direto da SNR a posteriori, como parâmetro básico, e propuseram a substituio desse parâmetro pela SNR a priori [5]. O resultado foi um conjunto de técnicas com baixa carga computacional, o que permite uma implementao direta em tempo real, e com resultados comparáveis aos obtidos com a técnica de EM. A metodologia usada para estimar a SNR a priori nesses trabalhos é denominada de deciso-dirigida e foi proposta inicialmente por McAulay e Malpass [6]. A partir de 1998 várias propostas têm surgido, buscando-se a melhoria na estimao da SNR a priori [7], [8], [9] e [10]. Recentemente, C. Plapous e outros [8] propuseram uma nova metodologia de reduo de ruído, que inclui um modelo diferenciado para estimar a SNR a priori, aqui denominada, de acordo com o original, de Two-Step Noise Reduction Technique (TSNR).

Com a ateno despertada para o uso desse parâmetro, decidiu-se, neste trabalho, estudar e avaliar a influncia da SNR a priori no resultados de algumas técnicas que nela são baseadas. Uma primeira pergunta que dever ser respondida é: como será o resultado de algumas dessas técnicas se for usada a SNR a priori real? Partindo do princípio de que as técnicas são para aplicaões em sinais de voz contaminados por ruído aditivo, foi possível elaborar um experimento completo, com diferentes técnicas e sinais, usando tanto a SNR a priori real como a estimada. Os resultados são reveladores e mostram que o uso da SNR a priori real pouco contribui para uma melhoria da qualidade do sinal processado. Esta concluso é importante, pois mostra que a melhoria dessas técnicas não passa por uma simples melhoria na metodologia de estimao da SNR a priori. Entende-se, neste caso, que a SNR a priori real é o limite teórico para a SNR a priori estimada. Nos estudos realizados, foram implementadas e avaliadas a técnica SE, a técnica de EM e TSNR.

Para cumprir os objetivos propostos, neste trabalho apresenta-se inicialmente uma reviso básica das técnicas de reduo de ruído baseadas na SNR a priori (Seo II), destacando também a estimao da mesma; na seqüência são discutidas a implementao e os resultados da avaliao (Seo III); por fim, são apresentadas as conclusões e as expectativas para trabalhos futuros (Seo IV).

## II. TÉCNICAS DE REDUÇÃO DE RUÍDO BASEADAS NA SNR A PRIORI

As técnicas mais usadas na redução de ruído em sinais de voz são baseadas na captação de um sinal por um único microfone, pois as mesmas são as mais viáveis computacionalmente para operações em tempo real. Além disso, considera-se que o ruído é aditivo e, tanto ruído como voz, são processos aleatórios estacionários independentes. Para um sinal de voz, isto só é válido para curtos intervalos de tempo, tipicamente da ordem de algumas dezenas de milissegundos.

Seja, por exemplo, um sinal de voz puro  $v(t)$ . Se a este sinal for adicionado o ruído  $r(t)$ , o novo sinal ruidoso  $y(t)$  será dado por:

$$y(t) = v(t) + r(t) \quad (1)$$

Partindo da equação (1), são desenvolvidos os diferentes métodos de redução de ruído baseados na captação do sinal por um único microfone. Neste ponto, é importante destacar que técnicas baseadas em mais de um microfone não têm conseguido espaço devido à natureza básica da maioria das aplicações.

### A. Subtração Espectral (SE)

Aplicando-se a Transformada de Fourier na equação (1), obtém-se:

$$Y(\omega) = V(\omega) + R(\omega) \quad (2)$$

Usando a hipótese de que o ruído é aditivo e decorrelacionado com o sinal de voz [3], da equação (1) mostra-se que:

$$|V(\omega)|^2 = |Y(\omega)|^2 - |R(\omega)|^2 \quad (3)$$

A equação (3) indica que é possível recuperar o espectro de amplitude do sinal original, desde que se conheça o espectro do ruído. Observa-se também que o espectro do sinal de voz recuperado pode ser negativo, o que é algo errado, tanto do ponto de vista físico, como também matemático. A solução para resolver tal incoerência não é única e a primeira foi proposta por Boll [3]. A partir de então surgiram diferentes metodologias e técnicas de implementação da equação (3), definindo assim diferentes métodos de redução de ruído [1], [4] e [7]. A partir do trabalho desenvolvido por [5], a equação (3) passou a ser implementada de maneira mais eficiente explorando-se SNR.

Da equação (3) pode-se definir, genericamente, um filtro redutor de ruído com função de transferência dada por:

$$|H(\omega)| = \frac{\sqrt{|Y(\omega)|^2 - E\{|R(\omega)|^2\}}}{|Y(\omega)|} \quad (4)$$

O fator  $E\{|R(\omega)|^2\}$  representa uma estimativa estatística do ruído.

Da equação (3), considerando também que o sinal de voz a ser recuperado é apenas uma estimativa do sinal de voz original tem-se:

$$\frac{E\{|V(\omega)|^2\}}{E\{|R(\omega)|^2\}} = \frac{|Y(\omega)|^2}{E\{|R(\omega)|^2\}} - 1 \quad (5)$$

O primeiro termo da equação (5) é denominado de SNR a priori e o segundo termo de SNR a posteriori, denotados com seguim:

$$\xi(\omega) = \frac{E\{|V(\omega)|^2\}}{E\{|R(\omega)|^2\}} \quad e \quad \gamma(\omega) = \frac{|Y(\omega)|^2}{E\{|R(\omega)|^2\}} \quad (6)$$

É importante salientar que os termos definidos e apresentados na equação (6) foram introduzidos primeiramente por McAulay e Malpass [6].

Uma combinação adequada das equações (4), (5) e (6) permite obter um filtro de Wiener generalizado, que permite também a implementação da subtração espectral direta, cujo parâmetro básico a SNR a priori. Assim, tem-se:

$$|\hat{H}_w(\omega)| = \left[ \frac{\xi(\omega)}{\xi(\omega) + \delta} \right]^\lambda \quad (7)$$

Para valores tais que  $\delta = 1$  e  $\lambda = 0,5$ , o filtro obtido é exatamente o filtro básico da subtração espectral.

Como será visto posteriormente, o maior problema a ser resolvido na implementação desse tipo de filtro é a obtenção de  $\xi(\omega)$  que, por definição, torna o filtro não-causal.

### B. Filtro proposto por Ephraim & Malah (EM)

Nesta técnica, os sinais de voz e ruído, além de se serem considerados processos aleatórios estacionários e independentes, também são modelados estatisticamente. Os autores consideraram que os coeficientes da expansão em Série de Fourier dos sinais são variáveis aleatórias gaussianas estatisticamente independentes. A base matemática do modelo é o teorema do limite central, que considera que cada coeficiente de Fourier é, no final, uma soma ponderada (ou integral) de variáveis aleatórias resultantes de um grande número de amostras do processo [11]. Assume-se também que os processos possuem média igual à zero.

De acordo com a equação (1), no domínio da frequência os sinais de voz puro e ruidoso podem ser escritos como:

$$V(\omega) = |V(\omega)| \cdot e^{j\alpha(\omega)} \quad (8)$$

$$Y(\omega) = |Y(\omega)| \cdot e^{j\theta(\omega)} \quad (9)$$

O objetivo é estimar o espectro de amplitude do sinal de voz  $|V(\omega)|$  a partir do sinal ruidoso  $y(t)$ , dentro de um determinado intervalo de observação (0 - T).

Assumindo que as componentes espectrais são estaticamente independentes, obtém-se um estimador diretamente do sinal ruidoso como segue:

$$|\hat{V}(\omega)| = E\{|V(\omega)| \mid y(t)\} \quad , \quad 0 \leq t \leq T \quad (10)$$

ou

$$|\hat{V}(\omega)| = E\{|V(\omega)| \mid Y_0, Y_1, \dots\} = E\{|V(\omega)| \mid Y_k\} \quad (11)$$

onde  $Y_k$  é a amostra de  $|Y(\omega)|$  e  $k$  representa a frequência específica analisada dentro do intervalo observado.

Da equação 11 tem-se [1]

$$|\hat{V}(\omega)| = \frac{\int_0^{\infty} \int_0^{2\pi} v(\omega) \cdot p\{Y(\omega)/v(\omega), \alpha(\omega)\} \cdot p\{v(\omega), \alpha(\omega)\} \cdot d\alpha(\omega) dv(\omega)}{\int_0^{\infty} \int_0^{2\pi} p\{Y(\omega)/v(\omega), \alpha(\omega)\} \cdot p\{v(\omega), \alpha(\omega)\} \cdot d\alpha(\omega) dv(\omega)} \quad (12)$$

onde  $p(\cdot)$  representa a função densidade de probabilidade e  $v(\omega)$  representa a variável aleatória do espectro de amplitude  $V(\omega)$  do sinal de voz. A solução da equação (12) é apresentada em detalhes em [11]. A resposta em frequência do filtro obtido, em termos de magnitude, é dada por:

$$|\hat{H}_{EM}(\omega)| = \Gamma(1,5) \frac{\sqrt{\eta(\omega)}}{\gamma(\omega)} \exp\left[\frac{-\eta(\omega)}{2}\right] \cdot \left\{ [1 + \eta(\omega)] I_0\left[\frac{\eta(\omega)}{2}\right] + \eta(\omega) I_1\left[\frac{\eta(\omega)}{2}\right] \right\} \quad (13)$$

Na equação (13) tem-se  $\eta(\omega) = \frac{\xi(\omega)}{1 + \xi(\omega)} \cdot \gamma(\omega)$ .

Observa-se que o filtro proposto por Ephraim e Malah é função tanto da SNR a priori como também da SNR a posteriori. Do ponto de vista de aplicação, o maior desafio na implementação desse filtro é, além da estimação da SNR a priori, a elevada carga computacional.

### C. Estimação da SNR a priori e a posteriori

Como mencionado, a implementação dos filtros redutores de ruído passa por uma estimação da potência do ruído e da SNR a priori e a posteriori.

Considerando que o sinal a ser processado possui intervalos de silêncio, o ruído é estimado durante os mesmos através de uma filtragem recursiva de primeira ordem, como segue:

$$E\{|R(w_i, f_k)|^2\} = \beta \cdot E\{|R(w_{i-1}, f_k)|^2\} + (1 - \beta) \cdot |Y(w_i, f_k)|^2 \quad (14)$$

onde  $|Y(w_i, f_k)|^2$  representa a potência do sinal ruidoso e  $\beta$  é uma constante que define a dependência do ruído na janela atual de análise em relação às janelas anteriores. Considerando uma implementação através da STFT (Short

Time Fourier Transform), na equação (14)  $w_i$  indica a  $i$ -ésima janela do sinal e  $f_k$  indica  $k$ -ésima frequência.

Esta estimação é fundamental para um bom desempenho do filtro redutor de ruído. Por exemplo, uma variação abrupta de potência de uma janela a outra pode contribuir fortemente para o surgimento do ruído musical. Nesses casos, um valor adequado de  $\beta$  suaviza essa diferença.

A partir da estimação do ruído nos intervalos de silêncio, estima-se, inicialmente, a SNR a posteriori, como segue:

$$\gamma(w_i, f_k) = \frac{|Y(w_i, f_k)|^2}{E\{|R(w_i, f_k)|^2\}} \quad (15)$$

A estimação da SNR a priori de acordo com a equação (6) depende do próprio sinal de voz estimado, o que caracteriza um sistema não-causal. Para resolver este problema, considera-se que a variação de potência do sinal de voz de uma janela para outra não é significativa. Estas considerações têm justificado o uso do estimador recursivo proposto por Ephraim e Malah, denominado de decisão dirigida, que é dado por:

$$\xi(w_i, f_k) = \alpha \cdot \frac{|\hat{V}(w_{i-1}, f_k)|^2}{E\{|R(w_i, f_k)|^2\}} + (1 - \alpha) \cdot T \cdot [\gamma(w_i, f_k) - 1] \quad (16)$$

onde  $T[\bullet]$  indica uma transformação sobre  $[\gamma(w_i, f_k) - 1]$  e  $\alpha$  representa o grau de dependência da  $\xi(w_i, f_k)$  com relação a  $\gamma(w_i, f_k)$ .

Nos estudos realizados, como também no trabalho original de Ephraim e Malah [1], verificou-se que fortes variações em  $\gamma(w_i, f_k)$  podem afetar o desempenho da  $\xi(w_i, f_k)$ . Entretanto, se for definido  $\alpha = 1$ , perde-se o ajuste fino na estimação da  $\xi(w_i, f_k)$ . Um valor típico para  $\alpha$  encontrado em vários trabalhos é  $\alpha = 0,98$ .

### D. Two-Step Noise Reduction Technique (TSNR)

Na equação (16), a estimação da relação sinal-ruído a priori introduz um atraso, pois o parâmetro  $\alpha$  é próximo de um. Assim, o ganho espectral estimado para a janela que está sendo processada coincide com o ganho da janela anterior. Visando melhorar a estimação da SNR, o trabalho apresentado em [8] propõe estimar o ganho espectral da janela atual usando o algoritmo decisão dirigida, para depois estimar a SNR a priori e processar, efetivamente, a janela  $w_i$ . Isto leva a um algoritmo em duas etapas. Na primeira etapa, usando o algoritmo decisão dirigida, é estimado um ganho espectral  $H_{DD}(w_i, f_k)$  como descrito na equação (7). Na segunda etapa, este ganho é usado para estimar a SNR a priori que será usada no novo filtro, como segue:

$$\xi(w_i, f_k)^{TSNR} = \alpha \cdot \frac{|H_{DD}(w_i, f_k) Y(w_i, f_k)|^2}{E\{|R(w_i, f_k)|^2\}} + (1 - \alpha) T \cdot [\gamma(w_i, f_k) - 1] \quad (17)$$

onde  $\alpha'$  tem as mesmas características de  $\alpha$ , mas são independentes.

No cálculo da equação (17), os autores propõem eliminar a contribuição da SNR a posteriori e assumem  $\alpha=1$ . Numa análise detalhada do algoritmo decisão dirigida para a SNR a priori como discutido em [2], isto significa que o ruído musical é reduzido ao nível mais baixo permitido. Porém, isto poderá gerar distorções no sinal processado. Segundo os autores, o valor  $\alpha=1$  é válido somente para a segunda etapa, a fim de aperfeiçoar a estimação da primeira etapa. Para a primeira etapa, utiliza-se  $\alpha = 0,98$ .

A nova relação sinal/ruído a priori estimada pode ser aplicada a qualquer filtro. Em particular, para os objetivos deste trabalho, foi escolhida a filtragem de Wiener que, com o novo parâmetro, gera o seguinte filtro redutor de ruído:

$$H_{TSNR}(w_i, f_k) = \frac{\xi^{TSNR}(w_i, f_k)}{1 + \xi^{TSNR}(w_i, f_k)} \quad (18)$$

### III. IMPLEMENTAÇÃO E AVALIAÇÃO

O objetivo principal deste trabalho é mostrar as limitações impostas pelas próprias técnicas que usam SNR a priori como parâmetro principal, considerando, porém, as implementações clássicas baseadas na STFT. As configurações para as diferentes técnicas foram as mesmas. Os sinais utilizados foram amostrados a uma taxa de 16 KHz e degradados por um ruído aditivo (sinal limpo + ruído) para gerar sinais ruidosos com diferentes SNR (0, 5 e 10 dB). Os cálculos de SNR para avaliação neste trabalho são simples e baseados diretamente na relação entre amostras de trechos somente de voz e trechos somente de ruído (amplitudes quadráticas médias de 4096 amostras, tomadas sempre nos mesmos intervalos dos sinais de voz e ruído puros). O ruído aditivo utilizado é o de um carro pequeno (Renault 25) rodando a 100 km/hm [12]. Os resultados apresentados foram obtidos usando-se o software de simulação MATLAB. Como ilustração, na figura 1 é apresentada a forma de onda de um dos sinais usados no testes (S1). Cada sinal foi processado com janela de Hanning com intervalos de sobreposição de 50 %, 512 amostras por janela e FFT (Fast Fourier Transform) de 512 pontos. O fator de atualização do ruído, presente na equação (14), foi  $\beta = 0,8$ . Na equação (16), foi adotada a seguinte transformação:

$$T \cdot [\gamma(w_i, f_k) - 1] = \begin{cases} \gamma(w_i, f_k) - 1, & \text{se } \gamma(w_i, f_k) \geq 1 \\ 0,01 & \text{se } \gamma(w_i, f_k) < 1 \end{cases} \quad (19)$$

O valor 0,01 usado na transformação limita a obtenção da SNR a priori em  $-73$  dB. Teste práticos mostram que esse limite gera melhores resultados [11] do que uma simples retificação de meia-onda, como proposto em [3].

#### A. Comparação entre os Valores de SNR a priori Estimado e Real e os respectivos Ganhos

Uma primeira análise a ser feita é uma comparação direta entre a SNR a priori estimada e real. Como se sabe, os valores de SNR existem para cada frequência analisada e processada, de acordo com a FFT usada. Porém, uma análise da SNR média em cada janela analisada já é suficiente para uma comparação básica. Assim, na figura 2A são apresentadas curvas médias de SNR a priori estimada e real. Verifica-se que as curvas são idênticas, mas há uma diferença nos valores que fica próxima de 20 dB. Portanto, espera-se que o uso de uma SNR a priori real leve a resultados bem melhores do que o uso da SNR a priori estimada. Porém, analisando agora as curvas médias de ganho apresentadas na figura 2B, verifica-se que a diferença é quase desprezível. Para uma melhor visualização dessas curvas, na figura 3 é apresentado apenas um pequeno intervalo de tempo. Os resultados mostram que os filtros testados são limitados quanto à variações na estimação da SNR a priori. Tal limitação fica mais evidente considerando-se a avaliação dos sinais processados, como apresentado no item seguinte.

#### B. Avaliação da qualidade dos sinais processados

Para verificar com precisão a diferença na qualidade dos sinais, foi usada a medida objetiva PESQ (*Perceptual Evaluation of Speech Quality*) padronizada pelo ITU [13]. Seu índice de qualidade varia entre 0 e 4,5 e uma diferença de 0,1 apresenta variação auditiva quase desprezível.

Na tabela 1 são apresentados os resultados para o sinal ruidoso processado pelas técnicas apresentadas neste artigo, com as diferentes SNR a priori. Foram usados três sinais: S1, frase em português; S2, frase em francês; S3, frase em inglês. O resultado obtido em todos os casos são reveladores. Por exemplo, para o sinal S1, com SNR igual a 5 dB, processado pela técnica da Subtração Espectral, a variação foi de apenas 0,273. Analisando todas as técnicas para o sinal S1, a variação média entre os índices de qualidade para os sinais processados com a SNR a priori estimada e real foi de 0,355. Observa-se também que a diferença de resultados para os diferentes níveis de SNR é praticamente a mesma.

O que se esperava nesse caso eram variações superiores às obtidas, o que não ocorreu. Essa pequena diferença é, de certa forma, incoerente, mas pode significar um limite teórico das técnicas avaliadas, principalmente se forem observados os sinais com SNR de 10dB.

Os resultados mostram que os procedimentos clássicos adotados na estimação da SNR a priori são suficientes para gerar bons resultados pelas diferentes técnicas. Porém, podem ser estudadas novas possibilidades. Por exemplo, observando as equações (14), (15) e (16), nota-se que os valores de alfa e beta são os mesmos para qualquer frequência. Nos estudos em andamento, estão sendo verificadas novas possibilidades para o uso de alfa e beta, o que pode gerar novas técnicas de redução de ruído.

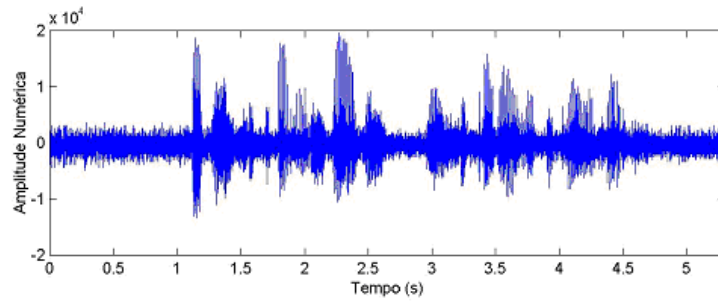


Figura 1 – Forma de onda do sinal S1 (SNR = 5dB)

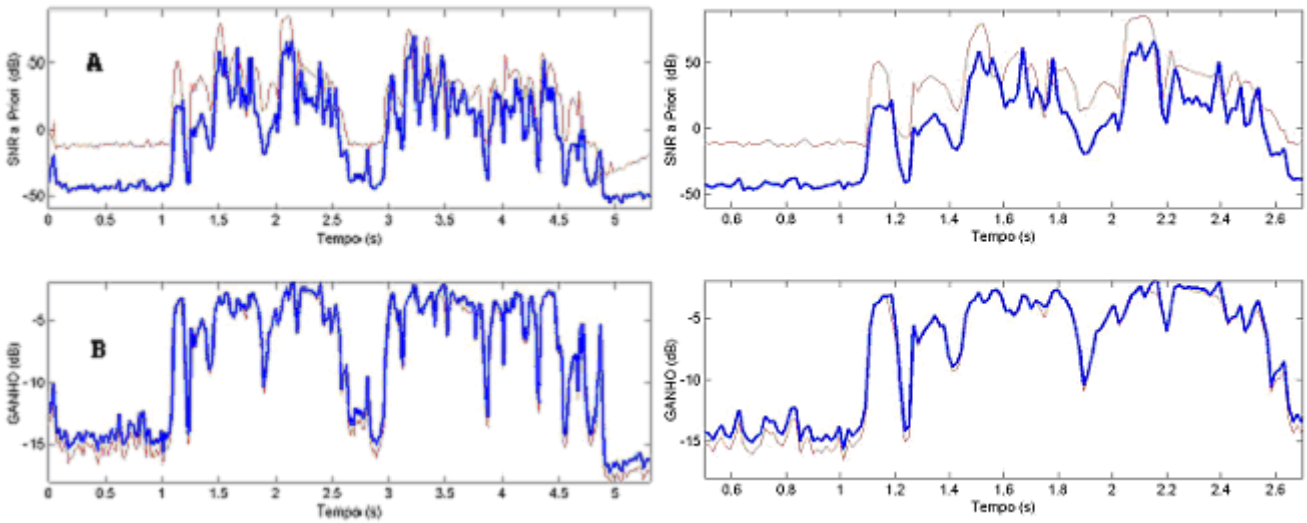


Figura 2 –Curvas de SNR e Ganho (curvas em *bold* foram obtidas com a SNR\_prio estimada)

Figura 3 – Detalhes da figura 2

TABELA I

ÍNDICE DE QUALIDADE PESQ PARA DIFERENTES TÉCNICAS E SNR A PRIORI.

Sinal	Método	SNR (dB)					
		0		5		10	
		SNRprio estimada	SNRprio real	SNRprio estimada	SNRprio real	SNRprio estimada	SNRprio real
S1	SE	2.135	2.417	2.502	2.775	2.840	3.119
	EM	2.108	2.621	2.538	2.973	2.896	3.328
	TSNR	2.069	2.417	2.457	2.775	2.803	3.119
S2	SE	2.381	2.645	2.658	2.871	2.892	3.139
	EM	2.090	2.797	2.426	3.001	2.680	3.289
	TSNR	2.362	2.645	2.629	2.871	2.864	3.139
S3	SE	2.027	2.281	2.375	2.561	2.641	2.834
	EM	1.853	2.418	2.191	2.713	2.510	3.017
	TSNR	2.003	2.281	2.344	2.561	2.606	2.834

## IV. CONCLUSÕES

Neste trabalho foram avaliadas diferentes técnicas de redução de ruído em sinais de voz baseadas na SNR a priori. As avaliações partiram, inicialmente, de uma comparação direta entre valores reais e estimados da SNR a priori. Os resultados mostram uma diferença média de 20 dB. Porém, os ganhos dos filtros, em média, são praticamente os mesmos. Para validar os estudos, foram avaliados diferentes sinais, processados por diferentes técnicas, e usando tanto a SNR a priori real como a estimada. Os resultados mostram que a diferença é quase insignificante do ponto de vista auditivo. Para alcançar os objetivos pretendidos, não foi necessário usar uma base de dados específicas para voz, já que existe uma comparação direta das técnicas usando os mesmos sinais. Como conclusão básica, entende-se que as técnicas usadas são limitadas e a busca de novos métodos de estimação para a SNR a priori pouco irá contribuir para uma melhoria global dessas técnicas. Novos caminhos podem, entretanto, serem seguidos. Um deles é a modificação na estimação da potência do ruído. O outro é a modificação na forma de calcular a SNR a priori que, até o momento, é a mesma para todas as frequências do sinal. Tais mudanças podem contribuir para o surgimento de novas técnicas de redução de ruído em sinais de voz.

## REFERÊNCIAS

- [1] Y. Ephraim and D. Malah. "Speech Enhancement Using Minimum Mean Square Error Short-Time Spectral Amplitude Estimator", IEEE Trans. On Acoust., Speech and Signal Processing, vol. 32, N° 6, December, 1984.
- [2] O. Cappé, "Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor", IEEE Trans. On Acoust., Speech and Signal Processing (TASSP), April 1994.
- [3] S.F. Boll, "Suppression of Acoustic Noise in Speech using Spectral Subtraction", IEEE Trans. Acoust., Speech, and Signal Processing (ASSP), vol. 29, p113-120, April-1979.
- [4] J.Vieira Filho, P. Scalart e J. G. Chiquito, "Redução de Ruído em Sinais de Voz: análise e avaliação das técnicas clássicas baseadas na subtração espectral a curto-terminos," Anais do XIII Simpósio Brasileiro de Telecomunicações (SBT), pp. 685-689, Águas de Lindóia-SP, Setembro de 1995.
- [5] P. Scalart and J. Vieira Filho, "Speech Enhancement Based on a Priori Signal to Noise Estimation", in Proc. 21<sup>st</sup> IEEE Int. Conf. Acoust. Speech Signal Processing, Atlanta, GA, May 1996, pp. 629-632.
- [6] J. R. McAulay, and M. Malpass, "Speech Enhancement Using a Soft-Decision Noise Suppression Filter", IEEE Trans. on Acoust., Speech and Signal Processing, vol. 28, N° 2, April, 1980.
- [7] J.E Porter and S.F. Boll, "Optimal Estimators for Spectral Restoration of Noisy Speech, Proc. IEEE ICASPP, 1984.
- [8] C. Plapous, C. Marro, and P. Scalart "Improved Signal-to-Noise Ratio Estimation for Speech Enhancement", IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 6, November, 2006, pp. 2098-2108.
- [9] I. Cohen, "Speech enhancement using a noncausal a priori SNR estimator", IEEE Signal Process. Lett., vol.11, no.9, pp.725-728, Sept. 2004.
- [10] J. Vieira-Filho, "Redução de Ruído em Sinais de Voz usando Critérios Psicoacústicos", Anais do XVIII Simpósio Brasileiro de Telecomunicações (CD), Belém, PA, 2000.
- [11] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Bell Laboratories, USA, 1978.
- [12] J. Vieira-Filho, "Redução de ruído em sinais de voz nos sistemas rádio móveis veiculares, Tese de Doutorado, UNICAMP, 1996.
- [13] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, 2001.