

Redução de Ruído em Sinais de Voz Usando Filtros de Kalman de Tempo e Frequência Discretos Combinados com Subtração Espectral de Potência e/ou Wavelets

Leandro Aureliano da Silva e Marcelo Basílio Joaquim

Resumo - Este artigo tem a finalidade de apresentar e comparar técnicas de redução de ruído na reconstrução de sinais de voz degradados por ruído. As técnicas propostas utilizam-se dos filtros de Kalman de tempo discreto e frequência discreta, ambos combinados com as técnicas de subtração espectral de potência e/ou wavelets. Os desempenhos dos algoritmos foram avaliados, utilizando-se a relação sinal/ruído segmentada e a distância de Itakura-Saito.

Palavras-Chave - Filtro de Kalman, Supressão de Ruído, Wavelets e Subtração Espectral.

Abstract - The goal of this paper has the purpose to present and to compare techniques of reduction of noise in the reconstruction of speech signals degraded by noise. The proposed methods are based on Kalman filtering in the time domain and frequency domain combined with spectral subtraction and wavelets. The performances of the algorithms had been evaluated using segmental signal-to-noise ratio and Itakura-Saito distance.

Keywords - Kalman's filter, Noise Suppression, Spectral Subtraction and Wavelets.

I. Introdução

Em muitas aplicações de processamento de sinais de voz e áudio, a presença de ruído aditivo pode degradar a qualidade e a inteligibilidade dos sinais. Com base nessas informações, muitas pesquisas têm sido realizadas nessa área, e conseqüentemente muitos algoritmos de melhoria na qualidade de sinais de áudio e voz têm surgido com o objetivo de eliminar ou mesmo reduzir a intensidade do ruído aditivo. Algumas técnicas se destacam, dentre elas: psicoacústica, subtração espectral, filtros de Wiener, filtros de Kalman combinados com redes neurais, etc. Sendo assim, este artigo apresenta o uso dos filtros de Kalman de tempo e frequência discretos, combinados com as técnicas de subtração espectral e/ou wavelets para redução de ruído em sinais de voz.

As técnicas da subtração espectral e wavelets são as responsáveis pela pré-filtragem dos sinais. Esses sinais pré-filtrados, por sua vez, serão utilizados como condições iniciais para ambos os filtros de Kalman. Já os filtros de Kalman são responsáveis pela minimização da distorção espectral, causada pelas técnicas de subtração e/ou wavelets.

As comparações de ambos os algoritmos são realizadas tendo como parâmetros de medida: a distância de Itakura-Saito e a relação sinal/ruído segmentada.

II. Descrição dos Algoritmos

A - Filtro de Kalman de Tempo Discreto (FKT)

O filtro de Kalman foi desenvolvido em 1960, por Rudolf Emil Kalman, que publicou o seu famoso artigo descrevendo uma solução recursiva para o problema da filtragem linear de dados discretos [1]. Desde então, devido aos grandes avanços da computação digital, o filtro de Kalman tem sido uma ferramenta muito importante nas áreas da navegação, da monitoração de processos, da economia e da reconstrução de sinais contaminados por ruído.

O desenvolvimento do filtro inicia-se a partir da modelagem do sinal de voz por um processo autorregressivo de ordem P ($AR(P)$) da seguinte forma:

$$x(n) = \sum_{k=1}^P a_p(k)x(n-k) + w(n) \quad (1)$$

em que: $a_p(k)$ são os coeficientes de predição linear de ordem P , $w(n)$ é o erro de predição, que está associado à excitação do modelo fonte-filtro de produção da fala e $x(n)$ o sinal de voz.

Sendo a maior parte dos sinais medidos na presença de ruído aditivo, tem-se:

$$y(n) = x(n) + v(n) \quad (2)$$

em que: $y(n)$ é o sinal de voz contaminado por ruído e $v(n)$ o ruído aditivo gaussiano.

Observando a equação 1, percebe-se que esta pode ser representada por um modelo de espaço de estado. Dessa forma, as equações 1 e 2 podem ser reescritas da seguinte maneira:

$$x(n) = A(n-1)x(n-1) + w(n) \quad (3)$$

$$y(n) = H(n)x(n) + v(n) \quad (4)$$

nas quais: $x(n)$ é um vetor de estado de dimensão $P \times 1$ de instante n ; $A(n-1)$ é a matriz transição de estados com dimensão $P \times P$, que relaciona o instante atual n com os instantes passados $n-1$; $w(n)$ é um vetor de excitação de entrada da equação de estado de ordem $P \times 1$ modelado por um ruído branco; $y(n)$ é o vetor de observação de dimensão $M \times 1$; $H(n)$ é uma matriz de distorção de canal de ordem $M \times P$;

$v(n)$ é um vetor de ruído branco aditivo de dimensão $M \times 1$ [8].

Admitindo os processos $w(n)$ e $v(n)$ como sendo ruídos brancos e independentes, as equações de correlação podem ser escritas como:

$$E[v(n)v^T(k)] = \begin{cases} R(n); & k = n \\ 0; & k \neq n \end{cases} \quad (5)$$

$$E[w(n)w^T(k)] = \begin{cases} Q(n); & k = n \\ 0; & k \neq n \end{cases} \quad (6)$$

em que: $R(n)$ e $Q(n)$ são elementos da diagonal das matrizes de covariâncias, respectivamente relacionadas ao ruído aditivo e erro de predição.

O filtro de Kalman faz as estimativas de um processo utilizando uma forma de controle de realimentação: o filtro estima o estado do processo em um dado momento, e então, a realimentação é obtida na forma de uma nova medida. Como tal, segundo Brown e Hwang [2], as equações para o filtro de Kalman são divididas em dois grupos: equações de atualização de tempo (predição) e equações de atualização da medida (correção). As equações de atualização de tempo são dadas por:

$$\hat{x}(n/n-1) = A(n-1)\hat{x}(n-1/n-1) \quad (7)$$

E as equações de atualização de medida são dadas por:

$$K(n) = P(n/n-1)H^T(n) \times [H(n)P(n/n-1)H^T(n) + R(n)]^{-1} \quad (8)$$

$$\hat{x}(n/n) = \hat{x}(n/n-1) + K(n)[y(n) - H(n)\hat{x}(n/n-1)] \quad (9)$$

$$P(n/n) = [I - K(n)H(n)]P(n/n-1) \quad (10)$$

em que: $P(n/n)$ é a matriz erro de covariância no instante n ; $K(n)$ é o ganho de Kalman responsável pela minimização dos elementos da diagonal de $P(n)$ que contém os erros médio quadráticos e $\hat{x}(n/n)$ a estimativa do sinal no instante n , dada as observações passadas de $y(n)$.

B - Filtro de Kalman de Frequência Discreta (FKF)

Seja o l -ésimo quadro definido por uma janela de análise em que $X(k,l)$ denota o espectro complexo do sinal sem ruído $x(n,l)$ e $v(n,l)$ o ruído branco gaussiano. Contaminando $x(n,l)$ com ruído branco tem-se:

$$y(n,l) = x(n,l) + v(n,l) \quad (11)$$

Reescrevendo o sinal $y(n,l)$ em função da transformada discreta de Fourier inversa (IDFT) de $X(k,l)$, tem-se:

$$y(n,l) = \frac{1}{N} \sum_{k=0}^{N-1} X(k,l) \exp\left(j2\pi \frac{kn}{N}\right) + v(n,l) \quad (12)$$

A equação 12 pode ser expandida em termos matriciais adquirindo a seguinte forma:

$$y(n,l) = \frac{1}{N} \underbrace{\begin{pmatrix} 1 \\ \exp\left(j2\pi \frac{n}{N}\right) \\ \vdots \\ \exp\left(j2\pi \frac{(N-1)n}{N}\right) \end{pmatrix}^T}_{F_n} \underbrace{\begin{pmatrix} X(0,l) \\ X(1,l) \\ \vdots \\ X(N-1,l) \end{pmatrix}}_{X_l} + v(n,l) \quad (13)$$

Assim, a equação 13 resulta em:

$$y(n,l) = F_n X_l + v(n,l) \quad (14)$$

em que: n representa o tempo dentro do quadro l e N o número de amostras no quadro. Já F_n é o vetor que contém a base da IDFT.

Na equação 14, X_l é um vetor que possui um espectro complexo para cada elemento do l -ésimo quadro. Como o tempo n não tem significado, a matriz transição de estados tem valor unitário. Assim, o algoritmo de FKF é representado pelas seguintes equações [3]:

$$K(n,l) = P_{(n-1,l)} F_n^{*T} \times [F_n P_{(n-1,l)} F_n^{*T} + R(n,l)]^{-1} \quad (15)$$

$$\hat{X}(n,l) = \hat{X}(n-1,l) + K(n,l) [y(n,l) - F_n \hat{X}(n-1,l)] \quad (16)$$

$$P(n,l) = P_{(n-1,l)} - K(n,l) F_n P_{(n-1,l)} \quad (17)$$

$$\hat{X}(0,l) = \hat{S}(e^{j\omega}) \quad (18)$$

$$P(0,l) = \begin{pmatrix} R(0,l) & 0 & \dots & 0 \\ 0 & R(1,l) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R(N-1,l) \end{pmatrix} \quad (19)$$

O algoritmo FKF é dado pelas equações 15, 16 e 17. A equação 15 traz o termo $*T$, que representa o complexo conjugado da matriz transposta. O vetor

complexo do filtro de Kalman na frequência é dado por:

$$\hat{X}_{(n,l)} = [\hat{X}(0,l), \hat{X}(1,l), \dots, \hat{X}(N-1,l)]^T \quad (20)$$

Tal vetor é uma estimativa de $X(k,l)$ e pode ser reconstruído pela aplicação da (*IDFT*).

Já os valores iniciais para o filtro são obtidos pelas equações 18 e 19. Na equação 18, o valor estimado $\hat{S}(e^{j\omega})$ é encontrado após a subtração espectral ou após a técnica de *wavelets*, mas ambos no domínio da frequência.

C - Subtração Espectral de Potência (SEP)

Dado $|Y(e^{j\omega_k})|^2$ como o espectro de potência do sinal contaminado por ruído, $|\mu(e^{j\omega_k})|^2$ como a média do espectro de potência do ruído avaliado em trechos de silêncio [4], a subtração espectral de potência é dada por:

$$|\hat{S}(e^{j\omega_k})|^2 = |Y(e^{j\omega_k})|^2 - \alpha |\mu(e^{j\omega_k})|^2 \quad (21)$$

em que: $|\hat{S}(e^{j\omega_k})|^2$ é uma estimativa do espectro de potência do sinal sem ruído. Já o parâmetro α controla o quanto de ruído é subtraído do sinal contaminado.

Devido à natureza aleatória do ruído, a subtração espectral pode gerar valores negativos, que vão decrescer a relação sinal/ruído (SNR). Para suplantarmos este problema é usada uma retificação dada por:

$$|\hat{X}(e^{j\omega_k})|^2 = \begin{cases} |\hat{S}(e^{j\omega_k})|^2, & \text{se } |\hat{S}(e^{j\omega_k})|^2 > \beta |\mu(e^{j\omega_k})|^2 \\ \beta |\mu(e^{j\omega_k})|^2, & \text{caso contrário} \end{cases} \quad (22)$$

em que: $0 \leq \beta \ll 1$ é o limite espectral mínimo. Já o parâmetro α depende da SNR e pode ser calculado como [4]:

$$\alpha = \alpha_0 - \frac{3}{20} SNR \quad -5dB \leq SNR \leq 20 \quad (23)$$

em que: α_0 tem seu valor igual a 4 [4].

Uma vez processado, o sinal estimado no domínio do tempo é obtido, utilizando-se a *IDFT* juntamente com a fase do sinal contaminado por ruído, conforme mostrado abaixo:

$$\hat{x}(n) = \frac{1}{N} \sum_{k=0}^{N-1} |\hat{X}(e^{j\omega_k})| e^{j\theta_Y(e^{j\omega_k})} e^{-j\omega_k n} \quad (24)$$

em que: $\omega_k = \frac{2\pi}{N}k$ é a frequência discreta da transformada e $\theta_Y(e^{j\omega_k})$ é a fase do sinal contaminado por ruído.

D - Algoritmo de Redução de Ruído usando Wavelets.

A Transformada *Wavelet* (TW) de um sinal $f(t)$ é definida como [5]:

$$Wf(a,b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt \quad (25)$$

Para um sinal discreto de N pontos, a integral acima pode ser aproximada por um somatório, tal que:

$$Wf(a,b) = \sum_{t=0}^{N-1} f(t) \psi_{a,b}(t) \quad (26)$$

A função $\psi_{a,b}(t)$, chamada *wavelet*, é derivada de uma função $\psi(t)$ através da seguinte transformação:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (27)$$

em que: “b” representa posição ou translação da *wavelet* e “a” chamado parâmetro de escala, o qual está associado à largura da janela.

Há uma ampla gama de escolha para função $\psi(t)$, denominada “*wavelet mãe*”, dentre elas: *Daubechies*, *symlets*, *coiflet*, etc. Neste artigo, adotou-se a *coiflet* suporte 6.

O princípio básico da redução de ruído utilizando transformada *wavelet* (TW) consiste em realizar decomposições no sinal original em coeficientes de aproximação e detalhe, gerando a árvore de decomposição. Nesse artigo, usou-se a decomposição até o nível ($m = 3$). Os coeficientes de aproximação (A_m) trazem as informações de baixa frequência associadas a *wavelet* adotada, no caso (*coiflet* suporte 6). Já os de detalhe (D_m) trazem as informações de alta frequência. Assim, a idéia básica é escolher quais coeficientes serão mantidos para preservar a informação e, em quais coeficientes serão aplicados um limiar, cujo objetivo é eliminar ou reduzir a intensidade do ruído. Então, a partir de observações dos coeficientes de aproximação e detalhe, nota-se que o coeficiente de detalhe 1 (D_1) é o que possui maior intensidade de ruído. Dessa forma, o limiar será aplicado diretamente nesse coeficiente.

Um dos limiares utilizados na literatura é o *Hard Thresholding*, que consiste em substituir os coeficientes menores que o limiar por zero. Neste artigo, a partir de vários testes, adotou-se o seguinte limiar [9]:

$$D_1(n) = \begin{cases} D_1(n), & \text{se } D_1(n) \geq 0,3 \times \text{Máx}(D_1(n)) \\ 0, & \text{Caso Contrário} \end{cases} \quad (28)$$

Após o processo de corte, utiliza-se a transformada de *wavelet* inversa (TWI), para obter o sinal pré-filtrado a ser usado como condição inicial em ambos os filtros de Kalman.

III. Resultados

Para a avaliação dos algoritmos apresentados nas seções anteriores, foram utilizados diferentes sinais de voz, amostrados numa taxa de 22050 Hz com 16 bits. Esses sinais foram contaminados por ruído branco. A janela aplicada nos algoritmos é a de Hamming com 512 amostras e com sobreposição de 50%. No procedimento de filtragem, o sinal é primeiramente filtrado pelos algoritmos de subtração espectral ou *wavelet*, e em seguida aplicado a um dos filtros de Kalman (FKT ou FKF). As figuras 1 e 2 representam o procedimento usado para o FKT, e as figuras 3 e 4 o procedimento usado para o FKF.

O diagrama em blocos da figura 1 representa a técnica de FKT combinada com subtração espectral de potência. Nessa técnica, utilizou-se um modelo AR de ordem 10 para o modelo da voz. Já para subtração espectral de potência, utilizou-se um valor fixo de $\beta = 0,002$ e o valor de α , determinado de acordo com a equação 23.

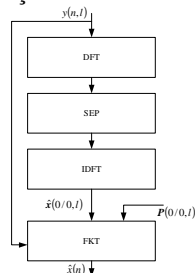


Figura 1 - FKT combinado com Subtração Espectral de Potência.

Para o diagrama em blocos da figura 2, os parâmetros para o FKT foram os mesmos adotados no anterior. Já para o algoritmo de *wavelets* adotou-se a decomposição dos coeficientes até o nível 3, sendo que a *wavelet* adotada foi a *coiflet* suporte 6.

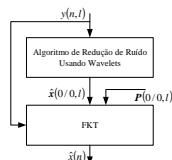


Figura 2 - FKT combinado com *wavelets*.

Já o diagrama em blocos da figura 3 mostra o processamento envolvendo FKF combinado com subtração espectral. Os parâmetros da subtração espectral foram os mesmos adotados no diagrama da figura 1 e a técnica FKF como descrita anteriormente.

O diagrama em blocos da figura 4 mostra o processamento envolvendo FKF combinado com *wavelets*. Os parâmetros da *wavelet* foram os

mesmos adotados no diagrama da figura 2. Nesse, o sinal após o processamento por *wavelets* é transformado para o domínio da frequência via *DFT*, e então entregue como condição inicial para FKF, que processa o sinal e devolve a estimativa do sinal sem ruído $\hat{x}(n)$.

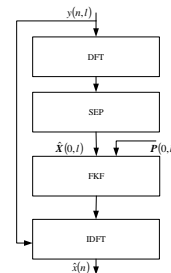


Figura 3 - FKF combinado com Subtração Espectral de Potência.

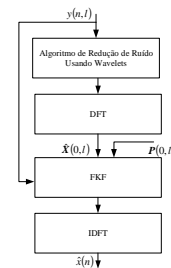


Figura 4 - FKF combinado com *wavelets*.

Para avaliar e comparar os resultados, dos algoritmos descritos acima, foram utilizadas as medidas de relação sinal/ruído segmentada (SNRseg) e a distância de Itakura-Saito ($d(a,b)$).

A SNRseg é uma medida mais efetiva que pode ser calculada em segmentos curtos de sinal de voz de forma a equilibrar os pesos atribuídos aos trechos de maior e menor intensidades do sinal. Essa medida é calculada, utilizando-se a equação 29 a seguir [6]:

$$SNR_{seg} = \frac{10}{M} \sum_{j=0}^{M-1} \log_{10} \left[\frac{\sum_{n=mj-N+1}^{mj} x^2(n)}{[x(n) - \hat{x}(n)]^2} \right] \quad (29)$$

em que: mj representa o limites de cada um dos M quadros de tamanho N .

A SNRseg não fornece uma medida significativa de desempenho quando dois sinais diferem em seus espectros. Entretanto, as medidas de distância são sensíveis às variações de espectro. Nesse caso, a distância de Itakura-Saito fornece melhores resultados e pode ser calculada utilizando os parâmetros de predição linear (LPC) conforme abaixo [7].

$$d(a,b) = \log \left[\frac{a R a^T}{b R b^T} \right] \quad (30)$$

em que: “a” é o vetor de coeficientes LPC do sinal original; “R” é a matriz de autocorrelação do sinal original e “b” é o vetor de coeficientes LPC do sinal estimado.

No primeiro teste, utilizou-se o diagrama em blocos da figura 1. Nesse teste o sinal de voz foi contaminado com ruído branco, obtendo uma relação sinal/ruído de entrada (SNRI) de 3 dB. O sinal sem ruído, o sinal contaminado e o sinal após o processamento são mostrados nas figuras 5, 6 e 7.

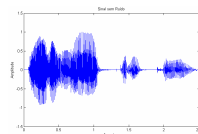


Figura 5 - Sinal sem ruído da palavra elétrica

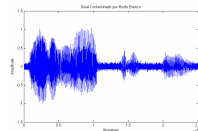


Figura 6 - Sinal contaminado por ruído branco

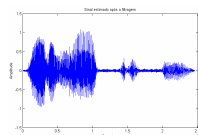


Figura 7 - Sinal após a filtragem usando FKT combinado com subtração espectral de potência.

Analisando a figura 7 em relação à figura 6, percebe-se uma considerável redução de ruído principalmente em instantes de silêncio. Esse processamento resultou em uma relação sinal/ruído segmentada de saída (SNRO) de 8 dB e uma distorção espectral medida pela distância de Itakura-Saito de 0,3250.

Outro parâmetro utilizado também para avaliar os resultados oriundos do processamento é o espectrograma. As figuras 8, 9 e 10 correspondem aos espectrogramas do sinal sem ruído, sinal contaminado e sinal estimado.

Conforme análise dos três espectrogramas, percebe-se na figura 10 que nas frequências entre 2000 e 3000 Hz, a intensidade do sinal foi ressaltada pelo algoritmo de filtragem se comparado com o espectrograma da figura 9. Pode-se perceber também que apesar de uma redução da intensidade do ruído, um ruído residual permanece.

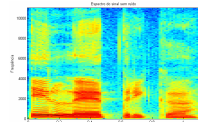


Figura 8 - Espectrograma do sinal sem ruído da palavra elétrica.

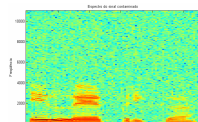


Figura 9 - Espectrograma do sinal contaminado com ruído branco.

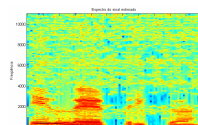


Figura 10 - Espectrograma do sinal após o processamento do algoritmo FKT combinado com subtração espectral de potência.

Apesar desse algoritmo apresentar uma redução de ruído considerável, após o processamento, surge um fenômeno conhecido como ruído musical causado pela subtração espectral. Este, por sua vez, permanece no sinal reconstruído e sua eliminação é quase impossível.

O segundo teste utilizou o algoritmo de FKT combinado com *wavelet* cujo diagrama em blocos é mostrado na figura 2. Nesse processamento a palavra elétrica também foi contaminada por ruído branco obtendo uma SNRI de 3 dB. A figura 11 mostra a forma de onda do sinal após a filtragem.

Comparando a figura 11 com a figura 6, percebe-se nos instantes de silêncio uma redução de ruído considerável. Para este teste a SNRO foi de 10 dB e a distância de Itakura-Saito de 0,3289. A figura 12 mostra o espectrograma do sinal filtrado.

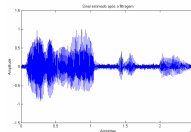


Figura 11 - Sinal estimado após o processamento de FKT combinado com *wavelet*.

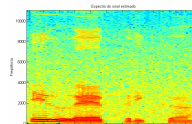


Figura 12 - Espectrograma do sinal estimado após o processamento usando FKT combinado com *wavelet*.

Pela análise do espectrograma da figura 12, percebe-se que as frequências entre 2000 e 3000 Hz, foi ressaltada quando comparado com o espectrograma da figura 9. Mas ainda assim, permanece um ruído residual no sinal.

Já no terceiro teste, utilizou-se o algoritmo de FKF combinado com subtração espectral. Nesse teste a palavra elétrica foi contaminada com ruído branco que resultou em uma SNRI de 3 dB. A figura 13, mostra o sinal filtrado.

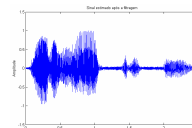


Figura 13 - Sinal filtrado pelo algoritmo de FKF combinado com subtração espectral de potência.

Nesse teste, obteve-se uma SNRO de 8 dB como no primeiro teste. Entretanto, a distância de Itakura-Saito teve seu valor igual a 0,3782 mostrando ser menos eficiente em relação à distorção do primeiro e do segundo teste.

O espectrograma do sinal filtrado neste teste é mostrado na figura 14.

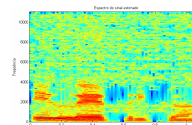


Figura 14 - Espectrograma após a filtragem por FKF combinado com subtração espectral.

Pela análise do espectrograma da figura 14, percebe-se que as frequências entre 2000 e 3000 Hz foram ressaltada quando comparadas com o espectrograma da figura 9. Mas ainda assim, permanece um ruído residual no sinal e o fenômeno conhecido como ruído musical.

No quarto teste utilizou-se o algoritmo de FKF combinado com *wavelet*. Aqui, a palavra elétrica foi contaminada com ruído branco que resultou em uma SNRI de 3 dB. A figura 15 mostra o sinal filtrado.

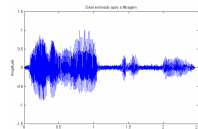


Figura 15 - Sinal após a filtragem por FKF combinado com *wavelet*.

Observando a forma de onda da figura 15 em relação à figura 6, percebe-se a redução de ruído envolvida. Esse teste resultou em uma SNRO de 10 dB. Já a distância de Itakura-Saito foi de 0,3401. O espectrograma do sinal filtrado pode ser visto na figura 16.

Analisando o espectrograma da figura 16, percebe-se que o algoritmo de filtragem realçou bem as frequências entre 2000 e 3000 Hz, quando comparado com o espectrograma do sinal contaminado da figura 9. Mas também, verifica-se a presença de um ruído residual.

Para verificar qual das técnicas apresentadas acima provoca menor distorção na reconstrução do sinal, é mostrado na figura 17, um gráfico que relaciona a distância de Itakura-Saito com relação à SNRI para a palavra usada como teste (“elétrica”) neste artigo.

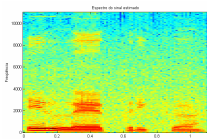


Figura 16 - Espectrograma do sinal filtrado pelo algoritmo FKF combinado com *wavelet*.

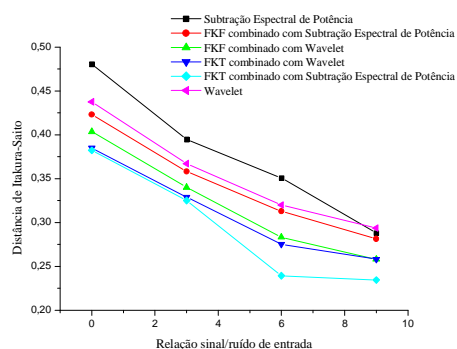


Figura 17 - Distância de Itakura-Saito em função da relação sinal/ruído de entrada (SNRI).

A partir da figura 17, observa-se que as técnicas de FKT com SE e FKT com *wavelets* são as técnicas que apresentam menor distorção espectral na reconstrução do sinal.

Para comprovar a eficiência dessas técnicas com relação à distorção espectral, a figura 18 mostra o

resultado médio de 15 palavras de todas as técnicas, relacionando distância de Itakura-Saito com relação sinal/ruído de entrada de 0, 3 e 6 dB.

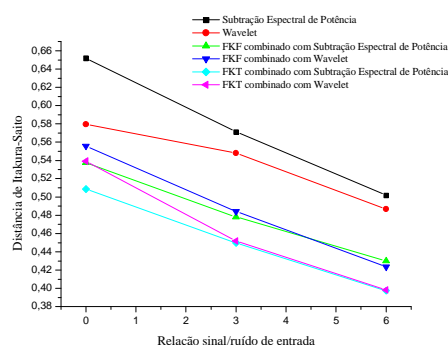


Figura 18 - Distância de Itakura-Saito em função da relação sinal/ruído de entrada (SNRI), para 15 palavras.

A partir da figura 18, verifica-se que as técnicas de FKT combinada com SE e Wavelets são as técnicas que apresentam a menor distorção na reconstrução do sinal. Com relação a SNRO, os resultados são próximos.

IV. Conclusão

Este trabalho apresentou o estudo e a implementação dos algoritmos de Kalman no tempo e frequência discretos, os quais usaram como condições iniciais, as técnicas de subtração espectral de potência e/ou *wavelet*. Para a comprovação da eficiência destas técnicas utilizou-se as medidas de distância de Itakura-Saito e a relação sinal/ruído segmentada. Após testes realizados, conclui-se que a técnica FKT, quando combinada com subtração espectral ou *wavelet*, apresenta melhores resultados na reconstrução do sinal contaminado por ruído em relação as outras técnicas descritas neste artigo.

V. Referências Bibliográficas

- [1] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transaction of the ASME-Journal of Basic Engineering, pp. 35-45, March 1960.
- [2] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley & Sons, Inc, 1997.
- [3] M. Fujimoto and Y. Ariki, *Noisy Speech Recognition Using noise reduction method based on Kalman Filter*, IEEE Proceedings of the International Conference on Acoustics Speech and Signal Processing, ICASSP'00, Vol. 3, pp. 1727-1730.
- [4] M. Berouti, R. Schwartz and J. Makhoul, *Enhancement of speech corrupted by acoustic noise*, Proc. IEEE Int. conf. Acoust., Speech, Signal Process., pp. 208-211, Apr. 1979.
- [5] Misiti, M; Misiti, Y; Oppenheim, G.; Poggi, J. M; *Wavelet Toolbox User'S Guide*, The Mathworks: Natick, 1996.
- [6] J. R. Deller, J. G. Proakis and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*, Prentice Hall, 1993.
- [7] L. R. Rabiner and R. W. Schaffer, *Digital processing of speech signals*, Prentice- Hall, 1978.
- [8] S. V. Vaseghi, *Advance Digital Signal Processing and Noise Reduction*, John Wiley & Sons, 2000.
- [9] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.