

3

Ajuste das LSF's usando a distância LSF euclidiana ponderada (d_{LSF})

Neste capítulo será abordado o método de ajuste de LSF's proposto em [5]. Esse método busca contribuir na resolução do problema de se obter parâmetros LSF mais precisos, que resultem em uma melhor qualidade de voz sintetizada no decodificador. O ajuste de LSF's é feito com base na distância LSF ponderada euclidiana, d_{LSF} , entre os valores interpolados e os referenciais (valores reais calculados realizando-se a análise LPC e a transformação para LSF para cada sub-quadro individualmente). No processo de ajuste busca-se minimizar a d_{LSF} de modo a tornar os parâmetros LSF interpolados mais próximos possível dos referenciais. O processo de ajuste de LSF's possui como parâmetros de entrada os coeficientes de otimização. Esses coeficientes são obtidos através de treinamento e permitem que o ajuste de LSF's seja direcionado, de modo a minimizar não somente a d_{LSF} , mas também a distorção espectral, DE.

3.1

Interpolação linear das LSF's e a distância LSF euclidiana ponderada (d_{LSF})

A relação entre o sinal de voz e o conjunto de 10 coeficientes LSF é ilustrada na Figura 3.1. Esta relação reside no fato de que a variação dos coeficientes LSF é função da variação do sinal de voz do qual são extraídos. É possível observar no gráfico superior quatro frases distintas separadas por instantes de silêncio, totalizando 10 s de fala. Embaixo temos a variação dos coeficientes espectrais, da 1ª a 10ª ordem, que ocorre em função da variação do sinal de voz.

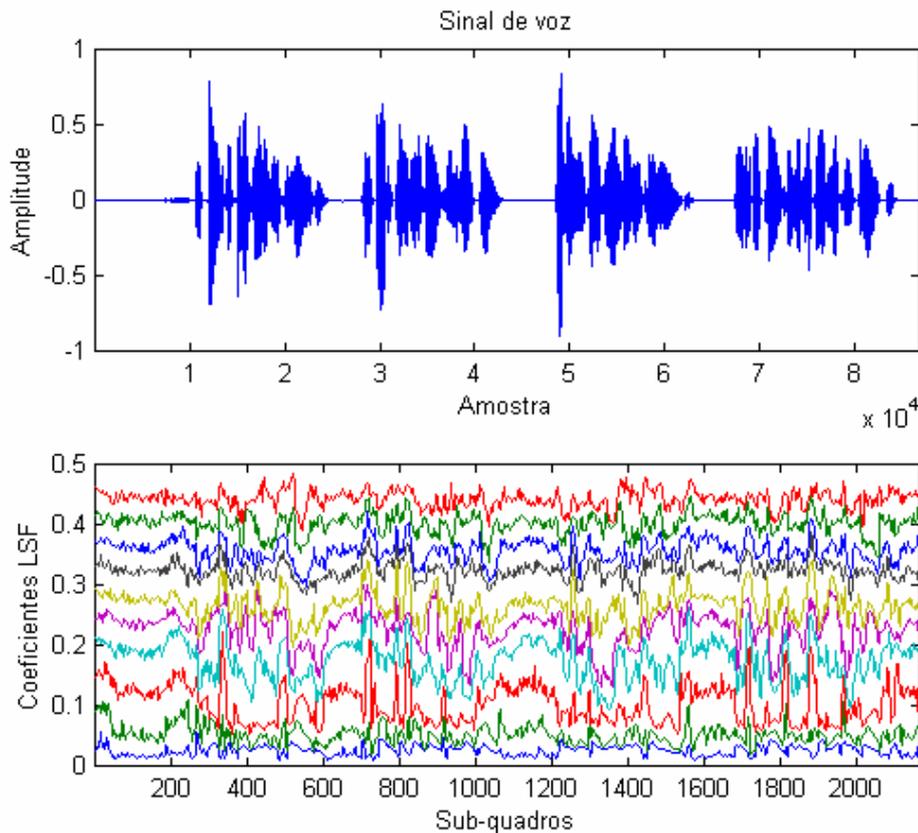


Figura 3.1 – Relação entre os coeficientes LSF e o sinal de voz.

A interpolação de LSF's entre dois quadros sucessivos é feita entre os últimos sub-quadros do quadro atual e do anterior, gerando as LSF's interpoladas. As LSF's de referência são obtidas realizando o cálculo das LSF's para todos os sub-quadros. A Figura 3.2 exibe uma parte dos coeficientes (dos 400 primeiros sub-quadros) mostrados na Figura 3.1. O gráfico acima corresponde às LSF's de referência e abaixo às interpoladas. É possível perceber claramente que as LSF's de referência apresentam uma variação maior. Isso se deve ao fato de que a interpolação linear tem um efeito de “suavização” da envoltória espectral [9].

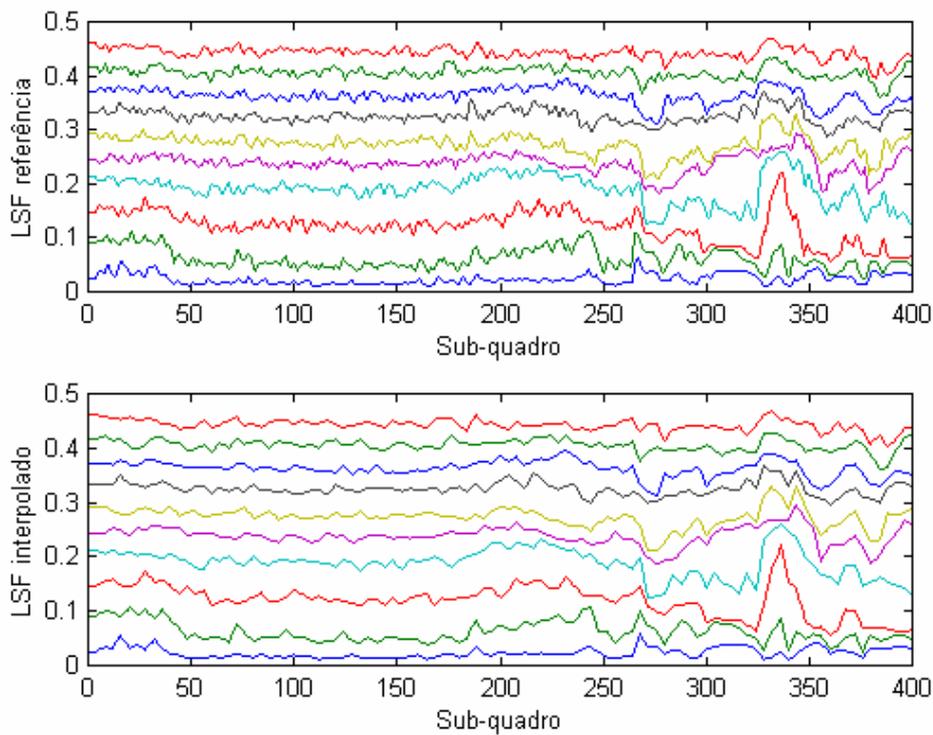


Figura 3.2 – Comparação entre os coeficientes LSF referências e interpolados.

A distância LSF euclidiana ponderada (d_{LSF_j}) [6] é definida, para um sub-quadro j qualquer, da seguinte forma:

$$d_{LSF_j} = \sum_{i=1}^p [c_i w_i^{(j)} (\omega_i^{(j)} - \hat{\omega}_i^{(j)})]^2 \quad (3-1)$$

onde

$$c_i = \begin{cases} 1 \rightarrow 1 \leq i \leq 8 \\ 0,8 \rightarrow i = 9 \\ 0,4 \rightarrow i = 10 \end{cases} \quad (3-2)$$

É importante observar q c_i também é proveniente da definição da d_{LSF_j} contida em [6]. Além disso, temos também o esquema de ponderação [10]

$$w_i^{(j)} = \frac{1}{\omega_i^{(j)} - \omega_{i-1}^{(j)}} + \frac{1}{\omega_{i+1}^{(j)} - \omega_i^{(j)}} \quad (3-3)$$

onde p é a ordem do preditor (que no caso do codificador utilizado é igual a 10), a variável $\omega_i^{(j)}$ é o i -ésimo elemento do vetor LSF de referência, e $\hat{\omega}_i^{(j)}$ o elemento correspondente do vetor interpolado.

3.2

Ajuste sem atraso

Definindo a distância total em um quadro, d_{TOT} , a soma das distâncias de todos os sub-quadros, temos que

$$d_{TOT} = \sum_{j=1}^I f_j d_{LSF_j}(\omega^{(j)}, \hat{\omega}^{(j)}) = \sum_{j=1}^I f_j \sum_{i=1}^p [c_i w_i^{(j)} (\omega_i^{(j)} - \hat{\omega}_i^{(j)})]^2 \quad (3-4)$$

onde f_j são constantes de ponderação introduzidas com o objetivo de possibilitar a otimização do processo de interpolação através de minimização das distâncias, e I é o número total de sub-quadros. Essas constantes são obtidas de modo experimental, através de rotinas de otimização não-lineares e que são funções próprias do Matlab. No codificador de voz utilizado temos quatro sub-quadros, portanto, $I = 4$. Também é importante notar que $\omega^{(j)}$ e $\hat{\omega}^{(j)}$ representam os vetores LSF e $\omega_i^{(j)}$ e $\hat{\omega}_i^{(j)}$ os i -ésimos elementos dos vetores, referência e interpolado, respectivamente.

A fórmula geral de interpolação linear é dada por

$$\hat{\omega}^{(j)} = (1 - \alpha_j) \tilde{\omega}^{(-1)} + \alpha_j \tilde{\omega}^{(0)} \quad (3-5)$$

onde $\alpha_j = \frac{j}{I}$, $\tilde{\omega}^{(-1)}$ e $\tilde{\omega}^{(0)}$ são vetores LSF associados aos últimos sub-quadros do quadro anterior e do atual, respectivamente.

Desenvolvendo a Equação (3-4), temos

$$d_{TOT} = \sum_{j=1}^I f_j \sum_{i=1}^p [c_i^2 w_i^{(j)2} (\omega_i^{(j)2} - 2\omega_i^{(j)} \hat{\omega}_i^{(j)} + \hat{\omega}_i^{(j)2})] \quad (3-6)$$

Aplicando a Equação (3-5) na (3-6), chegamos a

$$d_{TOT} = \sum_{j=1}^I f_j \sum_{i=1}^p \left\{ c_i^2 w_i^{(j)2} \left[\omega_i^{(j)2} - 2\omega_i^{(j)}(1-\alpha_j)\tilde{\omega}_i^{(-1)} - 2\omega_i^{(j)}\alpha_j\tilde{\omega}_i^{(0)} + \right. \right. \\ \left. \left. + \underbrace{\left((1-\alpha_j)\tilde{\omega}_i^{(-1)} + \alpha_j\tilde{\omega}_i^{(0)} \right)^2}_{(*)} \right] \right\} \quad (3-7)$$

$$(*) = (1-\alpha_j)^2 \tilde{\omega}_i^{(-1)2} + 2(1-\alpha_j)\tilde{\omega}_i^{(-1)}\alpha_j\tilde{\omega}_i^{(0)} + \alpha_j^2 \tilde{\omega}_i^{(0)2}$$

Para minimizar d_{TOT} , deriva-se em relação aos coeficientes LSF do último sub-quadro do quadro atual, $\tilde{\omega}_k^{(0)}$, sendo $k = 1, \dots, p$ (lembrando que $p = 10$ é a ordem do preditor), e iguala-se o resultado a zero. A derivação é feita em relação aos coeficientes do último sub-quadro porque este é exatamente o que se deseja ajustar.

$$\frac{d}{d\tilde{\omega}_k^{(0)}}(d_{TOT}) = \sum_{j=1}^I f_j \left\{ c_k^2 w_k^{(j)2} \left[-2\omega_k^{(j)}\alpha_j + 2(1-\alpha_j)\tilde{\omega}_k^{(-1)}\alpha_j + 2\alpha_j^2\tilde{\omega}_k^{(0)} \right] \right\} = 0 \quad (3-8)$$

$$\sum_{j=1}^I f_j \left\{ c_k^2 w_k^{(j)2} \left[-2\omega_k^{(j)}\alpha_j + 2(1-\alpha_j)\tilde{\omega}_k^{(-1)}\alpha_j \right] \right\} = \\ = -\sum_{j=1}^I f_j \left(c_k^2 w_k^{(j)2} 2\alpha_j^2 \tilde{\omega}_k^{(0)} \right) \quad (3-9)$$

$$\underbrace{\sum_{j=1}^I f_j c_k^2 w_k^{(j)2} 2\alpha_j \left[(1-\alpha_j)\tilde{\omega}_k^{(-1)} - \omega_k^{(j)} \right]}_{b_k} = - \underbrace{\left[\sum_{j=1}^I f_j \left(c_k^2 w_k^{(j)2} \alpha_j^2 \right) \right]}_{a_k} 2\tilde{\omega}_k^{(0)} \quad (3-10)$$

$$a_k = \sum_{j=1}^I f_j \left(c_k^2 w_k^{(j)2} \alpha_j^2 \right) \quad (3-11)$$

$$b_k = \sum_{j=1}^I f_j c_k^2 w_k^{(j)2} 2\alpha_j \left[(1-\alpha_j)\tilde{\omega}_k^{(-1)} - \omega_k^{(j)} \right] \quad (3-12)$$

Ou seja,

$$\tilde{\omega}_k^{(0)} = -\frac{b_k}{2a_k} \quad (3-13)$$

É importante observar na Equação (3-12) que o novo valor das LSF's para o último sub-quadro do quadro atual é calculado em função das diferenças entre as LSF's do último sub-quadro do quadro anterior e as LSF's referências.

Esse método, portanto, não utiliza nenhuma informação à frente do quadro atual. Por isso ele é chamado de “no lookahead”.

3.3

Ajuste com atraso

Levando em conta que um atraso de até 100 ms é aceitável em uma conversa telefônica, por exemplo, e que no codificador utilizado cada quadro tem 20 ms, podemos pensar em utilizar informação além do quadro atual. Sendo assim, no cálculo da distância total entra agora não só o somatório referente aos sub-quadros do quadro atual, mas também os sub-quadros à frente que forem utilizados. Assim,

$$d_{TOT} = \sum_{j=1}^L f_j d_{LSF_j}(\omega^{(j)}, \hat{\omega}^{(j)}) + \sum_{j=1}^L l_j d_{LSF_j}(\omega_N^{(j)}, \hat{\omega}_N^{(j)}) \quad (3-14)$$

onde L é o número de sub-quadros à frente utilizados, e as variáveis $\omega_N^{(j)}$ e $\hat{\omega}_N^{(j)}$ representam, respectivamente, os vetores LSF de referência e interpolado para os sub-quadros à frente. Se estabelecermos como limite apenas um quadro à frente, teremos L podendo assumir valores entre 1 e 4. Vale ressaltar também que para os sub-quadros à frente teremos também um outro conjunto de fatores de otimização, representado por $\{l_j, j = 1, \dots, L\}$.

Seguindo o mesmo desenvolvimento anteriormente mostrado, chegamos a

$$d_{TOT} = \sum_{j=1}^L f_j \sum_{i=1}^p [c_i w_i^{(j)} (\omega_i^{(j)} - \hat{\omega}_i^{(j)})]^2 + \sum_{j=1}^L l_j \sum_{i=1}^p [c_i w_{N,i}^{(j)} (\omega_{N,i}^{(j)} - \hat{\omega}_{N,i}^{(j)})]^2 \quad (3-15)$$

A primeira parte dessa equação é igual à (3-4), e representa a d_{LSF} do quadro atual. O segundo par de somatórios representa a d_{LSF} do quadro à frente.

A variável $w_{N,i}^{(j)}$ representa o esquema de ponderação, como apresentado na Equação (3-3). É importante observar que $w_{N,i}^{(j)}$, assim como as outras variáveis $(L, l_j, \omega_N^{(j)}, \hat{\omega}_N^{(j)})$, se refere ao quadro à frente.

A interpolação do quadro à frente é obtida através de

$$\hat{\omega}_N^{(j)} = (1 - \beta_j) \tilde{\omega}^{(0)} + \beta_j \tilde{\omega}^{(1)} \quad (3-16)$$

onde

$$\beta_j = \frac{j}{I} \quad (3-17)$$

Substituindo a Equação (3-16) na (3-15), temos

$$\begin{aligned} d_{TOT} = & \sum_{j=1}^I f_j \sum_{i=1}^p \left[c_i^2 w_i^{(j)2} \left(\omega_i^{(j)2} - 2\omega_i^{(j)} \hat{\omega}_i^{(j)} + \hat{\omega}_i^{(j)2} \right) \right] + \\ & + \sum_{j=1}^L l_j \sum_{i=1}^p \left[c_i^2 w_{N,i}^{(j)2} \left(\underbrace{\omega_{N,i}^{(j)2} - 2\omega_{N,i}^{(j)} \hat{\omega}_{N,i}^{(j)} + \hat{\omega}_{N,i}^{(j)2}}_{(*)} \right) \right] \end{aligned} \quad (3-18)$$

Desenvolvendo, resulta

$$\begin{aligned} (*) &= \omega_{N,i}^{(j)2} - 2\omega_{N,i}^{(j)} (1 - \beta_j) \tilde{\omega}_i^{(0)} - 2\omega_{N,i}^{(j)} \beta_j \tilde{\omega}_i^{(1)} + \underbrace{\left((1 - \beta_j) \tilde{\omega}_i^{(0)} + \beta_j \tilde{\omega}_i^{(1)} \right)^2}_{(**)} \\ (** &= (1 - \beta_j)^2 \tilde{\omega}_i^{(0)2} + 2(1 - \beta_j) \tilde{\omega}_i^{(0)} \beta_j \tilde{\omega}_i^{(1)} + \beta_j^2 \tilde{\omega}_i^{(1)2} \end{aligned}$$

Pode-se notar que a primeira parte da Equação (3-18) é igual a (3-6). Isso permitirá uma simplificação adiante. A minimização é feita derivando d_{TOT} e igualando a zero. Entretanto, esse procedimento já foi realizado para a primeira parte da Equação (3-18) em (3-8). Lembrando que $\tilde{\omega}_k^{(0)}$ representa os coeficientes LSF do último sub-quadro do quadro atual, sendo $k = 1, \dots, p$ (onde $p = 10$ é a ordem do predictor), temos

$$+ \sum_{j=1}^L l_j \left\{ c_k^2 w_{N,k}^{(j)2} \left[-2\omega_{N,k}^{(j)} (1 - \beta_j) + 2(1 - \beta_j)^2 \tilde{\omega}_k^{(0)2} + 2(1 - \beta_j) \beta_j \tilde{\omega}_k^{(1)} \right] \right\} = 0 \quad (3-19)$$

De fato, é possível observar que a primeira parte da Equação (3-19) é igual a (3-8).

$$\underbrace{\sum_{j=1}^I f_j c_k^2 w_k^{(j)2} 2\alpha_j [(1 - \alpha_j) \tilde{\omega}_k^{(-1)} - \omega_k^{(j)}]}_{b_k} + \underbrace{\sum_{j=1}^L l_j c_k^2 w_{N,k}^{(j)2} 2(1 - \beta_j) (\beta_j \tilde{\omega}_k^{(1)} - \omega_{N,k}^{(j)})}_{d_k} =$$

$$= - \left[\underbrace{\sum_{j=1}^I f_j c_k^2 w_k^{(j)2} \alpha_j^2}_{a_k} + \underbrace{\sum_{j=1}^L l_j c_k^2 w_{N,k}^{(j)2} (1 - \beta_j)^2}_{c_k} \right] 2\tilde{\omega}_k^{(0)} \quad (3-20)$$

$$c_k = a_k + \sum_{j=1}^L l_j c_k^2 w_{N,k}^{(j)2} (1 - \beta_j)^2 \quad (3-21)$$

$$d_k = b_k + \sum_{j=1}^L l_j c_k^2 w_{N,k}^{(j)2} 2(1 - \beta_j) (\beta_j \tilde{\omega}_k^{(1)} - \omega_{N,k}^{(j)}) \quad (3-22)$$

É importante ressaltar que a_k e b_k foram definidos nas equações (3-11) e (3-12), respectivamente. Finalmente, chegamos a um ajuste de LSF's para o último sub-quadro do quadro atual, levando em conta também sub-quadros à frente. Em função disso, esse método é chamado de “lookahead” e as LSF's ajustadas são expressas por

$$\tilde{\omega}_k^{(0)} = -\frac{d_k}{2c_k} \quad (3-23)$$

onde c_k e d_k são dadas pelas equações (3-21) e (3-22), respectivamente.

3.4

A distorção espectral (DE) e os coeficientes de otimização

A distorção espectral, associada ao i -ésimo sub-quadro, é uma medida objetiva de desempenho. É representada matematicamente na forma da equação (3-24).

$$DE_i = \sqrt{\frac{1}{F_s} \int_0^{F_s} \left[10 \log_{10} \frac{S_i(f)}{\hat{S}_i(f)} \right]^2 df} \quad (3-24)$$

onde a DE_i é dada em dB , F_s é a frequência de amostragem, e $S_i(f)$ e $\hat{S}_i(f)$ são as envoltórias espectrais LPC do sinal de entrada em suas versões referencial e interpolada, respectivamente. Essas funções são expressas por

$$S_i(f) = \frac{1}{A_i \left(e^{\frac{j2\pi f}{F_s}} \right)} \quad (3-25)$$

$$\hat{S}_i(f) = \frac{1}{\hat{A}_i \left(e^{\frac{j2\pi f}{F_s}} \right)} \quad (3-26)$$

onde $A_i(z)$ e $\hat{A}_i(z)$ são os polinômios LPC, original e interpolado, respectivamente.

A distorção espectral média é definida por

$$DE = \frac{1}{N} \sum_{i=1}^N DE_i \quad (3-27)$$

onde N é o número total de sub-quadros em análise.

Os coeficientes de otimização, $\{f_j\}$ e $\{l_j\}$, apresentados nas equações (3-4) e (3-14), respectivamente, são obtidos de modo experimental, através de um processo de treinamento. Esse processo consiste, conforme já dito, em executar

uma rotina de otimização não-linear que é função própria do Matlab. Como entrada essa função recebe um conjunto inicial de coeficientes, definido como $\{f_j = 1, j = 1, \dots, I\}$. O conjunto de coeficientes é aplicado na rotina de ajuste de LSF's, e em função da saída, que é a DE média, a função de otimização determina qual será o novo conjunto de coeficientes para a iteração seguinte. Ao longo das iterações a função de otimização busca determinar coeficientes que permitam ao ajuste de LSF's convergir para a menor DE possível. Atingir esse limite depende do número de iterações realizadas. O fim do processo de otimização ocorre quando a diferença entre a DE média calculada em uma iteração e a calculada na iteração anterior é menor que uma determinada tolerância. Nesse caso foi utilizada a tolerância padrão do Matlab, que é de 10^{-4} . O mesmo procedimento é aplicado também utilizando a d_{LSF} como saída da rotina de ajuste de LSF's. Com isso obtêm-se conjuntos de coeficientes que minimizam cada uma das medidas separadamente.

Os coeficientes de otimização utilizados em [11], são mostrados na Tabela 3.1.

Tabela 3.1 – Coeficientes de otimização apresentados em [11].

Medida minimizada	Método	Coeficientes de otimização				
		f_1	f_2	f_3	f_4	l_1
DE	No lookahead	0	0	0	1	
	Lookahead	0,2	0	0,2	1	0,4
d_{LSF}	No lookahead	0,6	0,3	0,5	1	
	Lookahead	1,2	0,8	1	1	2

Note-se que, para o método “no lookahead”, quando se deseja minimizar a DE, tem-se o coeficiente do 4º sub-quadro igual a 1 e todos os demais iguais a zero. Tal combinação de coeficientes é equivalente a uma interpolação linear convencional, sem nenhum ajuste das LSF's do último sub-quadro. Isso significa que não se obterá melhoria na DE realizando o ajuste com esses coeficientes.

3.5

Condições experimentais e resultados de simulações

Foram utilizados como conjunto de teste 10 arquivos de áudio contendo cada um 10 frases pronunciadas por 10 locutores diferentes, sendo 5 homens e 5 mulheres. Cada arquivo tem em média 30 s de duração. Os resultados apresentados são médias das medidas realizadas para cada arquivo individualmente. Foram avaliadas a distorção espectral média (DE), os percentuais de outliers entre 2 e 4% e maiores que 4%, e a d_{LSF} . As medidas de distorção foram realizadas no codificador descrito no capítulo 2 antes e depois da quantização.

Além das medidas de distorção, também foi aplicado o teste PESQ (Perceptual Evaluation of Speech Quality) [7][8]. O PESQ é uma medida objetiva que tenta avaliar a qualidade de voz que seria percebida por um ouvinte em um teste subjetivo de escuta. Essa medida foi padronizada na recomendação P.862 do ITU-T. Devido a essa padronização, atualmente o PESQ é muito usado como método para avaliação de qualidade de voz.

Assim como as medidas de distorção entre LSF's, o que o teste PESQ faz é comparar dois sinais, um referência, e outro modificado. Nesse caso, a diferença é que o sinal referência é o sinal original, sem ter passado pelo codificador, nem nenhum outro processo que o altere. O sinal modificado é o sinal sintetizado resultante da saída do decodificador.

Como saída, o teste PESQ retorna uma nota, em uma escala que varia de 1 a 4,5, de modo semelhante um teste subjetivo, como o MOS (Mean Opinion Score), por exemplo.

De modo a permitir uma comparação através da qual se possa avaliar a eficiência das melhorias propostas, foram inicialmente feitas medidas utilizando-se o codificador sem nenhuma alteração, isto é, sem a aplicação do método de ajuste de LSF's. Os resultados obtidos são apresentados na tabela 3.2. Nessa tabela apresenta-se os valores de DE, outliers e PESQ antes e após a quantização dos parâmetros LSF.

Tabela 3.2 – Medidas referenciais realizadas sem ajuste de LSF's. “NSA” significa “não se aplica”. “Q” significa “quantizar”.

	DE média	%2-4	%>4	d_{LSF}	PESQ
Antes de Q	1,1249	14,3329%	0,7476%	0,6007	NSA
Depois de Q	1,5892	23,1301%	1,2527%	0,9682	2,3342

3.5.1

Simulações com os coeficientes apresentados em [11]

Os resultados obtidos como os coeficientes apresentados em [11] são mostrados na Tabela 3.3 para as medidas realizadas antes da quantização. Não foram apresentadas medidas para o método “no lookahead” minimizando a DE porque, conforme explicado anteriormente, nesse caso os resultados são os mesmos obtidos na interpolação linear convencional. Pode-se observar que, conforme já era esperado, o método “lookahead” apresenta um melhor desempenho. Além disso, é possível notar também que a minimização da d_{LSF} , apesar de apresentar uma DE média maior, proporciona um melhor resultado em relação aos outliers.

Tabela 3.3 – Medidas antes do quantizador utilizando coeficientes apresentados em [11].

Medida minimizada	Método	DE média	%2-4	%>4	d_{LSF}
DE	Lookahead	1,1164	11,3467%	0,4774%	0,5025
d_{LSF}	No lookahead	1,1627	10,0917%	0,4084%	0,4808
	Lookahead	1,1356	7,8649%	0,2208%	0,4231

Para as medidas realizadas depois da quantização, foram obtidos os resultados mostrados na Tabela 3.4. Uma observação importante que pode ser feita quando se compara as Tabelas 3.3 e 3.4 com a 3.2 é a de que, na minimização da d_{LSF} , apesar de se ter uma DE média antes da quantização maior que a referência, depois da quantização, a medida é menor. Também chama a atenção uma pequena melhoria obtida no teste PESQ utilizando-se o método “lookahead”. Além disso, apesar da DE média no modo “lookahead” após a

quantização ser menor minimizando a d_{LSF} do que minimizando a DE, o PESQ para a minimização da DE é maior.

Tabela 3.4 – Medidas após a quantização utilizando coeficientes apresentados em [11].

Medida minimizada	Método	DE média	%2-4	%>4	d_{LSF}	PESQ
DE	Lookahead	1,5257	20,1671%	0,9066%	0,8614	2,3614
d_{LSF}	No lookahead	1,5658	20,9424%	1,0327%	0,9117	2,3243
	Lookahead	1,4928	17,6767%	0,6344%	0,7818	2,3408

3.5.2

Obtenção dos novos coeficientes de otimização

Utilizando os procedimentos descritos na Seção 3.4, foram determinados novos coeficientes de otimização. Para isso, foi utilizado um conjunto de treinamento formado a partir de 8 conjuntos de 10 frases, pronunciadas por 8 locutores diferentes, chegando-se a um total de 80 frases, em uma duração de aproximadamente 240s. Foram feitos treinos de modo a obter coeficientes de otimização que minimizassem a DE e a d_{LSF} . No caso dos coeficientes que minimizam a DE, pode-se observar que os mesmos apresentam um comportamento bem determinado, conforme é possível ver na Figura 3.3. No eixo y dessa figura tem-se os valores dos coeficientes, e no x, os sub-quadros aos quais se referem, sendo que o sub-quadro 1 é o primeiro do quadro atual e o 5 é o primeiro do quadro à frente. As curvas representam os modos de ajuste “no lookahead” e “lookahead”. A curva cujos pontos estão marcados com um sinal “+”, que vai somente até o 4º sub-quadro é relativa ao modo “no lookahead”. As demais curvas descrevem os coeficientes do modo “lookahead” utilizando de 1 a 4 sub-quadros à frente, conforme indicado na legenda.

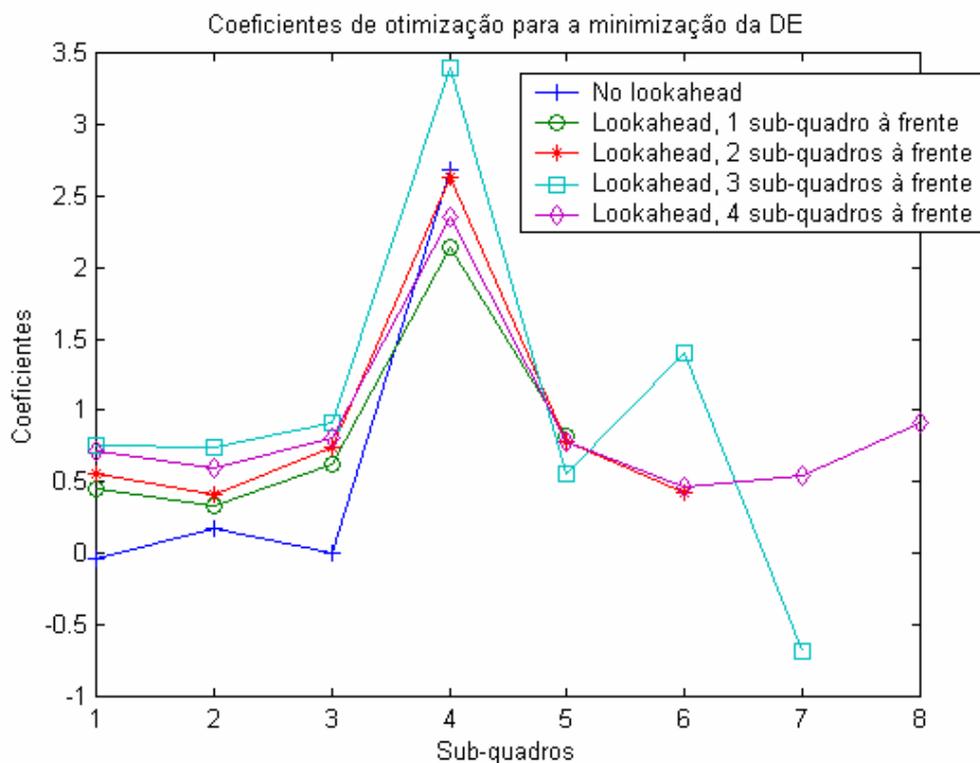


Figura 3.4 – Tendência dos novos coeficientes de otimização.

Para o caso dos coeficientes de otimização que minimizam a d_{LSF} não foi observada uma tendência definida, ficando a impressão de que os mesmos situam-se na gama de valores possíveis de modo praticamente aleatório, ou seja, se fosse feito um gráfico com esses coeficientes o resultado seria um conjunto de pontos dispersos.

As Tabelas 3.5 e 3.6 mostram os coeficientes de otimização obtidos para a minimização da DE e da d_{LSF} , respectivamente. Em cada uma dessas tabelas existe uma linha para cada número de sub-quadros à frente sendo utilizados. A linha em que tal número é igual a zero, significa que o modo é o “no lookahead”.

Tabela 3.5 – Novos coeficientes de otimização que minimizam a DE.

Número de sub-quadros	f_1	f_2	f_3	f_4	l_1	l_2	l_3	l_4
0	-0,0313	0,1757	0,0013	2,6795				
1	0,4541	0,3308	0,6166	2,1348	0,8182			
2	0,5528	0,4104	0,7483	2,6263	0,7765	0,4317		
3	0,7566	0,7417	0,9155	3,3966	0,5547	1,3983	-0,6867	
4	0,7182	0,5928	0,8121	2,3474	0,7844	0,4603	0,5484	0,9113

Tabela 3.6 – Novos coeficientes de otimização que minimizam a d_{LSF} .

Número de sub-quadros	f_1	f_2	f_3	f_4	l_1	l_2	l_3	l_4
0	0,2906	1,0781	0,1280	1,4228				
1	1,0624	0,8867	1,1000	0,7934	1,5589			
2	1,1145	0,9231	1,1616	0,8186	1,2913	0,7215		
3	1,2716	1,1423	1,2045	1,2105	0,8622	1,5503	-0,2931	
4	1,0912	1,1062	1,0697	1,1002	1,0540	0,5701	0,9464	1,0262

3.6

Desempenho dos novos coeficientes minimizando a DE

3.6.1

Análise dos resultados obtidos antes da quantização

Tabela 3.7 – Medidas antes de quantizar minimizando DE utilizando novos coeficientes.

Modo	Sub-quadros	DE média	%2-4	%>4	d_{LSF}
Sem ajuste		1,1249	14,3329%	0,7476%	0,6007
No lookahead		1,13460	13,71301%	0,65975%	0,57371
Lookahead	1	1,10801	10,18787%	0,37904%	0,46950
	2	1,10835	10,20664%	0,36361%	0,46665
	3	1,09848	9,10952%	0,30996%	0,43855
	4	1,08877	8,35608%	0,23453%	0,41558

Comparando com os resultados das medidas referenciais, sem ajuste de LSF's, é possível notar que no modo “no lookahead”, apesar de ter-se uma DE média ligeiramente maior, obteve-se menores valores para os outliers, tanto de 2 à 4%, quanto maiores que 4%. Também a d_{LSF} apresentou um melhor resultado.

Já no modo “lookahead” em todas as medidas realizadas obteve-se uma melhoria mais significativa, principalmente para os outliers e, dentre eles, os maiores que 4%. Tal avaliação confirma que o ajuste é eficaz em aproximar os coeficientes LSF dos seus valores referenciais, na medida em que minimiza a distorção.

Nas Figuras 3.5 a 3.8 são apresentados alguns gráficos que ajudam a visualizar os resultados apresentados. No eixo x temos o número de sub-quadros à

frente sendo utilizados. O zero significa que o método empregado foi o “no lookahead”. No eixo y temos a medida de distorção avaliada. Em todos os gráficos existem uma reta, que representa a medida referencial, sem ajuste, e uma curva, que descreve a evolução da medida em função do ajuste.

Distorção espectral - DE média

Na Figura 3.5 tem-se a representação da DE média. É possível observar, conforme já mencionado, que de início, no modo “no lookahead”, tem-se um valor maior que o referencial. Entretanto, ocorre uma redução quando se passa para o modo “lookahead” com utilizando 1 sub-quadro à frente. Com 2 sub-quadros essa redução não se mantém, havendo um ligeiro aumento. A partir de 3 é retomada a queda, atingindo um mínimo em 4.

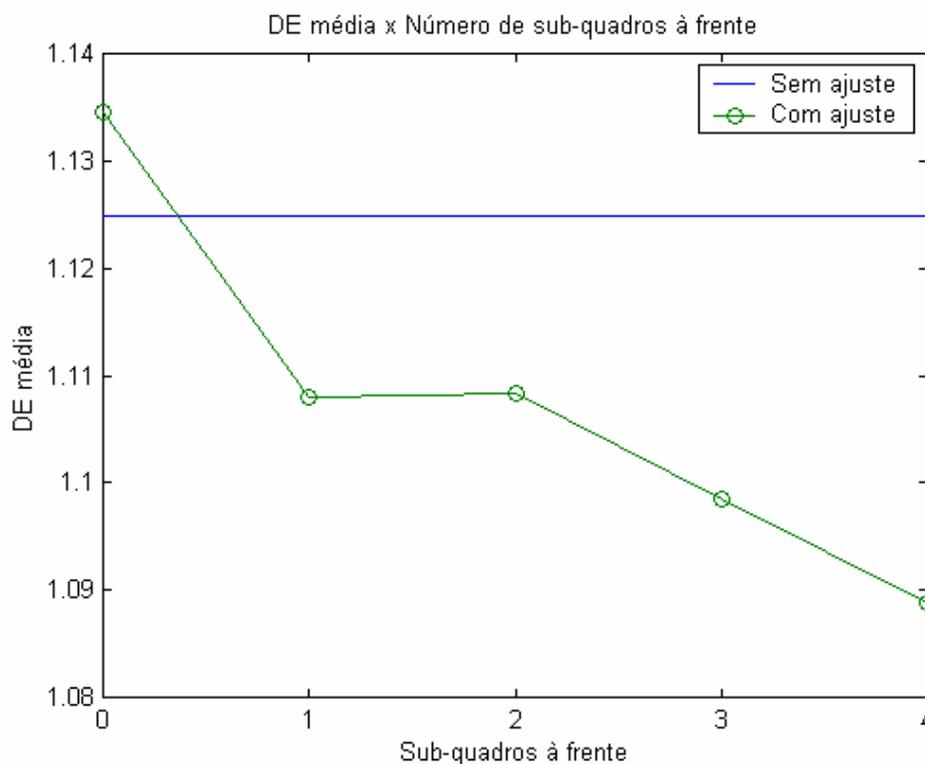


Figura 3.5 – DE média antes da quantização utilizando os novos coeficientes.

Outliers

Os dois gráficos seguintes (Figuras 3.6 e 3.7) mostram os outliers de 2 a 4% e maiores que 4%, respectivamente. Pode-se notar que nesse caso, ao contrário do que aconteceu com a DE média, já se começa com um valor abaixo do referencial. Isso indica que em se tratando de redução de outliers, mesmo o método “no lookahead” já é capaz de apresentar um bom desempenho. No modo “lookahead”, na passagem da utilização de 1 para 2 sub-quadros à frente não se observa melhoria. A redução volta a ocorrer em 3, até 4, onde se tem o mínimo.

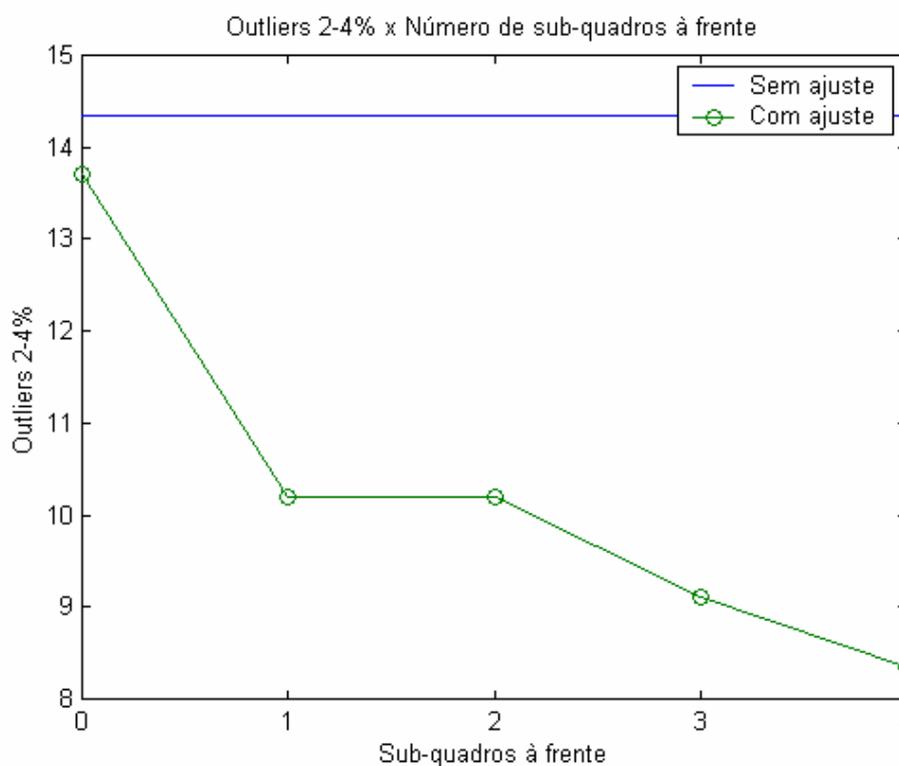


Figura 3.6 – Outliers de 2-4% antes da quantização utilizando os novos coeficientes.

Para os outliers maiores que 4% também é possível observar uma melhoria desde o início, no modo “no lookahead”. A diferença aqui é que ocorre redução para todas as quantidades de sub-quadros sendo utilizados, mesmo entre 1 e 2, ainda que não tão significativa nesse intervalo. O gráfico confirma a idéia de que o ajuste tem uma ação mais efetiva sobre os outliers do que sobre a DE média, pois atua de modo a tornar os coeficientes LSF mais próximos de seu valor real.

Vale lembrar que os outliers são picos de distorção e, por isso, sua redução contribui de maneira significativa para a melhoria da qualidade do sinal.

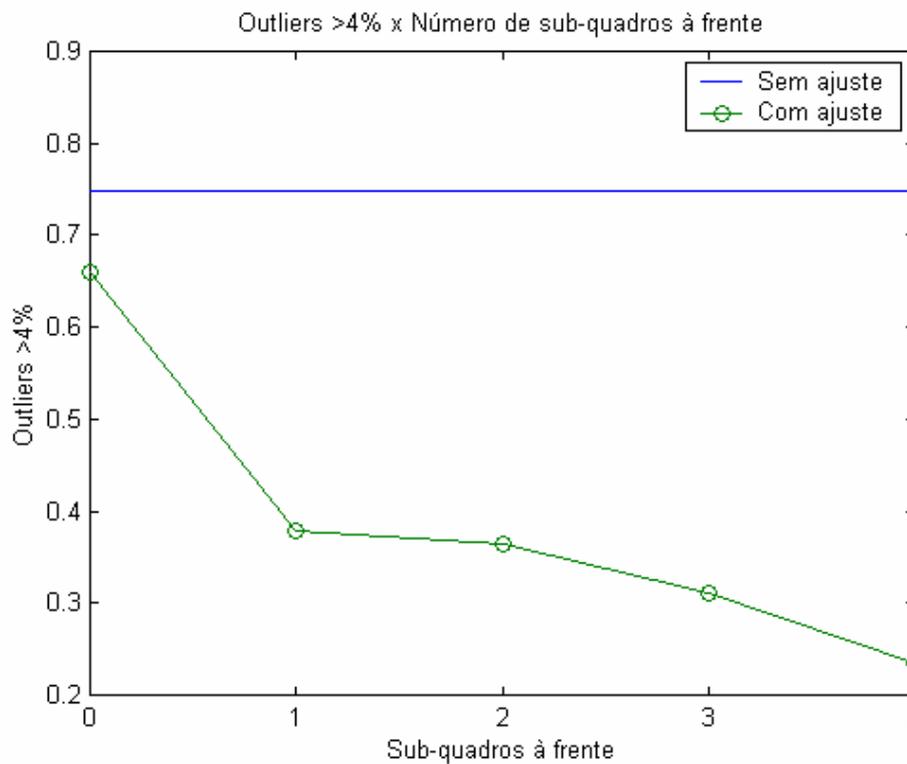


Figura 3.7 – Outliers >4% antes da quantização utilizando os novos coeficientes.

Distância euclidiana ponderada - d_{LSF}

Para a d_{LSF} tem-se um comportamento semelhante ao dos outliers maiores que 4%. É interessante notar que, apesar de se aplicar um método visando minimizar a DE, a d_{LSF} também sofre uma redução.

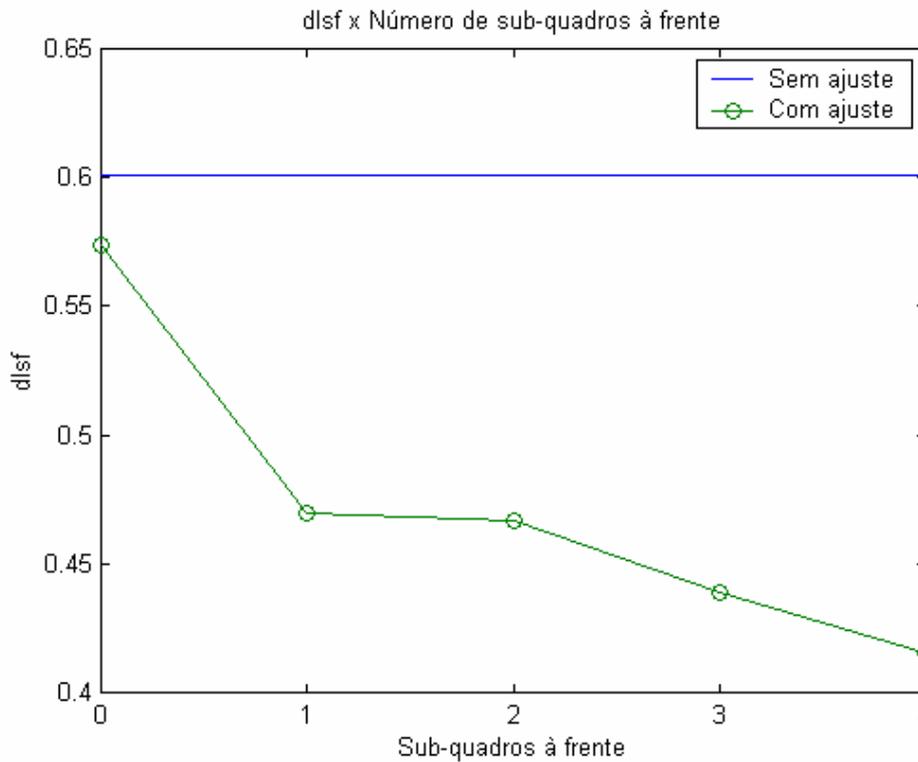


Figura 3.8 – d_{LSF} antes da quantização utilizando os novos coeficientes.

3.6.2

Análise dos resultados obtidos após a quantização

Os resultados apresentados até agora referem-se a medidas realizadas antes da quantização. Para depois do quantizador, temos os resultados apresentados na Tabela 3.8.

Tabela 3.8 - Medidas após quantizar minimizando a DE utilizando os novos coeficientes.

Modo	Sub-quadros	DE média	%2-4	%>4	d_{LSF}
Sem ajuste		1,5892	23,1301%	1,2527%	0,9682
No lookahead		1,57862	22,58240%	1,20047%	0,94475
Lookahead	1	1,51321	19,49065%	0,81433%	0,83456
	2	1,51591	19,52485%	0,82803%	0,83328
	3	1,50789	18,73486%	0,76365%	0,81564
	4	1,50392	18,28930%	0,72963%	0,80318

Os coeficientes LSF obtidos na saída do quantizador são os que efetivamente serão enviados ao decodificador para serem sintetizados e gerar o sinal de voz de saída. As medidas de distorção após o quantizador permitem, portanto, uma avaliação mais realista da melhoria obtida.

Pode-se observar que a DE média sofre uma redução a partir do modo “no lookahead”. O mesmo acontece para os outliers e a d_{LSF} . A seguir são apresentados gráficos para os resultados apresentados.

Distorção espectral – DE média

Na Figura 3.9 é possível notar uma redução mais acentuada na DE média quando se passa do modo “no lookahead” para o “lookahead”. Entretanto, já no modo “lookahead”, não se percebe uma redução muito significativa à medida em que aumenta o número de sub-quadros à frente sendo utilizados. Na verdade, ocorre até um pequeno aumento na passagem de 1 para 2 sub-quadros. Em todas as situações avaliadas o resultado obtido é sempre melhor que o do referencial.

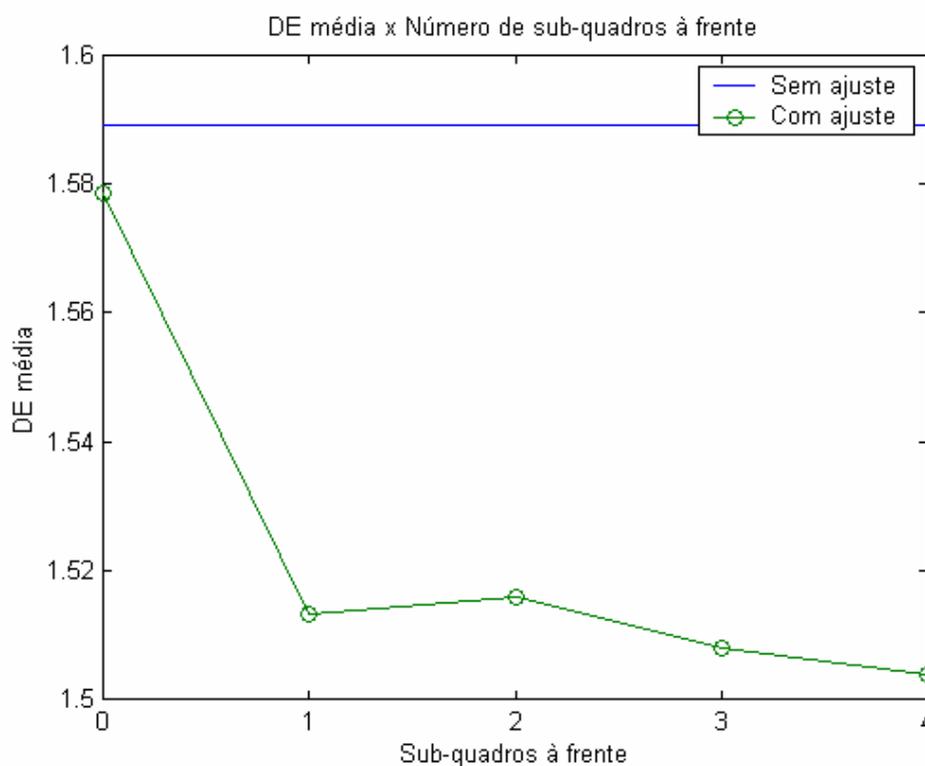


Figura 3.9 - DE média após a quantização utilizando os novos coeficientes.

Outliers

Os outliers (Figuras 3.10 e 3.11) apresentam curvas muito parecidas graficamente. Ambas as faixas, tanto de 2 a 4%, quanto maiores que 4%, têm uma redução acentuada do modo “no lookahead” para o “lookahead”, assim como a DE média. No modo “lookahead”, ocorre uma ligeira elevação na mudança de 1 para 2 sub-quadros à frente. Ainda nesse modo, pode-se dizer que a redução até o mínimo, que acontece com 4 sub-quadros, é menos significativa.

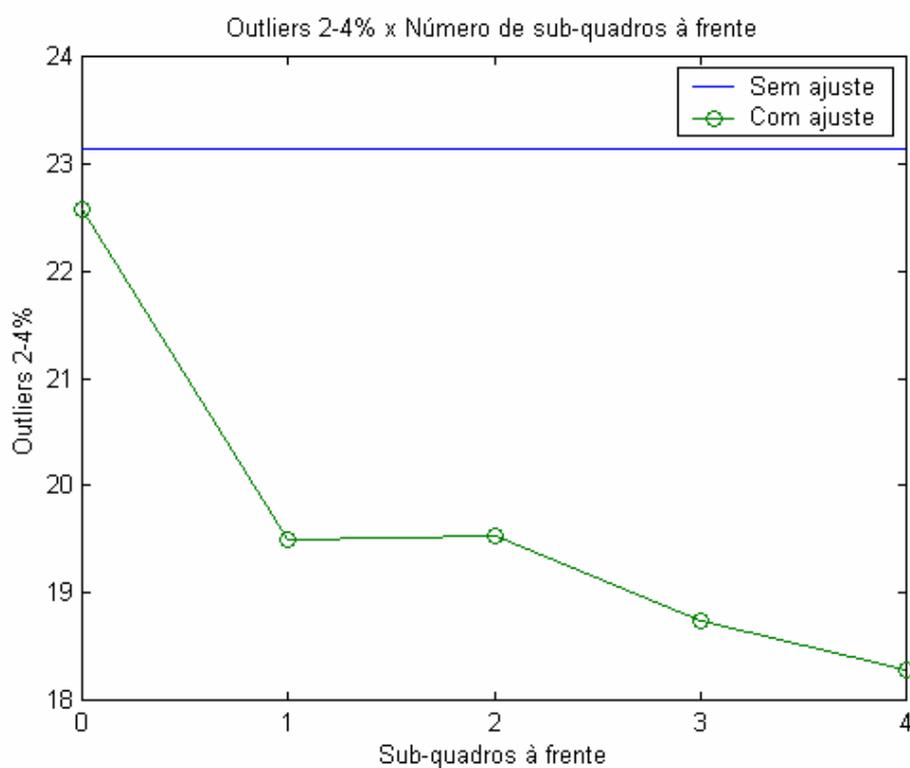


Figura 3.10 – Outliers 2-4% após a quantização utilizando os novos coeficientes.

É possível que o pequeno aumento observado na passagem de 1 para 2 sub-quadros seja função do conjunto de treinamento adotado. Pode ser que com outro conjunto não se obtenha comportamento semelhante. Contudo, apesar de tal situação não ser a ideal, com uma redução contínua, o resultado obtido pode ser considerado satisfatório.

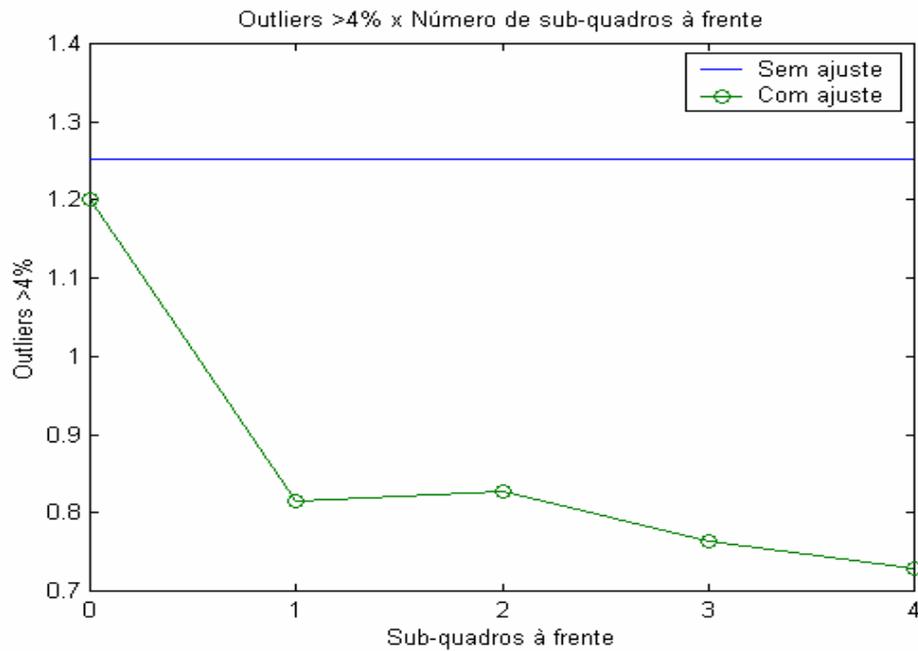


Figura 3.11 – Outliers >4% após a quantização utilizando os novos coeficientes.

Distância euclidiana ponderada - d_{LSF}

Diferentemente dos outliers, a d_{LSF} (Figura 3.12) apresenta redução contínua, mesmo na mudança de 1 para 2 sub-quadros, ainda que bem pequena nesse intervalo.

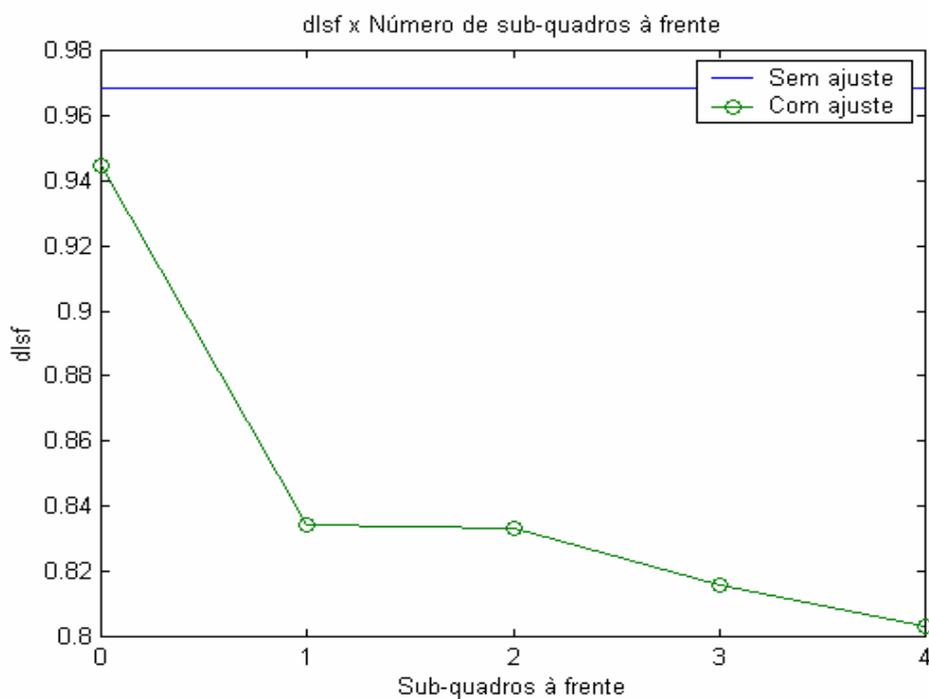


Figura 3.12 – d_{LSF} após a quantização utilizando os novos coeficientes.

3.6.3

Resultados da medida PESQ

Os resultados do teste PESQ são apresentados na a Tabela 3.9. É possível observar que em todos os casos obteve-se um resultado melhor utilizando-se o ajuste. Entretanto, não foi verificada uma característica que era esperada. A variação da medida parece ser aleatória, quando se imaginava que a mesma pudesse aumentar em função do número de sub-quadros à frente utilizados no processo de ajuste de LSF's. Tal fato talvez se deva à característica do codificador que, por ser de baixas taxas, apresenta um PESQ também baixo. Dessa forma, por já ter uma baixa qualidade de voz, pequenas melhorias que ocorram no sinal não se traduzem em um aumento coerente da medida PESQ.

Tabela 3.9 – Resultados da medida PESQ.

Modo	Sub-quadros	PESQ
Sem ajuste		2,3342
No lookahead		2,3524
Lookahead	1	2,3573
	2	2,3456
	3	2,3624
	4	2,3371

3.7

Desempenho dos novos coeficientes minimizando a d_{LSF}

Também foi avaliado o desempenho quando os coeficientes são treinados de modo a minimizar a d_{LSF} . Os resultados obtidos são apresentados a seguir.

3.7.1

Análise dos resultados obtidos antes da quantização

A Tabela 3.10 contém os valores medidos antes do quantizador. Pode-se observar que, apesar de a DE média ser maior do que quando se minimiza a DE (o que já era esperado), os outliers apresentam valores bem menores. Pode-se dizer até que se trata de uma redução bem significativa. A DE apresenta uma redução no modo “lookahead” com 4 sub-quadros.

Assim como quando foram apresentados os resultados para o minimização da DE, alguns gráficos serão mostrados a seguir. A organização visual mantém-se a mesma, número de sub-quadros à frente no eixo x, e medida de distorção no eixo y. A diferença é que agora tem-se uma espécie de comparação através da apresentação de duas curvas, uma nova, representando a medida para a minimização da d_{LSF} , e outra já conhecida para a minimização da DE. Além disso, também há a reta da medida referencial, sem ajuste.

Tabela 3.10 – Medidas antes de quantizar minimizando a d_{LSF} com novos coeficientes.

Modo	Sub-quadros	DE média	%2-4	%>4	d_{LSF}
Sem ajuste		1,1249	14,3329	0,7476	0,6007
No lookahead		1,16638	9,66618	0,37189	0,5053
Lookahead	1	1,13924	7,56613	0,18838	0,42283
	2	1,13958	7,55953	0,18938	0,46471
	3	1,12624	6,93233	0,16935	0,39586
	4	1,11139	6,44657	0,12083	0,38152

Distorção espectral – DE média

No gráfico da DE (Figura 3.13) podemos notar que a medida apresenta tendência de redução quando minimiza-se a d_{LSF} , mantendo-se, entretanto, acima da medida observada quando se minimiza a DE. O perfil das duas curvas é muito parecido, sendo quase como se fosse feito apenas um deslocamento vertical. Essa informação é extremamente importante para o desenvolvimento que se segue no próximo capítulo. A semelhança das curvas já fornece alguma indicação de que DE e d_{LSF} estão relacionadas. O que será feito no próximo capítulo é estabelecer uma expressão matemática para essa relação, dando base para a formulação de uma nova maneira de realizar o ajuste de LSF's.

A Figura 3.13 também deixa bem claro que quando se deseja uma redução da DE média, a minimização da d_{LSF} não é o melhor método. Pode-se observar que somente com o emprego de 4 sub-quadros à frente é possível obter uma

redução da DE em relação ao método usual, que não utiliza ajuste de LSF's. Mesmo assim, essa redução é bem pequena em relação ao modo sem ajuste.

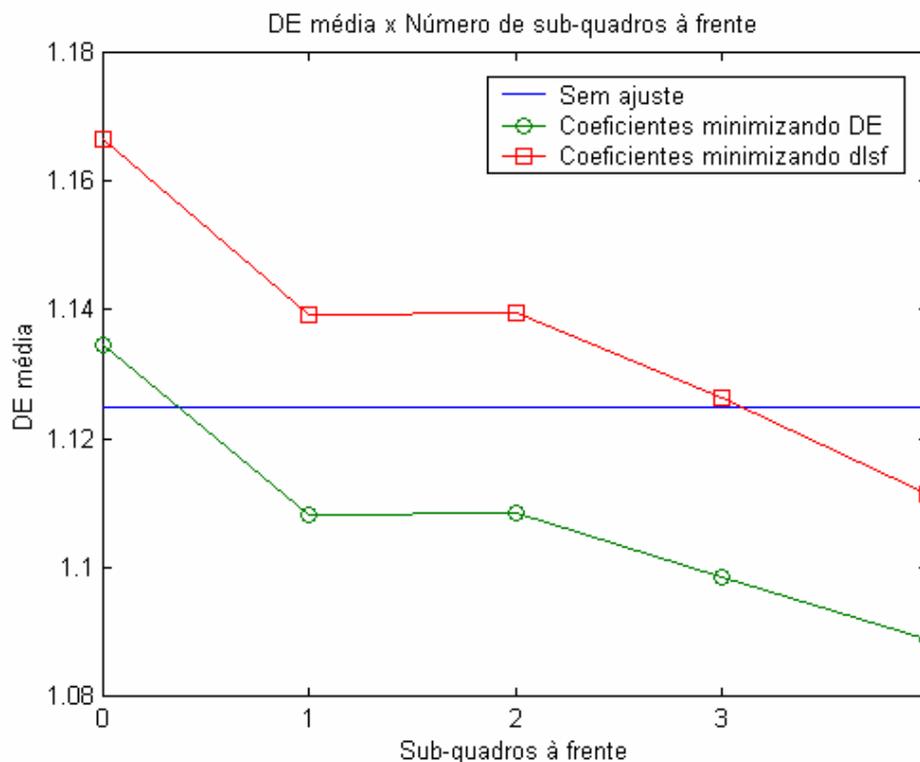


Figura 3.13 – DE média antes da quantização utilizando os novos coeficientes.

Outliers

Assim como no caso anterior, as duas curvas de outliers (figuras 3.14 e 3.15), minimizando a DE ou a d_{LSF} têm perfil semelhante. Porém, agora ocorre que a minimização da d_{LSF} apresenta o melhor resultado. Tanto os outliers de 2 a 4% como os maiores que 4% apresentam menores medidas. E essa diferença é significativa, ocorrendo em todas as situações, desde o modo “no lookahead”. Se comparado com o valor referencial pode-se notar, por exemplo, uma redução aproximadamente pela metade no percentual de outliers maiores que 4% no modo “no lookahead”.

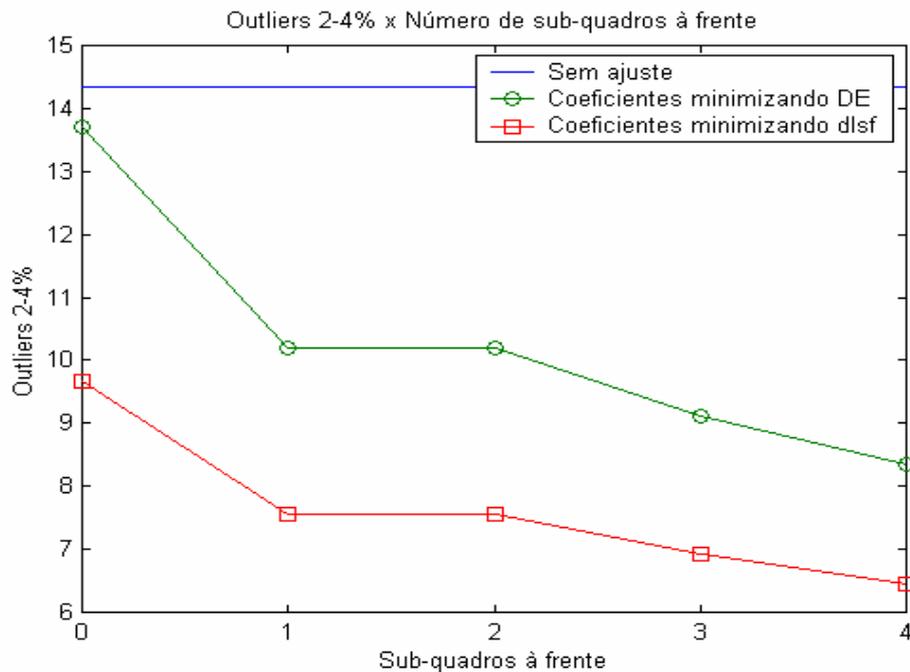


Figura 3.14 – Outliers 2-4% antes da quantização utilizando os novos coeficientes.

No gráfico dos outliers maiores que 4% percebe-se uma elevação muito pequena na medida na passagem de 1 para 2 sub-quadros quando se minimiza a d_{LSF} .

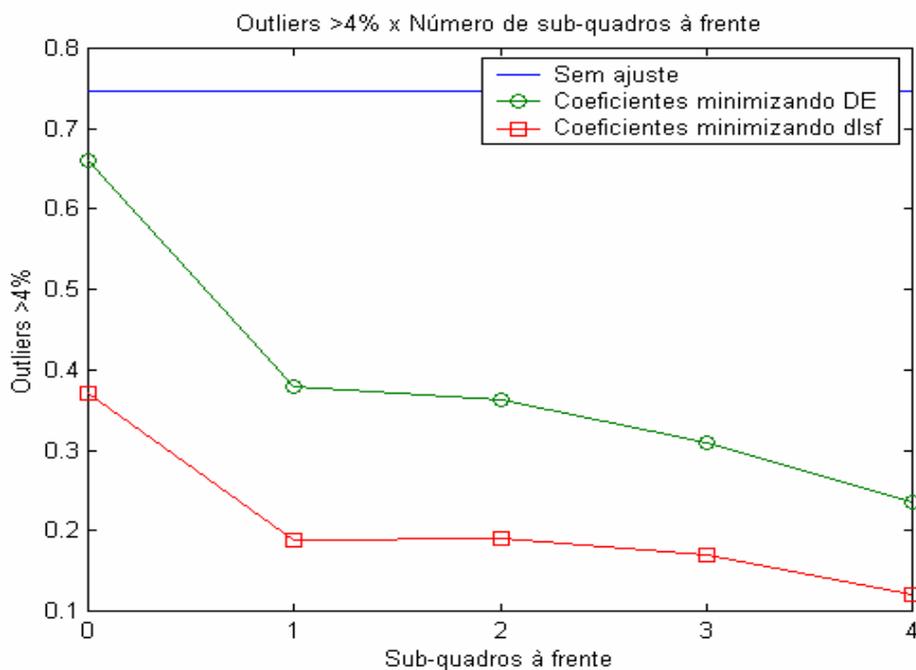


Figura 3.15 – Outliers >4% antes da quantização utilizando os novos coeficientes.

Distância euclidiana ponderada - d_{LSF}

Quando se minimiza a d_{LSF} , a própria apresenta um melhor resultado, como já era de se esperar. O gráfico da Figura 3.16, entretanto, apresenta um ponto não completamente aderente à essa lógica. No modo “lookahead”, utilizando-se 2 sub-quadros à frente, tem-se que a medida da d_{LSF} é muito próxima do seu valor quando se minimiza a DE. Esse “desvio” no comportamento previsto talvez possa ser consequência das características próprias das amostras de teste utilizadas. Lembrando do que foi dito sobre condições experimentais, os valores apresentados são médias de medidas realizadas com 10 amostras de modo independente. É perfeitamente normal que uma dessas amostras apresente um comportamento inesperado em determinadas condições. Essa anormalidade pode acabar por influenciar a média.

De modo geral, a d_{LSF} é sempre menor, apresentando redução ao longo do aumento do número de sub-quadros, desde o modo “no lookahead”.

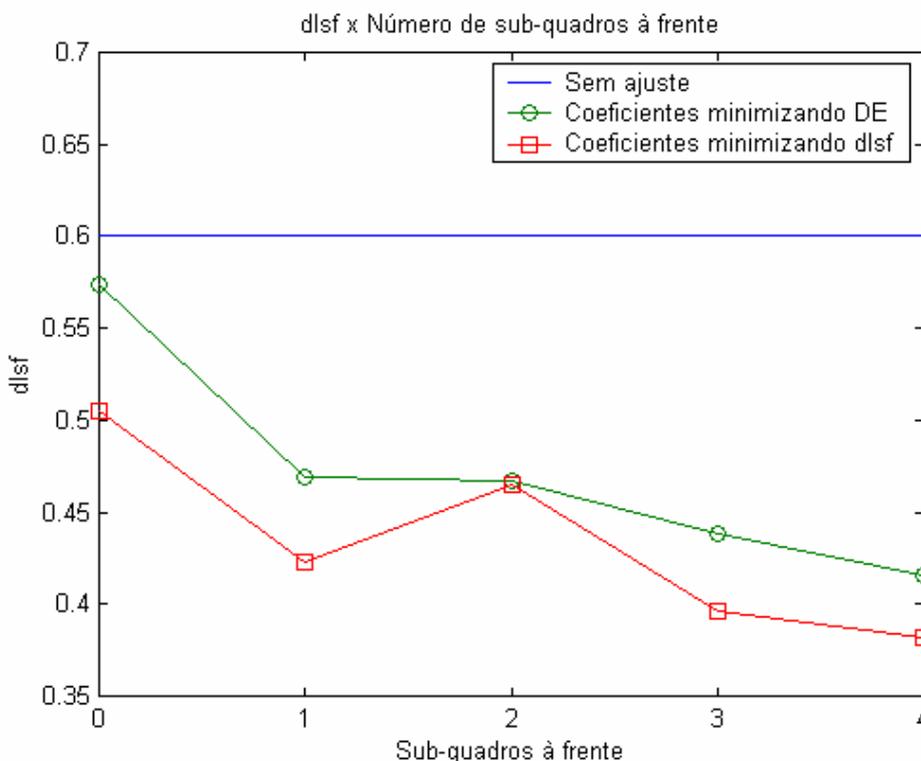


Figura 3.16 – d_{LSF} antes da quantização utilizando os novos coeficientes.

3.7.2

Análise dos resultados obtidos após a quantização

Para depois da quantização, utilizando coeficientes que minimizam a d_{LSF} , tem-se os resultados mostrados na Tabela 3.11.

Tabela 3.11 – Medidas após a quantização minimizando a d_{LSF} com novos coeficientes.

Modo	Sub-quadros	DE média	%2-4	%>4	d_{LSF}
Sem ajuste		1,5892	23,1301%	1,2527%	0,9682
No lookahead		1,56435	20,59314%	1,01608%	0,87338
Lookahead	1	1,50368	17,90572%	0,64297%	0,78215
	2	1,50679	17,93262%	0,64345%	0,77768
	3	1,51205	17,87777%	0,63038%	0,79410
	4	1,51301	17,79572%	0,62339%	0,80952

Distorção espectral – DE média

Pode-se notar na Tabela 3.11 e na Figura 3.17 que, após a quantização, na passagem do modo “no lookahead” para o “lookahead” tem-se uma redução significativa da DE média. Entretanto, conforme já havia sido observado no caso da minimização da DE, já no modo “lookahead” não se percebe uma melhoria significativa à medida que se aumenta o número de sub-quadros. Na verdade, no caso da minimização da d_{LSF} , ocorre até uma elevação contínua da DE.

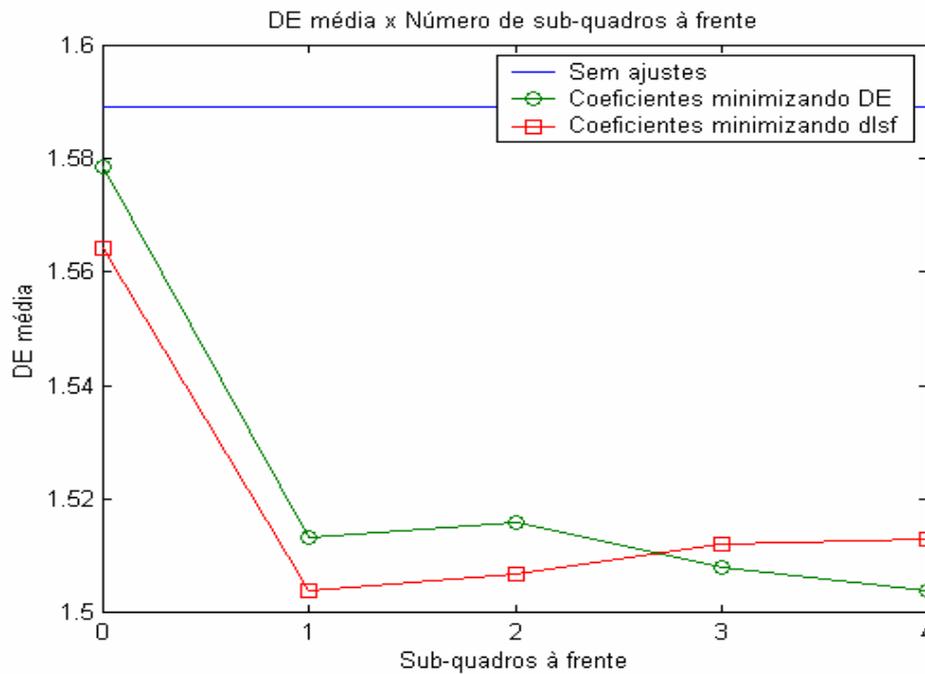


Figura 3.17 – DE média após a quantização utilizando novos coeficientes.

Outliers

Os outliers (Figuras 3.18 e 3.19) apresentam uma característica interessante que é uma variação muito pequena, quase que insignificante, no modo “lookahead”. Assim como quando é minimizada a DE, ocorre uma queda acentuada na passagem do modo “no lookahead” para o “lookahead”, atingindo um patamar que praticamente se mantém, independente do número de sub-quadros à frente que forem utilizados. De qualquer forma, a medida do outliers está sempre abaixo do referencial.

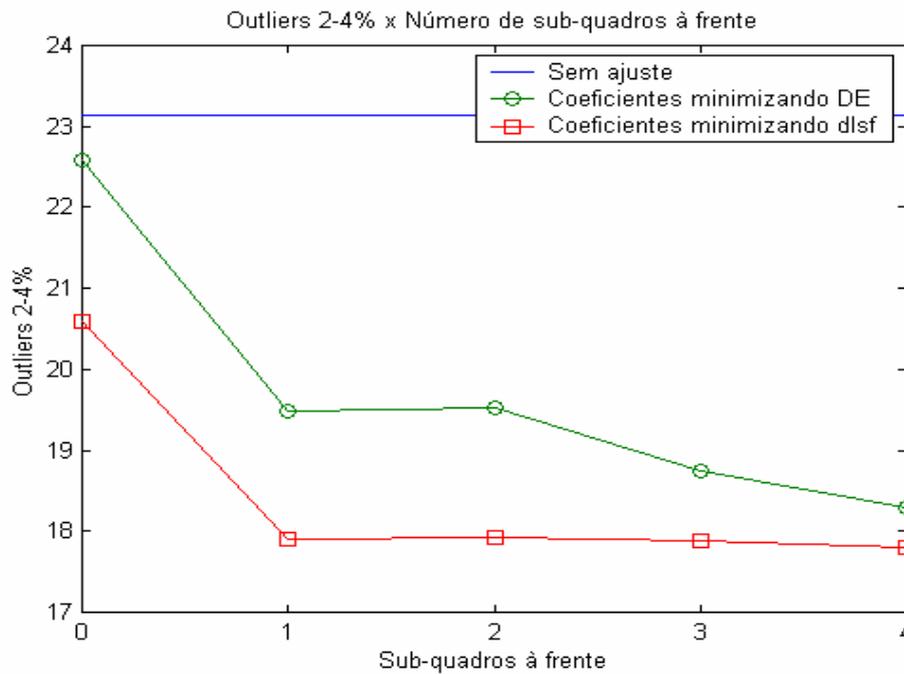


Figura 3.18 – Outliers 2-4% após a quantização utilizando novos coeficientes.

As características observadas valem tanto para os outliers de 2 a 4%, quanto para os maiores que 4%. No gráfico abaixo é possível observar mais uma vez o patamar que se estabelece para os outliers no modo “lookahead”.

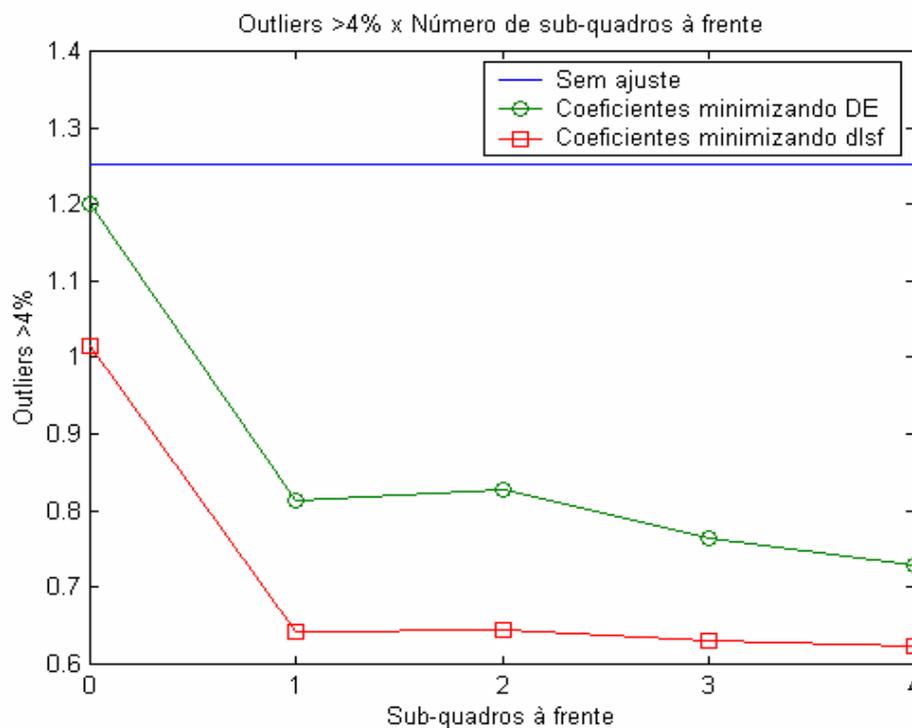


Figura 3.19 - Outliers >4% após a quantização utilizando novos coeficientes.

Distância euclidiana ponderada - d_{LSF}

A d_{LSF} (Figura 3.20) apresenta um comportamento que foge um pouco do esperado. Ela se mantém abaixo do seu valor quando a DE é minimizada, porém, no modo “lookahead”, com 4 sub-quadros, ocorre uma inversão de posições. Teoricamente, a medida deveria diminuir ou, na pior das hipóteses, manter o seu valor. Esse aumento no final pode ser visto também como um “desvio”, conforme comentado em seção anterior.

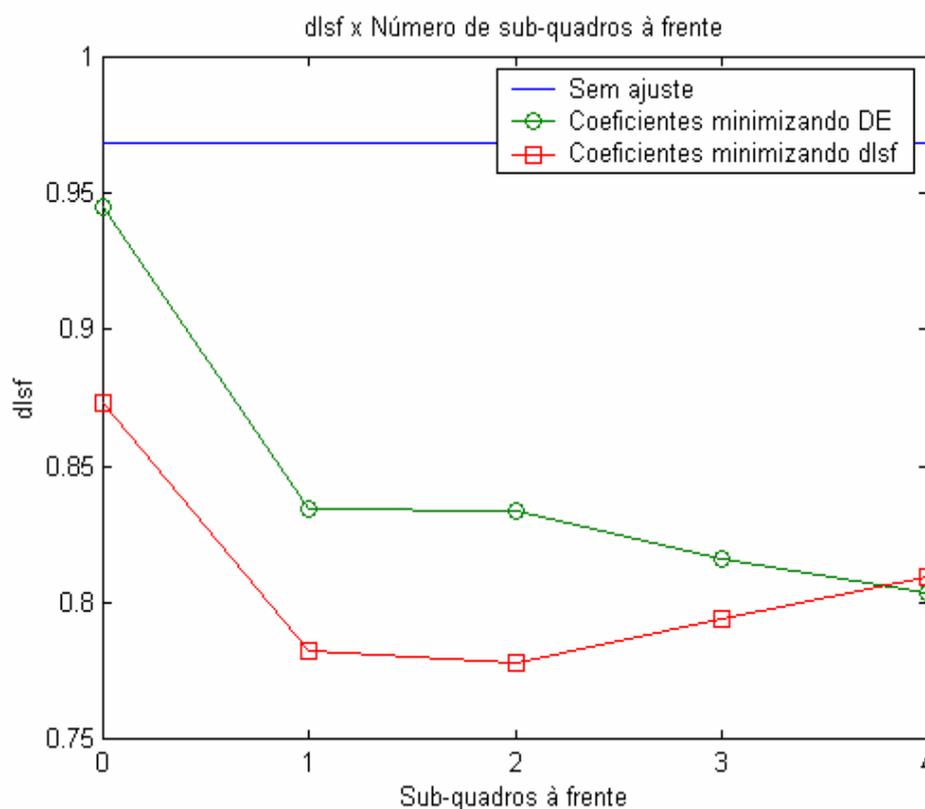


Figura 3.20 – d_{LSF} após a quantização utilizando novos coeficientes.

3.7.3

Resultados da medida PESQ

Mais uma vez, no caso do PESQ é apresentada apenas a tabela com os valores. Conforme pode-se verificar na Tabela 3.12, novamente, não é possível estabelecer uma relação entre o PESQ e o processo de ajuste de LSF's. O resultado foi pior que o modo sem ajuste tanto para o modo “no lookahead” como

para o “lookahead” com 1 sub-quadro. Mesmo nas outras situações, onde o resultado é melhor, ainda assim essa melhoria não é muito significativa.

Tabela 3.12 – Resultados da medida PESQ

Modo	Sub-quadros	PESQ
Sem ajuste		2,3342
No lookahead		2,3088
Lookahead	1	2,3293
	2	2,3497
	3	2,3463
	4	2,3409