



UNIVERSIDADE FEDERAL DE SERGIPE

Relatório Biochaves nº 1

Jônatas Cruz Santos
João Marcus Andrade Silveira



1 Objetivos

Realizar a estimação do pulso glotal utilizando o Iterative Adaptive Inverse Filtering (IAIF) [1], método clássico para estimação baseado em filtragem inversa, e realizar uma análise subjetiva dos resultados obtidos, considerando o formato de onda comum de um sinal de pulso glotal.

2 Metodologia

2.1 Predição Linear

A predição linear consiste na estimação do valor de uma amostra $x[n]$ do sinal a partir das k amostras anteriores, como representado a seguir:

$$x[n] = \sum_{m=1}^k a_m x[n-m] + r[n] \quad (1)$$

em que $r[n]$ representa a perturbação aleatória independente do sinal $x[n]$ e k é a ordem do filtro.

Assim, cada elemento de x pode ser representado como uma combinação linear entre k elementos anteriores, a menos de um erro de predição. O ajuste dos coeficientes é realizado de acordo com o método dos mínimos quadrados, utilizando pseudo-inversão.

Com os coeficientes encontrados, a partir de (1) por meio da transformada Z, chega-se à seguinte função de transferência:

$$H(z) = \frac{z^k}{z^k - a_1 z^{k-1} - \dots - a_{k-1} z - a_k} \quad (2)$$

em que os polos permitem estimar o contorno espectral do sinal.

2.2 Filtragem Inversa

A função de transferência em (2) pode representar um filtro digital. O processo de filtragem inversa consiste em filtrar o sinal $x[n]$, removendo as influências dos polos de $H(z)$, utilizando um filtro $G(z)$ que pode ser representado como segue:

$$G(z) = H(z)^{-1} \quad (3)$$

2.3 Iterative Adaptive Inverse Filtering

O IAIF é um método que emprega predição linear e filtragem inversa utilizando uma estrutura iterativa para estimar o pulso glotal. Esse método consiste em realizar uma filtragem passa-altas em um sinal de voz, $s[n]$, e aplicar, no sinal resultante, o procedimento ilustrado na Figura 1. Primeiramente é realizada a estimação da contribuição glotal, $H_g(z)$, obtida por predição linear de ordem 1, seguida pela remoção dessa por meio de filtragem inversa. Em seguida é estimada a contribuição do trato vocal, $H_v(z)$, utilizando predição linear de ordem v , sendo essa removida por filtragem inversa, seguida da eliminação do efeito da radiação dos lábios por meio de um

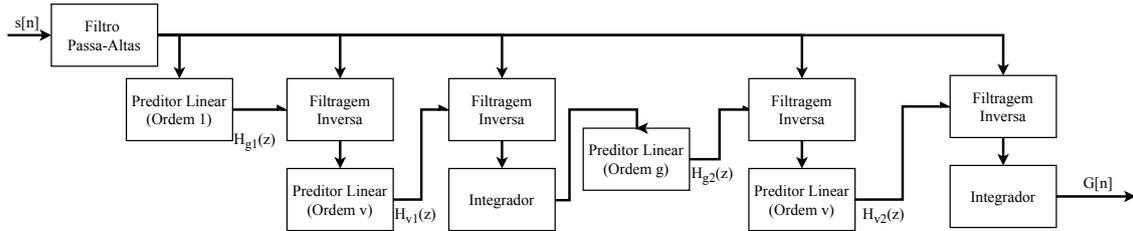


Figura 1. Diagrama de blocos do IAIF

integrador estimando assim o pulso glotal. Por fim, esse procedimento é executado mais uma vez, utilizando, entretanto, ordem g na estimação da contribuição.

3 Resultados

Na Figura 2 encontram-se representados um sinal $s[n]$ utilizado como exemplo e seu pulso glotal estimado $G[n]$, utilizando $g = 4$ e $v = 20$ (escolhidos de acordo com a percepção subjetiva dos autores), de acordo com o procedimento descrito anteriormente. Pode ser observado um comportamento coerente com o esperado, de acordo com as representações de pulso glotal na literatura. Entretanto, não é possível garantir a acurácia da estimação, sem que haja a comparação com uma representação válida do sinal glotal.

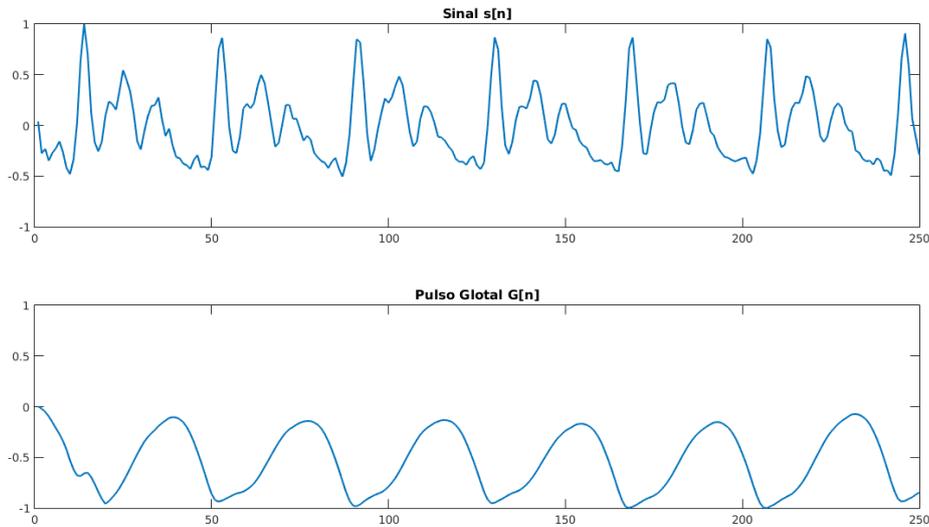


Figura 2. Sinal de voz $s[n]$ e seu pulso glotal estimado $G[n]$

Na figura 3 encontram-se representadas as contribuições espectrais do pulso glotal e trato vocal estimados nas duas iterações realizadas pelo IAIF em azul, e em vermelho o contorno espectral do processo de filtragem inversa. Nela, a estimação final da contribuição do trato vocal é representada por H_{v2} .

4 Conclusão e Próxima Etapa

O IAIF é um método de simples implementação e seus resultados apresentaram-se, aparentemente coerentes, de acordo com o referencial teórico. Entretanto, a

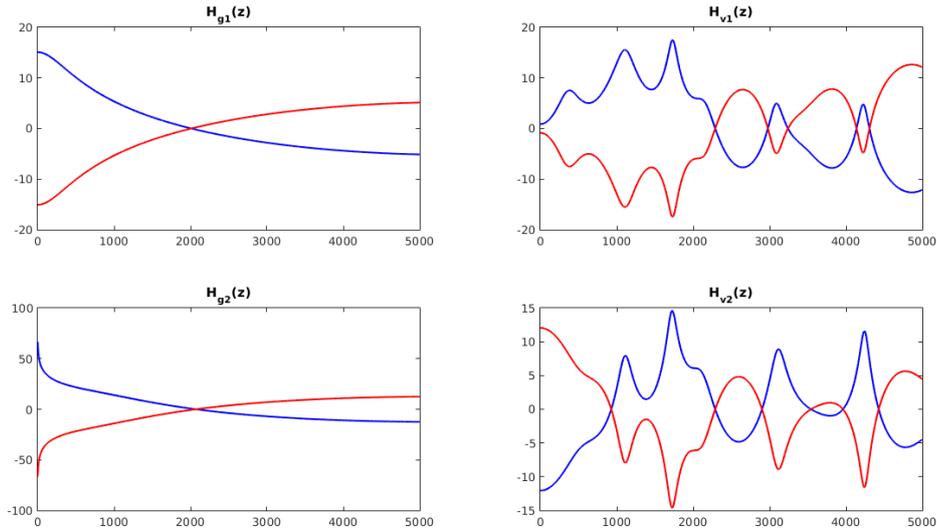


Figura 3. Contornos espectrais estimados para a contribuição glotal e a contribuição do trato vocal, no decorrer das etapas do IAIF

escolha dos parâmetros da predição linear são ajustados de maneira subjetiva, sendo necessário um método de ajuste mais objetivo, para garantir melhores resultados. Também se faz necessário realizar a comparação dos sinais glotais estimados com um sinal glotal confiável, para assegurar a acurácia do método.

Na próxima etapa do trabalho, pretende-se realizar a síntese de um sinal de voz, utilizando um pulso glotal e um trato vocal conhecidos, permitindo, assim, a comparação entre o sinal glotal estimado e o utilizado para a síntese. Também serão realizados testes de ajuste do modelo de Liljencrants-Fant [2] à derivada do sinal glotal estimado, analisando o erro e os parâmetros ajustados, de acordo com metodologia proposta por Strik e Boves [3].

Referências Bibliográficas.

- [1] P. Alku, E. Vilkmán, and U. Laine. Analysis of glottal waveform in different phonation types using the new iaif-method. In *Proc. 12th Int. Congress Phonetic Sciences*, volume 4, pages 362–365, 1991.
- [2] G. Fant, J. Liljencrants, and Q.-g. Lin. A four-parameter model of glottal flow. *STL-QPSR*, 4(1985):1–13, 1985.
- [3] H. Strik and L. Boves. On the relation between voice source parameters and prosodic features in connected speech. *Speech Communication*, 11(2-3):167–174, 1992.