

2.6 Modelo de síntese de Itakura-Saito

A figura 1 representa o modelo completo de síntese de Itakura-Saito para ordem de predição p . O período fundamental é p_0 , a sonoridade é V e a não-sonoridade é U , satisfazendo a relação complementar

$$V + U = 1.$$

O fator de ganho geral é, em primeira aproximação,

$$G = \sqrt{R_{dd}(0)},$$

isto é, é a raiz quadrada da variância do sinal de erro de predição $d(n)$.

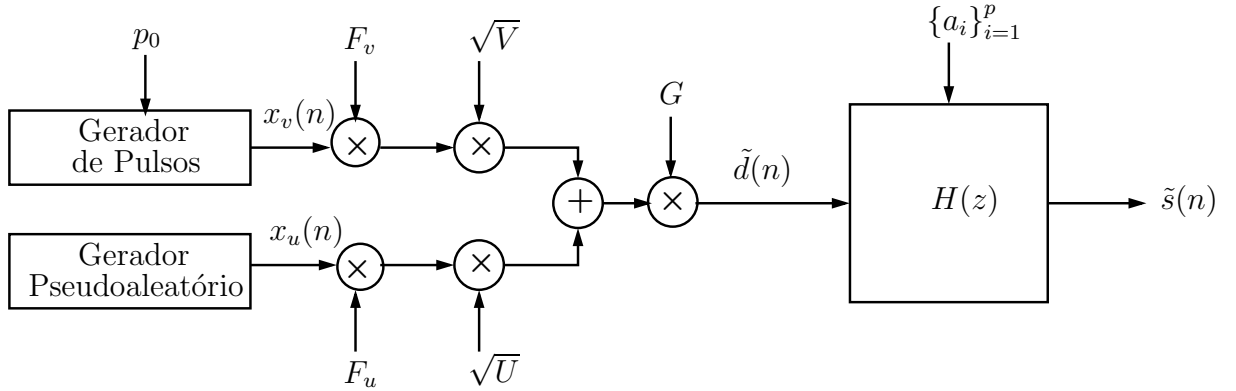


Figura 1: Modelo de síntese de Itakura-Saito

Supondo que a função de autocorrelação original do sinal de voz $s(n)$ tenha sido estimada pelo somatório total

$$R_{ss}(m) = \sum_{n=0}^{L-1} s(n)s(n+m)$$

para segmento de L amostras de comprimento, o fator sonoro é

$$F_v = \left[\frac{L}{p_0} \right]^{-1/2}$$

e o fator surdo é

$$F_u = \sqrt{\frac{1}{L}}.$$

Desta forma, o sinal de excitação pode ser descrito como

$$\tilde{d}(n) = G \left(F_v \sqrt{V} x_v(n) + F_u \sqrt{U} x_u(n) \right),$$

sendo a excitação sonora de amplitude unitária

$$x_v(n) = \sum_{i=0}^{\left\lceil \frac{L}{p_0} \right\rceil - 1} \delta(n - ip_0)$$

para $n = 0, 1, \dots, L - 1$ e tendo a excitação surda distribuída como

$$x_u(n) \sim N(0,1)$$

com função de autocorrelação

$$R_{x_u x_u}(m) = \delta(m).$$

O filtro de síntese

$$H(z) = \frac{1}{1 + \sum_{i=1}^p a_i z^{-i}}$$

recebe na entrada o sinal de excitação $\tilde{d}(n)$ e entrega na saída o sinal de voz reconstruído ou sintetizado $\tilde{s}(n)$.