

Técnicas de Codificação de Voz Aplicadas em Sistemas Móveis Celulares

Tópicos Especiais em Telecomunicações
Comunicações Móveis 1º/00

Daniilo Dias
&
João Luiz A. Carvalho

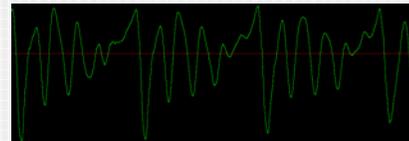
Tópicos Abordados

- Motivação
- Características do Sinal de Voz
- Codificadores de Forma de Onda
 - Amostragem
 - Quantização
 - Codificação: PCM, APCM, DPCM, ADPCM
- Codificadores Paramétricos: LPC
- Codificadores Híbridos
 - RPE-LTP (GSM)
 - CELP: VSELP (IS-54 e IS-136), ACELP (IS-136), QCELP (IS-95)
- Comparação entre as técnicas

Codificação de Voz e Telefonia Móvel Digital

- Diminuição da banda requerida
- Robustez e qualidade frente a ruído
 - Detecção digital
 - Codificação de canal
 - Estrutura do algoritmo usado
- A codificação deve ser eficiente:
 - Execução em tempo real em DSP's baratos e de baixo consumo
 - Prover boa qualidade de voz
 - Fácil inteligibilidade
 - Reconhecimento de interlocutor
 - Emoção e entonação

Características do Sinal de Voz



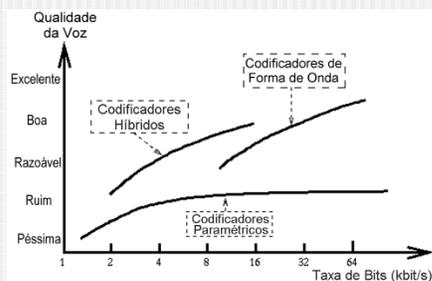
Segmento sonoro: fonema A (25ms)



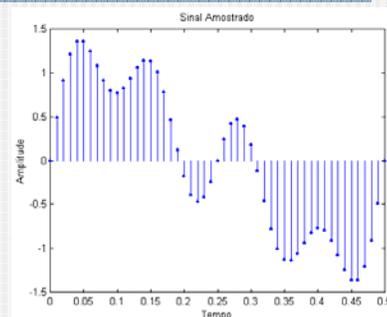
Segmento surdo: fonema S (25ms)

Características importantes: *pitch*, timbre, amplitude, correlação a curto termo, correlação a longo termo, espectro de frequência, distribuição de probabilidade da amplitude.

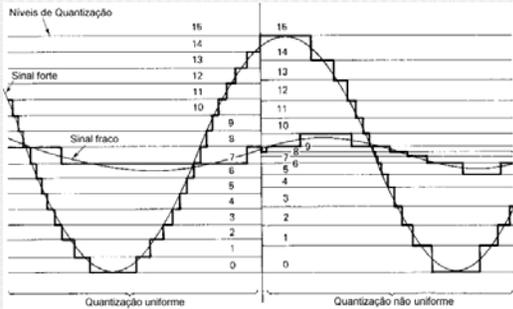
Classes de Codificadores



Codificadores de Forma de Onda: Amostragem



Codificadores de Forma de Onda: Quantização

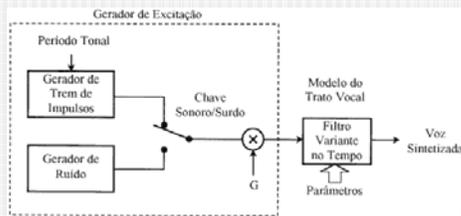


Codificadores de Forma de Onda

- PCM: *Pulse Code Modulation*
- APCM: *Adaptative PCM*
- DPCM: *Differential PCM*
- ADPCM: *Adaptative and Differential PCM*

PCM 64 kbps \Rightarrow ADPCM 24 a 48 kbps

Vocoder LPC (2,5 kbps) *Linear Prediction Coding*



$$H(z) = \frac{1}{A(z)}$$

$$A(z) = 1 - \sum_{i=1}^{10} \alpha_i z^{-i}$$

α_i : coeficientes preditores

Excitação Ideal

Geração do sinal de voz

$$s[n] = e[n] * h[n] \xrightarrow{\mathcal{F}} S(z) = E(z) \cdot H(z)$$

Passando o sinal de voz original pelo filtro inverso

$$e[n] \xleftarrow{\mathcal{F}^{-1}} S(z) \cdot A(z) = E(z)$$

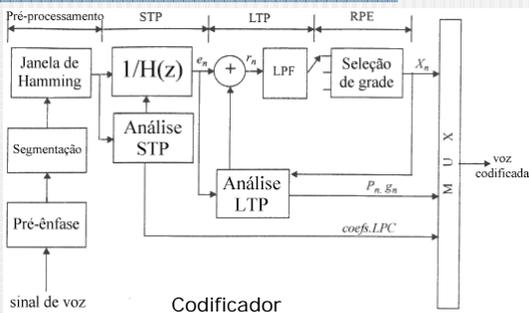
excitação ideal

Para melhorar a qualidade: melhorar a excitação

\Rightarrow Codificadores híbridos

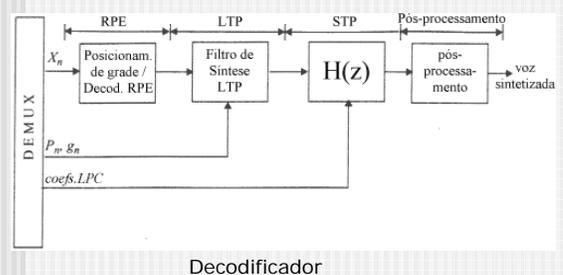
GSM: RPE-LTP (13 kbps)

Regular-Pulse Excited LPC with a Long-Term Predictor

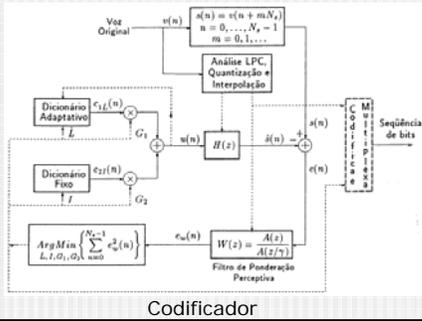


GSM: RPE-LTP (13 kbps)

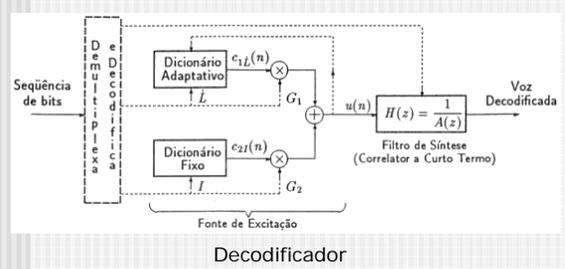
Regular-Pulse Excited LPC with a Long-Term Predictor



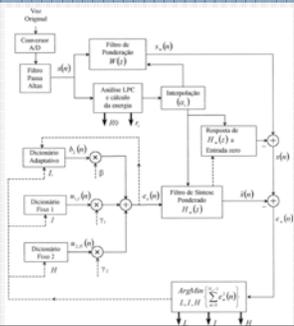
CELP (até 4 kbps) Code Excited Linear Predictive



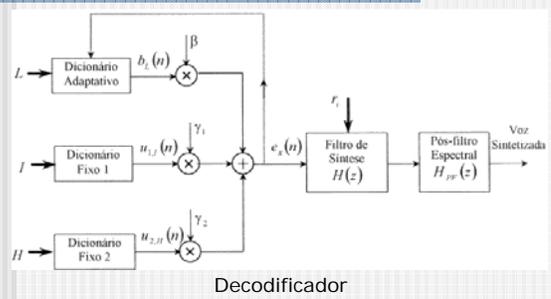
CELP (até 4 kbps) Code Excited Linear Predictive



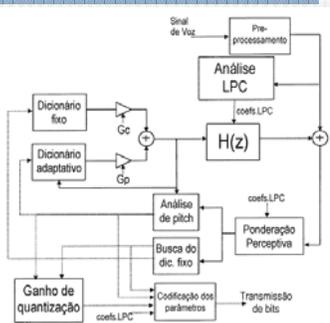
IS-54 e IS-136: VSELP Vectorial Sum Excited Linear Predictive (7,95 kbps)



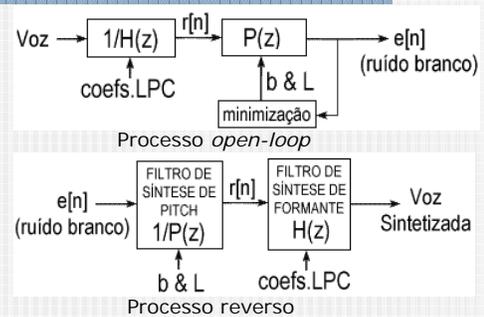
IS-54 e IS-136: VSELP Vectorial Sum Excited Linear Predictive (7,95 kbps)



IS-136: ACELP (7,95 kbps) Algebraic Code Excited Linear Predictive

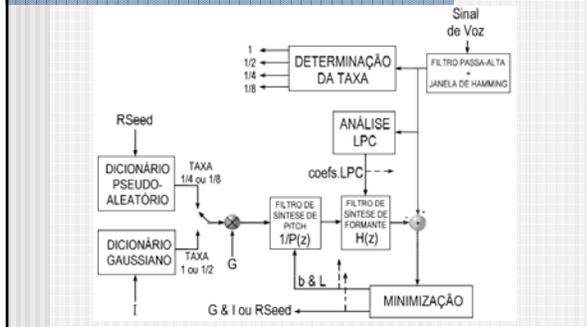


QCELP: Modelo Básico de Reconstrução da Voz



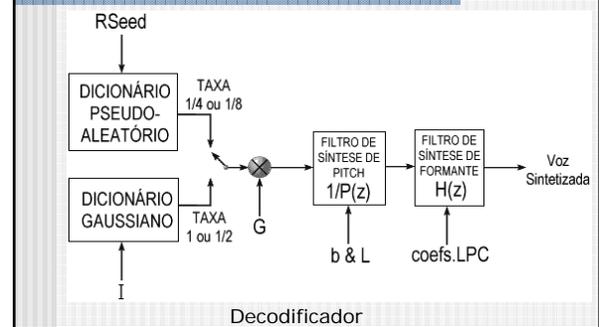
IS-95: QCELP

Qualcomm Code Excited Linear Predictive
Taxa Variável: 9,6 / 4,8 / 2,4 / 1,2 kbps



IS-95: QCELP

Qualcomm Code Excited Linear Predictive
Taxa Variável: 9,6 / 4,8 / 2,4 / 1,2 kbps



Conclusão

Comparação entre as técnicas

Técnica	Padrão	Taxa de bits	Excitação	Modelo do Trato Vocal	MOS rating
PCM		64 kbps	-	-	4,3
ADPCM		32 kbps	-	-	4,1
LPC		2,4 kbps	trem de pulsos / ruído	filtro LPC (análise STP)	2,5
CELP		8 kbps	1 dicionário adaptativo 1 dicionário fixo	filtro LPC (análise STP)	3,7
CELP		4,8 kbps	1 dic. adaptativo e 1 dic. fixo	filtro LPC (análise STP)	3,0
RPE-LTP	GSM	13 kbps	pulsos regulares e análise LTP	filtro LPC (análise STP)	3,54
VSELP	IS-54 e IS-136	7,95 kbps	1 dicionário adaptativo 2 dicionários fixos estruturados	filtro LPC (análise STP)	3,8
ACELP	IS-136	7,95 kbps	1 dicionário adaptativo 1 dicionário c/ estrutura algébrica	filtro LPC (análise STP)	3,9
QCELP	IS-95	1,2 / 2,4 kbps 4,8 / 9,6 kbps	1 dicionário pseudo-aleatório 1 dicionário gaussiano	filtro LPC (análise STP)	3,45 (full-rate)

Reconhecimento

- Prof. Lúcio M. Silva
- Prof. Sebastião Nascimento
- Prof. Paulo H. P. Carvalho

FIM

Danilo: danilods@zaz.com.br
João Luiz: joaolac@zaz.com.br

Cálculo dos coeficientes preditores do filtro LPC

$s_n = \sum_{i=1}^{10} \alpha_i s_{n-i} + e_n$

s_n : amostra atual
 s_{n-i} : 10 amostras anteriores
 e_n : erro de predição

$$H(z) = \frac{1}{1 - \sum_{i=1}^{10} \alpha_i z^{-i}}$$

Para minimizar a energia média E do sinal e_n :

$$E = \sum_{n=1}^N e_n^2 = \sum_{n=1}^N \left(\sum_{i=0}^{10} \alpha_i s_{n-i} \right)^2$$

$$\alpha_n = -1$$

Para minimizar E a respeito de um coeficiente α_m :

$$\frac{\partial E}{\partial \alpha_m} = \sum_{n=1}^N 2s_{n-m} \sum_{i=0}^{10} \alpha_i s_{n-i} = 0$$

$$= \sum_{i=0}^{10} \sum_{n=1}^N s_{n-m} s_{n-i} \alpha_i = 0$$

A soma mais interna pode ser vista como o coeficiente de correlação C_m :

$$\sum_{i=1}^{10} C_{im} \alpha_i = C_{0m}$$

\Rightarrow sistema linear de 10 equações e 10 variáveis