Prepoznavanje govora

LPC

Speaker dependant

- Postoje dva opšta pristupa kod prepoznavanja govora:
  - Speaker dependant
  - Speaker independant
- Kod speaker dependant prepoznavanja sistem treba da prepoznaje same reči, kao i govornika, i (eventualno) da prihvata prepoznavanje samo ako je reč izgovorio poznati govornik.
- Kod speaker independant sistema, prepoznavanje govora treba da funkcioniše nezavisno od govornika.

LPC

- Linear Predictive Coding
- Koristi se u kompresiji audio signala, kao I metod zapisa koeficijenata koji se često primenjuju kod speaker dependant sistema.
- Osnova je ideja da se trenutni sempl može predstaviti kao linearna kombinacija prethodnih semplova.
- Posmatrano sa druge strane, linearnom kombinacijom skupa semplova može da se izvrši predikcija budućeg sempla.
- Ovo posebno ima smisla ako se posmatra relativno kratak prozor (10-20ms) u kojem pretpostavljamo da se signal ponaša kao da je periodičan, iako to zapravo nije slučaj.

Osnovne LPC jednačine (1)

s(n) – sempl

ŝ(n) – predikcija sempla

p – preciznost (vrednost u opsegu 8 - 16, zavisno od frekvencije semplovanja)

α_k – koeficijent za k-ti prethodni sempl

$$\hat{s}(n) = \sum_{k=1}^{p} \alpha_k s(n-k)$$

Osnovne LPC jednačine (2)

N – veličina prozora

e(n) – greška predikcije

$$e(n) = s(n) - \hat{s}(n)$$

$$E = \sum_{n=1}^{N} e(n)^2 \Rightarrow \qquad E = \sum_{n=1}^{N} \left[s(n) - \sum_{k=1}^{p} \alpha_k s(n-k) \right]^2$$

- *E* predstavlja srednju kvadratnu grešku unutar prozora. Nju želimo da minimiziramo.

Minimiziranje

- Minimizaciju *E* vršimo parcijalnim diferenciranjem po svim α i izjednačavanjem sa 0. Dobićemo *p* jednačina sa *p* nepoznatih.

  i = 1, 2, …, *p*

$$\frac{\partial E}{\partial \alpha_i}=\sum_{n=1}^{N} 2\left[s(n)-\sum_{k=1}^{p} \alpha_k s(n-k)\right]\left[-s(n-i)\right]=0$$

$$\sum_{k=1}^{p} \alpha_k \sum_{n=1}^{N} s(n-k)\, s(n-i)=\sum_{n=1}^{N} s(n)\, s(n-i)$$

Autokorelacija

- Radi pojednostavljenja prethodnog zapisa, uvodi se autokorelaciona funkcija:

$$R(k)=\sum_{n=0}^{N-1} s(n)s(n-k)$$

$$R(k)=R(-k)$$

- Tipično se pre računanja *R* na ulazni niz *s(n)* primenjuje prozorska funkcija kao što su Hamingova ili Haningova.

Autokorelacioni metod

$$R(k)=\sum_{n=0}^{N-1} s(n)s(n-k)$$

$$\sum_{k=1}^{p} \alpha_k \sum_{n=1}^{N} s(n-k)s(n-i)=\sum_{n=1}^{N} s(n)s(n-i)$$

$$\sum_{k=1}^{p} \alpha_k R(|k-i|)=R(i)$$

- Ako ovo razvijemo po *i*, dobićemo *p* jednačina.

Autokorelacione jednačine

$$\alpha_1 R(0)+\alpha_2 R(1)+\cdots+\alpha_p R(p-1)=R(1)$$
$$\alpha_1 R(1)+\alpha_2 R(0)+\cdots+\alpha_p R(p-2)=R(2)$$
$$\vdots$$
$$\alpha_1 R(p)+\alpha_2 R(p-1)+\cdots+\alpha_p R(0)=R(p)$$

Matrični oblik

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-2) & R(p-1) \\ R(1) & R(0) & \cdots & R(p-3) & R(p-2) \\ \cdots & R(1) & \cdots & \cdots & \cdots \\ R(p-2) & \cdots & \cdots & R(0) & R(1) \\ R(p-1) & R(p-2) & \cdots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_{p-1} \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \cdots \\ R(p-1) \\ R(p) \end{bmatrix}$$