

**1- مقدمه**

اهداف درس:

آشنایی با ضرایب پیشگویی خطی

آشنایی با نحوه محاسبه ضرایب پیشگویی خطی

آشنایی با نحوه محاسبه ضرایب کپسترال از ضرایب پیشگویی خطی

**2- ضرایب پیشگویی خطی**

هدف اصلی پیشگویی خطی «تخمین دنباله خروجی است از یک ترکیب خطی از نمونه ها ورودی و خروجی های گذشته».

$$\hat{y}(n) = \sum_{j=0}^q b(j)x(n-j) - \sum_{i=1}^p a(i)y(n-i)$$

به  $a(i)$  و  $b(j)$  ضرایب پیش بینی کننده گفته می شود.

اغلب سیستم هایی که برای ما جالب توجه اند را می توان بوسیله معادلات دیفرانسیلی خطی با ضرایب ثابت توصیف کرد.

$$\sum_{i=0}^p a(i)y(n-i) = \sum_{j=0}^q b(j)x(n-j)$$

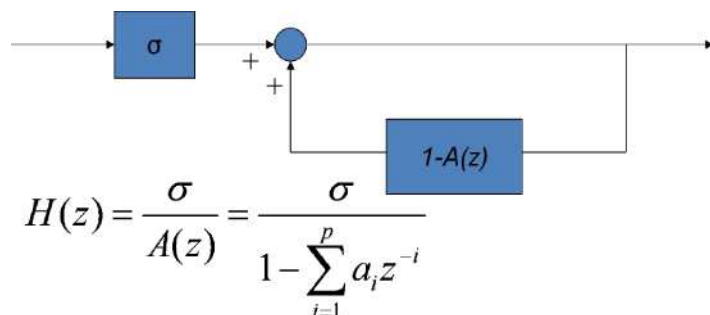
اگر  $H(z) = Y(z)/X(z)$  باشد که  $H(z)$  نسبت چندجمله ای های  $N(z)/D(z)$  باشد،

$$N(z) = \sum_{j=0}^q b(j)z^{-j} \text{ and } D(z) = \sum_{i=0}^p a(i)z^{-i}$$

می توان گفت که با داشتن ضرایب پیش بینی کننده ( $a$  و  $b$ ) می توان صفرها و قطب های  $H(z)$  را در اختیار داشت.

دو نوع مهم از پیشبینی کننده ها موجود است:

- مدل های تمام قطب (AllPole): در آمار به این مدل ها مدل های autoregressive (AR) گفته می شود. در این مدل ها  $N(z)$  یک عدد ثابت است. در تصویر 1 یک نمونه از مدل AR مشاهده می کنید.
- مدل های تمام صفر (All-zero): به این مدل ها، مدل های moving average (MA) گفته می شود. در این مدل ها مخرج  $D(z)$  عدد «یک» است.
- ترکیب دو مدل بالا را مدل autoregressive moving average (ARMA) می گویند.



تصویر 1- مدل پیش بینی کننده AR

با داشتن یک سیگنال  $y(n)$  با میانگین صفر در و با در نظر گرفتن مدل AR داریم:

$$\hat{y}(n) = -\sum_{i=1}^p a(i)y(n-i)$$

خطا به صورت زیر محاسبه می شود:

$$e(n) = y(n) - \hat{y}(n) = \sum_{i=0}^p a(i)y(n-i)$$

برای به دست آوردن پیش بینی کننده، از اصل تعامد استفاده می کنیم. این اصل بیان می کند که ضرایب مورد نظر ضرایبی هستند که باعث می شوند خطا متعامد بر نمونه ها  $y(n-1), y(n-2), \dots, y(n-p)$  شود.

پس نتیجه می گیریم که:

$$\langle y(n-j)e(n) \rangle = 0 \quad \text{for } j=1, 2, \dots, p$$

با به عبارت دیگر:

$$\left\langle y(n-j) \sum_{i=0}^p a(i)y(n-i) \right\rangle = 0$$

با تعویض عمل متوسط گیری و جمع کردن، به فرمول زیر می رسیم:  $\sum_{i=0}^p a(i) \sum_n y(n-i)y(n-j) = 0, j=1, \dots, p$

پیش بینی کننده های مورد نیاز بوسیله حل این معادلات به دست می آیند.

اصل تعامد همچنین بیان می کند که کمینه خطا بوسیله فرمول زیر به دست می آید:

$$\sum_{i=0}^p a(i) \sum_n y(n-i)y(n) = E$$

$$\sum_{i=0}^p a(i)r_{i-j} = 0, j=1, 2, \dots, p$$

می توان خطا را برای همه زمان ها کمینه کرد:

$$\sum_{i=0}^p a(i)r_i = E \quad r_i = \sum_{n=-\infty}^{\infty} y(n)y(n-i)$$

حال سؤال این است که چگونه مقدار مناسبی برای  $p$  انتخاب کنیم.

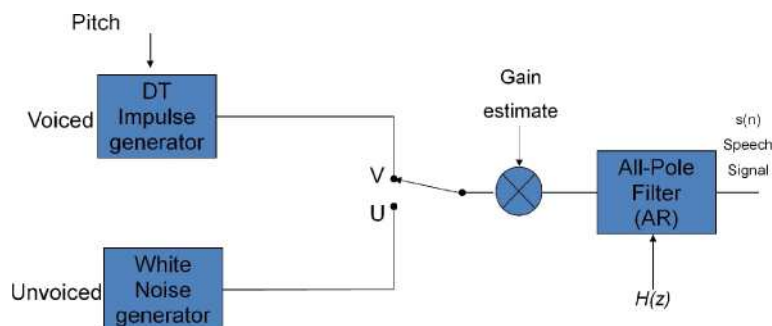
برای انتخاب تعداد ضریب مناسب ( $p$ ) روش های سرانگشتی موجود است.

$$p = \frac{2BW}{1000} + c$$

تعداد  $p$  بستگی مستقیمی به پهنای باند سیگنال گفتار دارد.

- برای یک مسیر صوتی معمولی، به صورت متوسط یک فرمونت در هر کیلوهرتز از پهنای باند وجود دارد.
- هر فرمونت به دو قطب  $\text{complex conjugate}$  نیاز دارد.
- در نتیجه برای هر فرمونت دو ضریب پیش بینی کننده نیاز است و یا دو ضریب برای هر کیلوهرتز از پهنای باند نیاز است.

در تصویر 2 نحوه مدل کردن گفتار به وسیله پیشگویی خطی را مشاهده می کنید.



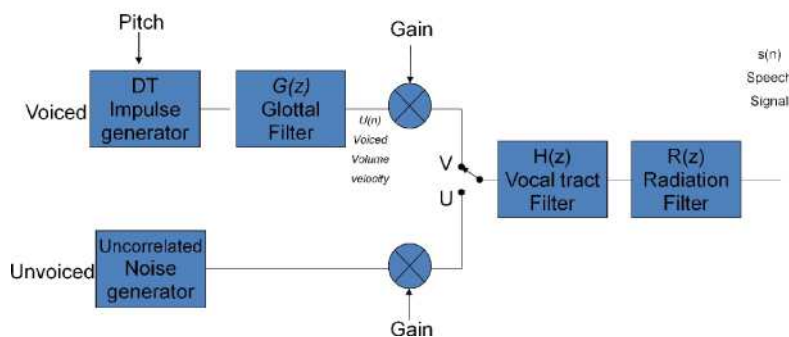
همان طور که در فص

نای بدون صدا هم

- تحریک:  $\zeta$
- از نویز به
- مسیر صوت

بیه تر به واقعیت

مدل ارائه شده در تص  
کرد. چنین مدلی را د



تصویر 3- مدل دقیق تر گفتار

بلوک های زیر اضافه شده اند:

- **Gain:** نشان دهنده بلندی و کمی صدا است و رابطه مستقیمی با انرژی دارد.
- **Glottal Filter:** تارهای صوتی انسان را مدل می کند. یعنی از ضربه ساده استفاده نمی کند بلکه به پالس تار صوتی انسان شکل طبیعی تر از ضربه می دهد.
- **Radiation Filter:** سعی می کند لب انسان را مدل کند. معمولاً این فیلتر به صورت مشتق گیر می باشد.

این بلوک ها به طبیعی تر شدن گفتار تولید شده کمک می کنند.

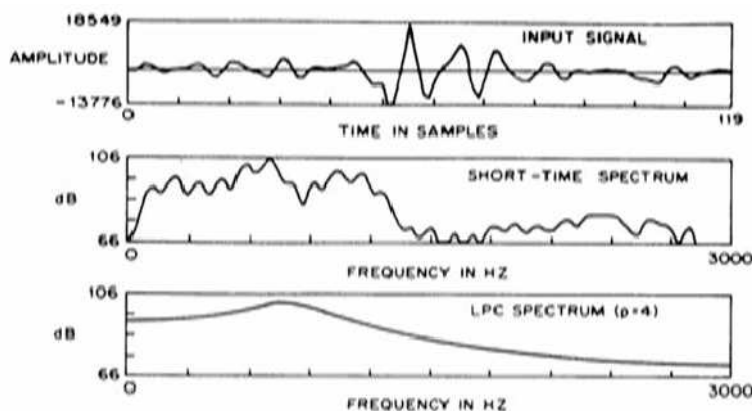
### 3- ضرایب پیشگویی خطی برای پردازش گفتار

مثالی از آنالیز LPC با 4 ضریب بر روی سیگنال گفتار را در تصویر 4 مشاهده می کنید.

همان طور از تصویر 4 وسط و پایین مشخص است، LPC پوش فرکانس را به دست می دهد.

به عبارتی روند کلی طیف سیگنال را مدل می کند ولی با تعداد پارامترهای خیلی کمتر (4 ضریب) در برابر نمونه 120 ورودی.

پس هم مدل سازی انجام داده ایم و هم فشرده سازی.



تصویر 4- نمونه ای از سیگنال در حوزه زمان (بالا) طیف فرکانسی (وسط) و پوش LPC (پایین)

در تصویر 5 سیگنال تصویر 4 با تعداد ضرایب مختلف مدل شده است.

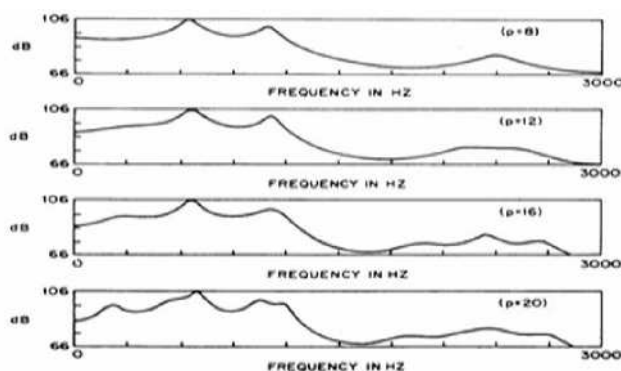
همان طور که مشخص است با تعداد ضرایب کم سیگنال خیلی صاف است و روند کلی را مدل می کند

ولی هرچه  $p$  افزایش می یابد جزئیات طیف سیگنال افزایش می یابد.

این طبیعی است زیرا هر چه  $p$  بیشتر می شود تعداد پارمترهای مدل بیشتر می شود و نتیجتاً مدل تصویر دقیق تری از واقعیت می شود.

البته افزایش تعداد  $p$  تا یک جایی مناسب است. زیرا بیش از یک حدی (حدود 20) تعداد پارامترها خیلی زیاد می شود و این برای پردازش های بعدی نامناسب است.

زیرا اطلاعات اضافی (بالاپایین شدن های ریز) هم وارد مدل می شود که طبیعتاً اطلاعات جدیدی به مدل اضافه نمی کنند و فقط پیچیدگی مدل را افزایش می دهند.



تصویر 5- اثر افزایش تعداد ضرایب پیش بینی کننده ( $p$ )

## 5- خلاصه و نتیجه گیری

در این فصل با مفهوم ضرایب پیشگویی خطی آشنا شدیم.

در این فصل با نحوه محاسبه ضرایب پیشگویی خطی آشنا شدیم.

در این فصل با نحوه محاسبه ضرایب کپسترال پیشگویی خطی آشنا شدیم.

## 6- منابع درس

- 1- Rabiner, "Fundamentals of Speech Recognition"
- 2- Huang, Acero, "Spoken Language Processing"
- 3- Deller, "Discrete-time processing of speech signals"

